

Cloud Search Service

User Guide

Issue 01
Date 2024-06-11



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.

Huawei Cloud Computing Technologies Co., Ltd.

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

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

Contents

1 Permissions Management.....	1
1.1 Creating a User and Granting Permissions.....	1
1.2 CSS Custom Policies.....	2
2 Elasticsearch.....	5
2.1 Creating a Cluster.....	5
2.1.1 Creating a Cluster in Security Mode.....	5
2.1.2 Creating a Cluster in Non-Security Mode.....	14
2.1.3 Cluster and Index Planning.....	21
2.1.4 Clusters in Security Mode.....	25
2.1.5 Changing the Billing Mode.....	28
2.1.6 Deploying a Cross-AZ Cluster.....	31
2.1.7 Creating a Cluster Using the Shared VPC.....	33
2.2 Importing Data.....	35
2.2.1 Using CDM to Import Data from OBS to Elasticsearch.....	35
2.2.2 Using DIS to Import Local Data to Elasticsearch.....	39
2.2.3 Using Logstash to Import Data to Elasticsearch.....	42
2.2.4 Using Kibana or APIs to Import Data to Elasticsearch.....	51
2.3 Migrating a Cluster Using Backup and Restoration.....	55
2.4 Accessing Elasticsearch Clusters.....	57
2.4.1 Accessing an Elasticsearch Cluster.....	57
2.4.2 Accessing a Cluster from a Public Network.....	58
2.4.3 Accessing a Cluster Using a VPC Endpoint.....	61
2.4.4 (Optional) Interconnecting with a Dedicated Load Balancer.....	64
2.4.4.1 Scenario.....	64
2.4.4.2 Connecting to a Dedicated Load Balancer.....	66
2.4.4.3 Sample Code for Two-Way Authentication During the Access to a Cluster.....	73
2.5 Index Backup and Restoration.....	76
2.5.1 Backup and Restoration Overview.....	76
2.5.2 Managing Automatic Snapshot Creation.....	77
2.5.3 Manually Creating a Snapshot.....	80
2.5.4 Restoring Data.....	83
2.5.5 Deleting a Snapshot.....	85
2.6 Cluster Specification Modification.....	85

2.6.1 Overview.....	86
2.6.2 Scaling Out a Cluster.....	87
2.6.3 Changing Specifications.....	89
2.6.4 Scaling in a Cluster.....	90
2.6.5 Removing Specified Nodes.....	93
2.6.6 Replacing a Specified Node.....	94
2.6.7 Adding Master/Client Nodes.....	96
2.6.8 Changing the Security Mode.....	97
2.6.9 Changing AZs.....	101
2.7 Upgrading the Cluster Version.....	104
2.8 Cluster Management.....	110
2.8.1 Cluster List Overview.....	110
2.8.2 Viewing Basic Cluster Information.....	111
2.8.3 Managing Tags.....	115
2.8.4 Managing Logs.....	117
2.8.5 Configuring YML Parameters.....	120
2.8.6 Viewing the Default Plugin List.....	121
2.8.7 Binding an Enterprise Project.....	122
2.8.8 Restarting a Cluster.....	124
2.8.9 Deleting a Cluster.....	125
2.9 Customizing Word Dictionaries.....	125
2.9.1 Managing Word Dictionaries.....	125
2.9.2 Example.....	129
2.10 Converting Between Simplified and Traditional Chinese (Using the Conversion Plugin).....	136
2.11 Using the Open Distro SQL Plug-in to Compile Queries.....	138
2.12 Using the Open Distro Alarm Plug-in to Configure SMN Alarms.....	143
2.12.1 (Optional) Authorizing CSS to Use SMN.....	143
2.12.2 Configuring SMN Alarms.....	143
2.13 Switching Hot and Cold Data.....	148
2.14 Managing Indexes.....	149
2.14.1 Creating and Managing Indexes.....	149
2.14.2 Changing Policies.....	152
2.15 Intelligent O&M.....	152
2.15.1 Overview of Intelligent O&M.....	152
2.15.2 Creating a Scan Task.....	153
2.15.3 Viewing Cluster Risk Items.....	154
2.15.4 Deleting a Scan Task.....	155
2.16 Kibana Platform.....	155
2.16.1 Logging In to Kibana.....	156
2.16.2 Accessing a Cluster from a Kibana Public Network.....	157
2.16.3 Creating a User and Granting Permissions by Using Kibana.....	160
2.16.4 Connecting User-Built Kibana to an Elasticsearch Cluster.....	167

3 Logstash	169
3.1 Creating a Cluster	169
3.2 Creating a Cluster in a Shared VPC	173
3.3 Configuring a Cluster	175
3.3.1 Configuration Center	175
3.3.2 Example Logstash Configuration File	179
3.3.3 Parameters for Configuring a System Template	181
3.4 Changing Cluster Configurations	188
3.4.1 Scaling Out a Cluster	188
3.4.2 Scaling in a Cluster	190
3.5 Viewing Basic Cluster Information	192
3.6 Managing Tags	196
3.7 Binding an Enterprise Project	197
3.8 Forcibly Restarting VMs in a Cluster	199
3.9 Deleting a Cluster	199
3.10 Managing Logs	200
3.11 Managing Certificates	202
4 OpenSearch	204
4.1 Creating a Cluster	204
4.2 Creating a Cluster in a Shared VPC	212
4.3 Accessing a Cluster	214
4.3.1 Quickly Accessing an OpenSearch Cluster	214
4.3.2 Accessing a Cluster from a Public Network	215
4.3.3 Accessing a Cluster Using a VPC Endpoint	218
4.3.4 (Optional) Interconnecting with a Dedicated Load Balancer	221
4.3.4.1 Scenario Description	221
4.3.4.2 Connecting to a Dedicated Load Balancer	223
4.3.4.3 Sample Code for Two-Way Authentication During the Access to a Cluster	230
4.4 Index Backup and Restoration	233
4.4.1 Backup and Restoration Overview	233
4.4.2 Managing Automatic Snapshot Creation	234
4.4.3 Manually Creating a Snapshot	237
4.4.4 Restoring Data	240
4.4.5 Deleting a Snapshot	242
4.5 Scaling In/Out a Cluster	243
4.5.1 Overview	243
4.5.2 Scaling Out a Cluster	244
4.5.3 Changing Specifications	246
4.5.4 Scaling in a Cluster	248
4.5.5 Removing Specified Nodes	250
4.5.6 Replacing a Specified Node	252
4.5.7 Adding Master/Client Nodes	253

4.5.8 Changing the Security Mode.....	254
4.5.9 Changing AZs.....	258
4.6 Managing Clusters.....	261
4.6.1 Viewing Basic Information About an Opensearch Cluster.....	261
4.6.2 Managing Tags.....	264
4.6.3 Managing Logs.....	266
4.6.4 Configuring YML Parameters.....	269
4.6.5 Viewing the Default Plugin List.....	271
4.6.6 Binding an Enterprise Project.....	272
4.6.7 Restarting a Cluster.....	274
4.6.8 Deleting a Cluster.....	275
4.7 Customizing Word Dictionaries.....	275
4.7.1 Managing Word Dictionaries.....	275
4.7.2 Example.....	279
4.8 Converting Between Simplified and Traditional Chinese (Using the Conversion Plugin).....	283
4.9 Configuring SMN Alarms.....	285
4.10 Switching Hot and Cold Data.....	291
4.11 Managing Indexes.....	292
4.11.1 Creating and Managing Index Policies.....	292
4.11.2 Changing an Index Policy.....	294
4.12 Intelligent O&M.....	296
4.12.1 Overview of Intelligent O&M.....	296
4.12.2 Creating a Scan Task.....	296
4.12.3 Viewing Cluster Risk Items.....	297
4.12.4 Deleting a Scan Task.....	298
4.13 OpenSearch Dashboards.....	299
4.13.1 Logging In to the OpenSearch Dashboards.....	299
4.13.2 Accessing a Cluster from a Kibana Public Network.....	300
4.13.3 Creating and Authorizing a User on the OpenSearch Dashboards.....	303
5 Viewing the Cluster Runtime Status and Storage Capacity Status.....	309
6 Enhanced Cluster Features.....	311
6.1 Vector Retrieval.....	311
6.1.1 Description.....	311
6.1.2 Cluster Planning for Vector Retrieval.....	312
6.1.3 Creating a Vector Index.....	313
6.1.4 Querying Vectors.....	320
6.1.5 Optimizing the Performance of Vector Retrieval.....	324
6.1.6 (Optional) Pre-Building and Registering a Center Point Vector.....	325
6.1.7 Managing the Vector Index Cache.....	327
6.1.8 Sample Code for Vector Search on a Client.....	327
6.1.9 Using PV_GRAPH to Search for Vector Indexes.....	330
6.2 Storage-Compute Decoupling.....	333

6.2.1 Context.....	333
6.2.2 Freezing an Index.....	333
6.2.3 Configuring Cache.....	341
6.2.4 Enhanced Cold Data Query Performance.....	343
6.2.5 Monitoring OBS Operations.....	346
6.3 Enhanced Import Performance.....	348
6.3.1 Context.....	348
6.3.2 Instructions.....	349
6.3.2.1 Bulk Route Optimization.....	349
6.3.2.2 Bulk Aggregation Optimization.....	350
6.3.2.3 Text Index Acceleration.....	350
6.3.2.4 Optimization of Other Parameters.....	351
6.3.3 Performance Data.....	351
6.4 Flow Control 2.0.....	352
6.4.1 Context.....	352
6.4.2 HTTP/HTTPS Flow Control.....	353
6.4.3 Memory Flow Control.....	355
6.4.4 Request Sampling.....	357
6.4.5 One-click Traffic Blocking.....	358
6.4.6 Access Statistics and Traffic Control Information Query.....	359
6.4.7 Temporary Access Statistics Logs.....	360
6.4.8 Recording Access Logs in Files.....	363
6.5 Flow Control 1.0.....	363
6.5.1 Context.....	363
6.5.2 HTTP/HTTPS Flow Control.....	365
6.5.3 Memory Flow Control.....	366
6.5.4 Global Path Whitelist for Flow Control.....	370
6.5.5 Request Sampling.....	371
6.5.6 Flow Control.....	372
6.5.7 Access Logs.....	375
6.5.8 CPU Flow Control.....	378
6.5.9 One-click Traffic Blocking.....	380
6.6 Large Query Isolation.....	380
6.6.1 Context.....	380
6.6.2 Procedure.....	380
6.7 Index Monitoring.....	385
6.7.1 Context.....	385
6.7.2 Enabling Index Monitoring.....	385
6.7.3 Checking the Index Read and Write Traffic.....	386
6.7.4 Checking Index Monitoring Information.....	388
6.7.5 kibana-monitor.....	392
6.8 Enhanced Cluster Monitoring.....	396

6.8.1 P99 Latency Monitoring.....	396
6.8.2 HTTP Status Code Monitoring.....	397
6.9 Enhanced Aggregation.....	399
6.9.1 Features.....	399
6.9.2 Grouping and Aggregation of Low-cardinality Fields.....	399
6.9.3 High-cardinality Field Histogram Aggregation.....	400
6.9.4 Low-cardinality and High-cardinality Field Mixing.....	401
6.10 Read/Write Splitting.....	402
6.10.1 Features.....	402
6.10.2 Instructions.....	403
6.10.2.1 Basic Settings.....	403
6.10.2.2 Synchronizing Specified Indexes.....	404
6.10.2.3 Matching Index Synchronization.....	405
6.10.2.4 Stopping Index Synchronization.....	406
6.10.2.5 Other Management APIs.....	406
6.10.3 Best Practices.....	408
7 Monitoring.....	409
7.1 Monitoring Metrics of Elasticsearch & OpenSearch Clusters.....	409
7.2 Monitoring Metrics of Logstash Clusters.....	434
7.3 Monitoring Metrics.....	438
7.4 Configuring Cluster Monitoring.....	446
8 Auditing.....	449
8.1 Key Operations Recorded by CTS.....	449
8.2 Querying Real-Time Traces.....	450

1 Permissions Management

1.1 Creating a User and Granting Permissions

This section describes how to use a group to grant permissions to a user. [Figure 1-1](#) shows the process for granting permissions.

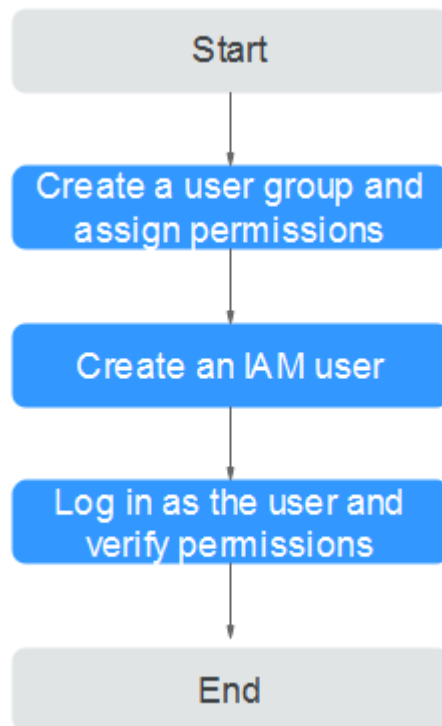
CSS has two types of user permissions: CSS administrator permission and read-only permission.

Prerequisites

Before assigning permissions to user groups, you have learned about the system policies listed in [Permissions Management](#).

Process Flow

Figure 1-1 Process of granting CSS permissions



1. **Create a user group and assign permissions.**
Create a user group on the IAM console, and assign the CSS permission to the group.
2. **Create an IAM user and add it to a user group.**
Create a user on the IAM console and add it to the user group created in 1.
3. **Log in** and verify permissions.
Log in to the console as the created user, switch to the authorized region, and verify the permissions.
 - Choose **Service List** > **Cloud Search Service**. Then click **Create Cluster** on the CSS console. If the cluster cannot be purchased (assuming that the current permissions include only **CSS ReadOnlyAccess**), the **CSS ReadOnlyAccess** policy has already taken effect.
 - Choose any other service from **Service List**. (Assume that the current policy contains only **CSS ReadOnlyAccess**.) If a message appears indicating insufficient permissions to access the service, the **CSS ReadOnlyAccess** policy has already taken effect.

1.2 CSS Custom Policies

Custom policies can be created to supplement the system-defined policies of CSS. For the actions supported for custom policies, see [Permissions Policies and Supported Actions](#).

You can create custom policies in either of the following ways:

- Visual editor: Select cloud services, actions, resources, and request conditions. You do not need to have knowledge of the policy syntax.
- JSON: Create a JSON policy or edit based on an existing policy.

For details about how to create custom policies, see [Creating a Custom Policy](#). The following section provides examples of common CSS custom policies.

NOTE

To be compatible with the open-source ecosystem, IAM permissions and data plane cluster permissions are managed separately. To enable the security capability of the data plane, you need to use the security mode.

Example Custom Policies

NOTE

To let an IAM user access an OBS bucket, you need to grant the **GetBucketStoragePolicy**, **GetBucketLocation**, **ListBucket**, and **ListAllMyBuckets** permissions to the user.

Example 1: Allowing users to create a CSS cluster

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "css:cluster:create",
        "vpc:securityGroups:get",
        "vpc:securityGroups:create",
        "vpc:securityGroups:delete",
        "vpc:securityGroupRules:get",
        "vpc:securityGroupRules:create",
        "vpc:securityGroupRules:delete",
        "vpc:vpcs:list",
        "vpc:privatelps:list",
        "vpc:ports:get",
        "vpc:ports:create",
        "vpc:ports:update",
        "vpc:ports:delete",
        "vpc:quotas:list",
        "vpc:subnets:get",
        "ecs:cloudServerFlavors:get",
        "ecs:serverInterfaces:use",
        "ecs:cloudServers:addNics",
        "ecs:quotas:get",
        "evs:types:get",
        "evs:quotas:get"
      ],
      "Effect": "Allow"
    }
  ]
}
```

Example 2: Denying cluster deletion

A policy with only **Deny** permissions must be used in conjunction with other policies for it to take effect. If the permissions assigned to a user contain both **Allow** and **Deny**, the **Deny** permissions take precedence over the **Allow** permissions.

The following method can be used if you need to assign permissions of the **CSS Admin** policy to a user but you want to prevent the user from deleting clusters.

Create a custom policy for denying cluster deletion, and attach both policies to the group to which the user belongs. Then, the user can perform all operations on CSS except deleting clusters. The following is an example of a deny policy:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Deny",
      "Action": [
        "css:cluster:delete"
      ]
    }
  ]
}
```

Example 3: Defining permissions for multiple services in a policy

A custom policy can contain the actions of multiple services that are of the global or project-level type. The following is an example policy containing actions of multiple services:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "ecs:cloudServers:resize",
        "ecs:cloudServers:delete",
        "ecs:cloudServers:delete",
        "css:cluster:restart",
        "css:*:get*",
        "css:*:list*"
      ],
      "Effect": "Allow"
    }
  ]
}
```

2 Elasticsearch

2.1 Creating a Cluster

2.1.1 Creating a Cluster in Security Mode

This section describes how to create an Elasticsearch cluster in security mode.

 **NOTE**

Public IP address access and Kibana public access can be used only after security mode is enabled.

Context

- If you choose the pay-per-use or yearly/monthly billing mode, you can directly create a cluster.
- When creating a cluster, the number of nodes that can be added varies according to the node type. For details, see [Table 2-1](#).

Table 2-1 Number of nodes in different types

Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32

Node Type	Number
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

1. Log in to the CSS management console.
2. On the **Dashboard** page, click **Create Cluster** in the upper right corner. The **Create** page is displayed.
Alternatively, choose **Clusters > Elasticsearch** in the navigation tree on the left. Click **Create Cluster** in the upper right corner. The **Create** page is displayed.
3. Configure **Billing Mode** and **Required Duration**.

Table 2-2 Billing parameters

Parameter	Description
Billing Mode	<p>Select Yearly/Monthly or Pay-per-use.</p> <ul style="list-style-type: none"> • Yearly/monthly: You pay for the cluster by year or month, in advance. The service duration range is one month to three years. If you plan to use a cluster for more than nine months, you are advised to purchase a yearly package for a better price. A yearly package costs the same as a 10 monthly package. • Pay-per-use: You are billed by actual duration of use, with a billing cycle of one hour. For example, 58 minutes of usage will be rounded up to an hour and billed.
Required Duration	<p>The duration for which the purchased EIP will use. The duration must be specified if the Billing Mode is set to Yearly/Monthly.</p> <p>Configure automatic renewal as required.</p>

4. Specify **Region** and **AZ**.




Table 2-3 Parameter description for Region and AZ

Parameter	Description
Region	Select a region for the cluster from the drop-down list on the right.
AZ	<p>Select AZs associated with the cluster region.</p> <p>You can select a maximum of three AZs. For details, see Deploying a Cross-AZ Cluster.</p>

5. Configure basic cluster information.

Table 2-4 Description of basic parameters

Parameter	Description
Version	Select a cluster version from the drop-down list box.

Parameter	Description
Name	<p>Cluster name, which contains 4 to 32 characters. Only letters, numbers, hyphens (-), and underscores (_) are allowed and the value must start with a letter.</p> <p>NOTE After a cluster is created, you can modify the cluster name as required. Click the name of a cluster to be modified. On the displayed Basic Information page, click  next to the cluster name. After the modification is completed, click  to save the modification. If you want to cancel the modification, click .</p>

6. Configure cluster specifications.

Table 2-5 Parameter description

Parameter	Description
Nodes	<p>Number of nodes in a cluster. Select a number from 1 to 32. You are advised to configure three or more nodes to ensure high availability of the cluster.</p> <ul style="list-style-type: none"> • If neither a master node nor client node is enabled, the nodes specified by this parameter are used to serve as both the master node and client node. Nodes provide the cluster management, data storage, cluster access, and data analysis functions. To ensure data stability in the cluster, it is recommended that you set this parameter to a value no less than 3. • If only the master node function is enabled, nodes specified by this parameter are used to store data and provide functions of client nodes. • If both the master and client node functions are enabled, the nodes specified by this parameter are only used for storing data. • If only the client node function is enabled, nodes specified by this parameter are used to store data and provide functions of the master node.
CPU Architecture	The supported type is determined by the actual regional environment. You can select x86 or Kunpeng .
Node Specifications	Specifications of nodes in a cluster. You can select a specified specification based on your needs. Each cluster supports only one specification. For details, see ECS Types .

Parameter	Description
Node Storage Type	Select a storage type. Common I/O, High I/O, Ultra-high I/O, and Extreme SSD are supported. NOTE If the type of storage in use is not supported, the storage type is not displayed.
Node Storage Capacity	Storage space. Its value varies with node specifications. The node storage capacity must be a multiple of 20.
Master node	The master node manages all nodes in the cluster. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the master node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes. After enabling the master node, specify Node Specifications, Nodes, and Node Storage Type . The value of Nodes must be an odd number greater than or equal to 3. Up to nine nodes are supported. The value of Node Storage Capacity is fixed. You can select a storage type based on your needs.
Client node	The client node allows clients to access clusters and analyze data. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the client node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes. After enabling the client node, specify Node Specifications, Nodes and Node Storage Type . The value of Nodes ranges from 1 to 32. The value of Node Storage Capacity is fixed. You can select a storage type based on your needs.
Cold data node	The cold data node is used to store historical data, for which query responses can be returned in minutes. If you do not require a quick query response, store historical data on cold data nodes to reduce costs. After enabling cold data node, configure Node Specifications, Nodes, Node Storage Type, and Node Storage Capacity . The value of Nodes ranges from 1 to 32. Select Node Storage Type and Node Storage Capacity as required. After the cold data node is enabled, CSS automatically adds cold and hot tags to related nodes.

7. Set the enterprise project.

When creating a CSS cluster, you can bind an enterprise project to the cluster if you have enabled the enterprise project function. You can select an enterprise project created by the current user from the drop-down list on the right or click **View Enterprise Project** to go to the **Enterprise Project Management Service** page and create a new project or view existing projects.

8. Click **Next: Network**. Configure the cluster network.

Table 2-6 Network configuration parameters

Parameter	Description
VPC	<p>A VPC is a secure, isolated, and logical network environment.</p> <p>Select the target VPC. Click View VPC to enter the VPC management console and view the created or shared VPC names and IDs. If no VPCs are available, create one.</p> <p>NOTE The VPC must contain CIDRs. Otherwise, cluster creation will fail. By default, a VPC will contain CIDRs.</p>
Subnet	<p>A subnet provides dedicated network resources that are isolated from other networks, improving network security.</p> <p>Select the target subnet. You can access the VPC management console to view the existed or shared subnet names and IDs.</p>
Security Group	<p>A security group is a collection of access control rules for ECSs that have the same security protection requirements and are mutually trusted in a VPC. To view more details about the security group, click View Security Group.</p> <p>NOTE</p> <ul style="list-style-type: none"> • For cluster access purposes, ensure that the security group contains port 9200. • If your cluster version is 7.6.2 or later, ensure that all the ports used for communication between nodes in the same security group are allowed. If such settings cannot be configured, ensure at least the access to port 9300 is allowed. • After the port 9300 is enabled, if the cluster disk usage is high, delete expired data to release the disk storage space. For details, see How Do I Clear Expired Data to Release Storage Space?
Security Mode	<p>After the security mode is enabled, communication will be encrypted and authentication required for the cluster.</p> <ul style="list-style-type: none"> • The default administrator account is admin. • Set and confirm the Administrator Password. This password will be required when you access this cluster.

Parameter	Description
HTTPS Access	<p>HTTPS access can be enabled only after the security mode of the cluster is enabled. After HTTPS access is enabled, communication is encrypted when you access the cluster.</p> <p>NOTE A cluster in security mode uses HTTPS for communication and will provide deteriorated read performance when compared with a normal cluster using HTTP. Its performance may be 20% less than the performance of a normal HTTP cluster under high concurrency. If you want fast read performance and the permission provided by the security mode to isolate resources (such as indexes, documents, and fields), you can disable the HTTPS Access function. After HTTPS Access is disabled, HTTP protocol is used for cluster communication. In this case, data security cannot be ensured and public IP address cannot be used.</p>
Public IP Address	<p>If HTTPS Access is enabled, you can configure Public Network Access and obtain an IP address for public network access. This IP address can be used to access this security cluster through the public network. For details, see Accessing a Cluster from a Public Network.</p>

9. Click **Next: Advanced Settings**. Configure the automatic snapshot creation and other functions.
 - a. Configure **Cluster Snapshot**. Set basic configuration and snapshot configuration.

The cluster snapshot function is enabled by default. You can also disable this function as required. To store automatic snapshots in OBS, an agency will be created to access OBS. Additional cost will be incurred if snapshots are stored in standard storage.

Table 2-7 Cluster snapshot parameter

Parameter	Description
OBS bucket	<p>Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket.</p> <p>The created or existing OBS bucket must meet the following requirements:</p> <ul style="list-style-type: none"> ● Storage Class is Standard. ● Region must be the same as that of the created cluster.

Parameter	Description
Backup Path	<p>Storage path of the snapshot in the OBS bucket.</p> <p>The backup path configuration rules are as follows:</p> <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (.). • The backup path cannot contain more than 1,023 characters.
IAM Agency	<p>IAM agency authorized by the current account for CSS to access or maintain data stored in OBS. You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency.</p> <p>The created or existing IAM agency must meet the following requirements:</p> <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

Table 2-8 Automatic snapshot creation parameter

Parameter	Description
Snapshot Name Prefix	<p>The snapshot name prefix contains 1 to 32 characters and must start with a lowercase letter. Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. A snapshot name consists of a snapshot name prefix and a timestamp, for example, snapshot-1566921603720.</p>
Time Zone	<p>Time zone for the backup time, which cannot be changed. Specify Backup Started Time based on the time zone.</p>
Backup Start Time	<p>The time when the backup starts automatically every day. You can specify this parameter only in full hours, for example, 00:00 or 01:00. The value ranges from 00:00 to 23:00. Select a time from the drop-down list.</p>

Parameter	Description
Retention Period (days)	<p>The number of days that snapshots are retained in the OBS bucket. The value ranges from 1 to 90. You can specify this parameter as required. The system automatically deletes expired snapshots every hour at half past the hour.</p> <p>For example, if you set the automatic snapshot creation policy as shown in Figure 2-1, the system, at 00:30 35 days later, will automatically delete the automated snapshots that were created at 00:00.</p>

Figure 2-1 Setting parameters for automatic snapshot creation

Automatic Snapshot Creation

Snapshot Name Prefix ✕

Time Zone GMT+08:00

Backup Start Time ▾ ▾

Retained Snapshots ? - +

- b. Configure advanced settings for the cluster.
 - **Default:** The **VPC Endpoint Service**, **Kibana Public Access**, and **Tags** functions are disabled by default. You can manually enable these functions after the cluster is created.
 - **Custom:** You can enable the **VPC Endpoint Service**, **Kibana Public Access**, and **Tags** functions as required.

Table 2-9 Parameters for advanced settings

Parameter	Description
VPC Endpoint Service	<p>After enabling this function, you can obtain a private domain name for accessing the cluster in the same VPC. For details, see Accessing a Cluster Using a VPC Endpoint.</p> <p>NOTE The VPC endpoint service cannot be enabled for a shared VPC.</p>

Parameter	Description
Kibana Public Access	You can configure this parameter only when security mode is enabled for a cluster. After enabling this function, you can obtain a public IP address for accessing Kibana. For details, see Accessing a Cluster from a Kibana Public Network .
Tag	Adding tags to clusters can help you identify and manage your cluster resources. You can customize tags or use tags predefined by Tag Management Service (TMS). For details, see Managing Tags . If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.

- Click **Next: Confirm Configuration**. Check the configuration and click **Next** to create a cluster.
- Click **Back to Cluster List** to switch to the **Clusters** page. The cluster you created is listed on the displayed page and its status is **Creating**. If the cluster is successfully created, its status will change to **Available**.
If the cluster creation fails, create the cluster again.

Follow-up Operations

After an Elasticsearch cluster is created, you are advised to optimize the query performance of the cluster to improve efficiency by referring to [Cluster Performance Tuning](#).

2.1.2 Creating a Cluster in Non-Security Mode

This section describes how to create an Elasticsearch cluster in non-security mode.

Procedure

- Log in to the CSS management console.
- On the **Dashboard** page, click **Create Cluster** in the upper right corner. The **Create** page is displayed.
Alternatively, choose **Clusters > Elasticsearch** in the navigation tree on the left. Click **Create Cluster** in the upper right corner. The **Create** page is displayed.
- Configure **Billing Mode** and **Required Duration**.

Table 2-10 Billing parameters

Parameter	Description
Billing Mode	<p>Select Yearly/Monthly or Pay-per-use.</p> <ul style="list-style-type: none"> • Yearly/monthly: You pay for the cluster by year or month, in advance. The service duration range is one month to three years. If you plan to use a cluster for more than nine months, you are advised to purchase a yearly package for a better price. A yearly package costs the same as a 10 monthly package. • Pay-per-use: You are billed by actual duration of use, with a billing cycle of one hour. For example, 58 minutes of usage will be rounded up to an hour and billed.
Required Duration	<p>The duration for which the purchased EIP will use. The duration must be specified if the Billing Mode is set to Yearly/Monthly.</p> <p>Configure automatic renewal as required.</p>

4. Specify **Region** and **AZ**.




Table 2-11 Parameter description for Region and AZ

Parameter	Description
Region	Select a region for the cluster from the drop-down list on the right.
AZ	<p>Select AZs associated with the cluster region.</p> <p>You can select up to three general AZs. For details, see Deploying a Cross-AZ Cluster.</p>

5. Configure basic cluster information.

Table 2-12 Description of basic parameters

Parameter	Description
Version	Select a cluster version from the drop-down list box.

Parameter	Description
Name	<p>Cluster name, which contains 4 to 32 characters. Only letters, numbers, hyphens (-), and underscores (_) are allowed and the value must start with a letter.</p> <p>NOTE After a cluster is created, you can modify the cluster name as required. Click the name of a cluster to be modified. On the displayed Basic Information page, click  next to the cluster name. After the modification is completed, click  to save the modification. If you want to cancel the modification, click .</p>

6. Configure cluster specifications.

Table 2-13 Parameter description

Parameter	Description
Nodes	<p>Number of nodes in a cluster. Select a number from 1 to 32. You are advised to configure three or more nodes to ensure high availability of the cluster.</p> <ul style="list-style-type: none"> • If neither a master node nor client node is enabled, the nodes specified by this parameter are used to serve as both the master node and client node. Nodes provide the cluster management, data storage, cluster access, and data analysis functions. To ensure data stability in the cluster, it is recommended that you set this parameter to a value no less than 3. • If only the master node function is enabled, nodes specified by this parameter are used to store data and provide functions of client nodes. • If both the master and client node functions are enabled, the nodes specified by this parameter are only used for storing data. • If only the client node function is enabled, nodes specified by this parameter are used to store data and provide functions of the master node.
CPU Architecture	The supported type is determined by the actual regional environment. You can select x86 or Kunpeng .
Node Specifications	Specifications of nodes in a cluster. You can select a specified specification based on your needs. Each cluster supports only one specification. For details, see ECS Types .

Parameter	Description
Node Storage Type	<p>In the current version, the following options are available: Common I/O, High I/O, Extreme SSD, and Ultra-high I/O.</p> <p>NOTE If the type of storage in use is not supported, the storage type is not displayed.</p>
Node Storage Capacity	<p>Storage space. Its value varies with node specifications. The node storage capacity must be a multiple of 20.</p>
Master node	<p>The master node manages all nodes in the cluster. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the master node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes.</p> <p>After enabling the master node, specify Node Specifications, Nodes, and Node Storage Type. The value of Nodes must be an odd number greater than or equal to 3. Up to nine nodes are supported. The value of Node Storage Capacity is fixed. You can select a storage type based on your needs.</p>
Client node	<p>The client node allows clients to access clusters and analyze data. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the client node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes.</p> <p>After enabling the client node, specify Node Specifications, Nodes and Node Storage Type. The value of Nodes ranges from 1 to 32. The value of Node Storage Capacity is fixed. You can select a storage type based on your needs.</p>
Cold data node	<p>The cold data node is used to store historical data, for which query responses can be returned in minutes. If you do not require a quick query response, store historical data on cold data nodes to reduce costs.</p> <p>After enabling cold data node, configure Node Specifications, Nodes, Node Storage Type, and Node Storage Capacity. The value of Nodes ranges from 1 to 32. Select Node Storage Type and Node Storage Capacity as required.</p> <p>After the cold data node is enabled, CSS automatically adds cold and hot tags to related nodes.</p>

7. Set the enterprise project.

When creating a CSS cluster, you can bind an enterprise project to the cluster if you have enabled the enterprise project function. You can select an enterprise project created by the current user from the drop-down list on the right or click **View Project Management** to go to the **Enterprise Project Management** console and create a new project or view existing projects.

8. Set network specifications of the cluster.

Table 2-14 Parameter description

Parameter	Description
VPC	<p>A VPC is a secure, isolated, and logical network environment.</p> <p>Select the target VPC. Click View VPC to enter the VPC management console and view the created or shared VPC names and IDs. If no VPCs are available, create one.</p> <p>NOTE The VPC must contain CIDRs. Otherwise, cluster creation will fail. By default, a VPC will contain CIDRs.</p>
Subnet	<p>A subnet provides dedicated network resources that are isolated from other networks, improving network security.</p> <p>Select the target subnet. You can access the VPC management console to view the names and IDs of the existing subnets in the VPC.</p>
Security Group	<p>A security group is a collection of access control rules for ECSs that have the same security protection requirements and are mutually trusted in a VPC. To view more details about the security group, click View Security Group.</p> <p>NOTE</p> <ul style="list-style-type: none"> • For cluster access purposes, ensure that the security group contains port 9200. • If your cluster version is 7.6.2 or later, ensure that all the ports used for communication between nodes in the same security group are allowed. If such settings cannot be configured, ensure at least the access to port 9300 is allowed. • After the port 9300 is enabled, if the cluster disk usage is high, delete expired data to release the disk storage space. For details, see How Do I Clear Expired Data to Release Storage Space?
Security Mode	Security mode is disabled.

9. Click **Next: Configure Advanced Settings**. Configure the automatic snapshot creation and other functions.
 - a. Configure **Cluster Snapshot**. Set basic configuration and snapshot configuration.
The cluster snapshot function is enabled by default. You can also disable this function as required. To store automatic snapshots in OBS, an agency

will be created to access OBS. Additional cost will be incurred if snapshots are stored in standard storage.

Table 2-15 Cluster snapshot parameter

Parameter	Description
OBS bucket	<p>Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket.</p> <p>The created or existing OBS bucket must meet the following requirements:</p> <ul style="list-style-type: none"> • Storage Class is Standard. • Region must be the same as that of the created cluster.
Backup Path	<p>Storage path of the snapshot in the OBS bucket.</p> <p>The backup path configuration rules are as follows:</p> <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.
IAM Agency	<p>IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency.</p> <p>The created or existing IAM agency must meet the following requirements:</p> <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

Table 2-16 Automatic snapshot creation parameter

Parameter	Description
Snapshot Name Prefix	<p>The snapshot name prefix contains 1 to 32 characters and must start with a lowercase letter. Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. A snapshot name consists of a snapshot name prefix and a timestamp, for example, snapshot-1566921603720.</p>

Parameter	Description
Time Zone	Time zone for the backup time, which cannot be changed. Specify Backup Started Time based on the time zone.
Backup Start Time	The time when the backup starts automatically every day. You can specify this parameter only in full hours, for example, 00:00 or 01:00. The value ranges from 00:00 to 23:00. Select a time from the drop-down list.
Retention Period (days)	The number of days that snapshots are retained in the OBS bucket. The value ranges from 1 to 90. You can specify this parameter as required. The system automatically deletes expired snapshots every hour at half past the hour. For example, if you set the automatic snapshot creation policy as shown in Figure 2-2 , the system, at 00:30 35 days later, will automatically delete the automated snapshots that were created at 00:00.

Figure 2-2 Setting parameters for automatic snapshot creation

Automatic Snapshot Creation

Snapshot Name Prefix ✕

Time Zone GMT+08:00

Backup Start Time

Retained Snapshots ?

- b. Configure advanced settings for the cluster.
 - **Default:** The **VPC Endpoint Service**, **Kibana Public Access**, and **Tag** functions are disabled by default. You can manually enable these functions after the cluster is created.
 - **Custom:** You can enable the **VPC Endpoint Service** and **Tag** functions as required.

Table 2-17 Parameters for advanced settings

Parameter	Description
VPC Endpoint Service	<p>After enabling this function, you can obtain a private domain name for accessing the cluster in the same VPC. For details, see Accessing a Cluster Using a VPC Endpoint.</p> <p>NOTE The VPC endpoint service cannot be enabled for a shared VPC.</p>
Kibana Public Access	Clusters in non-security mode cannot access Kibana through the Internet.
Tag	<p>Adding tags to clusters can help you identify and manage your cluster resources. You can customize tags or use tags predefined by Tag Management Service (TMS). For details, see Managing Tags.</p> <p>If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.</p>

10. Click **Next: Confirm**. Check the configuration and click **Next** to create a cluster.
11. Click **Back to Cluster List** to switch to the **Clusters** page. The cluster you created is listed on the displayed page and its status is **Creating**. If the cluster is successfully created, its status will change to **Available**.
If the cluster creation fails, create the cluster again.

2.1.3 Cluster and Index Planning

In Cloud Search Service (CSS), you can select the cluster version, architecture, storage type, number of cluster nodes, storage capacity, and number of index shards. Configure them based on your service requirements for read and write requests, data storage and computing, and search and analytics.

You can configure the following specifications of a CSS cluster:

- [Cluster Version](#)
- [Cluster Architecture](#)
- [Storage Types](#)
- [Cluster Nodes](#)
- [Node Storage Capacity](#)
- [Number of Index Shards](#)

Cluster Version

CSS supports Elasticsearch versions 7.6.2 and 7.10.2. You are advised to select a version based on the following principles:

1. For a new Elasticsearch cluster, select version 7.10.2 or 7.6.2.
2. If Elasticsearch cluster migration and code reconstruction are required, select 7.10.2 or 7.6.2. Otherwise, you are advised to use the same version as the previous major version.

Cluster Architecture

CSS supports multiple architectures, such as read/write splitting, cold and hot isolation, decoupled storage and computing, role separation, and cross-AZ deployment. The following table shows their applicable scenarios.

Architecture	Scenario	Benefits
Read/Write splitting	Production services involving many read operations and only a few write operations. After data is written, it does not need to be accessed within 10s.	High concurrency, low latency
Cold and hot data separation	Log services that have low requirements on cold data query performance.	Low costs
Decoupled storage and compute	Log services that have low requirements on cold data query performance (10s+) and do not require cold data update. This architecture can be used together with the cold and hot data separation architecture to build three levels of storage: hot, warm, and cold.	Low costs
Separation of roles	A cluster that is large, has a large number of indexes, or is highly scalable.	High availability
Cross-AZ deployment	Production services that have high requirements on availability or use local disks.	High availability

Storage Types

CSS supports cloud and local disks.

- Cloud disk types include computing-intensive (CPU:memory = 1:2), general computing (CPU:memory = 1:4), and memory-optimized (CPU:memory = 1:8).
- Local disk types include disk-intensive (with HDDs attached) and ultra-high I/O (with SSDs attached).

The following table shows their applicable scenarios.

Table 2-18 Applicable scenarios of storage types

Type	Scenario
Computing-intensive	Recommended scenario: search from a small amount of data (less than 100 GB on a single node)
General computing	Common scenario: search and analysis when the data volume on a single node is in the range 100 GB to 1,000 GB, for example, medium-scale e-commerce search, social search, and log search
Memory-optimized	Common scenario: search and analysis when the data volume of a single node is in the range 100 GB to 2,000 GB Vector search: Large memory helps improve cluster performance and stability.
Disk-intensive	Logs: Cold data needs to be stored and updated, and the requirements on cold data query performance is low.
Ultra-high I/O - Kunpeng	Large-scale logs: hot data storage
Ultra-high I/O - x86	Large-scale search and analysis: High computing or disk I/O performance is required, such as public opinion analysis, patent search, and database acceleration.

Cluster Nodes

After the architecture and storage type of a CSS cluster are selected, determine the number of nodes in the cluster based on your performance requirements.

Table 2-19 Node quantity calculation methods

Type	Performance Baseline	Node Quantity Calculation Method	Example
Write node	For a node with a cloud disk, the write performance baseline of a single vCPU is 1 MB/s. For an ultra-high I/O node, the write performance baseline of a single vCPU is 1.5 MB/s.	Traffic during peak hours/Number of vCPUs on a single node/Write performance baseline of a single vCPU x Number of copies	If the peak write rate is 100 MB/s and a node has 16 vCPUs and 64 GB memory, 12 nodes (100/16/1 x 2) are required.

Type	Performance Baseline	Node Quantity Calculation Method	Example
Query node	The performance of the same node varies greatly in different scenarios. It is difficult to evaluate the performance baseline of a single node. The average query response time is used as the query performance baseline for calculation.	$\text{QPS} / (\text{Number of vCPUs on a single node} \times 3/2 / \text{Average query response time per second}) \times \text{Number of shards}$	If the query QPS is 1000, the average query response time is 100 ms, three index shards are planned, and a node has 16 vCPUs and 64 GB memory, about 12 nodes $(1000 / (16 \times 3/2 / 0.1) \times 3)$ are required.
Number of nodes	/	Number of nodes = Number of write nodes + Number of query nodes	Number of nodes = Number of write nodes + Number of query nodes = 24

If two clusters can achieve the same performance, you are advised to select the one using higher specifications and fewer nodes. For example, a cluster using 3 nodes with 32 vCPUs and 64 GB memory achieves the same performance as the one using 12 nodes with 8 vCPUs and 16 GB memory, but the former runs more stable and can be more easily scaled. For a high-specification cluster that reaches the performance bottleneck, you simply need to scale it out (by adding nodes); whereas for a low-specification cluster, you need to scale it up (by changing to higher specifications).

Node Storage Capacity

The disk space of each node in a CSS cluster is determined by multiple factors, such as the data volume, number of copies (often set to 1), data bloat rate, and disk space usage (often set to 70%). You can use the following formula to calculate the storage capacity of a cluster:

$$\text{Storage capacity} = \text{Source data} \times (1 + \text{Number of copies}) \times 1.25 \times (1 + \text{Reserved space}) \approx \text{Source data} \times 2 \times 1.25 \times 1.3 = \text{Source data} \times 3.25$$

Number of Index Shards

You are advised to plan the number of index shards in a CSS cluster based on the following principles:

1. The size of a single shard is in the range 10 GB to 50 GB.
2. A cluster has fewer than 30,000 shards.
3. It is recommended that 1 GB memory be used for 20 to 30 shards, and that a single node have no more than 1,000 shards.

- For a single index, it is recommended that the number of index shards be the same as or a multiple of the number of nodes.

2.1.4 Clusters in Security Mode

When creating an Elasticsearch cluster, you can enable the security mode for it. Identity authentication is required when users access a security cluster. You can also authorize and encrypt security clusters.

Identity Verification

To access a security cluster, you need to enter the username and password. The identity verification is required for the following two types of users:

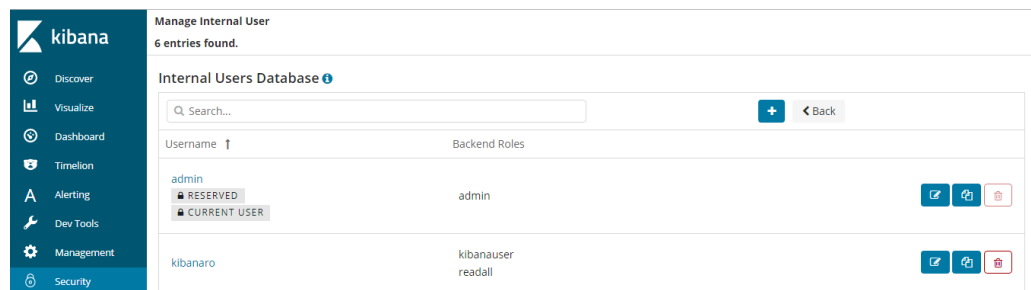
- Administrator: The default administrator username is **admin**, and the password is the one specified during cluster creation.
- Users: Enter the username and password created through Kibana.

Authorization

On the **Kibana** console, click **Security** to control user permissions in Elasticsearch clusters. You can configure hierarchical user permissions by cluster, index, document, and field. For details, see [Creating a User and Granting Permissions by Using Kibana](#).

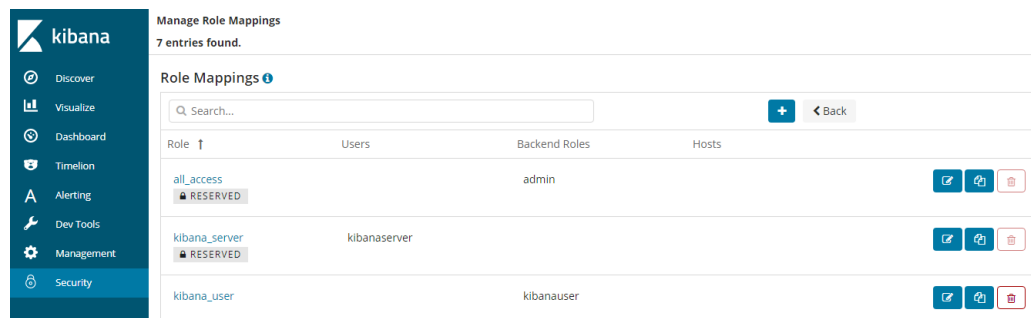
You can add or delete users, and map users to different roles for permissions control.

Figure 2-3 Configuring users



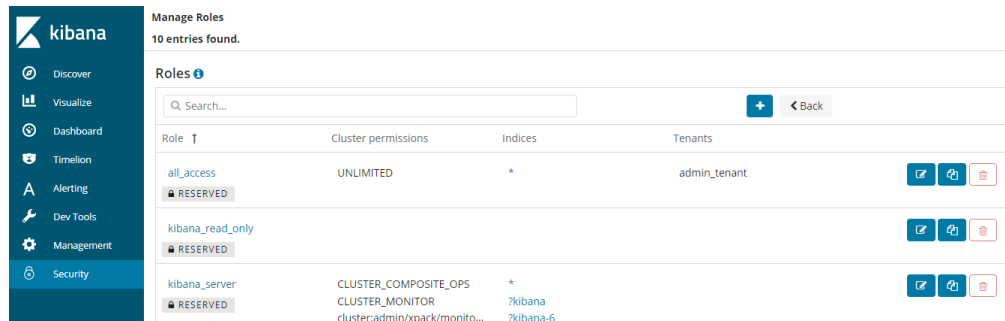
You can use role mapping to configure roles and map a user, backend role, and host name to a role.

Figure 2-4 Role mapping



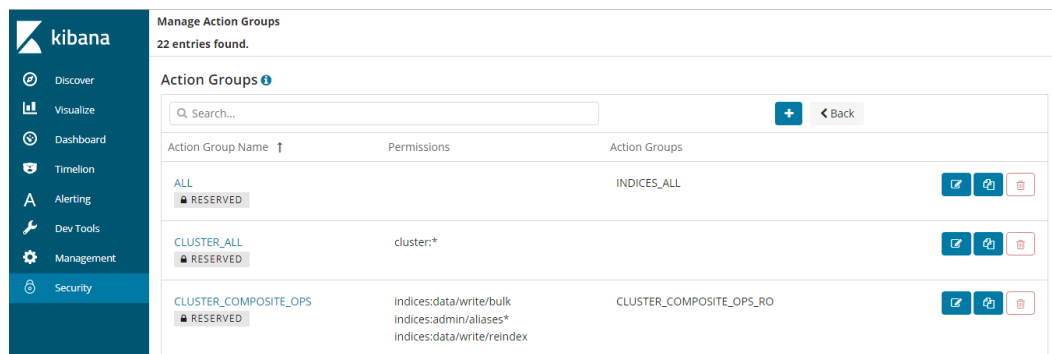
You can set permissions for each role to access clusters, indexes and documents and assign Kibana tenants different roles.

Figure 2-5 Configuring role permissions



You can set action groups, assign the groups to roles, and configure the roles' permission for accessing indexes and documents.

Figure 2-6 Configuring action groups



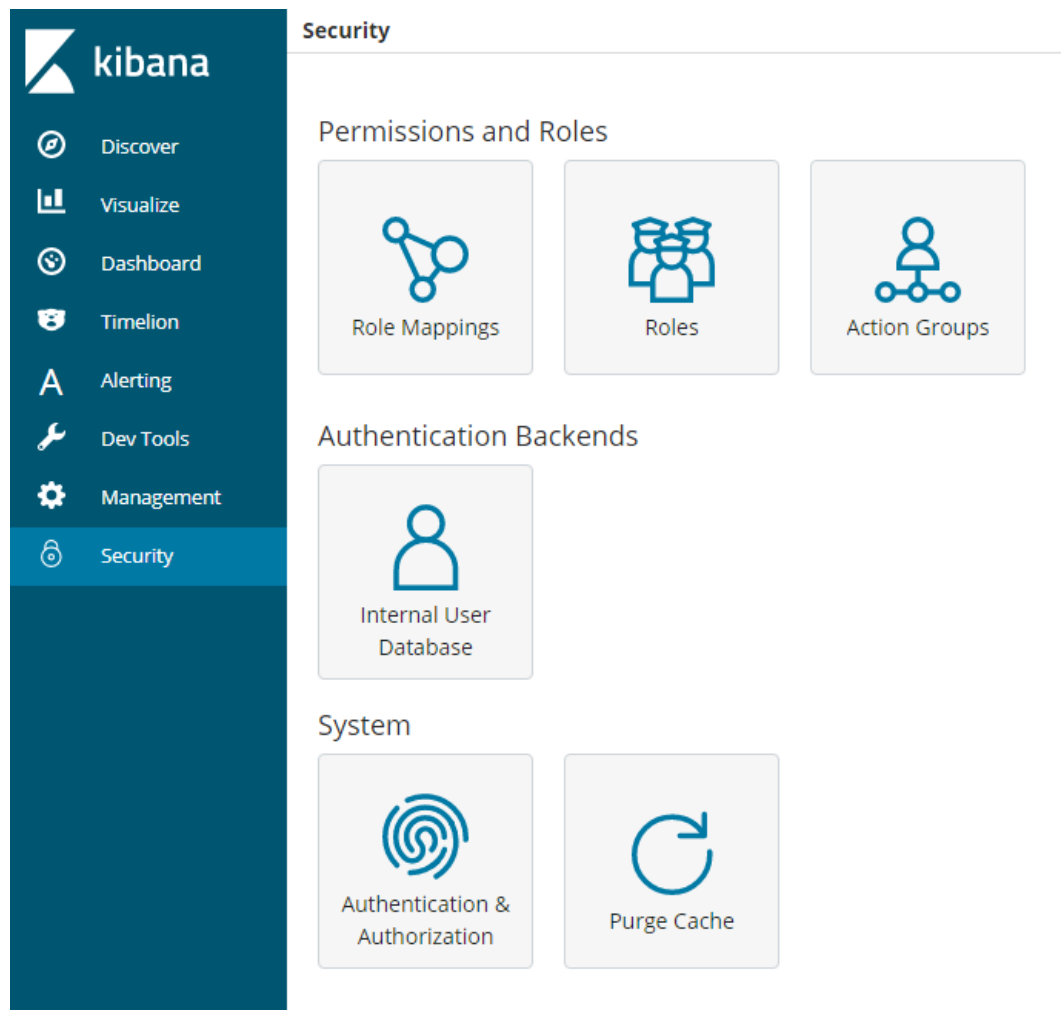
You can view the parameters of authentication and authorization for the current cluster. You can also run the **securityadmin** command to modify the configuration.

Figure 2-7 Viewing cluster parameters



You can also clear the security cache.

Figure 2-8 Clearing the security cache



Encryption

When key data is transferred between nodes or through the HTTP protocol, SSL/TLS encryption is used to ensure data security.

You can perform the preceding functions on Kibana, using **.yaml** files (not recommended), or by calling RESTful APIs. For more information about the security mode, see [Security](#).

Resetting the Administrator Password

If you want to change the administrator password of a security cluster or you have forgotten the password, reset the password.

1. On the **Clusters** page, locate the target cluster whose password you want to reset and click the cluster name. The **Cluster Information** page is displayed.
2. In the **Configuration** area, click **Reset** next to **Reset Password**.

 **NOTE**

- The password can contain 8 to 32 characters.
- The password must contain at least three of the following character types: uppercase letters, lowercase letters, digits, and special characters. The following special characters are supported: ~!@#\$\$%^&*()-_+=\|{};,:<.>/?
- Do not use the administrator name, or the administrator name spelled backwards.
- You are advised to change the password periodically.

Figure 2-9 Resetting passwords

Configuration	
Region	[blurred]
AZ	[blurred]
VPC	vpc- [blurred]
Subnet	subnet- [blurred]
Security Group	dws- [blurred] Change Security Group
Security Mode	Enabled
Reset Password	Reset
Enterprise Project	default
Public IP Address	100 [blurred] Disassociate
Access Control	Disabled Set
Bandwidth	22 Mbit/s Edit
HTTPS Access	Enabled Download Certificate
IPv4 Access Address	192. [blurred]

2.1.5 Changing the Billing Mode

CSS can be purchased in yearly/monthly mode. This mode is recommended for long-term users. The yearly/monthly billing mode can be configured during or after cluster creation.

 **NOTE**

If public access or Kibana public access is enabled for a yearly/monthly cluster, and the bandwidth is changed, you will be billed based on the new bandwidth. The yearly/monthly package cannot be unbound from the bandwidth before it expires. For details, see [Product Pricing Details](#).

Changing the Billing Mode from Pay-per-Use to Yearly/Monthly

1. On the **Clusters** page, select a cluster.
2. Choose **More > Switch to Yearly/Monthly** in the **Operation** column.

NOTE

If **Switch to Yearly/Monthly** is grayed out, you cannot change the billing mode in the current region.

3. On the **Change to Yearly/Monthly** page, configure **Required Duration** and **Auto-renew** as needed.
4. Click **Pay Now**. Confirm order details and click **Pay**.

Return to the **Clusters** page and check the current billing mode in the **Billing Mode** column.

Changing the Billing Mode from Yearly/Monthly to Pay-per-Use

1. On the **Clusters** page, select a cluster.
2. Choose **More > Change to Pay-per-Use Upon Expiration** in the **Operation** column.
3. In the **Change to Pay-per-Use Upon Expiration** dialog box, confirm the information and click **Yes**.

On the **Clusters** page and check the current billing mode in the **Billing Mode** column.

NOTE

- If the cluster billing mode is changed to pay-per-use during the yearly/monthly subscription period, the change will take effect when the yearly/monthly subscription expires.
- The billing mode cannot be changed for a cluster that has expired or been frozen.

Renewal

The yearly/monthly billing mode can be automatically or manually renewed. You can enable auto-renewal when you create a cluster. If auto-renewal is not enabled during cluster creation, you can manually renew a cluster. Perform the following steps:

1. On the **Clusters** page, select a cluster.
2. Choose **More > Renew** in the **Operation** column.
3. In the displayed dialog box, click **Yes** to switch to the **Renew** page.
On the **Renew** page, configure **Renewal Duration** and **Renewal Date**.
4. Click **Pay**. On the payment page, confirm order details and click **Pay**.

NOTE

A cluster that has expired or been frozen will be available after renewal.


Modifying Auto-Renew

Only yearly/monthly clusters can be renewed. If you use pay-per-use clusters, just ensure that your account has a valid payment method configured or a top-up account with a sufficient balance.

- Step 1** Log in to the CSS management console.
- Step 2** On the navigation pane in the left, select the target cluster type to view the cluster list.

For example, choose **Clusters > Elasticsearch** in the navigation pane to view the cluster list.
- Step 3** Select the cluster that you want to change auto-renew information and click **More > Change Auto-Renew** in the **Operation** column.
- Step 4** Confirm the cluster information in the dialog box and click **OK**.
- Step 5** In the **Modify Auto-Renew** page, modify the renew parameters by referring to [Table 2-20](#), and click **OK**.

Table 2-20 Modifying auto-renew

Parameter	Description
Instance Name/ID	Select the cluster that you want to modify auto-renew information. You can click  to view the cluster information.
Renewal Option	You can select Automatic and Manual .
New Auto-Renew Period	If the Renewal Option is set to Automatic , you need to set the New Auto-Renew Period based on your service requirements.
Auto-renewals	<p>If the Renewal Option is set to Automatic, you need to determine whether to reset the number of auto-renewal times.</p> <p>If you select Reset Auto-renewals, you also need to select Unlimited or Custom.</p> <ul style="list-style-type: none"> • Unlimited: The cluster can be automatically renewed for an unlimited number of times. • Custom: Enter a value from 1 to 1,000. The auto-renewal is disabled after the cluster is automatically renewed for the specified number of times.

----End

Unsubscription/Release

You can unsubscribe from a cluster billed on a yearly/monthly basis. The cluster will be released and all data will be permanently deleted.

1. On the **Clusters** page, select a cluster.

2. Choose **More > Unsubscribe/Release** in the **Operation** column.
3. In the text box, enter the name of the cluster to be unsubscribed from and click **Yes**.

On the **Unsubscribe** page, confirm the resource information and refund amount.

4. Select a reason for unsubscription, select the acknowledgement check boxes, and click **Unsubscribe**.

In the displayed confirmation dialog box, click **Yes**.

 **NOTE**

If the cluster is in **Available** state, an unsubscription order will be generated for refund, and the cluster will be deleted. If the cluster has expired or been frozen, the cluster will be directly deleted. For more information, see [Billing Center](#).

2.1.6 Deploying a Cross-AZ Cluster

To prevent data loss and minimize the cluster downtime in case of service interruption, CSS supports cross-AZ cluster deployment. When creating a cluster, you can select two or three AZs in the same region. The system will automatically allocate nodes to these AZs.

Allocating Nodes

If you select two or three AZs when creating a cluster, CSS automatically enables the cross-AZ HA function and properly allocates nodes to different AZs. [Table 2-21](#) describes how the nodes are allocated.

 **NOTE**

- When creating a cluster, ensure that the number of selected nodes is no less than the number of AZs. Otherwise, cross-AZ deployment is not supported.
- If you enable master nodes when deploying a cross-AZ cluster, the master nodes will also be distributed to different AZs.
- The node quantity difference between any two AZs is no more than one.

Table 2-21 Number of nodes and AZ distribution

Nodes	Single AZ	Two AZs		Three AZs		
	AZ1	AZ1	AZ2	AZ1	AZ2	AZ3
1	1	Not supported		Not supported		
2	2	1	1	Not supported		
3	3	2	1	1	1	1
4	4	2	2	2	1	1
...

Setting Replicas

Setting replicas enables clusters to effectively use the HA capability of AZs.

- In two-AZ deployment, if one AZ becomes unavailable, the other AZ continues to provide services. In this case, at least one replica is required. Elasticsearch has one replica by default. You can retain the default value if you do not require higher read performance.
- In three-AZ deployment, if one AZ becomes unavailable, the other AZs continue to provide services. In this case, at least one replica is required. Elasticsearch has one replica by default. If you need more replicas to improve the cluster's ability to handle queries, modify **settings** to change the number of replicas.

You can run the following command to modify the number of index replicas:

```
curl -XPUT http://ip:9200/{index_name}/_settings -d '{"number_of_replicas":2}'
```

Alternatively, run the following command to specify the number of replicas in the template:

```
curl -XPUT http://ip:9200/_template/templatename -d '{"template": "*", "settings": {"number_of_replicas": 2}}'
```

NOTE

- **ip**: private network address
- **index_name**: index name
- **number_of_replicas**: number of replicas after modification. The value in the preceding command indicates that two replicas are required.

Possible Service Interruptions

The following table describes the possible service interruptions when an AZ of a two- or three-AZ cluster is faulty.

Table 2-22 Possible service interruptions


AZs	Master Nodes	Service Interruption Analysis
2	0	<ul style="list-style-type: none"> • When the number of nodes is an even number: <ul style="list-style-type: none"> – If half of data nodes are faulty, replace one node in the faulty AZ before you select the master node. • When the number of nodes is an odd number: <ul style="list-style-type: none"> – If the faulty AZ contains one more node than the normal AZ, you need to replace one node in the faulty AZ before you select the master node. For details about how to replace nodes, contact technical support. – If the faulty AZ contains one less node than the normal AZ, services will not be interrupted and you can select the master node.

AZs	Master Nodes	Service Interruption Analysis
2	3	<p>There is a 50% possibility for service interruption. When two dedicated master nodes are allocated to one AZ and another master node is allocated to the other AZ:</p> <ul style="list-style-type: none"> • If service interruption happens in the AZ with one master node, you can select a master node from the AZ that has two dedicated master nodes. • If service interruption happens in the AZ with two dedicated master nodes, you have no choice in the remaining AZ, because it has only one dedicated master node. In this case, services will be interrupted and you need to contact technical support.
3	0	<p>If you configure four nodes in three AZs, each AZ will have at least one node. If the AZ with two nodes is faulty, the services will be interrupted. You are not advised configuring four nodes when selecting three AZs.</p> <p>Generally, service interruption will not occur.</p>
3	3	Service interruption does not occur.

2.1.7 Creating a Cluster Using the Shared VPC

A VPC subnet can be shared by multiple Huawei Cloud IAM accounts. You can create CSS clusters in a shared VPC subnet.

Step 1: Creating VPC Share

1. Log in to the [Huawei Cloud management console](#).
2. Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
3. Choose **Shared by Me > Resource Shares**.
4. Click **Create Resource Share** in the upper right corner.
5. On the displayed **Specify Resource Share Details** page, configure basic information and specify the subnet to be shared. Search for **vpc: subnet** and select the target subnet for sharing. Click **Next: Associate Permissions** in the lower right corner.

NOTE

When creating a resource share, you can specify up to 20 resources to share at a time. However, you can update the resource share you created to add more resources. For details, see [Updating a Resource Share](#).

6. On the **Associate Permissions** page, associate a RAM managed permission with each resource type, and then click **Next: Specify Principals** in the lower right corner.

RAM managed permissions available for your selection are system permissions predefined by RAM. Some resource types may have multiple

permissions available. You can select as needed. For the details of each permission, see [Viewing the RAM Permissions Library](#).

To create a CSS cluster in a shared VPC, you need to select the **default vpc subnet statement** permission.

7. On the **Grant Access to Principals** page, specify the principals that you want to have access to the resources, and then click **Next: Confirm** in the lower right corner.

In this step, you can select either **Allow sharing with any Huawei Cloud principal** or **Allow sharing only within your organization**. If you select the latter, choose any principals that are within your organization.

You can set **Principal Type** to **Organization** or **Huawei Cloud account ID**. The **Organization** option is available only when the toggle key **Sharing with Organizations** is turned on. For details, see [Enabling Sharing with Organizations](#).

8. Review and confirm the configuration details of your resource share and select **I have read and agree to Privacy Statement** on the **Confirm** page. Then, click **Submit** in the lower right corner.

After a resource share is created, RAM sends a sharing invitation to the specified principals. The principals can access and use the shared resources only after they accept the invitation. If the specified principals are within your organization and sharing with Organizations is enabled, the principals can access and use the shared resources without accepting the invitation.

 **NOTE**

Each principal can be shared with a maximum of 100 VPC subnets.

Step 2: Accepting VPC Share


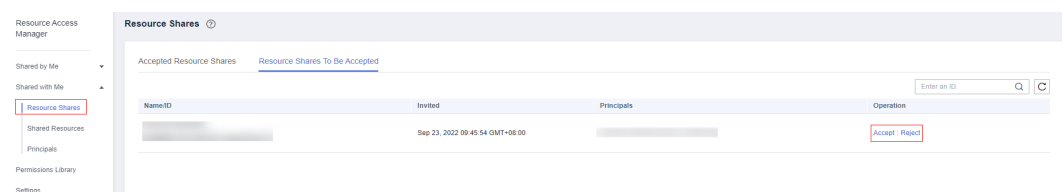
1. Log in to the [Huawei Cloud management console](#).
2. Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
3. Choose **Shared with Me > Resource Shares**.
4. Click the **Resource Shares To Be Accepted** tab, and select the resource share for which you are invited. Then, click **Accept** or **Reject** in the **Operation** column.

Figure 2-10 Responding to a resource sharing invitation



5. Click **OK** in the displayed dialog box.

After you accept invitations from certain resource shares, you can view them on the **Accepted Resource Shares** page. You can click a resource share name to view its configuration details.

 **NOTE**

Each principal can accept the invitations to resource shares involving a maximum of 100 VPC subnets.

Step 3: Creating a Cluster in the Shared VPC Subnet

1. Log in to the CSS console. In the navigation pane on the left, choose **Clusters** and select a cluster type.
For example, log in to the CSS console and choose **Clusters > Elasticsearch** in the navigation pane on the left.
2. On the **Clusters** page, click **Create Cluster**.
3. On the **Basic Configuration** page, configure the cluster parameters. For details, see [Creating an Elasticsearch Cluster](#), [Creating a Logstash Cluster](#), and [Creating an OpenSearch Cluster](#).

On the **Network Configuration** page, select the VPC and subnet that are shared with the current account for **VPC** and **Subnet** to create a cluster using the shared VPC.

- **VPC:** Select the name and ID of the VPC that is shared with the current account.
- **Subnet:** Select a subnet for your cluster. You can access the VPC service to view the shared subnet name and ID.

You can create a CSS cluster in the shared VPC subnet.

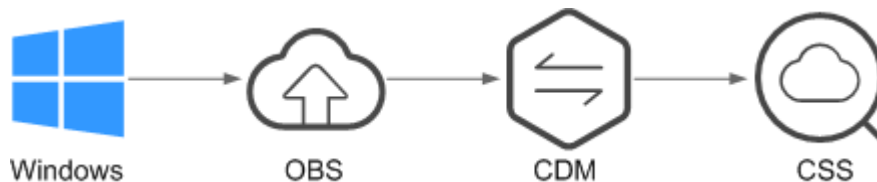
2.2 Importing Data

2.2.1 Using CDM to Import Data from OBS to Elasticsearch

You can use the CDM-provided wizard to import data stored in OBS to Elasticsearch in CSS. Data files can be in the JSON or CSV format.

[Figure 2-11](#) shows the data transmission process.

Figure 2-11 Process of using CDM to import OBS data to Elasticsearch



Procedure

1. Log in to the OBS management console.
2. Create an OBS bucket for storing data.
For details, see [Creating a Bucket](#) in the *Object Storage Service Console Operation Guide*.
The OBS bucket must be in the same region as the cluster.

3. Upload the data file to the OBS bucket.
For details, see [Uploading a File](#) in *Object Storage Service Console Operation Guide*.

For example, save the following data as a JSON file and upload the file to the OBS bucket.

```
{"productName":"Latest art shirts for women in autumn 2017","size":"L"}
{"productName":"Latest art shirts for women in autumn 2017","size":"M"}
{"productName":"Latest art shirts for women in autumn 2017","size":"S"}
{"productName":"Latest jeans for women in spring 2018","size":"M"}
{"productName":"Latest jeans for women in spring 2018","size":"S"}
{"productName":"Latest casual pants for women in spring 2017","size":"L"}
{"productName":"Latest casual pants for women in spring 2017","size":"S"}
```

4. Log in to the CSS management console.
5. In the navigation pane on the left, choose **Clusters > Elasticsearch** to switch to the **Clusters** page.
6. From the cluster list, locate the row that contains the cluster to which you want to import data, and click **Access Kibana** in the **Operation** column.
7. In the Kibana navigation pane on the left, choose **Dev Tools**.
8. Run the following command on the console to check whether the cluster has indexes:

```
GET _cat/indices?v
```

If there are indexes available in the cluster that you want to import data, you do not need to create an index. Go to step [10](#).

If there are no indexes available in the cluster that you want to import data, go to the next step and create an index.

9. On the **Console** page, run the related command to create an index for the data to be stored and specify a custom mapping to define the data type:

For example, on the **Console** page, run the following command to create index **demo** and specify a user-defined mapping to define the data type:

Versions earlier than 7.x

```
PUT /demo
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "products": {
      "properties": {
        "productName": {
          "type": "text",
          "analyzer": "ik_smart"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
}
```

Versions 7.x and later

```
PUT /demo
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "properties": {
```

```

"productName": {
  "type": "text",
  "analyzer": "ik_smart"
},
"size": {
  "type": "keyword"
}
}
}
}

```

The command is successfully executed if the following information is displayed.

```

{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "demo"
}

```

10. Log in to the CDM management console.
11. Purchase a CDM cluster.
For details, see [Creating a CDM Cluster](#) in *Cloud Data Migration User Guide*.
12. Create a link between CDM and CSS.
For details, see [Creating Links](#) in *Cloud Data Migration User Guide*.
13. Create a link between CDM and OBS.
For details, see [Creating Links](#) in *Cloud Data Migration User Guide*.
14. Create a job on the CDM cluster and migrate the data in the OBS bucket to the target cluster in CSS.
For details, see [Table/File Migration Jobs](#) in *Cloud Data Migration User Guide*.
15. On the **Console** page of Kibana, search for the imported data.
On the **Console** page of Kibana, run the following command to search for data. View the search results. If the searched data is consistent with the imported data, the data has been imported successfully.

```
GET demo/_search
```

The command is successfully executed if the following information is displayed.

```

{
  "took": 18,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 7,
    "max_score": 1,
    "hits": [
      {
        "_index": "demo",
        "_type": "products",
        "_id": "g6UepnEBuvdFwWkRmn4V",
        "_score": 1,
        "_source": {
          "size": "size:L",
          "productName": "Latest art shirts for women in autumn 2017"
        }
      }
    ]
  }
}

```

```

},
{
  "_index": "demo",
  "_type": "products",
  "_id": "hKUepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "M""""",
    "productName": """"{"productName": "Latest art shirts for women in autumn 2017""""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "haUepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "S""""",
    "productName": """"{"productName": "Latest art shirts for women in autumn 2017""""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "hqUepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "M""""",
    "productName": """"{"productName": "Latest jeans for women in autumn 2018""""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "h6UepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "S""""",
    "productName": """"{"productName": "Latest jeans for women in autumn 2018""""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "iKUepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "L""""",
    "productName": """"{"productName": "Latest casual pants for women in autumn 2017""""
  }
},
{
  "_index": "demo",
  "_type": "products",
  "_id": "iaUepnEBuudFwWkRmn4V",
  "_score": 1,
  "_source": {
    "size": """"size": "S""""",
    "productName": """"{"productName": "Latest casual pants for women in autumn 2017""""
  }
}
]
}
}

```

 **NOTE**

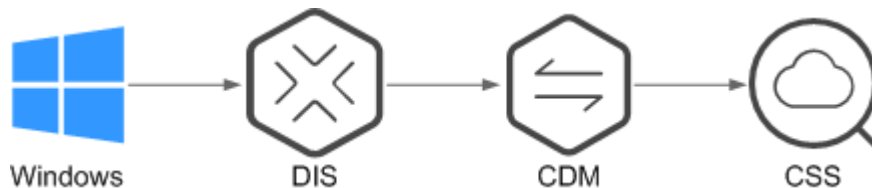
demo specifies the created index name. Set this parameter based on site requirements.

2.2.2 Using DIS to Import Local Data to Elasticsearch

You can use DIS to upload log data stored on the local Windows PC to the DIS queue and use CDM to migrate the data to Elasticsearch in CSS. In this way, you can efficiently manage and obtain logs through Elasticsearch. Data files can be in the JSON or CSV format.

[Figure 2-12](#) shows the data transmission process.

Figure 2-12 Process of using DIS to import local data to Elasticsearch



Procedure

1. Log in to the DIS management console.
2. Purchase a DIS stream.
For details, see [Creating a DIS Stream](#) in *Data Ingestion Service User Guide*.
3. Install and configure DIS Agent.
For details, see [Installing DIS Agent](#) and [Configuring DIS Agent](#) in *Data Ingestion Service User Guide*.
4. Start DIS Agent and upload the collected local data to the DIS queue.
For details, see [Starting DIS Agent](#) in the *Data Ingestion Service User Guide*.
For example, upload the following data to a DIS queue using the DIS Agent:

```
{"logName":"aaa","date":"bbb"}  
{"logName":"ccc","date":"ddd"}  
{"logName":"eee","date":"fff"}  
{"logName":"ggg","date":"hhh"}  
{"logName":"mmm","date":"nnn"}
```

5. Log in to the CSS management console.
6. In the navigation pane on the left, choose **Clusters > Elasticsearch** to switch to the **Clusters** page.
7. From the cluster list, locate the row that contains the cluster to which you want to import data, and click **Access Kibana** in the **Operation** column.
8. In the Kibana navigation pane on the left, choose **Dev Tools**.
9. Run the following command on the console to check whether the cluster has indexes:

```
GET _cat/indices?v
```


If there are indexes available in the cluster that you want to import data, you do not need to create an index. Go to step **11**.
If there are no indexes available in the cluster that you want to import data, go to the next step and create an index.
10. On the **Console** page, run the related command to create an index for the data to be stored and specify a custom mapping to define the data type:
For example, on the **Console** page, run the following command to create index **apache** and specify a custom mapping to define the data type:

Versions earlier than 7.x

```
PUT /apache
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "logs": {
      "properties": {
        "logName": {
          "type": "text",
          "analyzer": "ik_smart"
        },
        "date": {
          "type": "keyword"
        }
      }
    }
  }
}
```

Versions 7.x and later

```
PUT /apache
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "properties": {
      "logName": {
        "type": "text",
        "analyzer": "ik_smart"
      },
      "date": {
        "type": "keyword"
      }
    }
  }
}
```

The command is successfully executed if the following information is displayed.

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "apache"
}
```

11. Log in to the CDM management console.
12. Purchase a CDM cluster.
For details, see [Creating a CDM Cluster](#) in the *Cloud Data Migration User Guide*.
13. Create a link between CDM and CSS.
For details, see [Creating Links](#) in *Cloud Data Migration User Guide*.
14. Create a link between CDM and DIS.
For details, see [Creating Links](#) in *Cloud Data Migration User Guide*.
15. Create a job on the purchased CDM cluster and migrate the data in the DIS queue to the target cluster in CSS.
For details, see [Table/File Migration Jobs](#) in *Cloud Data Migration User Guide*.
16. On the **Console** page of Kibana, search for the imported data.

On the **Console** page of Kibana, run the following command to search for data. View the search results. If the searched data is consistent with the imported data, the data has been imported successfully.

```
GET apache/_search
```

The command is successfully executed if the following information is displayed.

```
{
  "took": 81,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 5,
    "max_score": 1,
    "hits": [
      {
        "_index": "apache",
        "_type": "logs",
        "_id": "txfbqnEBPuwwWJWL-qvP",
        "_score": 1,
        "_source": {
          "date": """"{"logName": "aaa"}""",
          "logName": """"date": "bbb"}""
        }
      },
      {
        "_index": "apache",
        "_type": "logs",
        "_id": "uBfbqnEBPuwwWJWL-qvP",
        "_score": 1,
        "_source": {
          "date": """"{"logName": "ccc"}""",
          "logName": """"date": "ddd"}""
        }
      },
      {
        "_index": "apache",
        "_type": "logs",
        "_id": "uRfbqnEBPuwwWJWL-qvP",
        "_score": 1,
        "_source": {
          "date": """"{"logName": "eee"}""",
          "logName": """"date": "fff"}""
        }
      },
      {
        "_index": "apache",
        "_type": "logs",
        "_id": "uhfbqnEBPuwwWJWL-qvP",
        "_score": 1,
        "_source": {
          "date": """"{"logName": "ggg"}""",
          "logName": """"date": "hhh"}""
        }
      },
      {
        "_index": "apache",
        "_type": "logs",
        "_id": "uxfbqnEBPuwwWJWL-qvP",
        "_score": 1,
        "_source": {
          "date": """"{"logName": "mmm"}""",
          "logName": """"date": "nnn"}""
        }
      }
    ]
  }
}
```

```
}  
}  
]  
}  
}
```

 NOTE

apache specifies the created index name. Set this parameter based on site requirements.

2.2.3 Using Logstash to Import Data to Elasticsearch

You can use Logstash to collect data and migrate collected data to Elasticsearch in CSS. This method helps you effectively obtain and manage data through Elasticsearch. Data files can be in the JSON or CSV format.

Logstash is an open-source, server-side data processing pipeline that ingests data from multiple sources simultaneously, transforms data, and then sends data to Elasticsearch. For details about Logstash, visit the following website: <https://www.elastic.co/guide/en/logstash/current/getting-started-with-logstash.html>

The following two scenarios are involved depending on the Logstash deployment:

- [Importing Data When Logstash Is Deployed on the External Network](#)
- [Importing Data When Logstash Is Deployed on an ECS](#)

Prerequisites

- To facilitate operations, you are advised to deploy Logstash on a host that runs the Linux operating system (OS).
- To download Logstash, visit the following website: <https://www.elastic.co/downloads/logstash-oss>

 NOTE

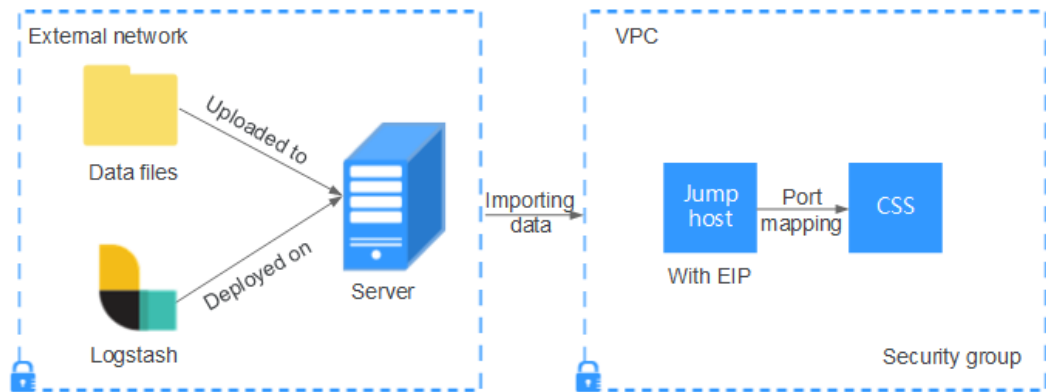
Logstash requires an OSS version same as the CSS version.

- After installing Logstash, perform the following steps to import data. For details about how to install Logstash, visit the following website: <https://www.elastic.co/guide/en/logstash/current/installing-logstash.html>
- The JDK must be installed before Logstash is installed. In Linux OS, you can run the **yum -y install java-1.8.0** command to install JDK 1.8.0. In Windows OS, you can download the required JDK version from the [official website of JDK](#), and install it by following the installation guide.
- In the [Importing Data When Logstash Is Deployed on an ECS](#) scenario, ensure that the ECS and the Elasticsearch cluster to which data is imported reside in the same VPC.

Importing Data When Logstash Is Deployed on the External Network

Figure 2-13 illustrates how data is imported when Logstash is deployed on an external network.

Figure 2-13 Importing data when Logstash is deployed on an external network



1. Create a jump host and configure it as follows:
 - The jump host is an ECS running the Linux OS and has been bound with an EIP.
 - The jump host resides in the same VPC as the CSS cluster.
 - SSH local port forwarding is configured for the jump host to forward requests from a chosen local port to port **9200** on one node of the CSS cluster.
 - Refer to [SSH documentation](#) for the local port forwarding configuration.
2. Use PuTTY to log in to the created jump host with the EIP.
3. Run the following command to perform port mapping and transfer the request sent to the port on the jump host to the target cluster:


```
ssh -g -L <Local port of the jump host:Private network address and port number of a node> -N -f root@<Private IP address of the jump host>
```

NOTE

- In the preceding command, *<Local port of the jump host>* refers to the port obtained in **1**.
- In the preceding command, *<Private network address and port number of a node>* refers to the private network address and port number of a node in the cluster. If the node is faulty, the command execution will fail. If the cluster contains multiple nodes, you can replace the value of **<private network address and port number of a node>** with the private network address and port number of any available node in the cluster. If the cluster contains only one node, restore the node and execute the command again.
- Replace *<Private IP address of the jump host>* in the preceding command with the IP address (with **Private IP**) of the created jump host in the **IP Address** column in the ECS list on the ECS management console.

For example, port **9200** on the jump host is assigned external network access permissions, the private network address and port number of the node are **192.168.0.81** and **9200**, respectively, and the private IP address of the jump host is **192.168.0.227**. You need to run the following command to perform port mapping:

```
ssh -g -L 9200:192.168.0.81:9200 -N -f root@192.168.0.227
```

4. Log in to the server where Logstash is deployed and store the data files to be imported on the server.

For example, data file `access_20181029_log` needs to be imported, the file storage path is `/tmp/access_log/`, and the data file includes the following data:

 **NOTE**

Create the `access_log` folder if it does not exist.

All	Heap used for segments		18.6403	MB
All	Heap used for doc values		0.119289	MB
All	Heap used for terms		17.4095	MB
All	Heap used for norms		0.0767822	MB
All	Heap used for points		0.225246	MB
All	Heap used for stored fields		0.809448	MB
All	Segment count		101	
All	Min Throughput	index-append	66232.6	docs/s
All	Median Throughput	index-append	66735.3	docs/s
All	Max Throughput	index-append	67745.6	docs/s
All	50th percentile latency	index-append	510.261	ms

5. In the server where Logstash is deployed, run the following command to create configuration file `logstash-simple.conf` in the Logstash installation directory:

```
cd /<Logstash installation directory>/
vi logstash-simple.conf
```

6. Input the following content in `logstash-simple.conf`:

```
input {
  Location of data
}
filter {
  Related data processing
}
output {
  elasticsearch {
    hosts => "<EIP of the jump host>:<Number of the port assigned external network access
permissions on the jump host>"
  }
}
```

- The **input** parameter indicates the data source. Set this parameter based on the actual conditions. For details about the **input** parameter and parameter usage, visit the following website: <https://www.elastic.co/guide/en/logstash/current/input-plugins.html>
- The **filter** parameter specifies the mode in which data is processed. For example, extract and process logs to convert unstructured information into structured information. For details about the **filter** parameter and parameter usage, visit the following website: <https://www.elastic.co/guide/en/logstash/current/filter-plugins.html>
- The **output** parameter indicates the destination address of the data. For details about the **output** parameter and parameter usage, visit <https://www.elastic.co/guide/en/logstash/current/output-plugins.html>. Replace `<EIP address of the jump host>` with the IP address (with **EIP**) of the created jump host in the **IP Address** column in the ECS list on the ECS management console. `<Number of the port assigned external network access permissions on the jump host>` is the number of the port obtained in **1**, for example, **9200**.

Consider the data files in the `/tmp/access_log/` path mentioned in [4](#) as an example. Assume that data import starts from data in the first row of the data file, the filtering condition is left unspecified (indicating no data processing operations are performed), the public IP address and port number of the jump host are **192.168.0.227** and **9200**, respectively, and the name of

the target index is **myindex**. Edit the configuration file as follows, and enter **:wq** to save the configuration file and exit.

```
input {
  file {
    path => "/tmp/access_log/*"
    start_position => "beginning"
  }
}
filter {
}
output {
  elasticsearch {
    hosts => "192.168.0.227:9200"
    index => "myindex"
  }
}
```

 **NOTE**

If a license error is reported, set **ilm_enabled** to **false**.

If the cluster has the security mode enabled, you need to download a certificate first.

- a. Download a certificate on the **Basic Information** page of the cluster.

Figure 2-14 Downloading a certificate

Configuration

Region	[Redacted]
AZ	[Redacted]
VPC	vpc [Redacted]
Subnet	subnet- [Redacted]
Security Group	dws [Redacted] Change Security Group
Security Mode	Enabled
Reset Password	Reset
Enterprise Project	default
Public IP Address	100. [Redacted] Disassociate
Access Control	Disabled Set
Bandwidth	22 Mbit/s Edit
HTTPS Access	Enabled Download Certificate
IPv4 Access Address	192 [Redacted]

- b. Store the certificate to the server where Logstash is deployed.
- c. Modify the **logstash-simple.conf** configuration file.

Consider the data files in the **/tmp/access_log/** path mentioned in 4 as an example. Assume that data import starts from data in the first row of the data file, the filtering condition is left unspecified (indicating no data processing operations are performed), and the public IP address and port number of the jump host are **192.168.0.227** and **9200**, respectively. The name of the index for importing data is **myindex**, and the certificate is stored in **/logstash/logstash6.8/config/CloudSearchService.cer**. Edit the configuration file as follows, and enter **:wq** to save the configuration file and exit.

```
input{
  file {
    path => "/tmp/access_log/"
    start_position => "beginning"
  }
}
filter {
}
output{
  elasticsearch{
    hosts => ["https://192.168.0.227:9200"]
  }
}
```

```

index => "myindex"
user => "admin"
password => "*****"
cacert => "/logstash/logstash6.8/config/CloudSearchService.cer"
manager_template => false
ilm_enabled => false
ssl => true
ssl_certificate_verification => false
}
}

```

NOTE

password: password for logging in to the cluster

7. Run the following command to import the data collected by Logstash to the cluster:

```
./bin/logstash -f logstash-simple.conf
```

NOTE

This command must be executed in the directory where the **logstash-simple.conf** file is stored. For example, if the **logstash-simple.conf** file is stored in **/root/logstash-7.1.1/**, go to the directory before running the command.

8. Log in to the CSS management console.
9. In the navigation pane on the left, choose **Clusters > Elasticsearch** to switch to the **Clusters** page.
10. From the cluster list, locate the row that contains the cluster to which you want to import data and click **Access Kibana** in the **Operation** column.
11. In the Kibana navigation pane on the left, choose **Dev Tools**.
12. On the **Console** page of Kibana, search for the imported data.

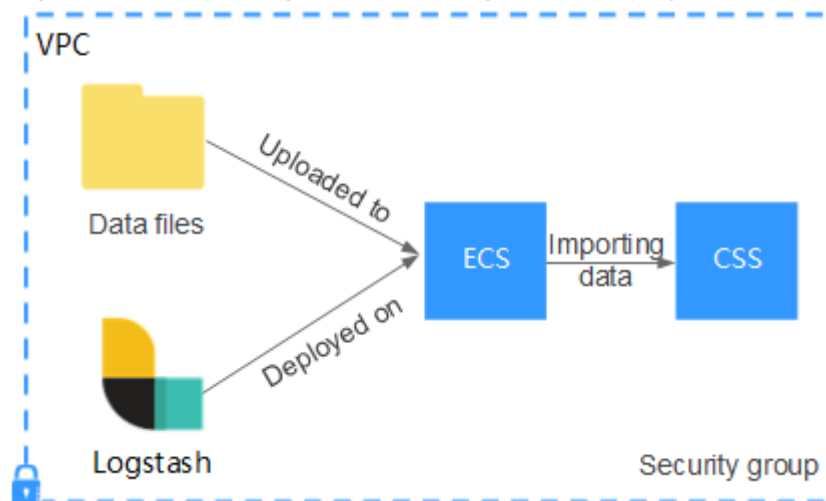
On the **Console** page of Kibana, run the following command to search for data. View the search results. If the searched data is consistent with the imported data, the data has been imported successfully.

```
GET myindex/_search
```

Importing Data When Logstash Is Deployed on an ECS

Figure 2-15 illustrates how data is imported when Logstash is deployed on an ECS that resides in the same VPC as the cluster to which data is to be imported.

Figure 2-15 Importing data when Logstash is deployed on an ECS



1. Ensure that the ECS where Logstash is deployed and the cluster to which data is to be imported reside in the same VPC, port **9200** of the ECS security group has been assigned external network access permissions, and an EIP has been bound to the ECS.

 **NOTE**

- If there are multiple servers in a VPC, you do not need to associate EIPs to other servers as long as one server is associated with an EIP. Switch to the node where Logstash is deployed from the node with which the EIP is associated.
 - If a private line or VPN is available, you do not need to associate an EIP.
2. Use PuTTY to log in to the ECS.

For example, data file **access_20181029_log** is stored in the **/tmp/access_log/** path of the ECS, and the data file includes the following data:

All	Heap used for segments		18.6403	MB
All	Heap used for doc values		0.119289	MB
All	Heap used for terms		17.4095	MB
All	Heap used for norms		0.0767822	MB
All	Heap used for points		0.225246	MB
All	Heap used for stored fields		0.809448	MB
All	Segment count		101	
All	Min Throughput	index-append	66232.6	docs/s
All	Median Throughput	index-append	66735.3	docs/s
All	Max Throughput	index-append	67745.6	docs/s
All	50th percentile latency	index-append	510.261	ms

3. Run the following command to create configuration file **logstash-simple.conf** in the Logstash installation directory:

```
cd /<Logstash installation directory>/
vi logstash-simple.conf
```

Input the following content in **logstash-simple.conf**:

```
input {
  Location of data
}
filter {
  Related data processing
}
output {
  elasticsearch{
    hosts => "<Private network address and port number of the node>"
  }
}
```

- The **input** parameter indicates the data source. Set this parameter based on the actual conditions. For details about the **input** parameter and parameter usage, visit the following website: <https://www.elastic.co/guide/en/logstash/current/input-plugins.html>
- The **filter** parameter specifies the mode in which data is processed. For example, extract and process logs to convert unstructured information into structured information. For details about the **filter** parameter and parameter usage, visit the following website: <https://www.elastic.co/guide/en/logstash/current/filter-plugins.html>
- The **output** parameter indicates the destination address of the data. For details about the **output** parameter and parameter usage, visit <https://www.elastic.co/guide/en/logstash/current/output-plugins.html>. *<private network address and port number of a node>* refers to the private network address and port number of a node in the cluster.

If the cluster contains multiple nodes, you are advised to replace the value of *<Private network address and port number of a node>* with the private network addresses and port numbers of all nodes in the cluster to

prevent node faults. Use commas (,) to separate the nodes' private network addresses and port numbers. The following is an example:

```
hosts => ["192.168.0.81:9200","192.168.0.24:9200"]
```

If the cluster contains only one node, the format is as follows:

```
hosts => "192.168.0.81:9200"
```

Consider the data files in the `/tmp/access_log/` path mentioned in [2](#) as an example. Assume that data import starts from data in the first row of the data file, the filtering condition is left unspecified (indicating no data processing operations are performed), the private network address and port number of the node in the cluster where data is to be imported are **192.168.0.81** and **9200**, respectively, and the name of the target index is **myindex**. Edit the configuration file as follows, and enter `:wq` to save the configuration file and exit.

```
input {
  file {
    path => "/tmp/access_log/*"
    start_position => "beginning"
  }
}
filter {
}
output {
  elasticsearch {
    hosts => "192.168.0.81:9200"
    index => "myindex"
  }
}
```

If the cluster has the security mode enabled, you need to download a certificate first.

- a. Download a certificate on the **Basic Information** page of the cluster.

Figure 2-16 Downloading a certificate

Configuration

Region	[Redacted]	
AZ	[Redacted]	
VPC	vpc	[Redacted]
Subnet	subnet-	[Redacted]
Security Group	dws	Change Security Group
Security Mode	Enabled	
Reset Password	Reset	
Enterprise Project	default	
Public IP Address	100. [Redacted]	Disassociate
Access Control	Disabled Set	
Bandwidth	22 Mbit/s	Edit
HTTPS Access	Enabled	Download Certificate
IPv4 Access Address	192	[Redacted]

- b. Store the certificate to the server where Logstash is deployed.
- c. Modify the **logstash-simple.conf** configuration file.

Consider the data files in the **/tmp/access_log/** path mentioned in step 2 as an example. Assume that data import starts from data in the first row of the data file, the filtering condition is left unspecified (indicating no data processing operations are performed), the public IP address and port number of the jump host are **192.168.0.227** and **9200**, respectively. The name of the index for importing data is **myindex**, and the certificate is stored in **/logstash/logstash6.8/config/CloudSearchService.cer**. Edit the configuration file as follows, and enter **:wq** to save the configuration file and exit.

```
input{
  file {
    path => "/tmp/access_log/"
    start_position => "beginning"
  }
}
filter {
}
output{
  elasticsearch{
    hosts => ["https://192.168.0.227:9200"]
  }
}
```

```
index => "myindex"  
user => "admin"  
password => "*****"  
cacert => "/logstash/logstash6.8/config/CloudSearchService.cer"  
manager_template => false  
ilm_enabled => false  
ssl => true  
ssl_certificate_verification => false  
}  
}
```

NOTE

password: password for logging in to the cluster

4. Run the following command to import the ECS data collected by Logstash to the cluster:

```
./bin/logstash -f logstash-simple.conf
```
5. Log in to the CSS management console.
6. In the navigation pane on the left, choose **Clusters > Elasticsearch** to switch to the **Clusters** page.
7. From the cluster list, locate the row that contains the cluster to which you want to import data and click **Access Kibana** in the **Operation** column.
8. In the Kibana navigation pane on the left, choose **Dev Tools**.
9. On the **Console** page of Kibana, search for the imported data.

On the **Console** page of Kibana, run the following command to search for data. View the search results. If the searched data is consistent with the imported data, the data has been imported successfully.

```
GET myindex/_search
```

2.2.4 Using Kibana or APIs to Import Data to Elasticsearch

You can import data in various formats, such as JSON, to Elasticsearch in CSS by using Kibana or APIs.

Importing Data Using Kibana

Before importing data, ensure that you can use Kibana to access the cluster. The following procedure illustrates how to use the **POST** command to import data.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters > Elasticsearch** to switch to the **Clusters** page.
3. Choose **Clusters** in the navigation pane. Locate the target cluster and click **Access Kibana** in the **Operation** column to log in to Kibana.
4. Click **Dev Tools** in the navigation tree on the left.
5. Run the following command on the Kibana page to check whether the cluster has indexes:

```
GET _cat/indices?v
```

If there are indexes available in the cluster that you want to import data, you do not need to create an index. Go to step 7.

If there are no indexes available in the cluster that you want to import data, go to the next step and create an index.

6. Run the related commands to create an index for storing data and specify a custom mapping to define the data type.

For example, on the **Console** page of Kibana, run the following command to create an index named **my_store** and specify a user-defined mapping to define the data type:

Versions earlier than 7.x

```
PUT /my_store
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "products": {
      "properties": {
        "productName": {
          "type": "text"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
}
```

Versions 7.x and later

```
PUT /my_store
{
  "settings": {
    "number_of_shards": 1
  },
  "mappings": {
    "properties": {
      "productName": {
        "type": "text"
      },
      "size": {
        "type": "keyword"
      }
    }
  }
}
```

7. Run commands to import data. For example, run the following command to import a piece of data:

Versions earlier than 7.x

```
POST /my_store/products/_bulk
{"index":{}}
{"productName":"Latest art shirts for women in 2017 autumn","size":"L"}
```

Versions 7.x and later

```
POST /my_store/_bulk
{"index":{}}
{"productName":"Latest art shirts for women in 2017 autumn","size":"L"}
```

The command output is similar to that shown in [Figure 2-17](#). If the value of the **errors** field in the result is **false**, the data is successfully imported.

Figure 2-17 Response message

```
1 {
2   "took": 42,
3   "errors": false,
4   "items": [
5     {
6       "index": {
7         "_index": "my_store",
8         "_type": "products",
9         "_id": "AWTGbHt7BwpN-hb3LKau",
10        "_version": 1,
11        "result": "created",
12        "_shards": {
13          "total": 2,
14          "successful": 2,
15          "failed": 0
16        },
17        "created": true,
18        "status": 201
19      }
20    }
21  ]
22 }
```

Importing Data Using APIs

You can call the bulk API using the cURL command to import a JSON data file.

NOTE

- You are advised to import a file smaller than 50 MB.
- This section uses a cluster in non-security mode as an example to describe how to run the cURL command to import data. For details about how to import data to a cluster in other modes, see [Accessing a Cluster Using cURL Commands](#).

1. Log in to the ECS that you use to access the cluster.
For details about how to log in to an ECS, see [Logging In to an ECS](#).
2. Upload the JSON data file to the ECS.
3. Run the following commands in the path where the JSON data file is stored in the ECS to import the JSON data to an Elasticsearch cluster.

In the command, replace the value of *{Private network address and port number of the node}* with the private network address and port number of a node in the cluster. If the node fails to work, the command will fail to be executed. If the cluster contains multiple nodes, you can replace the value of *{Private network address and port number of the node}* with the private network address and port number of any available node in the cluster. If the cluster contains only one node, restore the node and execute the command again. **test.json** indicates the JSON file whose data is to be imported.

```
curl -X PUT "http://{Private network address and port number of the node}/_bulk" -H 'Content-Type: application/json' --data-binary @test.json
```

 NOTE

The value of the **-X** parameter is a command and that of the **-H** parameter is a message header. In the preceding command, **PUT** is the value of the **-X** parameter and **'Content-Type: application/json' --data-binary @test.json** is the value of the **-H** parameter. Do not add **-k** between a parameter and its value.

Example: In this example, assume that you need to import data in the **test.json** file to an Elasticsearch cluster, where communication encryption is disabled and the private network address and port number of one node are **192.168.0.90** and **9200** respectively. The data in the **test.json** file is as follows:

Versions earlier than 7.x

```
{"index": {"_index": "my_store", "_type": "products"}}
{"productName": "Autumn new woman blouses 2019", "size": "M"}
{"index": {"_index": "my_store", "_type": "products"}}
{"productName": "Autumn new woman blouses 2019", "size": "L"}
```

Versions 7.x and later

```
{"index": {"_index": "my_store"}}
{"productName": "Autumn new woman blouse 2019", "size": "M"}
{"index": {"_index": "my_store"}}
{"productName": "Autumn new woman blouse 2019", "size": "L"}
```

Perform the following steps to import the data:

- a. Run the following command to create an index named **my_store**:

Versions earlier than 7.x

```
curl -X PUT http://192.168.0.90:9200/my_store -H 'Content-Type: application/json' -d '{
  {
    "settings": {
      "number_of_shards": 1
    },
    "mappings": {
      "products": {
        "properties": {
          "productName": {
            "type": "text"
          },
          "size": {
            "type": "keyword"
          }
        }
      }
    }
  }
}'
```

Versions 7.x and later

```
curl -X PUT http://192.168.0.90:9200/my_store -H 'Content-Type: application/json' -d '{
  {
    "settings": {
      "number_of_shards": 1
    },
    "mappings": {
      "properties": {
        "productName": {
          "type": "text"
        },
        "size": {
          "type": "keyword"
        }
      }
    }
  }
}'
```

- b. Run the following command to import the data in the **test.json** file:

```
curl -X PUT "http://192.168.0.90:9200/_bulk" -H 'Content-Type: application/json' --data-binary @test.json
```

In this case, if the following information is displayed, the data is successfully imported:

```
{
  "took": 204,
  "errors": false,
  "items": [
    {
      "index": {
        "_index": "my_store",
        "_type": "_doc",
        "_id": "DJQkBlwBbJvUd2769Wi-",
        "_version": 1,
        "result": "created",
        "_shards": {
          "total": 2,
          "successful": 1,
          "failed": 0,
          "_seq_no": 0,
          "_primary_term": 1,
          "status": "201"
        }
      }
    },
    {
      "index": {
        "_index": "my_store",
        "_type": "_doc",
        "_id": "DZQkBlwBbJvUd2769Wi_",
        "_version": 1,
        "result": "created",
        "_shards": {
          "total": 2,
          "successful": 1,
          "failed": 0,
          "_seq_no": 1,
          "_primary_term": 1,
          "status": "201"
        }
      }
    }
  ]
}
```

2.3 Migrating a Cluster Using Backup and Restoration

Overview

Data can be migrated between CSS Elasticsearch clusters by backing up and restoring cluster snapshots.

Application scenarios:

- Cluster upgrade: Migrate data from a cluster of an earlier version to a cluster of a later version.
- Cluster merge: Merge the index data of two clusters.

This section describes how to take a snapshot of a cluster and restore it to another cluster. Take Elasticsearch cluster **Es-1** and **Es-2** as an example.

Migration Duration

The number of nodes or index shards in the source and destination clusters determines how long the data migration will take. Data migration consists of two phases: data backup and restoration. The backup duration is determined by the source cluster and the restoration duration is determined by the destination cluster. The formula for calculating the total migration duration is as follows:

- If the number of index shards is greater than the number of nodes:

$$\text{Total duration (s)} = (800 \text{ GB} / 40 \text{ MB} / \text{Number of source cluster nodes} + 800 \text{ GB} / 40 \text{ MB} / \text{Number of destination cluster nodes}) \times \text{Number of indexes}$$

- If the number of index shards is smaller than the number of nodes:

$$\text{Total duration (s)} = (800 \text{ GB} / 40 \text{ MB} / \text{Number of shards of the source cluster index} + 800 \text{ GB} / 40 \text{ MB} / \text{Number of shards of the destination cluster index}) \times \text{Number of indexes}$$

NOTE

The migration duration estimated using the formula is the minimal duration possible (if each node transmits data at the fastest speed, 40 MB/s). The actual duration also depends on factors such as the network and resources condition.

Prerequisites

- The destination cluster (**Es-2**) and source cluster (**Es-1**) are available. You are advised to migrate a cluster during off-peak hours.
- Ensure that the destination cluster (**Es-2**) and source cluster (**Es-1**) are in the same region.
- Ensure that the version of the destination cluster (**Es-2**) is later than or same as that of the source cluster (**Es-1**).

- Ensure that the number of nodes in the destination cluster (**Es-2**) is greater than half of the number of nodes in the source cluster (**Es-1**).
- Ensure that the number of nodes in the destination cluster (**Es-2**) is greater than or equal to the number of shards in the source cluster (**Es-1**).
- Ensure that the CPU, memory, and disk configurations of the target cluster (**Es-2**) are greater than or equal to those of the source cluster (**Es-1**).

Procedure

1. Log in to the Cloud Search Service management console.
2. Choose **Clusters > Elasticsearch**. On the displayed page, click the source cluster name **Es-1** to go to the basic information page.
3. In the navigation pane, choose **Cluster Snapshots**, and set basic snapshot configurations.

Table 2-23 Basic configurations for a cluster snapshot

Parameter	Description
OBS Bucket	Select an OBS bucket for storing cluster snapshots.
Backup Path	Storage path of the cluster snapshot in the OBS bucket. You can retain the default value.
IAM Agency	Select an IAM agency to authorize CSS to access or maintain data stored in OBS. The IAM agency must have the OBS Administrator permission for project OBS in region Global service .

4. Click **Create**. In the dialog box that is displayed, configure the parameters and click **OK** to manually create a snapshot.

Table 2-24 Snapshot creation parameters

Parameter	Description
Snapshot Name	User-defined snapshot name. You can retain the default value.
Index	Enter the name of the index to be backed up. Use commas (,) to separate multiple indexes. Uppercase letters, spaces, and the following special characters are not allowed: "\< >/? If you do not specify this parameter, data of all indexes in the cluster is backed up by default. You can use the asterisk (*) to back up data of certain indexes. For example, if you enter index* , then data of indexes with the name prefix of index will be backed up.
Description	Snapshot description.

In the snapshot management list, if the snapshot status is **Available**, the snapshot has been created.

- In the snapshot management list, click **Restore** in the **Operation** column of the snapshot and configure restoration parameters to restore data to destination cluster **Es-2**.

Table 2-25 Snapshot restoration parameters

Parameter	Description
Index	Enter the name of the index to be restored. If this parameter is not specified, all index data will be restored. You can use the asterisk (*) to match multiple indexes. For example, index* indicates that all indexes with the prefix index in snapshots are restored.
Rename Pattern	Index name matching rule. The Rename Pattern and Rename Replacement take effect only when they are configured at the same time. You can configure them to rename matched indexes in snapshots.
Rename Replacement	Rule for renaming an index name. The Rename Pattern and Rename Replacement take effect only when they are configured at the same time. The default value restored_index_\$1 indicates that restored_ is added in front of the names of all restored indexes.
Cluster	Select a destination cluster, for example, Es-2 . NOTICE If the source and destination clusters have indexes with the same names, the indexes in the destination cluster will be overwritten by those in the source cluster after the restoration.

In the snapshot management list, if **Task Status** changes to **Restoration succeeded**, data in source cluster **Es-1** is successfully migrated to destination cluster **Es-2**.

2.4 Accessing Elasticsearch Clusters

2.4.1 Accessing an Elasticsearch Cluster

Elasticsearch clusters have built-in Kibana and Cerebro components. You can quickly access an Elasticsearch cluster through Kibana and Cerebro.

Access a Cluster Through Kibana

- Log in to the CSS management console.
- On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column to go to the Kibana login page.

- Non-security cluster: The Kibana console is displayed.
 - Security cluster: Enter the username and password on the login page and click **Log In** to go to the Kibana console. The default username is **admin** and the password is the one specified during cluster creation.
3. After the login is successful, you can access clusters through Kibana.

Accessing a Cluster Through Cerebro

1. Log in to the CSS management console.
2. On the **Clusters** page, locate the target cluster and click **More > Cerebro** in the **Operation** column to go to the Cerebro login page.
 - Non-security cluster: Click the cluster name on the Cerebro login page to go to the Cerebro console.
 - Security cluster: Click the cluster name on the Cerebro login page, enter the username and password, and click **Authenticate** to go to the Cerebro console. The default username is **admin** and the password is the one specified during cluster creation.
3. After the login is successful, you can access clusters through Cerebro.

2.4.2 Accessing a Cluster from a Public Network

You can access a security cluster (Elasticsearch clusters in version 6.5.4 or later support the security mode) that has the HTTPS access enabled through the public IP address provided by the system.

By default, CSS uses a shared load balancer for public network access. You can use a dedicated load balancer to improve performance. For details about its configuration, see [Connecting to a Dedicated Load Balancer](#).

NOTE

If public network access is enabled for CSS, then EIP and bandwidth resources will be used and billed.

Configuring Public Network Access

1. Log in to the CSS management console.
2. On the **Create Cluster** page, enable **Security Mode**. Set the administrator password and enable HTTPS access.
3. Select **Automatically assign** for **Public IP Address** and set related parameters.

Figure 2-18 Configuring public network access

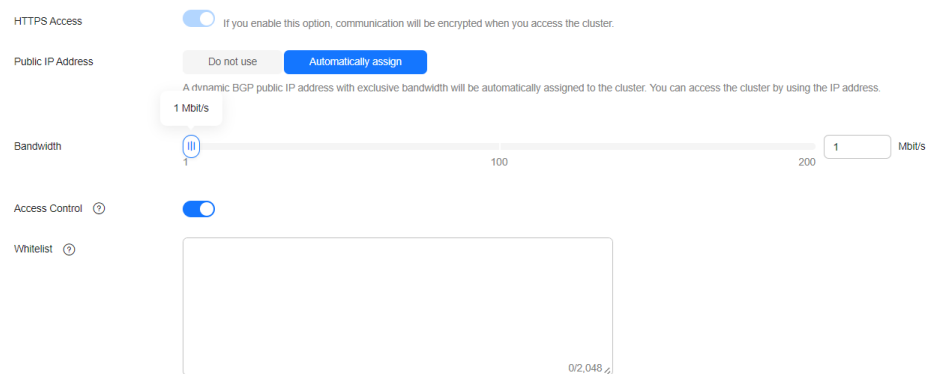


Table 2-26 Public network access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address
Access Control	If you disable this function, all IP addresses can access the cluster through the public IP address. If you enable access control, only IP addresses in the whitelist can access the cluster through the public IP address.
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled.

Managing Public Network Access

You can configure, modify, view the public network access of, or disassociate the public IP address from a cluster.

1. Log in to the CSS management console.
2. On the **Clusters** page, click the name of the target cluster. On the **Basic Information** page that is displayed, manage the public network access configurations.

Figure 2-19 Modifying public network access configurations

Configuration	
Region	[Redacted]
AZ	[Redacted]
VPC	vpc [Redacted]
Subnet	subnet [Redacted]
Security Group	dws [Redacted] Change Security Group
Security Mode	Enabled
Reset Password	Reset
Enterprise Project	default
Public IP Address	100 [Redacted] Disassociate
Access Control	Disabled Set
Bandwidth	22 Mbit/s Edit
HTTPS Access	Enabled Download Certificate
IPv4 Access Address	192. [Redacted]

- Configuring public network access
If you enabled HTTPS but did not configure the public network access during security cluster creation, you can configure it on the **Basic Information** page after configuring the cluster.
Click **Associate** next to **Public IP Address**, set the access bandwidth, and click **OK**.
If the association fails, wait for several minutes and try again.
- Modifying public network access
For a cluster for which you have configured public network access, you can click **Edit** next to **Bandwidth** to modify the bandwidth, or you can click **Set** next to **Access Control** to set the access control function and the whitelist for access.
- Viewing public network access
On the **Basic Information** page, you can view the public IP address associated with the current cluster.

- Disassociating a public IP address from a cluster
To disassociate the public IP address, click **Disassociate** next to **Public IP Address**.

Accessing a Cluster Through the Public IP Address

After configuring the public IP address, you can use it to access the cluster.

For example, run the following cURL commands to view the index information in the cluster. In this example, the public access IP address of one node in the cluster is **10.62.179.32** and the port number is **9200**.

- If the cluster you access does not have the security mode enabled, run the following command:

```
curl 'http://10.62.179.32:9200/_cat/indices'
```
- If the cluster you access has the security mode enabled, access the cluster using HTTPS and add the username, password and **-u** to the cURL command.

```
curl -u username:password -k 'https://10.62.179.32:9200/_cat/indices'
```

2.4.3 Accessing a Cluster Using a VPC Endpoint

If the VPC endpoint service is enabled, you can use a private domain name or node IP address generated by the endpoint to access the cluster. When the VPC endpoint service is enabled, a VPC endpoint will be created by default. You can select **Private Domain Name Creation** as required. VPC endpoint creation requires specific permissions. For details, see [VPCEP Permissions](#).

VPC Endpoint uses a shared load balancer for intranet access. If your workloads require quick access, you are advised to connect a dedicated load balancer to the cluster. For details, see [Connecting to a Dedicated Load Balancer](#).

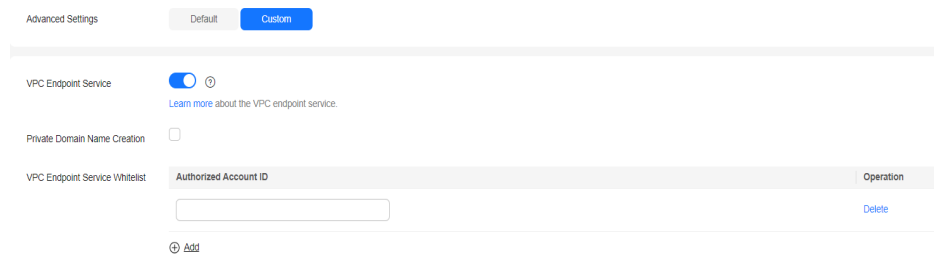
CAUTION

The public IP address access and VPC endpoint service share a load balancer. If you have configured a public access whitelist, public and private IP addresses that access the cluster through VPCEP are restricted because the public IP address access shares the load balancer with the VPC endpoint service. In this case, you need to add IP address **198.19.128.0/17** to the public access whitelist to allow traffic through VPCEP.

Enabling the VPC Endpoint Service

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner.
3. On the **Create Cluster** page, set **Advanced Settings** to **Custom**. Enable the VPC endpoint service.

Figure 2-20 Enabling the VPC endpoint service



- **Private Domain Name Creation:** If you enable this function, the system automatically creates a private domain name for you, which you can use to access the cluster.
- **VPC Endpoint Service Whitelist:** You can add an authorized account ID to the VPC endpoint service whitelist. Then you can access the cluster using the private domain name or the node IP address.
- You can click **Add** to add multiple accounts.
- Click **Delete** in the **Operation** column to delete the accounts that are not allowed to access the cluster.

NOTE

- If the authorized account ID is set to *, all users are allowed to access the cluster.
- You can view authorized account IDs on the **My Credentials** page.
- After the VPC endpoint service is enabled for a cluster, you will be billed per use. For more information, see [Billing Modes](#).

Managing VPC Endpoint Service

You can enable the VPC endpoint service while creating a cluster, and also enable it by performing the following steps after cluster creation.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
3. Click the **VPC Endpoint Service** tab, and turn on the button next to **VPC Endpoint Service**.

In the displayed dialog box, you can determine whether to enable the private domain name. Click **Yes** to enable the VPC endpoint service.

NOTE

- If the VPC endpoint service is enabled, you can use a private domain name or node IP address generated by the VPC endpoint to access the cluster. For details, see [Accessing the Cluster Using the Private Domain Name or Node IP Address](#).
 - If you disable the VPC endpoint service, none of the users can access the cluster using the private domain name.
4. (Optional) Click **Modify** next to **VPC Endpoint Service Whitelist** to update the existing whitelist.
 5. Manage VPC endpoints.

The **VPC Endpoint Service** page displays all VPC endpoints connected to the current VPC endpoint service.

Figure 2-21 Managing VPC endpoints

VPC Endpoint ID	Status	Max. Connections	Owner	Created	Operation
a24c-...	Accepted	3000		Mar 15, 2024 16:14:20 GMT+08:00	Accept Reject

Click **Accept** or **Reject** in the **Operation** column to change the node status. If you reject the connection with a VPC endpoint, you cannot access the cluster through the private domain name generated by that VPC endpoint.

Accessing the Cluster Using the Private Domain Name or Node IP Address

1. Obtain the private domain name or node IP address.

Log in to the CSS console, click the target cluster name and go to the **Cluster Information** page. Click the **VPC Endpoint Service** tab and view the private domain name.

Figure 2-22 Viewing the node IP address and private domain name

Cluster Information	
Name	cn-...
Private IP Address	192...
Private Domain Name	vpcep-...com
VPC Endpoint Service Whitelist	-- Modify

2. Run the cURL command to execute the API or call the API by using a program before accessing the cluster. For details about Elasticsearch operations and APIs, see the [Elasticsearch Reference](#).

The ECS must meet the following requirements:

- Sufficient disk space is allocated for the ECS.
- The ECS and the cluster must be in the same VPC. After enabling the VPC endpoint service, you can access the cluster from the ECS even when the cluster is not in the same VPC as the ECS.
- The security group of the ECS must be the same as that of the cluster.
If this requirement is not met, modify the ECS security group or configure the inbound and outbound rules of the ECS security group to allow the ECS security group to be accessed by all security groups of the cluster. For details, see [Configuring Security Group Rules](#).
- Configure security group rule settings of the target CSS cluster. Set **Protocol** to **TCP** and **Port Range** to **9200** or a port range including port **9200** for both the outbound and inbound directions.

For example, run the following cURL command to view the index information in the cluster. In this example, the private network address is **vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com** and port **9200** is used to access the cluster.

- If the cluster you access does not have the security mode enabled, run the following command:

```
curl 'http://vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com:9200/_cat/indices'
```

- If the cluster you access has the security mode enabled, access the cluster using HTTPS and add the username, password and `-u` to the cURL command.

```
curl -u username:password -k 'https://vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com:9200/_cat/indices'
```

2.4.4 (Optional) Interconnecting with a Dedicated Load Balancer

2.4.4.1 Scenario

CSS integrates shared load balancers and allows you to bind public network access and enable the VPC Endpoint service. Dedicated load balancers provide more functions and higher performance than shared load balancers. This section describes how to connect a cluster to a dedicated load balancer.

Advantages of connecting a cluster to a dedicated load balancer:

- A non-security cluster can also use capabilities of the Elastic Load Balance (ELB) service.
- You can use customized certificates for HTTPS bidirectional authentication.
- Seven-layer traffic monitoring and alarm configuration are supported, allowing you to view the cluster status at any time.

There are eight service forms for clusters in different security modes to connect to dedicated load balancers. [Table 2-27](#) describes the ELB capabilities for the eight service forms. [Table 2-28](#) describes the configurations for the eight service forms.

NOTICE

You are not advised connecting a load balancer that has been bound to a public IP address to a non-security cluster. Access from the public network using such a load balancer may bring security risks because non-security clusters can be accessed over HTTP without security authentication.

Table 2-27 ELB capabilities for different clusters

Security Mode	Service Form Provided by ELB for External Systems	ELB Load Balancing	ELB Traffic Monitoring	ELB Two-way Authentication
Non-security	No authentication	Yes	Yes	No
	One-way authentication	Yes	Yes	Yes
	Two-way authentication			

Security Mode	Service Form Provided by ELB for External Systems	ELB Load Balancing	ELB Traffic Monitoring	ELB Two-way Authentication
Security mode + HTTP	Password authentication	Yes	Yes	No
	One-way authentication + Password authentication	Yes	Yes	Yes
	Two-way authentication + Password authentication			
Security mode + HTTPS	One-way authentication + Password authentication Two-way authentication + Password authentication	Yes	Yes	Yes

Table 2-28 Configuration for interconnecting different clusters with ELB

Security Mode	Service Form Provided by ELB for External Systems	ELB Listener			Backend Server Group		
		Frontend Protocol	Port	SSL Parsing Mode	Backend Protocol	Health Check Port	Health Check Path
Non-security	No authentication	HTTP	9200	No authentication	HTTP	9200	/
	One-way authentication	HTTPS	9200	One-way authentication	HTTP	9200	
	Two-way authentication	HTTPS	9200	Two-way authentication	HTTP	9200	

Security mode + HTTP	Password authentication	HTTP	9200	No authentication	HTTP	9200	/_opendistro/_security/health
	One-way authentication + Password authentication	HTTPS	9200	One-way authentication	HTTP	9200	
	Two-way authentication + Password authentication	HTTPS	9200	Two-way authentication	HTTP	9200	
Security mode + HTTPS	One-way authentication + Password authentication	HTTPS	9200	One-way authentication	HTTPS	9200	
	Two-way authentication + Password authentication	HTTPS	9200	Two-way authentication	HTTPS	9200	

2.4.4.2 Connecting to a Dedicated Load Balancer

This section describes how to connect a CSS cluster to a dedicated load balancer.

(Optional) Preparing a Self-signed Certificate

If the target ELB listener uses the HTTP protocol, skip this step.

Prepare and upload a self-signed certificate.

 **NOTE**

You are advised to use a certificate purchased in Cloud Certificate Manager (CCM) or issued by an authoritative organization.

1. Log in to a Linux client where the OpenSSL tool and JDK are installed.
2. Run the following commands to create a self-signed certificate:

```
mkdir ca
mkdir server
mkdir client

#Use OpenSSL to create a CA certificate.
cd ca
#Create the OpenSSL configuration file ca_cert.conf for the CA certificate.
cat >ca_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt = no
```

```
[ req_distinguished_name ]
O           = ELB
EOF
#Create private key file ca.key for the CA certificate.
openssl genrsa -out ca.key 2048
#Create the CSR file ca.csr for the CA certificate.
openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf
#Create a self-signed CA certificate ca.crt.
openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key
#Convert the CA certificate format to p12.
openssl pkcs12 -export -clcerts -in ca.crt -inkey ca.key -out ca.p12
#Convert the CA certificate format to JKS.
keytool -importkeystore -srckeystore ca.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
ca.jks

#Use the CA certificate to issue a server certificate.
cd ../server
#Create the OpenSSL configuration file server_cert.conf for the server certificate. Change the CN
field to the domain name or IP address of the server as required.
cat >server_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt             = no

[ req_distinguished_name ]
O           = ELB
CN          = 127.0.0.1
EOF
#Create the private key file server.key for the server certificate.
openssl genrsa -out server.key 2048
#Create the CSR request file server.csr for the server certificate.
openssl req -out server.csr -key server.key -new -config ./server_cert.conf
#Use the CA certificate to issue the server certificate server.crt.
openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -
CAkey ../ca/ca.key
#Convert the server certificate format to p12.
openssl pkcs12 -export -clcerts -in server.crt -inkey server.key -out server.p12
#Convert the service certificate format to JKS.
keytool -importkeystore -srckeystore server.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
server.jks

#Use the CA certificate to issue a client certificate.
cd ../client
#Create the OpenSSL configuration file client_cert.conf for the client certificate. Change the CN field
to the domain name or IP address of the server as required.
cat >client_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt             = no

[ req_distinguished_name ]
O           = ELB
CN          = 127.0.0.1
EOF
#Create private key client.key for the client certificate.
openssl genrsa -out client.key 2048
#Create the CSR file client.csr for the client certificate.
openssl req -out client.csr -key client.key -new -config ./client_cert.conf
#Use the CA certificate to issue the client certificate client.crt.
openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -
CAkey ../ca/ca.key
#Convert the client certificate to a p12 file that can be identified by the browser.
openssl pkcs12 -export -clcerts -in client.crt -inkey client.key -out client.p12
#Convert the client certificate format to JKS.
keytool -importkeystore -srckeystore client.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
client.jks
```

3. Upload the self-signed certificate. For details, see [Configuring the Server Certificate and Private Key](#).

Creating a Dedicated Load Balancer

1. Log in to the ELB management console.
2. Create a dedicated load balancer. For details, see [Creating a Dedicated Load Balancer](#). [Table 2-29](#) describes the parameters required for connecting a CSS cluster with a dedicated load balancer.

Table 2-29 Parameters for interconnecting a CSS cluster with a dedicated load balancer

Parameter	Description	Example
Type	Load balancer type. Select Dedicated .	Dedicated
Billed By	Billing mode of the dedicated load balancer.	Pay-per-use
Region	Region where the CSS cluster is located.	-
IP as Backend Servers	A CSS cluster can be connected only after the cross-VPC backend is enabled.	Enabled
Network Type	Type of the network used by the load balancer to provide services for external systems.	Private IPv4 network
VPC	VPC where the load balancer works. This parameter is mandatory no matter which network type is selected. Select the VPC of the CSS cluster	-
Subnet	Subnet where the load balancer is to be created. This parameter is mandatory no matter which network type is selected. Select the subnet of the CSS cluster	-

Parameter	Description	Example
Specifications	You are advised to select Application load balancing (HTTP/HTTPS) , which provides better functions and performance.	Application load balancing (HTTP/HTTPS) Small I

Interconnecting with a Load Balancer

NOTE

A cluster in security mode with HTTPS access enabled does not support HTTP protocol authentication. If you need to enable HTTP protocol authentication, disable the security mode of the cluster.


Before changing the security mode, disable load balancing. After the security mode is changed, enable load balancing.

1. Log in to the CSS management console.
2. On the **Clusters** page, select the cluster you want to connect to the load balancer and click the cluster name. The cluster basic information page is displayed.
3. In the navigation pane, choose **Load Balancing**. Toggle on the load balancing switch and configure basic load balancing information.
 - **Load Balancer:** Select a created load balancer. You can also click **Create Load Balancer** to create one.
 - **Agency:** Select an agency name. If no agency is available, click **Create Agency** to create one. The selected agency must have the **ELB Administrator** and **ELB FullAccess** permissions.

Figure 2-23 Enabling load balancing

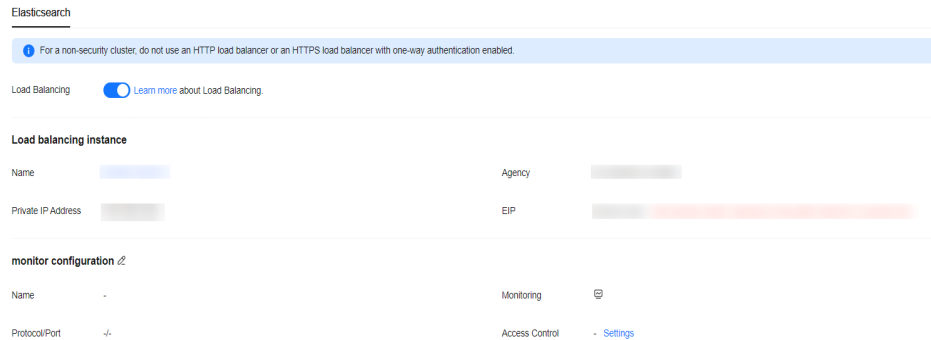
Basic Configuration

Load Balancer [Create Load Balancer](#) 

Agency [Create Agency](#) 

4. Click **OK**. The listener configuration page is displayed.

Figure 2-24 Creating a listener




- In the **Listener Configuration** area, click  to configure listener information.

Figure 2-25 Configuring a listener

monitor configuration

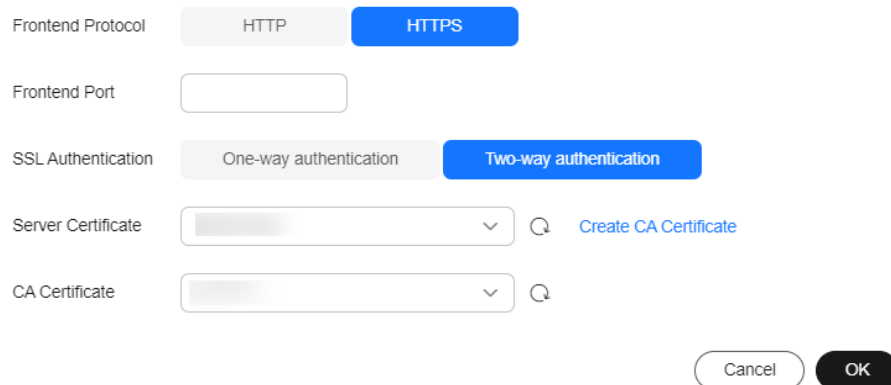


Table 2-30 Listener configuration information

Parameter	Description
Frontend Protocol	The protocol used by the client and listener to distribute traffic. Select a protocol as required.
Frontend Port	The port used by the client and listener to distribute traffic. For example, 9200. You need to specify this parameter as required.
SSL Authentication	Authentication mode for the client to access the server. Select a parsing mode as required.

Parameter	Description
Server Certificate	The server certificate is used for SSL handshake negotiation. The certificate content and private key must be provided. When SSL Authentication is set to Two-way authentication , this parameter is mandatory.
CA Certificate	Also called client CA public key certificate. It is used to verify the issuer of a client certificate. When the HTTPS two-way authentication is enabled, an HTTPS connection can be established only when the client can provide the certificate issued by a specified CA. This parameter is mandatory only when the Frontend Protocol is set to HTTPS .

- (Optional) In the **Connection Mode** area, you can click **Settings** next to **Access Control** to configure the IP addresses or network segments that are allowed to access the system. If you do not set the IP addresses or network segments, all IP addresses are allowed to access the system by default.

In the **Health Check** area, you can view the health check result of each node IP address. The following table describes the health check results.

Health Check Result	Description
Normal	The IP address of the node is properly connected.
Abnormal	The node IP address is connected and unavailable.

Accessing a Cluster Using the Curl Command

Run the following commands to check whether the dedicated load balancer can be connected to a cluster.

Table 2-31 Commands for accessing different clusters

Security Mode	Service Form Provided by ELB for External Systems	Curl Command for Accessing a Cluster
Non-security	No authentication	<code>curl http://IP:9200</code>

Security Mode	Service Form Provided by ELB for External Systems	Curl Command for Accessing a Cluster
	One-way authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200</code>
	Two-way authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200</code>
Security mode + HTTP	Password authentication	<code>curl http://IP:9200 -u user:pwd</code>
	One-way authentication + Password authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
	Two-way authentication + Password authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
Security mode + HTTPS	One-way authentication + Password authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
	Two-way authentication + Password authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>

Table 2-32 Variables

Variable	Description
IP	ELB IP address
user	Username for accessing the CSS cluster
pwd	Password of the user

If the Elasticsearch cluster information is returned, the connection is successful. For example, if a security cluster using the HTTPS protocol is connected to a load balancer using two-way authentication, the information shown in [Figure 2-26](#) is returned.

Figure 2-26 Accessing a cluster

```
root@~# curl --cacert ./ca/ca.crt --cert ./client.crt --key ./client.key https://127.0.0.1:9200 -u admin:admin
{"name": "css-test1-ess-asn-1-1",
 "cluster_name": "css-test1",
 "cluster_uuid": "nX81L1jT_2CMBFe1bgmA",
 "version": {
  "number": "7.10.2",
  "build_flavor": "oss",
  "build_type": "tar",
  "build_hash": "unknown",
  "build_date": "unknown",
  "build_snapshot": true,
  "lucene_version": "8.7.0",
  "minimum_wire_compatibility_version": "6.7.0",
  "minimum_index_compatibility_version": "6.0.0-beta1"
 },
 "tagline": "You Know, for Search"
}
```

2.4.4.3 Sample Code for Two-Way Authentication During the Access to a Cluster

This section provides the sample code for two-way authentication during the access to a cluster from a Java client.

ESSecuredClientWithCerDemo Code

```
import org.apache.commons.io.IOUtils;
import org.apache.http.auth.AuthScope;
import org.apache.http.auth.UsernamePasswordCredentials;
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.client.BasicCredentialsProvider;
import org.apache.http.HttpHost;
import org.apache.http.nio.conn.ssl.SSLIOSessionStrategy;
import org.elasticsearch.action.search.SearchRequest;
import org.elasticsearch.action.search.SearchResponse;
import org.elasticsearch.client.RequestOptions;
import org.elasticsearch.client.RestClient;
import org.elasticsearch.client.RestClientBuilder;
import org.elasticsearch.client.RestHighLevelClient;
import org.elasticsearch.index.query.QueryBuilders;
import org.elasticsearch.search.SearchHit;
import org.elasticsearch.search.SearchHits;
import org.elasticsearch.search.builder.SearchSourceBuilder;
import java.io.FileInputStream;
import java.io.IOException;
import java.security.KeyStore;
import java.security.SecureRandom;
import javax.net.ssl.HostnameVerifier;
import javax.net.ssl.KeyManagerFactory;
import javax.net.ssl.SSLContext;
import javax.net.ssl.SSLSession;
import javax.net.ssl.TrustManagerFactory;
public class ESSecuredClientWithCerDemo {
    private static final String KEY_STORE_PWD = "";
    private static final String TRUST_KEY_STORE_PWD = "";
    private static final String CA_JKS_PATH = "ca.jks";
    private static final String CLIENT_JKS_PATH = "client.jks";
    private static final String ELB_ADDRESS = "127.0.0.1";
    private static final int ELB_PORT = 9200;
    private static final String CSS_USERNAME = "user";
    private static final String CSS_PWD = "";
    public static void main(String[] args) {
        // Create a client.
        RestHighLevelClient client = initESClient(ELB_ADDRESS, CSS_USERNAME, CSS_PWD);
        try {
            // Search match_all, which is equivalent to {"query":{"match_all":{"}}}.

```

```
SearchRequest searchRequest = new SearchRequest();
SearchSourceBuilder searchSourceBuilder = new SearchSourceBuilder();
searchSourceBuilder.query(QueryBuilders.matchAllQuery());
searchRequest.source(searchSourceBuilder);
// query
SearchResponse searchResponse = client.search(searchRequest, RequestOptions.DEFAULT);
System.out.println("query result: " + searchResponse.toString());
SearchHits hits = searchResponse.getHits();
for (SearchHit hit : hits) {
    System.out.println(hit.getSourceAsString());
}
System.out.println("query success");
Thread.sleep(2000L);
} catch (InterruptedException | IOException e) {
    e.printStackTrace();
} finally {
    IOUtils.closeQuietly(client);
}
}
private static RestHighLevelClient initESClient(String clusterAddress, String userName, String password) {
    final CredentialsProvider credentialsProvider = new BasicCredentialsProvider();
    credentialsProvider.setCredentials(AuthScope.ANY, new UsernamePasswordCredentials(userName,
password));
    SSLContext ctx = null;
    try {
        KeyStore ks = getKeyStore(CLIENT_JKS_PATH, KEY_STORE_PWD, "JKS");
        KeyManagerFactory kmf = KeyManagerFactory.getInstance("SunX509");
        kmf.init(ks, KEY_STORE_PWD.toCharArray());
        KeyStore tks = getKeyStore(CA_JKS_PATH, TRUST_KEY_STORE_PWD, "JKS");
        TrustManagerFactory tmf = TrustManagerFactory.getInstance("SunX509");
        tmf.init(tks);
        ctx = SSLContext.getInstance("SSL", "SunJSSE");
        ctx.init(kmf.getKeyManagers(), tmf.getTrustManagers(), new SecureRandom());
    } catch (Exception e) {
        e.printStackTrace();
    }
    SSLIOSessionStrategy sessionStrategy = new SSLIOSessionStrategy(ctx, new HostnameVerifier() {
        @Override
        public boolean verify(String arg0, SSLSession arg1) {
            return true;
        }
    });
    SecuredHttpClientConfigCallback httpClientConfigCallback = new
SecuredHttpClientConfigCallback(sessionStrategy,
credentialsProvider);
    RestClientBuilder builder = RestClient.builder(new HttpHost(clusterAddress, ELB_PORT, "https"))
.setHttpClientConfigCallback(httpClientConfigCallback);
    RestHighLevelClient client = new RestHighLevelClient(builder);
    return client;
}
private static KeyStore getKeyStore(String path, String pwd, String type) {
    KeyStore keyStore = null;
    FileInputStream is = null;
    try {
        is = new FileInputStream(path);
        keyStore = KeyStore.getInstance(type);
        keyStore.load(is, pwd.toCharArray());
    } catch (Exception e) {
        e.printStackTrace();
    } finally {
        IOUtils.closeQuietly(is);
    }
    return keyStore;
}
}
```

SecuredHttpClientConfigCallback Code

```
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.nio.client.HttpAsyncClientBuilder;
import org.apache.http.nio.conn.ssl.SSLIOStrategy;
import org.elasticsearch.client.RestClientBuilder;
import org.elasticsearch.common.Nullable;
import java.util.Objects;

class SecuredHttpClientConfigCallback implements RestClientBuilder.HttpClientConfigCallback {
    @Nullable
    private final CredentialsProvider credentialsProvider;
    /**
     * The {@link SSLIOStrategy} for all requests to enable SSL / TLS encryption.
     */
    private final SSLIOStrategy sslStrategy;
    /**
     * Create a new {@link SecuredHttpClientConfigCallback}.
     *
     * @param credentialsProvider The credential provider, if a username/password have been supplied
     * @param sslStrategy The SSL strategy, if SSL / TLS have been supplied
     * @throws NullPointerException if {@code sslStrategy} is {@code null}
     */
    SecuredHttpClientConfigCallback(final SSLIOStrategy sslStrategy,
        @Nullable final CredentialsProvider credentialsProvider) {
        this.sslStrategy = Objects.requireNonNull(sslStrategy);
        this.credentialsProvider = credentialsProvider;
    }
    /**
     * Get the {@link CredentialsProvider} that will be added to the HTTP client.
     *
     * @return Can be {@code null}.
     */
    @Nullable
    CredentialsProvider getCredentialsProvider() {
        return credentialsProvider;
    }
    /**
     * Get the {@link SSLIOStrategy} that will be added to the HTTP client.
     *
     * @return Never {@code null}.
     */
    SSLIOStrategy getSSLStrategy() {
        return sslStrategy;
    }
    /**
     * Sets the {@linkplain HttpAsyncClientBuilder#setDefaultCredentialsProvider(CredentialsProvider)
    credential provider},
     *
     * @param httpClientBuilder The client to configure.
     * @return Always {@code httpClientBuilder}.
     */
    @Override
    public HttpAsyncClientBuilder customizeHttpClient(final HttpAsyncClientBuilder httpClientBuilder) {
        // enable SSL / TLS
        httpClientBuilder.setSSLStrategy(sslStrategy);
        // enable user authentication
        if (credentialsProvider != null) {
            httpClientBuilder.setDefaultCredentialsProvider(credentialsProvider);
        }
        return httpClientBuilder;
    }
}
```

pom.xml Code

```
<?xml version="1.0" encoding="UTF-8"?>
<project xmlns="http://maven.apache.org/POM/4.0.0"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/
```

```
maven-4.0.0.xsd">
  <modelVersion>4.0.0</modelVersion>
  <groupId>1</groupId>
  <artifactId>ESClient</artifactId>
  <version>1.0-SNAPSHOT</version>
  <name>ESClient</name>

  <properties>
    <maven.compiler.source>8</maven.compiler.source>
    <maven.compiler.target>8</maven.compiler.target>
    <project.build.sourceEncoding>UTF-8</project.build.sourceEncoding>
    <elasticsearch.version>7.10.2</elasticsearch.version>
  </properties>
  <dependencies>
    <dependency>
      <groupId>org.elasticsearch.client</groupId>
      <artifactId>transport</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>org.elasticsearch</groupId>
      <artifactId>elasticsearch</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>org.elasticsearch.client</groupId>
      <artifactId>elasticsearch-rest-high-level-client</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>commons-io</groupId>
      <artifactId>commons-io</artifactId>
      <version>2.11.0</version>
    </dependency>
  </dependencies>
</project>
```

2.5 Index Backup and Restoration

2.5.1 Backup and Restoration Overview

You can back up index data in clusters. If data loss occurs or you want to retrieve data of a specified duration, you can restore the index data. Index backup is implemented by creating cluster snapshots. When creating a backup for the first time, you are advised to back up data of all indexes.

- **Managing Automatic Snapshot Creation:** Snapshots are automatically created at a specified time each day according to the rules you create. You can enable or disable the automatic snapshot creation function and set the automatic snapshot creation policy.
- **Manually Creating a Snapshot:** You can manually create a snapshot at any time to back up all data or data of specified indexes.
- **Restoring Data:** You can use existing snapshots to restore the backup index data to a specified cluster.
- **Deleting a Snapshot:** Delete snapshots you do not require and release resources.

2.5.2 Managing Automatic Snapshot Creation

Snapshots are automatically created at a specified time according to the rules you create. You can enable or disable the automatic snapshot creation function and set the automatic snapshot creation policy.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- When creating a backup for the first time, you are advised to back up data of all indexes.
- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.
- Before creating a snapshot, you need to perform basic configurations, including configuring the OBS bucket for storing snapshots, the snapshot backup path, and IAM agency used for security authentication.
- If there are available snapshots in the snapshot list when you configure the OBS bucket for storing cluster snapshots for the first time, you cannot change the bucket for snapshots that are subsequently created automatically or manually. Exercise caution when you configure the OBS bucket.
- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, enable the snapshot function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.
- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- The first snapshot of a cluster is a full snapshot, and subsequent snapshots are incremental snapshots. CSS snapshot files depend on each other.

Managing Automatic Snapshot Creation

1. In the CSS navigation pane on the left, click **Clusters**.
2. On the **Clusters** page that is displayed, click the name of the target cluster. In the navigation pane on the left, choose **Cluster Snapshots**.

Alternatively, on the **Clusters** page, locate the row that contains the target cluster and click **More > Back Up and Restore** in the **Operation** column to switch to the **Cluster Snapshots** page.



3. On the displayed **Cluster Snapshots** page, click the icon to the right of **Cluster Snapshot** to enable the cluster snapshot function.
4. Enable the cluster snapshot function. OBS buckets and IAM agencies are automatically created to store snapshots. The automatically created OBS bucket and IAM agency are displayed on the page. You can also click  on the right of **Basic Configuration** to edit the configuration.

Table 2-33 Cluster snapshot parameter

Parameter	Description
OBS bucket	Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket . The created or existing OBS bucket must meet the following requirements: <ul style="list-style-type: none"> • Storage Class is Standard. • Region must be the same as that of the created cluster.
Backup Path	Storage path of the snapshot in the OBS bucket. The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \\:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency . The created or existing IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

5. Enable the automatic snapshot creation function. The **Configure Automatic Snapshot Creation** dialog box is displayed. If the automatic snapshot creation function is enabled, you can click  on the right of **Automatic Snapshot Creation** to modify the snapshot policy.
 - **Snapshot Name Prefix:** Enter a maximum of 32 characters starting with a lowercase letter. Only lowercase letters, digits, hyphens (-), and

underscores (_) are allowed. A snapshot name consists of a snapshot name prefix and a timestamp, for example, **snapshot-2018022405925**.

- **Time Zone:** indicates the time zone for the backup time. Specify backup start time based on the time zone.
- **Backup Start Time:** indicates the time when the backup starts automatically every hour, every day, or a specified day of a week. You can specify this parameter at the top of the hour every day or on a specified day of a week, for example, **00:00** or **01:00**. The value ranges from 00:00 to 23:00. Select the backup time from the drop-down list box.
- **Index:** Enter an index name. You can select an index for backup. Use commas (,) to separate multiple indexes. Uppercase letters, spaces, and special characters "`<|>/?`" are not allowed. If you do not specify this parameter, data of all indexes in the cluster is backed up by default. You can use the asterisk (*) to back up data of certain indexes. For example, if you enter **index***, then data of indexes with the name prefix of **index** will be backed up.

Run the **GET /_cat/indices** command in Kibana to query the names of all indexes in the cluster.

Figure 2-27 Automatically creating a snapshot

Configure Automatic Snapshot Creation

Snapshot Name Prefix

Time Zone

Backup Start Time

Retained Snapshots

Index

If you are sure you want to set the currently selected automatic backup policy, enter CONFIRM.

6. Click **OK** to save the snapshot policy.

Snapshots that are automatically created according to the snapshot policy are displayed in the snapshot list, along with manually created snapshots. You can distinguish them by the **Snapshot Type** setting. In the upper right corner of the snapshot list, enter the keyword of the snapshot name or snapshot ID to search for the desired snapshots.

7. (Optional) Disable the automatic snapshot creation function.

After you disable the automatic snapshot creation function, the system stops automatic creation of snapshots. If the system is creating a snapshot based on the automatic snapshot creation policy and the snapshot is not yet displayed in the snapshot list, you cannot disable the automatic snapshot creation function. In this case, if you click the button next to **Automatic Snapshot Creation**, a message is displayed, indicating that you cannot disable the function. You are advised to disable the function after the system completes automatic creation of the snapshot, and the created snapshot is displayed in the snapshot list.

When disabling the automatic snapshot creation function, you can choose whether to delete the snapshots that have been automatically created by selecting **Delete automated snapshots** in the displayed dialog box. By default, automatically created snapshots are not deleted.

- If you do not select **Delete automated snapshots**, automatically created snapshots are not deleted when you disable the automatic snapshot creation function. You can manually delete them later. For details, see [Deleting a Snapshot](#). If you do not manually delete the automatically created snapshots and enable the automatic snapshot creation function again, then all snapshots with **Snapshot Type** set to **Automated** in the snapshot list of the cluster can only be automatically deleted by the system. The system automatically deletes snapshots based on the policy configured when the automatic snapshot creation function is enabled. For example, if the number of retained snapshots is set to **10** in this policy and more than 10 snapshots are created, the system automatically deletes the excess snapshots on the half hour.
- If you select **Delete automated snapshots**, all snapshots with **Snapshot Type** set to **Automated** in the snapshot list will be deleted when you disable the automatic snapshot creation function.

NOTE

If snapshots are disabled, existing snapshots will not be automatically deleted. If you need to delete the snapshots, manage the bucket that stores snapshots on the OBS console.

2.5.3 Manually Creating a Snapshot

You can manually create a snapshot at any time to back up all data or data of specified indexes.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- When creating a backup for the first time, you are advised to back up data of all indexes.

- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.
- Before creating a snapshot, you need to perform basic configurations, including configuring the OBS bucket for storing snapshots, the snapshot backup path, and IAM agency used for security authentication.
- If there are available snapshots in the snapshot list when you configure the OBS bucket for storing cluster snapshots for the first time, you cannot change the bucket for snapshots that are subsequently created automatically or manually. Exercise caution when you configure the OBS bucket.
- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, enable the snapshot function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.
- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- The first snapshot of a cluster is a full snapshot, and subsequent snapshots are incremental snapshots. CSS snapshot files depend on each other.

Manually Creating a Snapshot


1. In the CSS navigation pane on the left, click **Clusters**.
2. On the **Clusters** page that is displayed, click the name of the target cluster. In the navigation pane on the left, choose **Cluster Snapshots**.
Alternatively, on the **Clusters** page, locate the row that contains the target cluster and click **More > Back Up and Restore** in the **Operation** column to switch to the **Cluster Snapshots** page.
3. On the displayed **Cluster Snapshots** page, click the icon to the right of **Cluster Snapshot** to enable the cluster snapshot function.
4. Enable the cluster snapshot function. OBS buckets and IAM agencies are automatically created to store snapshots. The automatically created OBS bucket and IAM agency are displayed on the page. You can also click  on the right of **Basic Configuration** to edit the configuration.

Table 2-34 Cluster snapshot parameter

Parameter	Description
OBS bucket	Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket . The created or existing OBS bucket must meet the following requirements: <ul style="list-style-type: none"> • Storage Class is Standard.
Backup Path	Storage path of the snapshot in the OBS bucket. The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*\?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency . The created or existing IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

- After basic configurations are completed, click **Create**.
 - Name** indicates the name of the manually created snapshot, which can contain 4 to 64 characters and must start with a lowercase letter. Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. For snapshots you create manually, you can specify the snapshot name. The system will not automatically add the time information to the snapshot name.
 - Index**: Enter an index name. You can select an index for backup. Use commas (,) to separate multiple indexes. Uppercase letters, spaces, and the following special characters are not allowed: "\<>/? If you do not specify this parameter, data of all indexes in the cluster is backed up by default. You can use the asterisk (*) to back up data of certain indices. For example, if you enter **index***, then data of indices with the name prefix of **index** will be backed up.
Run the **GET /_cat/indices** command in Kibana to query the names of all indexes in the cluster.

- **Description:** indicates the description of the created snapshot. The value contains 0 to 256 characters, and certain special characters (<>) are not allowed.

Figure 2-28 Manually creating a snapshot

Create Snapshot

Snapshot Name

Index 0/1,024 ↕

Description 0/256 ↕

6. Click **OK**.

After the snapshot is created, it will be displayed in the snapshot list. The status **Available** indicates that the snapshot is created successfully, along with manually created snapshots. You can distinguish them by the **Snapshot Type** setting. In the upper right corner of the snapshot list, enter the keyword of the snapshot name or snapshot ID to search for the desired snapshots.

2.5.4 Restoring Data

You can use existing snapshots to restore the backup index data to a specified cluster.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.
- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, enable the snapshot

function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.

- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- Cluster data cannot be queried during snapshot restoration.
- If you restore a CSS cluster snapshot to another cluster, indexes with the same name in the destination cluster will be overwritten. If the snapshot and the destination cluster use different shards, the indexes with the same name will not be overwritten.
- The version of the destination cluster used for restoration must be the same as or higher than that of the source cluster.

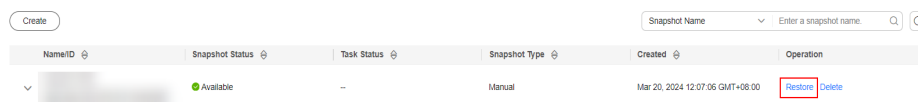
Restoring Data

You can use snapshots whose **Snapshot Status** is **Available** to restore cluster data. The stored snapshot data can be restored to other clusters.

Restoring data will overwrite current data in clusters. Therefore, exercise caution when restoring data.

1. In the **Snapshots** area, locate the row that contains the snapshot you want to restore and click **Restore** in the **Operation** column.

Figure 2-29 Selecting a snapshot



2. On the **Restore** page, set restoration parameters.

Index: Enter the name of the index you want to restore. If you do not specify any index name, data of all indexes will be restored. The value can contain 0 to 1,024 characters. Uppercase letters, spaces, and certain special characters (including "`<|>/?`") are not allowed. You can use the asterisk (*) to match multiple indexes. For example, **index*** indicates that all indexes with the prefix **index** in snapshots are restored.

Rename Pattern: Enter a regular expression. Indexes that match the regular expression are restored. The default value **index_(.+)** indicates restoring data of all indexes. The value contains 0 to 1,024 characters. Uppercase letters, spaces, and certain special characters (including "`<|>/?,)`") are not allowed.

Rename Replacement: Enter the index renaming rule. The default value **restored_index_\$1** indicates that **restored_** is added in front of the names of all restored indexes. The value contains 0 to 1,024 characters. Uppercase letters, spaces, and certain special characters (including "`<|>/?,)`") are not allowed.

 **NOTE**

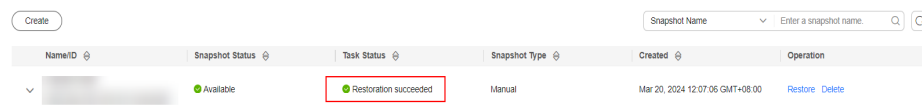
The **Rename Pattern** and **Rename Replacement** take effect only when they are configured at the same time.

Cluster: Select the cluster that you want to restore. You can select the current cluster or others. However, you can only restore the snapshot to clusters whose status is **Available**. If the status of the current cluster is **Unavailable**, you cannot restore the snapshot to the current cluster. When you restore data to another cluster, the version of the target cluster must be later than or equal to that of the current cluster. If the target cluster you selected has an index with the same name as the original cluster, data in the index will be overwritten after the restoration. Exercise caution when performing this operation.

Overwrite Index Shards of the Buckets with the Same Name in the Target Cluster: By default, the shards are not overwritten. Data is restored using snapshots by overwriting the snapshot files. After the index with the same name in the target cluster is overwritten, the index data in the target cluster may be lost. Exercise caution when performing this operation.

3. Click **OK**. If restoration succeeds, **Task Status** of the snapshot in the snapshot list will change to **Restoration succeeded**, and the index data is generated again according to the snapshot information.

Figure 2-30 Successful restoration



Name/ID	Snapshot Status	Task Status	Snapshot Type	Created	Operation
[Redacted]	Available	Restoration succeeded	Manual	Mar 20, 2024 12:07:06 GMT+08:00	Restore Delete

2.5.5 Deleting a Snapshot

If you no longer need a snapshot, delete it to release storage resources. If the automatic snapshot creation function is enabled, snapshots that are automatically created cannot be deleted manually, and the system automatically deletes these snapshots on the half hour after the time specified by **Retention Period (days)**. If you disable the automatic snapshot creation function while retaining the automated snapshots, then you can manually delete them later. If you do not manually delete the automatically created snapshots and enable the automatic snapshot creation function again, then all snapshots with **Snapshot Type** set to **Automated** in the snapshot list of the cluster can only be automatically deleted by the system.

 **NOTE**

After a snapshot is deleted, its data cannot be restored. Exercise caution when deleting a snapshot.

1. In the snapshot list, locate the snapshot that you want to delete.
2. Click **Delete** in the **Operation** column. In the dialog box that is displayed, confirm the snapshot information and click **OK**.

2.6 Cluster Specification Modification

2.6.1 Overview

You can scale in or out a cluster and change cluster specifications. In this way, you can improve cluster efficiency and reduce O&M costs.

Scaling Out a Cluster

- If a data node (ess) processes many data writing and querying requests and responds slowly, you can expand its storage capacity to improve its efficiency. If some nodes turn unavailable due to the excessive data volume or misoperations, you can add new nodes to ensure the cluster availability.
- **Cold data nodes** (ess-cold) are used to share the workload of data nodes. To prevent cold data loss, you can expand the storage capacity of the cold data node or add new ones.

Changing Specifications

- If the allocation of new indexes or shards takes too long or the node coordination and scheduling are inefficient, you can change the master node (ess-master) specifications.
- If too many tasks need to be distributed or too many results have been aggregated, you can change the client node (ess-client) specifications.
- If the speed of data writing and query decreases suddenly, you can change the data node (ess) specifications.
- If cold data query becomes slow, you can change the cold node (ess-cold) specifications.

Scaling in a Cluster

- If a cluster can process existing data without fully using its resources, you can scale in the cluster to reduce costs.

Removing Specified Nodes

- If a cluster can process existing data without fully using its nodes, you can remove one or more specified nodes from the cluster to reduce costs.

Replacing a Specified Node

- If a node in the cluster is faulty, you can create a new node with the same specifications to replace it.

Adding Master/Client Nodes

- If the workloads on the data plane of a cluster increase, you can dynamically scale the cluster by adding master/client nodes.

Changing the Security Mode

After a cluster is created, its security mode can be changed using the following methods:

- Change a non-security cluster to a security cluster that uses HTTP or HTTPS protocol.
- Change a security cluster that uses HTTP or HTTPS protocol to a non-security cluster.
- Change the protocol of a security cluster.

Changing AZs

You can **Add AZ** or **Migrate AZ**.

- **Add AZ:** Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster to improve cluster availability.
- **Migrate AZ:** Completely migrate data from the current AZ to another AZ that has sufficient resources.

2.6.2 Scaling Out a Cluster

If the workloads on the data plane of a cluster change, you can scale out the cluster by increasing the number or capacity of its nodes. Services are not interrupted during cluster scale-out.

Prerequisites

- The target cluster is available and has no tasks in progress.
- The target cluster has sufficient quotas available.

Constraints

- The **Node Specifications** cannot be modified during scale-out. You can modify **Node Specifications** by referring to [Changing Specifications](#).
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- The number of nodes and node storage capacity cannot be expanded at the same time for a yearly/monthly cluster.
- The quota of nodes in different types varies. For details, see [Table 2-35](#).

Table 2-35 Number of nodes in different types

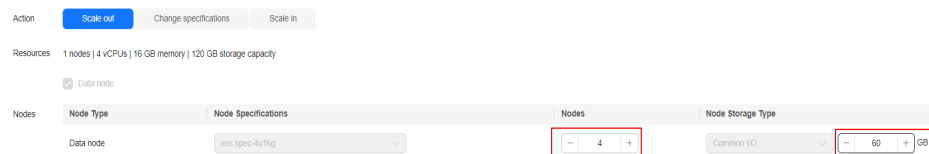
Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32

Node Type	Number
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale out** to set parameters.
 - **Action**: Select **Scale out**.
 - **Resource**: The changed amount of resources.
 - **Nodes**: The number of nodes and node storage capacity of the default data node.
 - **Nodes**: For details, see [Table 2-35](#).
 - The value range of **Node Storage Type** depends on the **Node Specifications**. The value must be a multiple of 20.

Figure 2-31 Scaling out a cluster



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling out**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled out.

2.6.3 Changing Specifications

If the workloads on the data plane of a cluster change, you can change its node specifications as needed.

Prerequisites

- The target cluster is available and has no tasks in progress.
- The target cluster has sufficient quotas available.
- When changing the node specifications, ensure that all service data has copies so the services will not be interrupted.

Run the **GET _cat/indices?v** command in Kibana. If the returned **rep** value is greater than **0**, the data has copies. If the returned **rep** value is **0**, the data has no copies. In this case, create snapshot for the cluster by referring to [Manually Creating a Snapshot](#).

- If the data volume is large, it may take long to modify the node specifications. You are advised to modify specifications during off-peak hours.

Constraints

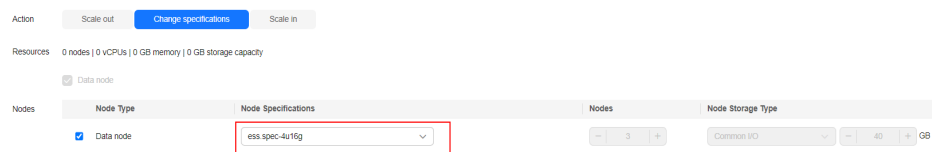
- The number of nodes and the capacity of node storage cannot be changed. You can add nodes and increase the node storage capacity by referring to [Scaling Out a Cluster](#). For details about how to reduce the number of nodes, see [Scaling in a Cluster](#).
- After decreasing cluster specifications, the cluster performance will deteriorate and service capabilities will be affected. Exercise caution when performing this operation.
- If a cluster has multiple node types, you can change the specifications of only one type at a time. After the change, nodes in other types still maintain their original specifications.
- Kibana is unavailable during specification change.
- During the specification modification, the nodes are stopped and restarted in sequence. It is a rolling process.

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.

3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Change Specifications** to set parameters.
 - **Action:** select **Change specifications**.
 - **Resources:** The changed amount of resources.
 - **Nodes:** Specifications of the default data nodes. Select the required specifications from the **Node Specifications** drop-down list and select the node that you want to change the specifications.
 - If a cluster has master nodes, client nodes, or cold data nodes, you can change their specifications.

Figure 2-32 Changing cluster specifications



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. In the dialog box that is displayed, confirm whether to select **Verify index copies** and **Cluster status check** and click **OK** to start the specifications change.

Index copy verification:

By default, CSS checks for indexes that do not have copies. You can skip this step, but the lack of index copies may affect services during a cluster specifications change.

- If you selected **Verify index copies** and the cluster has no master node, indexes must have at least one copy and the cluster must have at least three nodes.
- If you selected **Verify index copies** and the cluster has no master node, indexes must have at least one copy.

Cluster status check:

The cluster status is checked before the specifications change by default. The specifications of nodes are changed one by one to ensure success and data security. If a cluster is overloaded and services are faulty, the request for a specifications change will not be delivered. In this case, you can disable cluster status check. If you ignore the cluster status check before the specifications change, the cluster may be faulty and services may be interrupted. Exercise caution when performing this operation.

8. Click **Back to Cluster List** to switch to the **Clusters** page. The **Cluster Status** is **Configuration modified**. When **Cluster Status** changes to **Available**, the cluster specifications have been successfully modified.

2.6.4 Scaling in a Cluster

If a cluster can process existing data without fully using its resources, you can scale in the cluster to reduce costs. You are advised to scale in clusters during off-peak hours.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Only the number of nodes can be modified during cluster scale-in. The node specifications and node storage capacity cannot be modified. You can modify node specifications by referring to [Changing Specifications](#). You can modify node storage capacity by referring to [Scaling Out a Cluster](#).
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- Ensure that the disk usage after scale-in is less than 80% and each AZ of each node type has at least one node.
- When scaling in a cluster, the data in the node to be deleted is migrated to other nodes. The timeout threshold for data migration is five hours. If data migration is not complete within 5 hours, the cluster scale-in fails. You are advised to perform scale-in for multiple times when the cluster has huge amounts of data.
- For a cluster without master nodes, the number of remaining data nodes (including cold data nodes and other types of nodes) after scale-in must be greater than half of the original node number, and greater than the maximum number of index replicas.
- For a cluster with master nodes, the number of removed master nodes in a scale-in must be fewer than half of the original master node number. After scale-in, there has to be an odd number of master nodes, and there has to be at least three of them.
- A cluster with two nodes cannot be scaled in. You can create a cluster using a single node and then [Migrating Cluster Data Through Backup and Restoration](#).
- The quota of nodes in different types varies. For details, see [Table 2-36](#).

Table 2-36 Number of nodes in different types

Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32

Node Type	Number
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

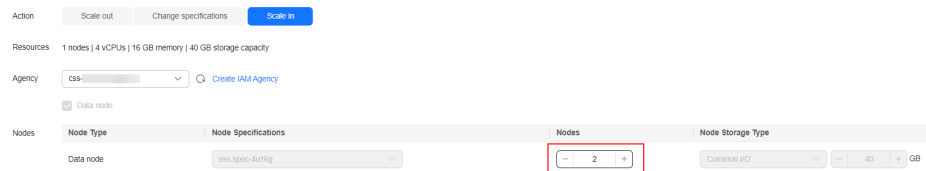
1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale in** to set parameters.
 - **Action**: Select **Scale in**.
 - **Resources**: The changed amount of resources.
 - **Agency**: Select an IAM agency to grant the current account the permission to switch AZs.
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

 **NOTE**

The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.

- **Nodes:** The number of the default data nodes. For details about the value range that can be changed, see [Table 2-36](#).

Figure 2-33 Scaling in a cluster



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling in**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled in.

2.6.5 Removing Specified Nodes

If a cluster can process existing data without fully using its nodes, you can remove one or more specified nodes from the cluster to reduce costs. You are advised to scale in clusters during off-peak hours.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Ensure that the disk usage after scale-in is less than 80% and each AZ of each node type has at least one node.
- In a cross-AZ cluster, the difference between the numbers of the same type nodes in different AZs cannot exceed 1.
- For a cluster without master nodes, the number of removed data nodes and cold data nodes in a scale-in must be fewer than half of the original number of data nodes and cold data nodes, and the number of remaining data nodes and cold data nodes after a scale-in must be greater than the maximum number of index replicas.
- For a cluster with master nodes, the number of removed master nodes in a scale-in must be fewer than half of the original master node number. After scale-in, there has to be an odd number of master nodes, and there has to be at least three of them.

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.

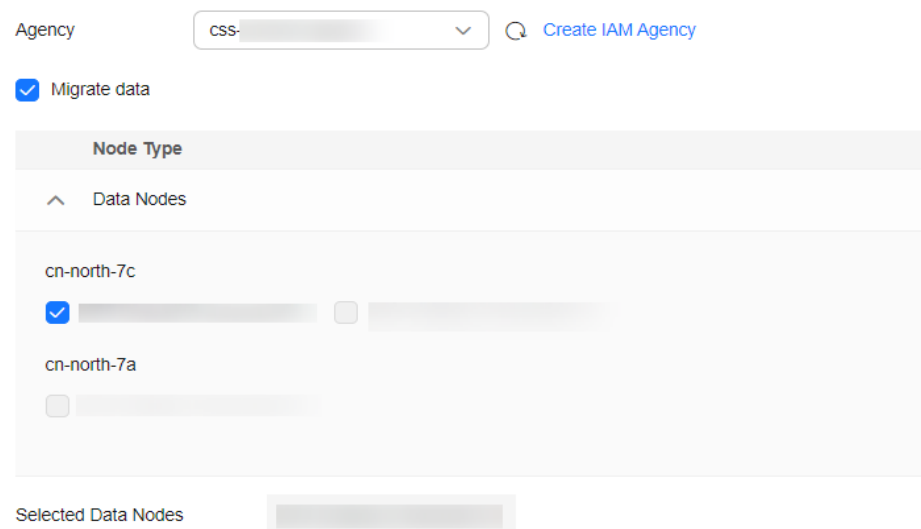
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, click the **Scale In** tab.
5. On the **Scale In** tab page, set the following parameters:
 - **Agency:** Select an IAM agency to grant the current account the permission to switch AZs.
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

 **NOTE**

The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.

- **Whether to perform data migration:** If this option is selected, data migration is performed. If the target node contains disabled indexes or indexes that have no replicas, this option must be selected.
- In the data node table, select the node to be scaled in.

Figure 2-34 Deleting specified nodes



6. Click **Next**.
7. Confirm the information and click **Submit**.
8. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling in**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled in.

2.6.6 Replacing a Specified Node

If a node in the cluster is faulty, you can create a new node with the same specifications to replace it. During the replacement of a specified node, data of that node will be migrated in advance and will not be lost.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Only one node can be replaced at a time.
- The ID, IP address, specifications, and AZ of the new node will be the same as those of the original one.
- The configurations you modified manually will not be retained after node replacement. For example, if you have manually added a return route to the original node, you need to add it to the new node again after the node replacement is complete.
- If the node you want to replace is a data node (ess) or cold data node (ess-cold), pay attention to the following precautions:
 - a. Before a data node or cold data node is replaced, its data needs to be migrated to other nodes. To properly store the data, ensure the maximum sum of replicas and primary shards of an index is smaller than the total number of data nodes (ess and ess-cold nodes) in the cluster. The node replacement duration depends heavily on the migration speed.
 - b. Clusters whose version is earlier than 7.6.2 cannot have closed indexes. Otherwise, data nodes or cold data nodes cannot be replaced.
 - c. The AZ of the node to be replaced must have two or more data nodes (including ess and ess-cold).
 - d. If the cluster of the node to be replaced does not have a master node (ess-master), the number of available data nodes (including ess and ess-cold) in the cluster must be greater than or equal to 3.
 - e. The preceding precautions do not apply if you are replacing a master node (ess-master) or client node (ess-client).
 - f. The precautions 1 to 4 do not apply if you are replacing a faulty node, regardless of its type. Faulty nodes are not included in `_cat/nodes`.

Procedure

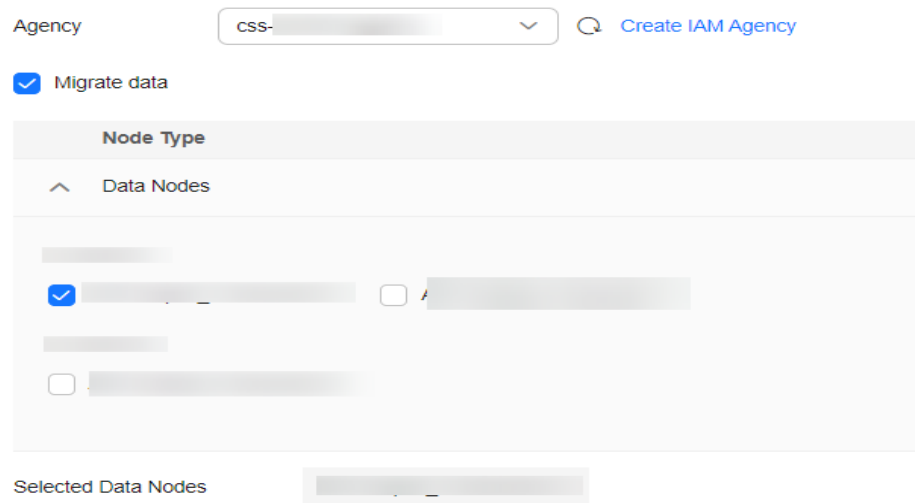
1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, click the **Replace Node** tab.
5. On the **Replace Node** tab page, set the following parameters:
 - **Agency**: Select an IAM agency to grant the current account the permission to switch AZs.
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

NOTE

- The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.
- **Whether to perform data migration**: If this option is selected, data migration is performed. If the target node has disabled indexes or indexes that have no replicas, this option must be selected.

- Select the node to be replaced in the data node table.

Figure 2-35 Replacing a specified node



6. Click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Upgrading**. When **Cluster Status** changes to **Available**, the node has been successfully replaced.

2.6.7 Adding Master/Client Nodes

If workloads on the data plane of a cluster increase, you can add master or client nodes as needed. Services are not interrupted while they are added.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- If a cluster already has master and client nodes, the **Add Master/Client Node** tab is not displayed on the **Modify Configuration** page. In this case, you need to add the master or client nodes by referring to [Scaling Out a Cluster](#).
- When you add master or client nodes, the number of nodes that can be configured varies depending on the node type. For details, see [Table 2-37](#).

Table 2-37 Number of nodes in different types

Node Flavor	Number
Master node	An odd number ranging from 3 to 9
Client node	1 to 32

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Add Master/Client Node** tab.
5. Select the target node type and set the node specifications, quantity, and storage.
 - Master and client nodes cannot be added at the same time.
 - If a cluster already has a master or client node, you can only add nodes of the other type.

Figure 2-36 Adding a master or client node

The screenshot shows a configuration interface for adding nodes. At the top, there are two radio buttons: 'Master node' (selected) and 'Client node'. Below this is a table with four columns: 'Type', 'Node Specifications', 'Nodes', and 'Node Storage Type'. The 'Type' column contains 'Master Node'. The 'Node Specifications' column contains a dropdown menu with 'ess.spec-4u8g(sold out)' selected. The 'Nodes' column contains a text input with '3' and minus/plus buttons. The 'Node Storage Type' column contains a dropdown menu with 'Common I/O' selected and '40 GB' displayed.

6. Click **Next**.
 7. Confirm the information and click **Submit**.
- Return to the cluster list page. The **Task Status** of the cluster is **Scaling out**.
- If you added a master node and **Cluster Status** changed to **Available**, the master node has been successfully added.

NOTICE

If the cluster version is earlier than 7.x, when the **Cluster Status** changes to **Available**, you need to restart all data nodes and cold data nodes in the cluster to make the new node take effect. If the data nodes and cold data nodes are not restarted, the cluster may be reported as unavailable. (The cluster services still run properly.) For details, see [Restarting a Cluster](#).

- If you added a client node and **Cluster Status** changed to **Available**, the client node has been added. You can restart data nodes and cold data nodes to shut down Cerebro and Kibana processes on the nodes.

2.6.8 Changing the Security Mode

After a cluster is created, its security mode can be changed using the following methods:

- [Switching from the Non-Security Mode to Security Mode](#)
- [Switching from the Security to Non-Security Mode](#)
- [Switching the Protocol of Security Clusters](#)

Context

You can create clusters in multiple security modes. For details about the differences between security modes, see [Table 2-38](#).

Table 2-38 Cluster security modes

Security Mode	Scenario	Advantage	Disadvantage
Non-Security Mode	Intranet services and test scenarios	Simple. Easy to access.	Poor security. Anyone can access such clusters.
Security Mode + HTTP Protocol	User permissions can be isolated, which is applicable to scenarios sensitive to cluster performance.	Security authentication is required for accessing such clusters, which improves cluster security. Accessing a cluster through HTTP protocol can retain the high performance of the cluster.	Cannot be accessed from the public network.
Security Mode + HTTPS Protocol	Scenarios that require high security and public network access.	Security authentication is required for accessing such clusters, which improves cluster security. HTTPS protocol allows public network to access such clusters.	The performance of clusters using HTTPS is 20% lower than that of using HTTP.

Prerequisites

- You are advised to back up data before changing the cluster security mode.
- The target cluster is available and has no tasks in progress.

Constraints

- Only clusters (whose version is 6.5.4 or later) created after November 2022 support security mode switching.
- A cluster automatically restarts when its security mode is being changed. Services are interrupted during the restart. The authentication mode for invoking the cluster will change after the restart, and client configurations need to be adjusted accordingly.

- If a cluster has already opened the Kibana session box, a session error message will be displayed after you change the cluster security mode. In this case, clear the cache and open Kibana again.

Switching from the Non-Security Mode to Security Mode

You can change a non-security cluster to a security cluster that uses HTTP or HTTPS. After a cluster's security mode is enabled, security authentication is required for accessing the cluster.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters > Elasticsearch**. The Elasticsearch cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. Choose the **Configure Security Mode** tab.
5. Enable the security mode. Enter and confirm the administrator password of the cluster.

Figure 2-37 Enabling the security mode

6. Enable or disable **HTTPS Access**.
 - If you enable **HTTPS Access**: The HTTPS protocol is used to encrypt cluster communication and you can configure public networks to access the cluster.
 - If you disable **HTTPS Access**: The HTTP protocol is used and you cannot configure public networks to access the cluster.
7. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

Switching from the Security to Non-Security Mode

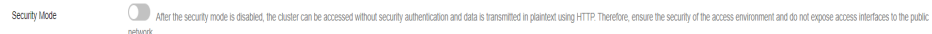
You can change a security cluster that uses HTTP or HTTPS to a non-security cluster. After a cluster's security mode is disabled, security authentication is no longer required for accessing the cluster.

NOTICE

- Clusters in non-security mode can be accessed without security authentication, and HTTP protocol is used to transmit data. Ensure the security of the cluster access environment and do not expose the access interface to the public network.
- During the switchover from the security mode to the non-security mode, the indexes of the original security cluster will be deleted. Back up data before disabling the security mode.
- If a security cluster has been bound to a public IP address, unbind it before changing the security mode.
- If a security cluster has enabled Kibana public network access, disable it before changing the security mode.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters**. On the displayed **Clusters** page, locate the target cluster and choose **More > Modify Configuration** in the **Operation** column.
3. Choose the **Configure Security Mode** tab.
4. Disable the security mode.

Figure 2-38 Disabling the security mode



5. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

Switching the Protocol of Security Clusters

You can change the protocol of a security cluster.

NOTICE

If a security cluster has been bound to a public IP address, you need to unbind it before changing HTTPS protocol to HTTP.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters**. On the displayed **Clusters** page, locate the target cluster and choose **More > Modify Configuration** in the **Operation** column.
3. Choose the **Configure Security Mode** tab.
4. Enable or disable **HTTPS Access**.

Figure 2-39 Configuring the protocol

Security Mode If you enable this option, you will need to verify your identity to access the cluster

Administrator Username

Administrator Password

Confirm Password

HTTPS Access If you enable this option, communication will be encrypted when you access the cluster.

- If you enable **HTTPS Access**:
HTTPS protocol is used to encrypt cluster communication and you can configure public network access.
 - If you disable **HTTPS Access**: An alarm message is displayed. Click **OK** to disable the function.
Cluster communication is no longer encrypted and the public network access function cannot be enabled.
5. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

2.6.9 Changing AZs

CSS supports cross-AZ deployment. You can add an AZ to obtain more resources or improve cluster availability, and can migrate your current AZ to one with higher specifications. This section describes how to add or migrate your AZs.

Description

You can **Add AZ** or **Migrate AZ**.

- **Add AZ**: Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster to improve cluster availability.
- **Migrate AZ**: Completely migrate data from the current AZ to another AZ that has sufficient resources.

Prerequisites

- Ensure that an AZ with sufficient resources exists.
- The target cluster is available and has no tasks in progress.
- Make sure that no non-standard operations have been performed in the cluster. If you have made non-standard modifications, such as modifying return routes, system parameters, and Kibana configurations, these modifications will be lost after the AZ change and your services may be affected.

Constraints

- To ensure service continuity, the total number of data nodes and cold data nodes in a cluster must be greater than or equal to 3.
- During the change, nodes are brought offline one by one and then new nodes are created. Ensure that the disk capacity of other nodes can store all the data of the node after a single node is brought offline.
- To prevent backup allocation failures after a node is brought offline during the change, ensure that the maximum number of primary and standby index shards of an index can be allocated to the remaining data nodes and cold data nodes. That is, the maximum number of primary and standby shards of an index plus 1 is less than or equal to the total number of data nodes and cold data nodes in the current cluster.
- You are advised to back up data before the change to prevent data loss caused by upgrade faults.
- Before a change completes, some nodes may have been moved to a new AZ. In this case, the AZs before and after the change are both displayed. After the change succeeds, the new AZs and their nodes will be displayed properly.
- When adding AZs, the current AZ must be retained in the change. When adding one or two AZs to a single-AZ cluster, you must change AZs for all nodes at the same time. When adding an AZ to a dual-AZ cluster, you can change AZs for a single type of nodes or all nodes in a cluster at a time. For example, in a cluster using the dual-AZ architecture, you can use the three-AZ architecture for master nodes alone. During HA modification, the nodes with the smallest configurations are modified to rebuild the cluster. After the HA modification is complete, the YML configuration of the nodes that are not modified is also updated. You need to restart the cluster to make the modification take effect.
- When migrating an AZ, you can select only one target AZ. You can migrate AZs for a single type of nodes or all nodes in a cluster at a time. For example, in a cluster with two AZs, you can migrate the AZ of the master node to the other AZ. After adding AZs, you need to restart the cluster to make the modification take effect.

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. Click the **Change AZ** tab.
5. On the **Change AZ** page, set parameters.

Table 2-39 Parameters for changing AZs

Parameter	Description
Operation Type	<ul style="list-style-type: none"> • Add AZ: Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster. During HA modification, the nodes with the smallest configurations are modified to rebuild the cluster. After the HA modification is complete, the YML configuration of the nodes that are not modified is also updated. You need to restart the cluster to make the modification take effect. • Migrate AZ: Migrate data from one AZ to another. After adding AZs, you need to restart the cluster to make the modification take effect.
Node Type	Select a type of node or All nodes to change their AZ. NOTE When adding one or two AZs to a single-AZ cluster, you can only select All nodes to change AZs for all nodes at a time.
Current AZ	Current AZ of a cluster
Target AZ	Target AZ. <ul style="list-style-type: none"> • Add AZ: Select up to three AZs, which must include all your current AZs. • Migrate AZ: Select only one target AZ, which cannot be your current AZ.
Agency	Select an IAM agency to grant the current account the permission to change AZs. If no agencies are available, click Create IAM Agency to go to the IAM console and create an agency. NOTE The selected agency must be authorized with the Tenant Administrator or VPC Administrator policy.

6. Click **Submit**. Determine whether to check for the backup of all indexes and click **OK** to start the change.

Figure 2-40 Checking full index snapshots

Check Full Index Snapshot

You are advised to check the full index snapshot before the change.
 If you have backed up data, check whether the backup period exceeds one month. If yes, back up the latest data.
 Currently, only the index name can be matched, and the specific content and backup time cannot be verified.

Check full index snapshot

Cancel **OK**

7. The current AZ change task is displayed in the task list. If the task status is **Running**, expand the task list and click **View Progress** to view the progress details.

If the task status is **Failed**, you can retry or terminate the task.

- Retry a task: Click **Retry** in the **Operation** column of a task.
- Terminate a task: Click **Terminate** in the **Operation** column of a task.

If the AZ of the original node is not changed after the task is terminated, you can recover the node by referring to [Replacing a Specified Node](#).

NOTE

If the AZ of some nodes has been changed and the AZ form of the cluster has changed, stopping the switchover task may make the delivery of the previous switchover request fail. Exercise caution when stopping the switchover task.

2.7 Upgrading the Cluster Version

Same-version upgrade, cross-engine upgrade, and cross-version upgrade are supported. Same-version upgrade is to upgrade the kernel patch of a cluster to fix problems or optimize performance. Cross-engine upgrade is to upgrade an Elasticsearch cluster to an OpenSearch cluster. Cross-version upgrade is to upgrade the cluster version to enhance functions or incorporate versions.

Description

Principle

Nodes in the cluster are upgraded one by one so that services are not interrupted. The upgrade process is as follows: bring a node offline, migrate its data to another node, create a new node of the target version, and mount the NIC ports of the offline node to the new node to retain the node IP address. After a new node is added to the cluster, other nodes will be updated in the same way in sequence. If there is a large amount of data in a cluster, the upgrade duration depends on the data migration duration.

Process

1. [Pre-Upgrade Check](#)
2. [Creating a Snapshot](#)
3. [Creating an Upgrade Task](#)

Version Restrictions

The supported target version varies according to the current cluster version. For details, see [Table 2-40](#).

Table 2-40 Version restrictions

Current Version	Target Version
Elasticsearch: 6.2.3	Elasticsearch: 6.5.4 or 6.8.23
Elasticsearch: 6.5.4	Elasticsearch: 6.8.23
Elasticsearch: 6.8.23	Elasticsearch: 7.6.2 or 7.10.2
Elasticsearch: 7.1.1	Elasticsearch: 7.6.2 or 7.10.2

Current Version	Target Version
Elasticsearch: 7.6.2	Elasticsearch: 7.10.2
Elasticsearch: 7.9.3	Elasticsearch: 7.10.2
Elasticsearch: 7.10.2	OpenSearch: 1.3.6
<p>Note:</p> <ul style="list-style-type: none"> • Elasticsearch 7.6.2 and 7.10.2 are mainstream cluster versions. You are advised to upgrade your clusters to these two versions. The supported target versions are displayed in the drop-down list of Target Image. • Elasticsearch clusters of version 5.X cannot be upgraded across versions. Elasticsearch clusters of versions 6.2.3 and 6.5.4 can be upgraded to 6.8.23 and then to 7.X.X. • Currently, only Elasticsearch clusters of version 7.10.2 can be upgraded to OpenSearch clusters of version 1.3.6 across engines. 	

Constraints

- A maximum of 20 clusters can be upgraded at the same time. You are advised to perform the upgrade during off-peak hours.
- Clusters that have ongoing tasks cannot be upgraded.
- Once started, an upgrade task cannot be stopped until it succeeds or fails.
- During the upgrade, nodes are replaced one by one. Requests sent to a node that is being replaced may fail. In this case, you are advised to access the cluster through the VPC Endpoint service or a dedicated load balancer.
- During the upgrade, the Kibana and Cerebro components will be rebuilt and cannot be accessed. Different Kibana versions are incompatible with each other. During the upgrade, you may fail to access Kibana due to version incompatibility. A cluster can be accessed after it is successfully upgraded.

Pre-Upgrade Check

To ensure a successful upgrade, you must check the items listed in the following table before performing an upgrade.

Table 2-41 Pre-upgrade checklist

Check Item	Check Method	Description	Normal Status
Cluster status	System check	After an upgrade task is started, the system automatically checks the cluster status. Clusters whose status is green or yellow can provide services properly and have no unallocated primary shards.	The cluster status is Available .
Node quantity	System check	After an upgrade task is started, the system automatically checks the number of nodes. The total number of data nodes and cold data nodes in a cluster must be greater than or equal to 3 so that services will not be interrupted.	The total number of data nodes and cold data nodes in a cluster must be greater than or equal to 3.
Disk capacity	System check	After an upgrade task is started, the system automatically checks the disk capacity. During the upgrade, nodes are brought offline one by one and then new nodes are created. Ensure that the disk capacity of all the remaining nodes can process all data of the node that has been brought offline.	After a node is brought offline, the remaining nodes can contain all data of the cluster.
Data backup	System check	Check whether the maximum number of primary and standby shards of indexes in a cluster can be allocated to the remaining data nodes and cold data nodes. Prevent backup allocation failures after a node is brought offline during the upgrade.	Maximum number of primary and standby index shards plus 1 must be less than or equal to the total number of data nodes and cold data nodes before the upgrade.
Data backup	System check	Before the upgrade, back up data to prevent data loss caused by upgrade faults. When submitting an upgrade task, you can determine whether to enable the system to check for the backup of all indexes.	Check whether data has been backed up.

Check Item	Check Method	Description	Normal Status
Resources	System check	After an upgrade task is started, the system automatically checks resources. Resources will be created during the upgrade. Ensure that resources are available.	Resources are available and sufficient.
Custom plugins	System and manual check	<p>Perform this check only when custom plugins are installed in the source cluster. If a cluster has a custom plugin, upload all plugin packages of the target version on the plugin management page before the upgrade. During the upgrade, install the custom plugin in the new nodes. Otherwise, the custom plugins will be lost after the cluster is successfully upgraded. After an upgrade task is started, the system automatically checks whether the custom plugin package has been uploaded, but you need to check whether the uploaded plugin package is correct.</p> <p>NOTE If the uploaded plugin package is incorrect or incompatible, the plugin package cannot be automatically installed during the upgrade. As a result, the upgrade task fails. To restore a cluster, you can terminate the upgrade task and restore the node that fails to be upgraded by Replacing a Specified Node.</p> <p>After the upgrade is complete, the status of the custom plugin is reset to Uploaded.</p>	The plugin package of the cluster to be upgraded has been uploaded to the plugin list.
Custom configurations	System check	During the upgrade, the system automatically synchronizes the content of the cluster configuration file elasticsearch.yml .	Clusters' custom configurations are not lost after the upgrade.

Check Item	Check Method	Description	Normal Status
Non-standard operations	Manual check	Check whether non-standard operations are contained in the upgrade. Non-standard operations refer to manual operations that are not recorded. These operations cannot be automatically transferred during the upgrade, for example, modification of the Kibana.yml configuration file, system configuration, and route return.	Some non-standard operations are compatible. For example, the modification of a security plugin can be retained through metadata, and the modification of system configuration can be retained using images. Some non-standard operations, such as the modification of the kibana.yml file, cannot be retained, and you must back up the file in advance.
Compatibility check	System and manual check	After a cross-version upgrade task is started, the system automatically checks whether the source and target versions have incompatible configurations. If a custom plugin is installed for a cluster, the version compatibility of the custom plugin needs to be manually checked.	Configurations before and after the cross-version upgrade are compatible.

Creating an Upgrade Task

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters**. On the cluster list page that is displayed, click the name of a cluster.
3. On the displayed basic cluster information page, click **Version Upgrade**.
4. On the displayed page, set upgrade parameters.

Table 2-42 Upgrade parameters

Parameter	Description
Upgrade Type	<ul style="list-style-type: none"> • Same-version upgrade: Upgrade the kernel patch of the cluster. The cluster version number remains unchanged. • Cross-version upgrade: Upgrade the cluster version. • Cross-engine upgrade: Upgrade an Elasticsearch cluster to an OpenSearch cluster. Currently, only the Elasticsearch cluster of version 7.10.2 can be upgraded to the OpenSearch cluster of version 1.3.6.
Target Image	<p>Image of the target version. When you select an image, the image name and target version details are displayed.</p> <p>The supported target versions are displayed in the drop-down list of Target Image. If the target image cannot be selected, the possible causes are as follows:</p> <ul style="list-style-type: none"> • The current cluster is of the latest version. • The current cluster is created before 2023 and has vector indexes. • The new version images have not been added at the current region.
Agency	<p>Select an IAM agency to grant the upgrade permission to the current account.</p> <p>If no agency is available, click Create Agency to go to the IAM console and create an agency.</p> <p>NOTE The selected agency must be assigned the Tenant Administrator or VPC Administrator policy.</p>

5. After setting the parameters, click **Submit**. Determine whether to enable **Check full index snapshot** and **Perform cluster load detection** and click **OK**.
If a cluster is overloaded, the upgrade task may suspend or fail. Enabling **Cluster load detection** can effectively avoid failures. If any of the following situations occurs during the detection, wait or reduce the load. If you urgently need to upgrade the version and you have understood the upgrade failure risks, you can disable the **Cluster load detection** function. The cluster load detection items are as follows:
 - **nodes.thread_pool.search.queue < 1000**: check whether the maximum number of search queues is less than 1000.
 - **nodes.thread_pool.write.queue < 200**: Check whether the maximum number of write queues is less than 200.
 - **nodes.process.cpu.percent < 90**: Check whether the maximum CPU usage is less than 90%.
 - **nodes.os.cpu.load_average/Number of CPU cores < 80%**: Check whether the ratio of the maximum load to the number of CPU cores is less than 80%.
6. View the upgrade task in the task list. If the task status is **Running**, you can expand the task list and click **View Progress** to view the upgrade progress.

If the task status is **Failed**, you can retry or terminate the task.

- Retry the task: Click **Retry** in the **Operation** column.
- Terminate the task: Click **Terminate** in the **Operation** column.

NOTICE

- Same version upgrade: If the upgrade task status is **Failed**, you can terminate the upgrade task.
- Cross version upgrade: You can stop an upgrade task only when the task status is **Failed** and no node has been upgraded.

After an upgrade task is terminated, the **Task Status** of the cluster is rolled back to the status before the upgrade, and other tasks in the cluster are not affected.

2.8 Cluster Management

2.8.1 Cluster List Overview

The cluster list displays all CSS clusters. If there are a large number of clusters, these clusters will be displayed on multiple pages. You can view clusters of all statuses from the cluster list.

Clusters are listed in chronological order by default in the cluster list, with the most recent cluster displayed at the top. [Table 2-43](#) shows the cluster parameters.




In the upper right corner of the cluster list, you can enter the cluster name, enterprise project, private IP address, or cluster ID and click  to search for a cluster. You can also click  in the upper right corner to refresh the cluster list. Click  to download the cluster list.

Table 2-43 Cluster list parameter description

Parameter	Description
Name/ID	Name and ID of a cluster. You can click a cluster name to switch to the Basic Information page. The cluster ID is automatically generated by the system and uniquely identifies a cluster.
Cluster Status	Status of a cluster. For details about the cluster status, see Viewing the Cluster Runtime Status and Storage Capacity Status .
Task Status	Status of a task, such as cluster restart, cluster capacity expansion, cluster backup, and cluster restoration.
Version	Elasticsearch version of the cluster.


Parameter	Description
Created	Time when the cluster is created.
Enterprise Project	Enterprise project that a cluster belongs to.
Private Network Address	Private network address and port number of the cluster. You can use these parameters to access the cluster. If the cluster has multiple nodes, the private network addresses and port numbers of all nodes are displayed.
Billing Mode	Billing mode of a cluster. It can be Pay-per-use or Yearly/Monthly .
Operation	Operations that can be performed on a cluster, including accessing Kibana, checking metrics, restarting a cluster, and deleting a cluster. If an operation is not allowed, the button is gray.

2.8.2 Viewing Basic Cluster Information

On the **Cluster Information** page, you can view the information about a cluster, including the private network address, public IP address, version, and node.

1. Log in to the CSS management console.
2. Choose **Clusters > Elasticsearch**. The cluster list is displayed.
3. Click a cluster name to go to the **Cluster Information** page and view the basic information about the cluster.

Table 2-44 Parameters for configuring basic information

Type	Parameter	Description
Cluster Information	Name	Cluster name. The name can be customized. You can click  on the right to change the cluster name.
	ID	Unique ID of a cluster, which is automatically generated by the system. Each cluster in the same region has a unique ID.
	Version	Cluster version information. For details about how to upgrade the cluster version, see Upgrading the Cluster Version .
	Cluster Status	Current status of a cluster

Type	Parameter	Description
	Task Status	Current task status of a cluster. If no task is in progress, -- is displayed.
	Created	Time when a cluster was created
	Cluster Storage Capacity (GB)	Storage capacity of a cluster
	Used Cluster Storage (GB)	Used storage capacity of a cluster
Configuration	Region	Region where a cluster is located
	AZ	AZ where a cluster is located
	VPC	VPC to which the cluster belongs
	Subnet	Subnet to which the cluster belongs
	Security Group	<p>Security group to which a cluster belongs.</p> <p>To change the security group of a cluster, click Change Security Group on the right.</p> <p>NOTICE Before changing the security group, ensure that the port 9200 required for service access has been enabled. Incorrect security group configuration may cause service access failures. Exercise caution when performing this operation.</p>
	Security Mode	<p>Security mode of a cluster.</p> <ul style="list-style-type: none"> Enabled: The current cluster is a security cluster. Disabled: The current cluster is a non-security cluster. <p>For details about how to change the security mode of a cluster, see Changing the Security Mode.</p>

Type	Parameter	Description
	Reset Password	<p>This parameter is displayed only for security clusters.</p> <p>Click Reset to change the password of the administrator account admin of the security cluster.</p> <p>NOTE Requirements for administrator passwords:</p> <ul style="list-style-type: none"> • The password can contain 8 to 32 characters. • The password must contain at least three of the following character types: uppercase letters, lowercase letters, digits, and special characters. The following special characters are supported: ~!@#%&*()- _ = + \ [] ; : , < . > / ? • Do not use the administrator name, or the administrator name spelled backwards. • You are advised to change the password periodically.
	Enterprise Project	<p>Enterprise project to which a cluster belongs.</p> <p>You can click the project name to view the basic information about the enterprise project.</p>
	Public IP Address	<p>Public network access information, which is displayed only for clusters in security mode.</p> <ul style="list-style-type: none"> • For a security cluster with public network access enabled, the configured public network address is displayed. You can use this address to access the security cluster from the public network. • For a security cluster with public network access disabled, -- is displayed. <p>When using a public IP address to access a cluster, you are advised to enable access control and configure an access whitelist to improve cluster security. For details about how to configure the public network access, see Accessing a Cluster from a Public Network.</p>

Type	Parameter	Description
	Access Control	<p>Whether to set access control for a cluster. This parameter is displayed only for clusters with public network access enabled.</p> <ul style="list-style-type: none"> • Enabled: Only IP addresses in the whitelist can access the cluster through the public network. • Disabled: Any IP address can access the cluster through the public network. <p>Click Set to configure the access control and the whitelist.</p>
	Bandwidth	<p>The bandwidth for public network access. This parameter is displayed only for clusters with public network access enabled.</p> <p>Click Edit to change the bandwidth size.</p>
	HTTPS Access	<p>Whether to enable the HTTPS access protocol for a cluster.</p> <ul style="list-style-type: none"> • Disabled: The HTTP protocol is used for cluster access. • Enabled: The HTTPS protocol is used for cluster access. Only security clusters can enable this function. If HTTPS Access is enabled, you can click Download Certificate to obtain the CER security certificate for accessing the security cluster. Currently, the security certificate cannot be used in the public network environment. <p>For details about how to change the access mode of a cluster in security mode, see Switching the Protocol of Security Clusters.</p>

Type	Parameter	Description
	Private IPv4 Address	Private IP address and port number of a cluster, which can be used to access the cluster. If the cluster has only one node, the IP address and port number of only one node are displayed, for example, 10.62.179.32:9200 . If the cluster has multiple nodes, the IP addresses and port numbers of all nodes are displayed, for example, 10.62.179.32:9200,10.62.179.33:9200 .
Node	Node Specifications	Specifications of nodes in a cluster
	Node Storage Type	Storage capacity and storage type of nodes in a cluster
	Nodes	Number of nodes in a cluster

2.8.3 Managing Tags

Tags are cluster identifiers. Adding tags to clusters can help you identify and manage your cluster resources.

You can add tags to a cluster when creating the cluster or add them on the details page of the created cluster.

If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.

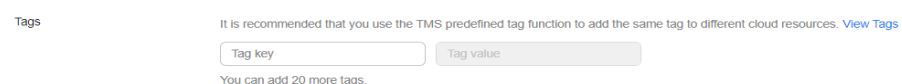
Managing Tags of a New Cluster

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner. The **Create Cluster** page is displayed.
3. On the **Create Cluster** page, set **Advanced Settings** to **Custom**. Add tags for a cluster.

You can select a predefined tag and set **Tag value** for the tag. You can click **View Predefined Tag** to switch to the TMS management console and view existing tags.

You can also create new tags by specifying **Tag key** and **Tag value**.

Figure 2-41 Adding tags during cluster creation



You can add a maximum of 10 tags for a CSS cluster. If the entered tag is incorrect, you can click **Delete** on the right of the tag to delete the tag.

Table 2-45 Naming rules for a tag key and value

Parameter	Description
Tag key	<p>Must be unique in a cluster.</p> <p>The value cannot contain more than 64 characters.</p> <p>It can contain only numbers, letters, Chinese characters, and the following special characters: <code>_.:=-@</code> The value cannot start or end with a space.</p> <p>Cannot be left blank.</p>
Tag value	<p>The value cannot contain more than 64 characters.</p> <p>It can contain only numbers, letters, Chinese characters, and the following special characters: <code>_.:=-@</code> The value cannot start or end with a space.</p> <p>Cannot be left blank.</p>

Managing Tags of Existing Clusters

You can modify, delete, or add tags for a cluster.

1. Log in to the CSS management console.
2. On the **Clusters** page, click the name of a cluster for which you want to manage tags.

The **Basic Information** page is displayed.

3. In the navigation pane on the left, choose the **Tags** tab. You can add, modify, or delete tags.
 - View

On the **Tags** page, you can view details about tags of the cluster, including the number of tags and the key and value of each tag.
 - Add

Click **Add** in the upper left corner. In the displayed **Add Tag** dialog box, enter the key and value of the tag to be added, and click **OK**.
 - Modify

You can only change the value of an existing tag.

In the **Operation** column of a tag, click **Edit**. In the displayed **Edit Tag** page, enter a new tag value and click **OK**.
 - Delete

In the **Operation** column of a tag, click **Delete**. After confirmation, click **Yes** on the displayed **Delete Tag** page.

Searching for Clusters by Tag

1. Log in to the CSS management console.

2. On the **Clusters** page, click **Search by Tag** in the upper right corner of the cluster list.
3. Select or enter the tag key and tag value you want to search for, and click **Add** to add the tag to the search text box.
You can select a tag key or tag value from their drop-down lists. The system returns a list of clusters that exactly match the tag key or tag value. If you enter multiple tags, the cluster that meets requirements of all the tags will be filtered.
You can add a maximum of 10 tags at one time.
4. Click **Search**.
The system searches for the target cluster by tag key and value.

2.8.4 Managing Logs

CSS provides log backup and search functions to help you locate faults. You can back up cluster logs to OBS buckets and download required log files to analyze and locate faults.

If logs are backed up in OBS buckets, extra fees are charged. For details, see the [Billing Modes](#).

Log Query

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. In the navigation pane on the left, choose **Log Management**.
4. Query logs on the log management page.
Select the node, log type, and log level you want to query, and then click



. The query result is displayed.

When you search for logs, the latest 10,000 logs are matched. A maximum of 100 logs are displayed.

Enabling Log Backup

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. Click the **Logs** tab and toggle on the **Log Management** switch.
4. In the **Edit Log Backup Configuration** dialog box, set the parameters.
In the displayed dialog box, **OBS Bucket**, **Backup Path**, and **IAM Agency** are automatically created for log backup. You can change the default value by referring to [Table 2-46](#).


If the **Log Management** function has been enabled for the cluster, you can click  on the right of **Log Backup Configuration** and modify the configuration in the displayed **Edit Log Backup Configuration** dialog box. For details, see [Table 2-46](#).

Table 2-46 Parameters for configuring log backup


Parameter	Description	Remarks
OBS Bucket	Select an OBS bucket from the drop-down list for storing logs. You can also click Create Bucket on the right to create an OBS bucket.	The OBS bucket and the cluster must be in the same region. NOTE To let an IAM user access an OBS bucket, you need to grant the GetBucketStoragePolicy , GetBucketLocation , ListBucket , and ListAllMyBuckets permissions to the user.
Backup Path	Storage path of logs in the OBS bucket	The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The total length of the backup path cannot exceed 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in the OBS bucket. You can also click Create IAM Agency on the right to create an IAM agency.	The IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • Mandatory policies: OBS Administrator

5. Back up logs.

- Automatically backing up logs

Click the icon on the right of **Auto Backup** to enable the auto backup function.

After the automatic backup function is enabled, set the backup start time in the **Configure Auto Backup** dialog box. When the scheduled time arrives, the system will back up logs automatically.

After the **Automatic Snapshot Creation** function is enabled, you can click  on the right of the parameter to change the backup start time.

- Manually backing up logs

On the **Log Backup** tab page, click **Back Up**. On the displayed page, click **Yes** to start backup.


If **Task Status** in the log backup list is **Successful**, the backup is successful.

 **NOTE**

All logs in the cluster are copied to a specified OBS path. You can view or download log files from the path of the OBS bucket.

6. Search for logs.

On the **Log Search** page, select the target node, log type, and log level, and

click  . The search results are displayed.

When you search for logs, the latest 10,000 logs are matched. A maximum of 100 logs are displayed.

Viewing Logs

After backing up logs, you can click **Backup Path** to go to the OBS console and view the logs.

Figure 2-42 Accessing OBS



Backed up logs mainly include deprecation logs, run logs, index slow logs, and search slow logs. [Table 2-47](#) lists the storage types of the OBS bucket.

Table 2-47 Log types

Log Name	Description
clustername_deprecation.log	Deprecation log
clustername_index_indexing_slowlog.log	Search slow log

Log Name	Description
clustername_index_search_slowlog.log	Index slow log
clustername.log	Elasticsearch run log
clustername_access.log	Access log

2.8.5 Configuring YML Parameters

You can modify the `elasticsearch.yml` file.

Modifying Parameter Configurations

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. Click **Parameter Configurations** and click **Edit** to modify module parameters as required.

Table 2-48 Module parameters

Module Name	Parameter	Description
Cross-domain Access	http.cors.allow-credentials	Whether to return the Access-Control-Allow-Credentials of the header during cross-domain access Value: true or false Default value: false
	http.cors.allow-origin	Origin IP address allowed for cross-domain access, for example, 122.122.122.122:9200
	http.cors.max-age	Cache duration of the browser. The cache is automatically cleared after the time range you specify. Unit: s Default value: 1,728,000
	http.cors.allow-headers	Headers allowed for cross-domain access, including X-Requested-With , Content-Type , and Content-Length . Use commas (,) and spaces to separate headers.
	http.cors.enabled	Whether to allow cross-domain access Value: true or false Default value: false

Module Name	Parameter	Description
	http.cors.allow-methods	Methods allowed for cross-domain access, including OPTIONS , HEAD , GET , POST , PUT , and DELETE . Use commas (,) and spaces to separate methods.
Reindexing	reindex.remote.whitelist	Configured for migrating data from the current cluster to the target cluster through the reindex API. The example value is 122.122.122.122:9200.
Custom Cache	indices.queries.cache.size	Cache size in the query phase Value range: 1 to 100 Unit: % Default value: 10%
Queue Size in a Thread Pool	thread_pool.force_merge.size	Queue size in the force merge thread pool. The value is an integer. Default value: 1
Customize	You can add parameters based on your needs.	Customized parameters NOTE <ul style="list-style-type: none"> Enter multiple values in the format as [value1, value2, value3...]. Separate values by commas (,) and spaces. Colons (:) are not allowed.

- After the modification is complete, click **Submit**. In the displayed **Submit Configuration** dialog box, select the box indicating "I understand that the modification will take effect after the cluster is restarted." and click **Yes**.
If the **Status** is **Succeeded** in the parameter modification list, the modification has been saved. Up to 20 modification records can be displayed.
- Return to the cluster list and choose **More > Restart** in the **Operation** column to restart the cluster and make the modification take effect.
 - You need to restart the cluster after modification, or **Configuration unupdated** will be displayed in the **Task Status** column on the **Clusters** page.
 - If you restart the cluster after the modification, and **Task Status** displays **Configuration error**, the parameter configuration file fails to be modified.

2.8.6 Viewing the Default Plugin List

CSS clusters have default plugins. You can view the default plugin information on the console or Kibana.

Viewing Plugins on the Console

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters**. Click the target cluster name and go to the **Basic Information** page of the cluster.
3. Click the **Plugins** tab.
4. On the **Default** page, view default plugins supported by the current version.

Viewing Plugins on the Kibana

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters**. Locate the target cluster and click **Access Kibana** in the **Operation** column to log in to Kibana.
 - Non-security cluster: The Kibana console is displayed.
 - Security cluster: Enter the username and password on the login page and click **Log In** to go to the Kibana console. The default username is **admin** and the password is the administrator password you specified during cluster creation.
3. Go to **Dev Tools** and run the following command to view the cluster plugin information:

```
GET _cat/plugins?v
```

The following is an example of the response body:

name	component	version
css-test-ess-esn-1-1	analysis-dynamic-synonym	7.6.2-xxxx-ei-css-v1.0.1
css-test-ess-esn-1-1	analysis-icu	7.6.2-xxxx-ei-css-v1.1.6
css-test-ess-esn-1-1	analysis-ik	7.6.2-xxxx-ei-css-v1.0.1
.....		

name indicates the cluster node name, **component** indicates the plugin name, and **version** indicates the plugin version.

2.8.7 Binding an Enterprise Project

You can create enterprise projects based on your organizational structure. Then you can manage resources across different regions by enterprise project, add users and user groups to enterprise projects, and grant different permissions to the users and user groups. This section describes how to bind a CSS cluster to an enterprise project and how to modify an enterprise project.

Prerequisites

To use the enterprise project function, you need to assign permissions to the corresponding account. You can [submit a service ticket](#) to apply for the permissions.

Before binding an enterprise project, you have [created an enterprise project](#).

Binding an Enterprise Project

When creating a cluster, you can bind an existing enterprise project to the cluster, or click **View Enterprise Project** to go to the enterprise project management console and create a new project or view existing projects.

Modifying an Enterprise Project

For a cluster that has been created, you can modify its enterprise project based on the site requirements.

1. Log in to the CSS management console.
2. In the navigation pane on the left, select a cluster type. The cluster management page is displayed.
3. In the cluster list on the displayed page, click the target cluster name to switch to the **Cluster Information** page.
4. On the **Cluster Information** page, click the enterprise project name on the right of **Enterprise Project**. The project management page is displayed.

Figure 2-43 Enterprise Project

Configuration

Region	
AZ	
VPC	vpc-
Subnet	subnet-
Security Group	dws- Change Security Group
Security Mode	Disabled
Enterprise Project	default
HTTPS Access	Disabled
IPv4 Access Address	192-

5. On the **Resources** tab page, select the region of the current cluster, and select **CSS** for **Service**. In this case, the corresponding CSS cluster is displayed in the resource list.

Figure 2-44 Filtering CSS clusters

The screenshot shows the 'Resources' tab with a grid of services. The 'CloudSearchService' is selected. Below the grid, the 'Resource Type' is set to 'Cluster' and the 'Selected' section shows 'Service: CloudSearchService' and 'Region: [Current Region]'. A 'Search' button is visible at the bottom right.

6. Select the cluster whose enterprise project you want to modify and click **Remove**.
7. On the **Remove Resource** page, specify **Mode** and select **Destination Enterprise Project**, and click **OK**.
8. After the resource is removed, you can view the modified enterprise project information on the **Clusters** page.

2.8.8 Restarting a Cluster

If a cluster becomes faulty, you can restart it to check if it can run normally.

Prerequisites

- The target cluster is not frozen and has no task in progress.
- If a cluster is available, ensure that it has stopped processing service requests (such as importing data and searching for data). Otherwise, data may be lost when the cluster is restarted. You are advised to perform this operation during off-peak hours.

Context

CSS supports quick restart and rolling restart.

Quick Restart

- All clusters support this function.
- If you select a node type for quick restart, all nodes of the selected type will be restarted together.
- If you select a node name for quick restart, only the specified node will be restarted.
- The cluster is unavailable during quick restart.

Rolling Restart

- Rolling restart is supported only when a cluster has at least three nodes (including master nodes, client nodes, and cold data nodes).
- Rolling restart can be performed only by specifying node types. If you select a node type for rolling restart, the nodes of the selected type will be restarted in sequence.
- During the rolling restart, only the nodes that are being restarted are unavailable and other nodes can run normally.
- When the data volume is large, rolling restart will take a long time.

Quick Restart

1. Log in to the CSS management console.
2. In the navigation tree on the left, select a cluster type. The cluster management list page is displayed.
3. In the **Operation** column of the target cluster, choose **More > Restart**.
4. On the **Restart Cluster** page, select **Quick Restart**.

You can quick restart nodes by **Node type** or **Node name**. If you select **Node type**, then you can select multiple node types and perform quick restart at

the time. If you select **Node name**, you can perform quick restart only on one node at a time.

5. Refresh the page and check the cluster status. During the restart, the cluster status is **Processing**, and the task status is **Restarting**. If the cluster status changes to **Available**, the cluster has been restarted successfully.

Rolling Restart

1. Log in to the CSS management console.
2. In the navigation tree on the left, select a cluster type. The cluster management list page is displayed.
3. In the **Operation** column of the target cluster, choose **More > Restart**.
4. On the **Restart Cluster** page, select **Rolling Restart**.
You can perform rolling restart by **Node type**. Select specific node types for restart.
5. Refresh the page and check the cluster status. During the restart, the cluster status is **Processing**, and the task status is **Restarting**. If the cluster status changes to **Available**, the cluster has been restarted successfully.

2.8.9 Deleting a Cluster

You can delete clusters that you no longer need.

NOTE

- If you delete a cluster, the cluster service data will be cleared. Exercise caution when performing this operation.
- The snapshots of a cluster stored in OBS are not deleted with the cluster. You can restore a deleted cluster using its snapshots stored in the OBS bucket. For details, see [Can I Restore a Deleted Cluster?](#)

Procedure

1. Log in to the CSS management console.
2. In the navigation tree on the left, select a cluster type. The cluster list page is displayed.
3. Locate the target cluster and click **More > Delete** in the **Operation** column.
4. In the displayed dialog box, enter the name of the cluster to be deleted and click **OK**.

2.9 Customizing Word Dictionaries

2.9.1 Managing Word Dictionaries

You can configure the custom word dictionary to identify the segments of specified words. For example, you can search for the keyword of company names, such as, Huawei, and network buzzwords.

 **NOTE**

- You cannot use the custom word dictionary function for clusters created before the function was released (March 10, 2018).
- Hot update is supported. The updated custom word dictionary can take effect without cluster restart.
- Custom word dictionaries are generally used for Chinese word segmentation. They can also be used to segment English words based on special characters except #&+-.@_

Context

Custom word dictionary uses the IK and synonym analyzer.

The IK analyzer has a main word dictionary and a stop word dictionary. The synonym analyzer has a synonym word dictionary. Before configuring a custom word dictionary, upload the prepared word dictionary file to OBS. For details, see [Uploading the Word Dictionary File to OBS](#).

The IK analyzer uses the **ik_max_word** and **ik_smart** word segmentation policies. The synonym analyzer uses the `ik_synonym` word segmentation policy.

- `ik_max_word`: splits the text at a fine granularity.
- `ik_smart`: splits the text at a coarse granularity.

Prerequisites

- To use the custom word dictionary, the account or IAM user used for logging in to the CSS management console must have both of the following permissions:
 - **OBS Administrator** for project **OBS** in region **Global service**
 - **Elasticsearch Administrator** in the current region
- Prepare the word dictionary file on the local PC as required by referring to [Uploading the Word Dictionary File to OBS](#).

Uploading the Word Dictionary File to OBS

Before configuring a custom word dictionary, upload the word dictionary to an OBS bucket.

1. Prepare the word dictionary file according to [Table 2-49](#).

Table 2-49 Dictionary description

Word Dictionary Type	Introduction	Requirement
Main Word Dictionary	Main words are the words on which users want to perform word segmentation. The main word dictionary is a collection of main words.	The main word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. Letters must be in lowercase. The maximum size of a main word dictionary file is 100 MB.
Stop Word Dictionary	Stop words are the words which users can ignore. A stop word dictionary is a collection of stop words.	The stop word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. The maximum size of a stop word dictionary file is 80 MB.
Synonym Dictionary	Synonyms are words with the same meaning. A synonym dictionary is a collection of synonyms.	The synonym dictionary file must be a text file encoded using UTF-8 without BOM, with a pair of comma-separated synonyms per line. The maximum size of a synonym dictionary file is 80 MB.

2. Upload the word dictionary file to an OBS bucket. For details, see [Uploading an Object](#). The OBS bucket to which data is uploaded must be in the same region as the cluster.

Configuring a Word Dictionary

1. On the CSS management console, choose a cluster type from the left navigation pane. The cluster management page is displayed.
2. On the **Clusters** page, click the name of the target cluster.
3. Click the **Word Dictionaries** tab.
4. On the displayed **Word Dictionaries** page, set the switch to enable or disable the custom word library function.
 - **OBS Bucket:** indicates the OBS bucket where the main word dictionary file, stop word dictionary file, and synonym dictionary file are stored. If no OBS bucket is available, create one by referring to [Creating a Bucket](#). The OBS bucket must be in the same region as the cluster.
 - **Main word dictionary object:** The main word dictionary file must be a text file encoded using UTF-8 without BOM. One subword occupies a line. Letters must be in lowercase. The maximum size of a main word dictionary file is 100 MB.

- Stop word dictionary object: The stop word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. The maximum size of a stop word dictionary file is 80 MB.
- Synonym word dictionary object: The synonym dictionary file must be a text file encoded using UTF-8 without BOM. One pair of comma-separated synonyms occupies a line. The maximum size of a synonym dictionary file is 80 MB.

Figure 2-45 Configuring a custom word dictionary

Word Dictionaries

OBS Bucket [Create Bucket](#) ?

Main Word Dictionary Stop Word Dictionary Synonym Dictionary

Main Word Dictionary ?

Stop Word Dictionary ?

Synonym Dictionary ?

5. Click **Save**. In the displayed **Confirm** dialog box, click **OK**. The word dictionary information is displayed in the lower part of the page. The word dictionary status is **Updating**. Wait for about one minute. After the word dictionary configuration is complete, the word dictionary status will change to **Succeeded**, indicating that the configured word dictionary has taken effect in the cluster.

Modifying a Word Dictionary

You can separately update the main word dictionary, the stop word dictionary, and the synonym dictionary.

On the **Word Dictionaries** page, modify **OBS Bucket**, **Main Word Dictionary**, **Stop Word Dictionary**, or **Synonym Word Dictionary**, and click **Save**. In the displayed dialog box, click **OK**. When the word dictionary status changes from **Updating** to **Successful**, the custom word dictionary is modified.

Figure 2-46 Configuring a custom word dictionary

Word Dictionaries

OBS Bucket [Create Bucket](#) ?

Main Word Dictionary Stop Word Dictionary Synonym Dictionary

Main Word Dictionary ?

Stop Word Dictionary ?

Synonym Dictionary ?

Disabling a Word Dictionary

You can disable your word dictionary when it is no longer in need.

On the **Word Dictionaries** page, disable the function and click **OK** in the displayed dialog box. After the word dictionary is disabled, the word dictionary configuration information will not be displayed.

2.9.2 Example

Application Scenarios

Configure a custom word dictionary for the cluster, set main words, stop words, and synonyms. Search for the target text by keyword and synonym and view the search results.

Step 1: Configure a Custom Word Dictionary

1. Prepare a word dictionary file (a text file encoded using UTF-8 without BOM) and upload it to the target OBS path.

Set the main word dictionary file, stop word dictionary file, and synonym word dictionary file.

NOTE

The default word dictionary contains common stop words such as **are** and **the**. You do not need to upload such stop words.

2. In the navigation pane on the left, choose **Clusters**.
3. On the **Clusters** page, click the name of the target cluster.
4. Click the **Word Dictionaries** tab. Configure the word dictionary file for the step 1 by referring to [Configuring a Word Dictionary](#).
5. After the word dictionary takes effect, return to the cluster list. Locate the target cluster and click **Kibana** in the **Operation** column to access the cluster.
6. On the Kibana page, click **Dev Tools** in the navigation tree on the left. The operation page is displayed.

7. Run the following commands to check the performance of different word segmentation policies.

- Use the `ik_smart` word segmentation policy to split the target text.

Example code:

```
POST /_analyze
{
  "analyzer": "ik_smart",
  "text": "Text used for word segmentation"
}
```

After the operation is completed, view the word segmentation result.

```
{
  "tokens": [
    {
      "token": "word-1",
      "start_offset": 0,
      "end_offset": 4,
      "type": "CN_WORD",
      "position": 0
    },
    {
      "token": "word-2",
      "start_offset": 5,
      "end_offset": 8,
      "type": "CN_WORD",
      "position": 1
    }
  ]
}
```

- Use the `ik_max_word` word segmentation policy to split the target text.

Example code:

```
POST /_analyze
{
  "analyzer": "ik_max_word",
  "text": "Text used for word segmentation"
}
```

After the operation is completed, view the word segmentation result.

```
{
  "tokens" : [
    {
      "token": "word-1",
      "start_offset" : 0,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 0
    },
    {
      "token" : "word-3",
      "start_offset" : 0,
      "end_offset" : 2,
      "type" : "CN_WORD",
      "position" : 1
    },
    {
      "token" : "word-4",
      "start_offset" : 0,
      "end_offset" : 1,
      "type" : "CN_WORD",
      "position" : 2
    },
    {
      "token" : "word-5",
      "start_offset" : 1,
      "end_offset" : 3,
      "type" : "CN_WORD",
      "position" : 3
    }
  ]
}
```

```
"position" : 3
},
{
  "token" : "word-6",
  "start_offset" : 2,
  "end_offset" : 4,
  "type" : "CN_WORD",
  "position" : 4
},
{
  "token" : "word-7",
  "start_offset" : 3,
  "end_offset" : 4,
  "type" : "CN_WORD",
  "position" : 5
},
{
  "token" : "word-2",
  "start_offset" : 5,
  "end_offset" : 8,
  "type" : "CN_WORD",
  "position" : 6
},
{
  "token" : "word-8",
  "start_offset" : 5,
  "end_offset" : 7,
  "type" : "CN_WORD",
  "position" : 7
},
{
  "token" : "word-9",
  "start_offset" : 6,
  "end_offset" : 8,
  "type" : "CN_WORD",
  "position" : 8
},
{
  "token" : "word-10",
  "start_offset" : 7,
  "end_offset" : 8,
  "type" : "CN_WORD",
  "position" : 9
}
}
]
```

Step 2: Use Keywords for Search

The commands for versions earlier than Elasticsearch 7.x are different from those for versions later than Elasticsearch 7.x. Examples are as follows.

- **Versions earlier than 7.x**
 - a. Create the **book** index and configure the word segmentation policy.

In this example, both **analyzer** and **search_analyzer** are set to **ik_max_word**. You can also use **ik_smart**.

```
PUT /book
{
  "settings": {
    "number_of_shards": 2,
    "number_of_replicas": 1
  },
  "mappings": {
    "type1": {
      "properties": {
        "content": {
```

```

        "type": "text",
        "analyzer": "ik_max_word",
        "search_analyzer": "ik_max_word"
      }
    }
  }
}

```

- b. Import the text information to the **book** index.

```

PUT /book/type1/1
{
  "content": "Imported text"
}

```

- c. Use a keyword to search for the text and view the search results.

```

GET /book/type1/_search
{
  "query": {
    "match": {
      "content": "Keyword"
    }
  }
}

```

Search result

```

{
  "took" : 20,
  "timed_out" : false,
  "_shards" : {
    "total" : 2,
    "successful" : 2,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 1.1507283,
    "hits" : [
      {
        "_index" : "book",
        "_type" : "type1",
        "_id" : "1",
        "_score" : 1.1507283,
        "_source" : {
          "content" : "Imported text"
        }
      }
    ]
  }
}

```

- **7.x and later versions**

- a. Create the **book** index and configure the word segmentation policy.

In this example, both **analyzer** and **search_analyzer** are set to **ik_max_word**. You can also use **ik_smart**.

```

PUT /book
{
  "settings": {
    "number_of_shards": 2,
    "number_of_replicas": 1
  },
  "mappings": {
    "properties": {
      "content": {
        "type": "text",
        "analyzer": "ik_max_word",
        "search_analyzer": "ik_max_word"
      }
    }
  }
}

```

```
}
  }
}
```

- b. Import the text information to the **book** index.

```
PUT /book/_doc/1
{
  "content": "Imported text"
}
```

- c. Use a keyword to search for the text and view the search results.

```
GET /book/_doc/_search
{
  "query": {
    "match": {
      "content": "Keyword"
    }
  }
}
```

Search result

```
{
  "took" : 16,
  "timed_out" : false,
  "_shards" : {
    "total" : 2,
    "successful" : 2,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 1.7260926,
    "hits" : [
      {
        "_index" : "book",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 1.7260926,
        "_source" : {
          "content" : "Imported text"
        }
      }
    ]
  }
}
```

Step 3: Use Synonyms for Search

The commands for versions earlier than Elasticsearch 7.x are different from those for versions later than Elasticsearch 7.x. Examples are as follows.

- **Versions earlier than 7.x**

- a. Create the **myindex** index and configure the word segmentation policy.

```
PUT myindex
{
  "settings": {
    "analysis": {
      "filter": {
        "my_synonym": {
          "type": "dynamic_synonym"
        }
      },
      "analyzer": {
```

```

    "ik_synonym": {
      "filter": [
        "my_synonym"
      ],
      "type": "custom",
      "tokenizer": "ik_smart"
    }
  }
},
"mappings": {
  "mytype": {
    "properties": {
      "desc": {
        "type": "text",
        "analyzer": "ik_synonym"
      }
    }
  }
}
}

```

- b. Import the text information to the **myindex** index.

```

PUT /myindex/mytype/1
{
  "desc": "Imported text"
}

```

- c. Conduct search based on the synonym and view the search results.

```

GET /myindex/_search
{
  "query": {
    "match": {
      "desc": "Keyword"
    }
  }
}

```

Search result

```

{
  "took" : 2,
  "timed_out" : false,
  "_shards" : {
    "total" : 5,
    "successful" : 5,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 0.49445358,
    "hits" : [
      {
        "_index" : "myindex",
        "_type" : "mytype",
        "_id" : "1",
        "_score" : 0.49445358,
        "_source" : {
          "desc" : "Imported text"
        }
      }
    ]
  }
}

```

- **7.x and later versions**

- a. Create the **myindex** index and configure the word segmentation policy.

```

PUT myindex
{
  "settings": {

```

```

"analysis": {
  "filter": {
    "my_synonym": {
      "type": "dynamic_synonym"
    }
  },
  "analyzer": {
    "ik_synonym": {
      "filter": [
        "my_synonym"
      ],
      "type": "custom",
      "tokenizer": "ik_smart"
    }
  }
},
"mappings": {
  "properties": {
    "desc": {
      "type": "text",
      "analyzer": "ik_synonym"
    }
  }
}

```

- b. Import the text information to the **myindex** index.

```

PUT /myindex/_doc/1
{
  "desc": "Imported text"
}

```

- c. Conduct search based on the synonym and view the search results.

```

GET /myindex/_search
{
  "query": {
    "match": {
      "desc": "Keyword"
    }
  }
}

```

Search result

```

{
  "took" : 1,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 0.1519955,
    "hits" : [
      {
        "_index" : "myindex",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 0.1519955,
        "_source" : {
          "desc" : "Imported text"
        }
      }
    ]
  }
}

```

```
}  
}
```

2.10 Converting Between Simplified and Traditional Chinese (Using the Conversion Plugin)

By default, a simplified-traditional Chinese conversion plugin is installed in CSS. The plugin implements conversion between simplified and traditional Chinese. With this plugin, you can search index data containing the corresponding simplified Chinese based on the traditional Chinese keyword, and vice versa.

The simplified-traditional Chinese conversion plugin can be used as the analyzer, tokenizer, token-filter, or char-filter.

The simplified-traditional Chinese conversion plugin provides the following two conversion types:

- s2t: converts simplified Chinese to traditional Chinese.
- t2s: converts traditional Chinese to simplified Chinese.

Example

1. Log in to the CSS management console.
2. In the navigation pane on the left, click **Clusters** to switch to the **Clusters** page.
3. In the cluster list, locate the row containing the cluster and click **Access Kibana** in the **Operation** column.
If the target cluster has the security mode enabled, enter the username and password you set when you created the cluster.
4. In the Kibana navigation pane on the left, choose **Dev Tools**.
5. On the **Console** page, run the following command to create index **stconvert** and specify a user-defined mapping to define the data type:

Versions earlier than 7.x

```
PUT /stconvert  
{  
  "settings": {  
    "number_of_shards": 1,  
    "number_of_replicas": 0,  
    "analysis": {  
      "analyzer": {  
        "ts_ik": {  
          "tokenizer": "ik_smart",  
          "char_filter": [  
            "tsconvert",  
            "stconvert"  
          ]  
        }  
      },  
      "char_filter": {  
        "tsconvert": {  
          "type": "stconvert",  
          "convert_type": "t2s"  
        },  
        "stconvert": {  
          "type": "stconvert",  
          "convert_type": "s2t"  
        }  
      }  
    }  
  }  
}
```



```

    }
  },
  "mappings": {
    "type": {
      "properties": {
        "desc": {
          "type": "text",
          "analyzer": "ts_ik"
        }
      }
    }
  }
}

```

Versions 7.x and later

```

PUT /stconvert
{
  "settings": {
    "number_of_shards": 1,
    "number_of_replicas": 0,
    "analysis": {
      "analyzer": {
        "ts_ik": {
          "tokenizer": "ik_smart",
          "char_filter": [
            "tsconvert",
            "stconvert"
          ]
        }
      },
      "char_filter": {
        "tsconvert": {
          "type": "stconvert",
          "convert_type": "t2s"
        },
        "stconvert": {
          "type": "stconvert",
          "convert_type": "s2t"
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "desc": {
        "type": "text",
        "analyzer": "ts_ik"
      }
    }
  }
}

```

The command output is similar to the following:

```

{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "stconvert"
}

```

6. On the **Console** page, run the following command to import data to index **stconvert**:

Versions earlier than 7.x

```

POST /stconvert/type/1
{
  "desc": "Text in traditional Chinese"
}

```

Versions 7.x and later

```
POST /stconvert/_doc/1
{
  "desc": "Text in traditional Chinese"
}
```

If the value of **failed** in the command output is **0**, the data is imported successfully.

7. On the **Console** page, run the following command to search for the keyword and view the search result:

```
GET /stconvert/_search
{
  "query": {
    "match": {
      "desc": "Keyword"
    }
  }
}
```

The command output is similar to the following:

```
{
  "took" : 15,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 0.5753642,
    "hits" : [
      {
        "_index" : "stconvert",
        "_type" : "type",
        "_id" : "1",
        "_score" : 0.5753642,
        "_source" : {
          "desc": "Text in traditional Chinese"
        }
      }
    ]
  }
}
```

2.11 Using the Open Distro SQL Plug-in to Compile Queries

For Elasticsearch 6.5.4 and later versions, Open Distro for Elasticsearch SQL lets you write queries in SQL rather than in the Elasticsearch query domain-specific language (DSL).

If you are already familiar with SQL and do not want to learn query DSL, this feature is a great option.

Basic Operations

- Kibana (recommended)
 - Log in to Kibana and send requests using request parameters or request body to **_opendistro/_sqlURI** in the **Dev Tools** page.

```
POST _opendistro/_sql
{
```

```
"query": "SELECT * FROM my-index LIMIT 50"
}
```

- By default, the result is returned in the JSON structure. If you want the result to be returned in the CSV format, run the following command:

```
POST _opendistro/_sql?format=csv
{
  "query": "SELECT * FROM my-index LIMIT 50"
}
```

When data is returned in the CSV format, each row corresponds to a document and each column corresponds to a field.

- cURL commands

You can also run cURL commands in ECS to execute SQL statements.

```
curl -XPOST https://localhost:9200/_opendistro/_sql -u username:password -k -d '{"query": "SELECT * FROM kibana_sample_data_flights LIMIT 10"}' -H 'Content-Type: application/json'
```

Supported Operations

Open Distro for Elasticsearch supports the following SQL operations: statements, conditions, aggregations, include and exclude fields, common functions, joins, and show.

- Statements

Table 2-50 Statements

Statement	Example
Select	SELECT * FROM my-index
Delete	DELETE FROM my-index WHERE _id=1
Where	SELECT * FROM my-index WHERE ['field']='value'
Order by	SELECT * FROM my-index ORDER BY _id asc
Group by	SELECT * FROM my-index GROUP BY range(age, 20,30,39)
Limit	SELECT * FROM my-index LIMIT 50 (default is 200)
Union	SELECT * FROM my-index1 UNION SELECT * FROM my-index2
Minus	SELECT * FROM my-index1 MINUS SELECT * FROM my-index2

NOTE

As with any complex query, large UNION and MINUS statements can strain or even crash your cluster.

- Conditions

Table 2-51 Conditions

Condition	Example
Like	SELECT * FROM my-index WHERE name LIKE 'j%'
And	SELECT * FROM my-index WHERE name LIKE 'j%' AND age > 21
Or	SELECT * FROM my-index WHERE name LIKE 'j%' OR age > 21
Count distinct	SELECT count(distinct age) FROM my-index
In	SELECT * FROM my-index WHERE name IN ('alejandro', 'carolina')
Not	SELECT * FROM my-index WHERE name NOT IN ('jane')
Between	SELECT * FROM my-index WHERE age BETWEEN 20 AND 30
Aliases	SELECT avg(age) AS Average_Age FROM my-index
Date	SELECT * FROM my-index WHERE birthday='1990-11-15'
Null	SELECT * FROM my-index WHERE name IS NULL

- Aggregations

Table 2-52 Aggregations

Aggregation	Example
avg()	SELECT avg(age) FROM my-index
count()	SELECT count(age) FROM my-index
max()	SELECT max(age) AS Highest_Age FROM my-index
min()	SELECT min(age) AS Lowest_Age FROM my-index
sum()	SELECT sum(age) AS Age_Sum FROM my-index

- Include and exclude fields

Table 2-53 Include and exclude fields

Pattern	Example
include()	SELECT include('a*'), exclude('age') FROM my-index

Pattern	Example
exclude()	SELECT exclude('*name') FROM my-index

- Functions

Table 2-54 Functions

Function	Example
floor	SELECT floor(number) AS Rounded_Down FROM my-index
trim	SELECT trim(name) FROM my-index
log	SELECT log(number) FROM my-index
log10	SELECT log10(number) FROM my-index
substring	SELECT substring(name, 2,5) FROM my-index
round	SELECT round(number) FROM my-index
sqrt	SELECT sqrt(number) FROM my-index
concat_ws	SELECT concat_ws(' ', age, height) AS combined FROM my-index
/	SELECT number / 100 FROM my-index
%	SELECT number % 100 FROM my-index
date_format	SELECT date_format(date, 'Y') FROM my-index

 **NOTE**

You must enable fielddata in the document mapping for most string functions to work properly.

- Joins

Table 2-55 Joins

Join	Example
Inner join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s JOIN school sc ON sc.name = s.school_name WHERE s.age > 20

Join	Example
Left outer join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s LEFT JOIN school sc ON sc.name = s.school_name
Cross join	SELECT s.firstname, s.lastname, s.gender, sc.name FROM student s CROSS JOIN school sc

For details about the restrictions, see [Joins](#).

- Show
Show commands display indexes and mappings that match an index pattern. You can use * or % for wildcards.

Table 2-56 Show

Show	Example
Show tables like	SHOW TABLES LIKE logs-*

Joins

Open Distro for Elasticsearch SQL supports inner joins, left outer joins and cross joins. Joins have the following constraints:

- You can only join two indexes.
- You must use an alias for an index (for example, people p).
- In an ON clause, you can only use the AND conditions.
- In a WHERE statement, do not combine trees that contain multiple indexes. For example, the following statement will work:
WHERE (a.type1 > 3 OR a.type1 < 0) AND (b.type2 > 4 OR b.type2 < -1)
- The following statement will not work:
WHERE (a.type1 > 3 OR b.type2 < 0) AND (a.type1 > 4 OR b.type2 < -1)
- You cannot use GROUP BY or ORDER BY to obtain results.
- LIMIT with OFFSET (for example, LIMIT 25 OFFSET 25) is not supported.

JDBC Driver

The Java Database Connectivity (JDBC) driver allows you to integrate Open Distro for Elasticsearch with your business intelligence (BI) applications.

For details about how to download and use JAR files, see [GitHub Repositories](#).

2.12 Using the Open Distro Alarm Plug-in to Configure SMN Alarms

2.12.1 (Optional) Authorizing CSS to Use SMN

Scenario Description

To use the OpenDistro alarm plugin (**opendistro_alerting**), authorize your Elasticsearch cluster to use SMN to send notifications. For details about how to use the OpenDistro alarm plugin, see [Configuring SMN Alarms](#).

Service authorization is to delegate CSS to use other cloud resources. For example, you can authorize CSS to use SMN to send notifications.

Constraints and Limitations

Only the SMN service can be authorized.

Procedure

1. Log in to the CSS management console as an administrator with IAM permissions.
2. In the navigation pane, choose **Service Authorization**.
3. On the **Service Authorization** page, click **Create Agency**. In the dialog box displayed, confirm that the agency is successfully created.
 - If an agency has been created, "css_smn_agency exist, no need to created." is displayed in the upper right corner.
 - If you do not have the creation permission, a message indicating that the current user does not have the permission and you need to check the account permission on IAM is displayed in the upper right corner.

2.12.2 Configuring SMN Alarms

Scenario Description

By default, the open-source OpenDistro alarm plugin (**opendistro_alerting**) is integrated into CSS to send notifications when data meets specific conditions. This plugin consists of three components: **Dashboard**, **Monitors**, and **Destinations**. CSS integrates the SMN service in the **Destinations** component and can send alarm messages only through the SMN service as the destination.

This section describes how to use the OpenDistro alarm plugin to configure SMN alarms for Elasticsearch clusters in Kibana.

NOTE

For details about the official guide of the open-source alarm plug-in **Opendistro Alerting**, visit [OpenDistro-Monitors](#).

Constraints and Limitations

The open-source OpenDistro alarm plugin is installed on Elasticsearch clusters of the versions 7.1.1, 7.6.2, and 7.10.2 by default.

Prerequisites

- The SMN service has been authorized. For details, see [\(Optional\) Authorizing CSS to Use SMN](#).
- You have created a topic on the SMN console. For details, see [Creating a Topic](#).

Procedure

1. Log in to the CSS management console.
2. On the **Cluster Management > Elasticsearch** page, select the target cluster and click **Access Kibana** in the **Operation** column.
3. On the Kibana page, choose **Open Distro for Elasticsearch > Alerting** in the navigation pane on the left.
4. Create an SMN destination to send alert messages.
 - a. On the **Alerting** page, click the **Destinations** tab and click **Add destination** to configure destination information.

Table 2-57 Destinations parameter description

Parameter	Description
Name	User-defined destination name
Type	Retain the default value SMN .
Topic	Select the SMN topic you have created for sending alarm messages. NOTE For the Elasticsearch cluster of version 7.1.1, you need to manually enter the topic name. Ensure that the topic name is the same as that in the SMN service.

Figure 2-47 Add destination

- b. Click **Create** to return to the destination list. If the created SMN destination is displayed in the list, the creation is complete.

Figure 2-48 Destination list

- 5. Create a monitoring task and configure the alarm triggering condition and monitoring frequency.
 - a. Click the **Monitors** tab on the **Alerting** page and click **Create monitors** to configure monitoring information.

Table 2-58 Monitor parameters

Parameter	Description
Monitor name	User-defined monitor name

Parameter	Description
Monitor state	Monitoring status. You are advised to keep this function enabled.
Method of definition	Select a method to define monitoring. You are advised to use Define using extraction query . <ul style="list-style-type: none"> • Define using visual graph: use visualized query statement • Define using extraction query: use specific query statement
Index	Index to be monitored
Time field	When Define using visual graph is selected, select a time field and define counting parameters such as count .
Frequency	Select the monitoring frequency and set the monitoring interval. The options include: <ul style="list-style-type: none"> • By interval • Daily • Weekly • Monthly • Custom cron expression

- b. Click **Create**. The **Create trigger** page is displayed.
- c. On the **Create trigger** page, set the alarm triggering conditions and actions to be triggered.

Table 2-59 Trigger parameters

Parameter	Description
Trigger name	User-defined trigger name
Severity level	Sensitivity of a trigger, that is, the number of alarms that are triggered before an alarm message is sent. 1 indicates the highest sensitivity.
Trigger condition	Trigger condition. An alarm is triggered when the trigger condition is hit.
Action name	Name of a trigger action
Destination	Select the SMN destination created in section 4 .
Message subject	Title of the alarm message. This parameter is required only when Elasticsearch clusters of version 7.10.2 is used.

Parameter	Description
Message	Body of an alarm message. By default, the subject and body are defined when the destination is an email. For details, see Message Publishing .
Action throttling	Message sending frequency. It limits the number of notification messages can be received in a specified period. For example, if this parameter is set to 10 minutes, SMN sends only one alarm notification in the next 10 minutes even if the trigger condition is hit for multiple times. After 10 minutes, SMN sends another alarm notification if the alarm condition is met.

Figure 2-49 Setting the destination of a trigger action

Configure actions

▼ SMN notification: test-action

Action name

test-action

Names can only contain letters, numbers, and special characters

Destination

test - (SMN) ▼

- d. Click **Send test message**. If a subscriber receives an email, as shown in [Figure 2-51](#), the trigger is configured successfully.

Figure 2-50 Sending test messages

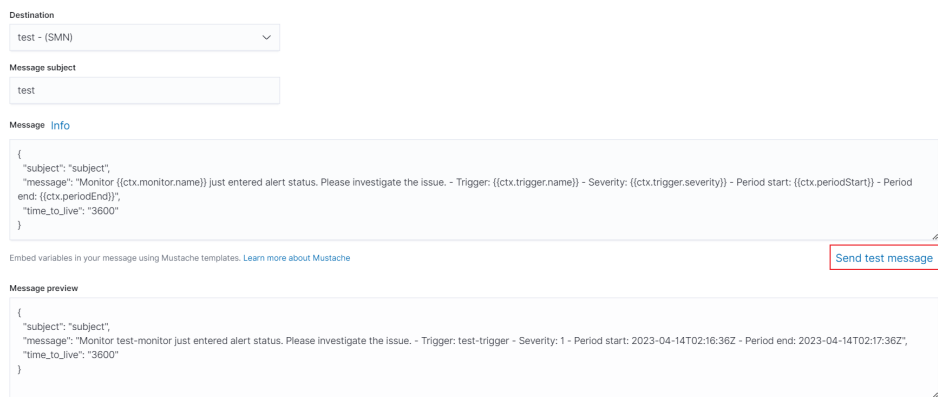
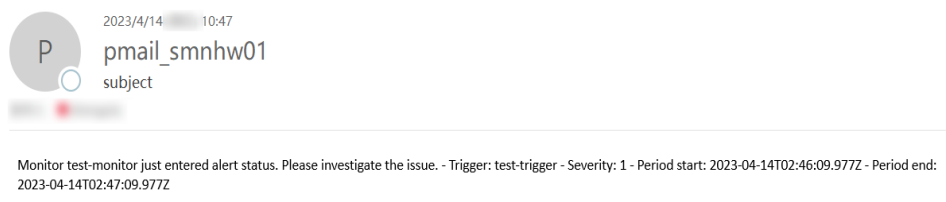


Figure 2-51 Email notification



- e. Click **Create** to return to the Monitor details page.

2.13 Switching Hot and Cold Data

CSS provides you with cold data nodes. You can store data that requires query response in seconds on high-performance nodes and store data that requires query response in minutes on cold data nodes with large capacity and low specifications.

NOTE

- When creating a cluster, you need to configure nodes as data nodes. When you enable the cold data node function, data nodes become hot nodes.
- You can enable the cold data node, master node, and client node functions at the same time.
- You can increase nodes and expand storage capacity of cold data nodes. The maximum storage capacity is determined by the node specifications. Local disks do not support storage capacity expansion.

Hot and Cold Data Node Switchover

If you enable cold data nodes when creating a cluster, the cold data nodes are labeled with **cold**. Other data nodes become hot nodes and are labeled with **hot**. You can specify indexes to allocate data to cold or hot nodes.

You can configure a template to store indices on the specified cold or hot node.

The following figure shows this process. Log in to the **Kibana Console** page of the cluster, modify the template by configuring the index starting with **myindex**, and store the indexes on the cold node. In this case, the **myindex*** data is stored on the cold data node by modifying the template.

- For the 5.x version, run the following command to create a template:

```
PUT _template/test
{
  "order": 1,
  "template": "myindex*",
  "settings": {
    "index": {
      "refresh_interval": "30s",
      "number_of_shards": "3",
      "number_of_replicas": "1",
      "routing.allocation.require.box_type": "cold"
    }
  }
}
```

- For 6.x or later versions, run the following command to create a template:

```
PUT _template/test
{
  "order": 1,
  "index_patterns": "myindex*",
  "settings": {
    "refresh_interval": "30s",
    "number_of_shards": "3",
    "number_of_replicas": "1",
    "routing.allocation.require.box_type": "cold"
  }
}
```

You can perform operations on the created index.

```
PUT myindex/_settings
{
  "index.routing.allocation.require.box_type": "cold"
}
```

You can cancel the configurations of hot and cold data nodes.

```
PUT myindex/_settings
{
  "index.routing.allocation.require.box_type": null
}
```

2.14 Managing Indexes

2.14.1 Creating and Managing Indexes

Clusters of version 7.6.2 or later support index status management. ISM is a plugin that allows you to automate periodic and administrative operations based on changes on the index age, index size, or number of documents. When using the ISM plug-in, you can define policies that automatically handle index rollovers or deletions based on your needs.

NOTE

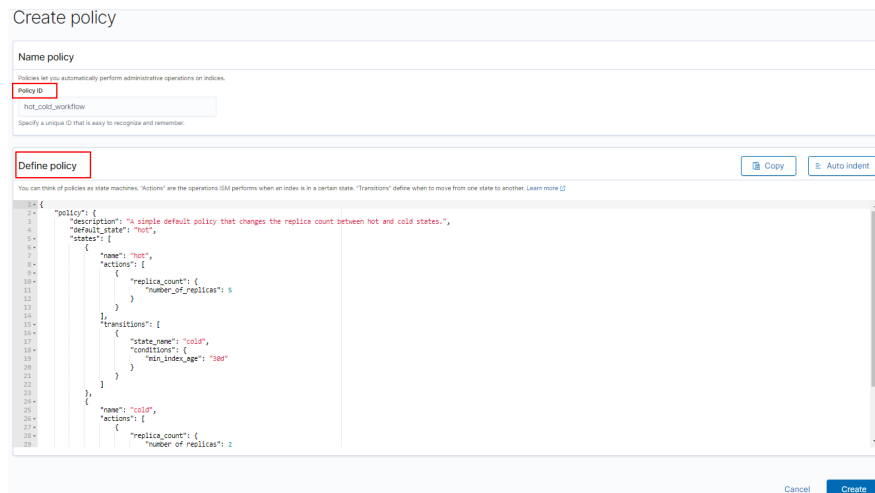
The following procedure uses Elasticsearch 7.6.2 as an example. The Kibana UI varies depending on the Kibana version, but their operations are similar.

Creating an Index Policy

1. Log in to Kibana and choose **IM** or **Index Management** on the left. The **Index Management** page is displayed.

2. Click **Create policy** to create an index policy.
3. Enter a policy ID in the **Policy ID** text box and enter your policy in the **Define policy** text box.

Figure 2-52 Configuring a policy



4. Click **Create**.

Attaching a Policy to an Index

You can attach a policy to one or more indexes and add the policy ID to an index template. When you create indexes using that index template pattern, the policy will be attached to all created indexes.

- **Method 1: Kibana commands**

On the **Dev Tools** page of Kibana, run the following command to associate a policy ID with an index template:

```
PUT _template/<template_name>
{
  "index_patterns": ["index_name-*"],
  "settings": {
    "opendistro.index_state_management.policy_id": "policy_id"
  }
}
```

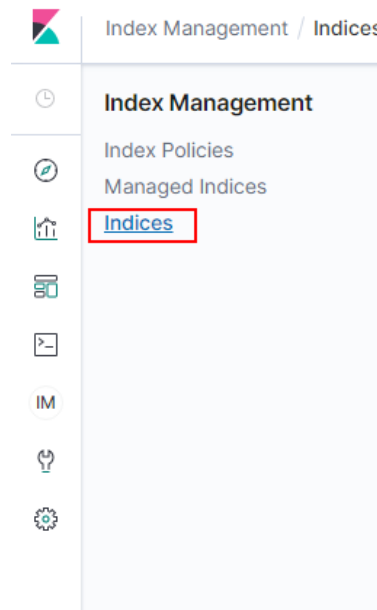
- **<template_name>**: Replace it with the name of a created index template.
- **policy_id**: Replace it with a custom policy ID.

For details about how to create an index template, see [Index Template](#).

- **Method 2: Kibana console**

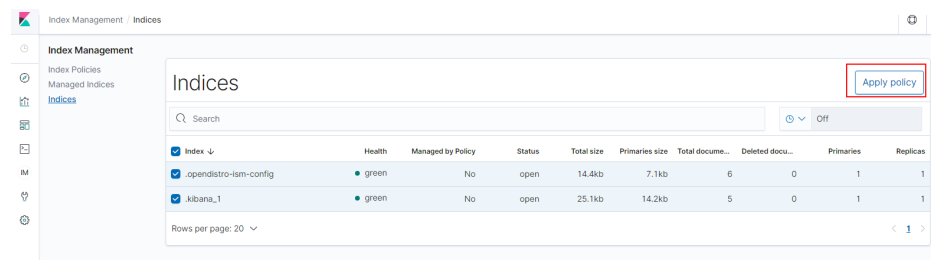
- a. On the **Index Management** page of Kibana, choose **Indices**.

Figure 2-53 Choosing Indexes



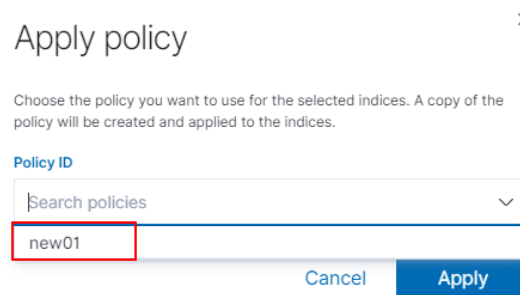
- b. In the **Indices** list, select the target index to which you want to attach a policy.
- c. Click **Apply policy** in the upper right corner.

Figure 2-54 Adding a policy



- d. Select the policy you created from the **Policy ID** drop-down list.

Figure 2-55 Selecting a policy



- e. Click **Apply**.
After you attach a policy to an index, ISM creates a job that runs every 5 minutes by default, to execute the policy, check conditions, and convert the index to different statuses.

Managing Index Policies

1. Click **Managed Indices**.
2. If you want to change the policy, click **Change policy**. For details, see [Changing Policies](#).
3. To delete a policy, select your policy, and click **Remove policy**.
4. To retry a policy, select your policy, and click **Retry policy**.

For details, see [Index State Management](#).

2.14.2 Changing Policies

You can change any managed index policy. ISM has constraints to ensure that policy changes do not break indexes.

If an index is stuck in its current status, never proceeding, and you want to update its policy immediately, make sure that the new policy includes the same status (same name, action, and order) as the old policy. In this case, ISM applies the new policy even if the policy is being executed.

If you update the policy without including an identical status, ISM updates the policy only after all actions in the current status finish executing. Alternatively, you can select a specific status in the old policy and have the new policy take effect.

To change a policy using Kibana, do the following:

1. Under **Managed Indices**, select the indexes to which you want to attach the new policy.
2. Click **Change policy** in the upper right corner. The **Choose managed indices** page is displayed. Configure parameters required for changing a policy.

Table 2-60 Parameters required for changing a policy

Parameter	Description
Managed indices	Select the indexes to which you want to attach the new policy. Multiple indexes can be selected.
State filters	Select an index status. When a status is selected, the new policy is attached to an index in this status.
New policy	Select a new policy.

3. After configuration is complete, click **Change**.

2.15 Intelligent O&M

2.15.1 Overview of Intelligent O&M

CSS provides intelligent O&M to detect potential cluster risks and provide risk handling suggestions.

Currently, the Elasticsearch clusters support intelligent O&M.

Intelligent O&M supports the following functions:

- **Creating a Scan Task**
Before using the intelligent O&M function, you need to create a scan task.
- **Viewing Cluster Risk Items**
After a scan task is started, you can view details about cluster risk items in the intelligent O&M list.
- **Deleting a Scan Task**
After processing all risk items found in a scan task, you can delete the scan task.

2.15.2 Creating a Scan Task

If the intelligent O&M function is enabled for CSS, you need to start a scan task.

Prerequisites

A CSS cluster has been created. For details, see [Creating an Elasticsearch Cluster in Security Mode](#).

Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click the name of the cluster for which you want to perform intelligent O&M. The basic information page of the cluster is displayed.
3. Choose **Intelligent O&M** from the navigation pane.
4. On the Intelligent O&M page, click **Scan** in the upper left corner.
5. In the dialog box, enter the basic information about the scan task and click **OK**.

Table 2-61 Detection task information

Parameter	Description
Name	Name of a scan task.
Description	Brief description of a scan task.
SMN Topic	This parameter is available if you select Send SMN notification upon task completion . If no SMN topic has been created, go to the SMN console to create one.
Notification Level	This parameter is available if you select Send SMN notification upon task completion . If the scan result contains a risk at this level or higher, SMN will send an alarm notification that lists all the risk items in the result.

After a scan task is created, you can view it in the intelligent O&M list.

Follow-up Operations

View cluster risks and diagnose the cluster health status. For details, see [Viewing Cluster Risk Items](#).

2.15.3 Viewing Cluster Risk Items

After a scan task is started, you can view details about cluster risk items in the intelligent O&M list.

Prerequisites



A scan task has been started. For details, see [Creating a Scan Task](#).

Check Items

The following items will be checked and the detected risks will be displayed in the intelligent O&M list:

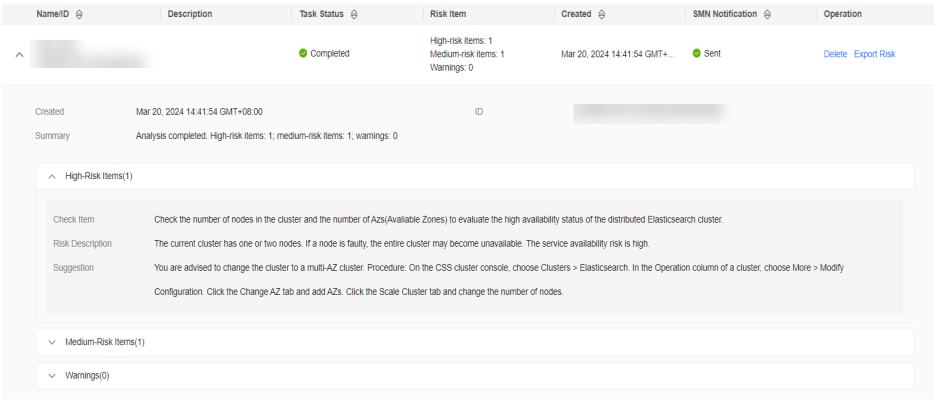
- Check the current health status of the cluster. Red: Some primary shards are not allocated. Yellow: Some secondary shards are not allocated. Green: that all shards are allocated.
- Check the number of nodes in the cluster and the number of AZs to evaluate the high availability status of the distributed Elasticsearch cluster.
- Check whether index replicas are enabled. If replicas are not enabled and a fault occurs, an index may be unavailable, and the data in a cluster using local disks may be lost.
- Check for Kibana index conflicts in clusters.
- Check disk usage. If the disk usage of a node is too high, new index shards may fail to be allocated to the node and the cluster performance may be affected.
- Check whether the storage usage of cluster data nodes or cold data nodes is balanced. Unbalanced storage distribution may result in unbalanced cluster loads and increase read/write latency.
- Check whether any node in the current cluster is disconnected or unavailable for 5 consecutive minutes.
- Check for nodes with too many shards. A large number of shards will consume too many node resources, increasing read/write latency and slowing down metadata update.
- Check the size of all shards. A large shard may affect performance deterioration, occupy too much node memory, and slow down shard restoration during scaling or fault recovery.
- Check whether the current cluster has an available new version.
- Check for snapshot creation failures and snapshot records in the cluster in the last seven days.

Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click a cluster name to go to the basic information page of the cluster.
3. Choose **Intelligent O&M** from the navigation pane.
4. On the intelligent O&M list page, select a started scan task. Click  on the left of the task name to view its creation time, summary, ID, and risk items. Click  on the left of a risk item to view its details, including the check item, risk description, and risk suggestion.

You can handle cluster risks in a timely manner based on the suggestions.

Figure 2-56 Risk items



NameID	Description	Task Status	Risk Item	Created	SMN Notification	Operation
...	...	Completed	High-risk items: 1 Medium-risk items: 1 Warnings: 0	Mar 20, 2024 14:41:54 GMT+...	Sent	Delete Export Risk

Created	Mar 20, 2024 14:41:54 GMT+08:00	ID	...
Summary	Analysis completed. High-risk items: 1, medium-risk items: 1, warnings: 0		
High-Risk Items(1) <ul style="list-style-type: none"> Check Item Check the number of nodes in the cluster and the number of AZs(Available Zones) to evaluate the high availability status of the distributed Elasticsearch cluster. Risk Description The current cluster has one or two nodes. If a node is faulty, the entire cluster may become unavailable. The service availability risk is high. Suggestion You are advised to change the cluster to a multi-AZ cluster. Procedure: On the CSS cluster console, choose Clusters > Elasticsearch. In the Operation column of a cluster, choose More > Modify Configuration. Click the Change AZ tab and add AZs. Click the Scale Cluster tab and change the number of nodes. 			
Medium-Risk Items(1)			
Warnings(0)			

2.15.4 Deleting a Scan Task

After processing all risk items found in a scan task, you can delete the scan task. After a scan task is deleted, the system deletes all diagnosis information corresponding to the scan task.

Prerequisites

A scan task has been started. For details, see [Creating a Scan Task](#).

Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click a cluster name to go to the basic information page of the cluster.
3. Choose **Intelligent O&M** from the navigation pane.
4. Locate a scan task you want to delete and click **Delete** in the **Operation** column.
5. In the dialog box, click **OK**.

2.16 Kibana Platform

2.16.1 Logging In to Kibana

After creating a CSS cluster, you can log in to Kibana through the console or public network.

Kibana Usage Restrictions

- You can customize the username, role name, and tenant name in Kibana. Chinese characters are not allowed.
- Kibana does not support Chinese.

Procedure

- Logging in to the console
 - a. Log in to the CSS management console.
 - b. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column to go to the Kibana login page.
 - Non-security cluster: The Kibana console is displayed.
 - Security cluster: Enter the username and password on the login page and click **Log In** to go to the Kibana console. The default username is **admin** and the password is the one specified during cluster creation.
 - c. After the login is successful, you can access the Elasticsearch cluster through Kibana.
- Logging in using a public IP address

If you have enabled Kibana public access during cluster creation, you can use the Kibana public IP address to log in to the cluster. For details, see [Kibana Public Access](#).
- Logging in using a private IP address

You can access Kibana using a private IP address. To obtain the internal IP address, perform the following steps:

 - a. Log in to the CSS management console.
 - b. Choose **Clusters > Elasticsearch**. The cluster list is displayed.
 - c. Click the cluster name to go to the basic information page and obtain the private network address of the cluster.

Change the port number **9200** of the intranet access address to **5601**, that is, the intranet IP address for accessing Kibana.

Figure 2-57 Obtaining the private IP address

Configuration	
Region	[Redacted]
AZ	[Redacted]
VPC	vpc [Redacted]
Subnet	subnet [Redacted]
Security Group	dws [Redacted] Change Security Group
Security Mode	Disabled
Enterprise Project	default
HTTPS Access	Disabled
IPv4 Access Address	192 [Redacted]

2.16.2 Accessing a Cluster from a Kibana Public Network

For CSS clusters that have security mode enabled, you can enable Kibana public access. After the configuration is complete, an IP address will be provided to access Kibana of this cluster over the Internet.

You can configure Kibana public access during cluster creation, or after a cluster in security mode is created.

NOTE

- You can enable **Security Mode** for clusters of version 6.5.4 and later versions.
- Kibana public access cannot be configured for Elasticsearch clusters created in security mode before this function was rolled out (before June 2020).
- The whitelist for Kibana public network access depends on the ELB whitelist. After you updated the whitelist, the new settings take effect immediately for new connections. For existing persistent connections using the IP addresses that have been removed from the whitelist, the new settings take effect about 1 minute after these connections are stopped.

Configuring Kibana Public Access When Creating a Cluster

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner. The **Create Cluster** page is displayed.
3. On the **Create Cluster** page, enable **Security Mode**.
4. Set **Advanced Settings** to **Custom**, enable **Kibana Public Access**, and set parameters.

Table 2-62 Kibana public access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address Value range: 1 to 100 Unit: Mbit/s
Access Control	If you disable this function, all IP addresses can access Kibana through the public IP address. If you enable this function, only IP addresses or IP address in the whitelist can access Kibana through the public IP address.
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled. You are advised to enable this function.

After the cluster is created, click the cluster name to go to the **Basic Information** page. On the **Kibana Public Access** page, you can view the Kibana public IP address.

Configuring Kibana Public Access for an Existing Cluster

You can enable, disable, modify, and view Kibana public access for an existing cluster that has security mode enabled.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
3. Click the **Kibana Public Access** tab. Turn on the **Kibana Public Access** switch to enable the Kibana public access function.
4. On the displayed page, set parameters.

Table 2-63 Kibana public access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address Value range: 1 to 100 Unit: Mbit/s
Access Control	If you disable this function, all IP addresses can access Kibana through the public IP address. If you enable this function, only IP addresses or IP address in the whitelist can access Kibana through the public IP address.

Parameter	Description
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled. You are advised to enable this function.

5. After you set the parameters, click **OK**.

Modifying Kibana Public Access

For clusters configured Kibana public access, you can modify its bandwidth and access control or disable this function.

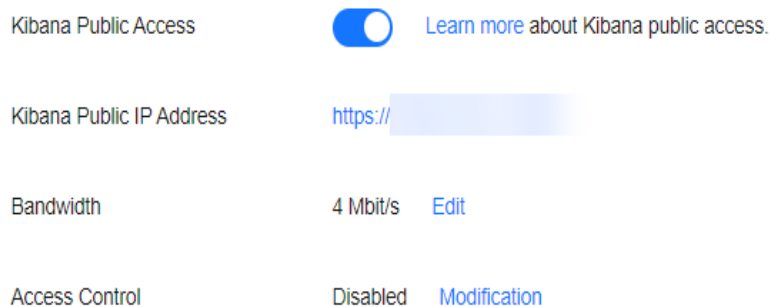
1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
3. Click the **Kibana Public Access** tab to modify the Kibana public access configuration.
 - Modifying bandwidth
Click **Modify** on the right of **Bandwidth**. On the **Modify Bandwidth** page, modify the bandwidth and click **OK**.
 - Modifying access control
Click **Modify** on the right of **Access Control**. On the **Modify Access Control** page, set **Access Control** and **Whitelist**, and click **OK**.
 - Disabling Kibana public access
Toggle off the **Kibana Public Access** switch.

Accessing Kibana with the Public IP Address

After configuring Kibana public access, you will obtain a public IP address that you can use to access Kibana of this cluster.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
3. Click the **Kibana Public Access** tab to obtain the Kibana public IP address.

Figure 2-58 Obtaining the Kibana public IP address



4. Use this IP address to access Kibana of this cluster through the Internet.

2.16.3 Creating a User and Granting Permissions by Using Kibana

CSS uses the `opendistro_security` plug-in to provide security cluster capabilities. The `opendistro_security` plug-in is built based on the RBAC model. RBAC involves three core concepts: user, action, and role. RBAC simplifies the relationship between users and actions, simplifies permission management, and facilitates permission expansion and maintenance. The following figure shows the relationship between the three.

Figure 2-59 User, action, and role



Table 2-64 Parameters

Parameter	Description
User	A user can send operation requests to Elasticsearch clusters. The user has credentials such as username and password, and zero or multiple backend roles and custom attributes.
Role	A role is a combination of permissions and action groups, including operation permissions on clusters, indexes, documents, or fields.
Permission	Single permission, for example, creating an index (for example, <code>indices:admin/create</code>)

Parameter	Description
Role mapping	A user will be assigned a role after successful authentication. Role mapping is to map a role to a user (or a backend role). For example, the mapping from kibana_user (role) to jdoh (user) means that John Doe obtains all permissions of kibana_user after being authenticated by kibana_user . Similarly, the mapping from all_access (role) to admin (backend role) means that any user with the backend role admin (from the LDAP/Active Directory server) has all the permissions of role all_access after being authenticated. You can map a role to multiple users or backend roles.
Action group	A group of permissions. For example, the predefined SEARCH action group grants roles to use _search and _msearchAPI .

In addition to the RBAC model, Elasticsearch has an important concept called tenant. RBAC is used to manage user authorization, and tenants are used for information sharing across tenants. In a tenant space, IAM users can share information such as dashboard data and index patterns.

This section describes how to use Kibana to create a user and grant permissions to the user. Kibana can be used to create users and grant permissions only when the security mode is enabled for the cluster.

 **NOTE**

- The Kibana UI varies depending on the Kibana version, but their operations are similar. This section takes Kibana 7.6.2 as an example to describe the procedure.
- You can customize the username, role name, and tenant name in Kibana. Chinese characters are not allowed.
- Step 1: [Logging in to Kibana](#)
- Step 2: [Creating a User](#)
- Step 3: [Creating a Role and Granting Permissions](#)
- Step 4: [Configuring a Role for a User](#)

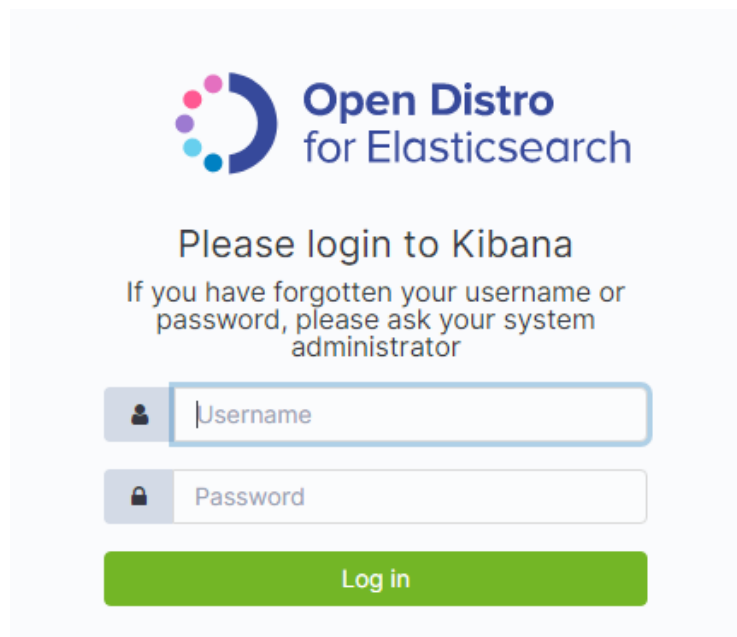
Logging in to Kibana

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.

Enter the administrator username and password to log in to Kibana.

- Username: **admin** (default administrator account name)
- Password: Enter the administrator password you set when creating the cluster in security mode.

Figure 2-60 Login page

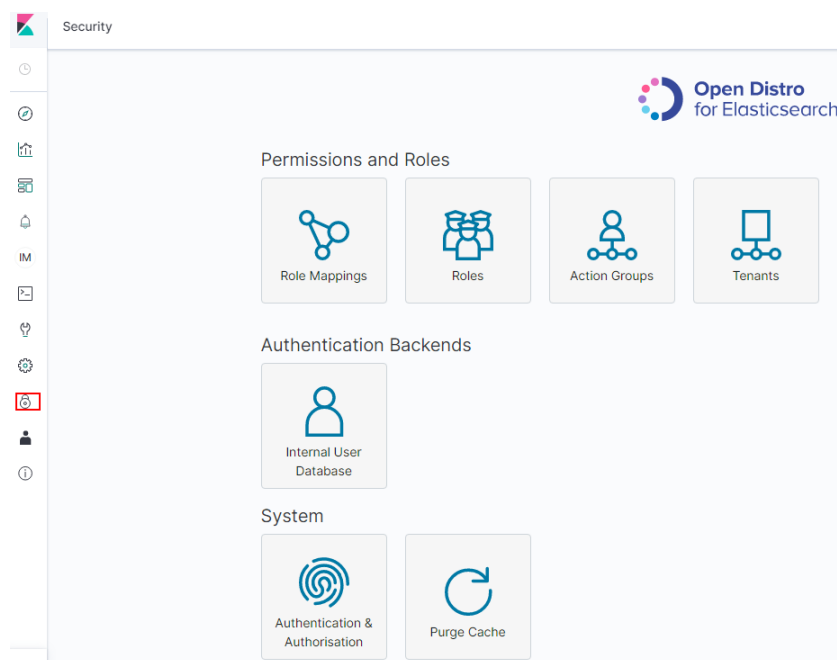


Creating a User

Log in to Kibana and create a user on the **Security** page.

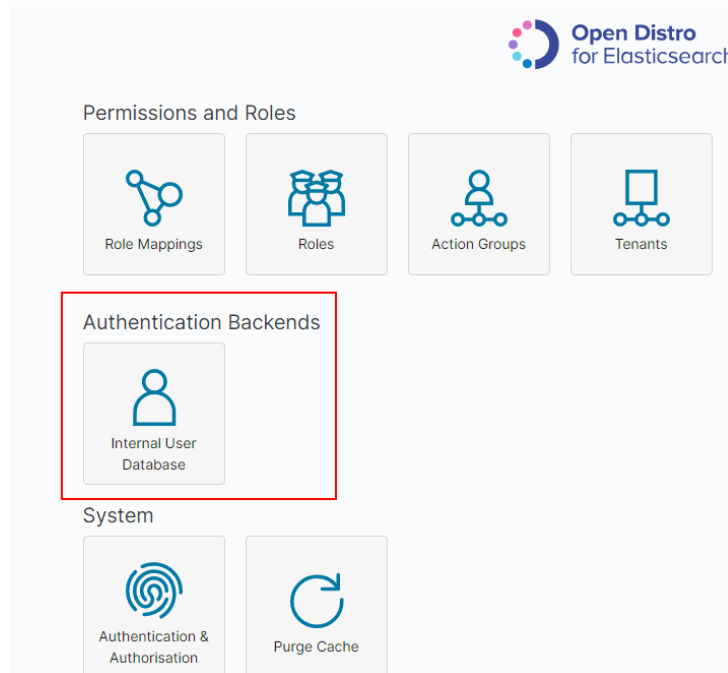
1. After a successful login, choose **Security** in the navigation tree on the left of the Kibana operation page. The **Security** page is displayed.

Figure 2-61 Accessing the Security page



2. Choose **Authentication Backends > Internal Users Database**.

Figure 2-62 Adding a user (1)




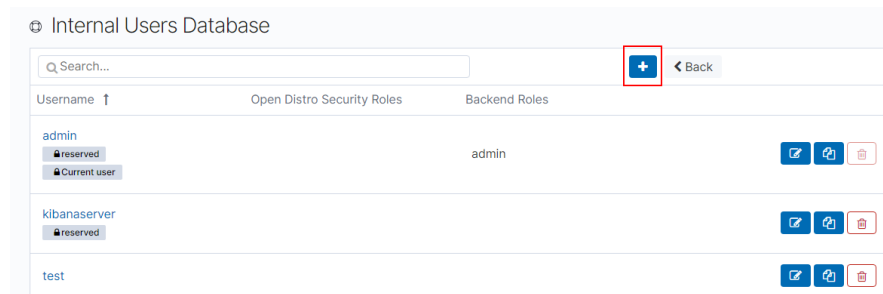
3. On the **Internal Users Database** page, choose  . The page for adding user information is displayed.

Figure 2-63 Adding a user (2)



4. On the user creation page, specify **Username**, **Password**, and **Repeatpassword**, and click **Submit**.

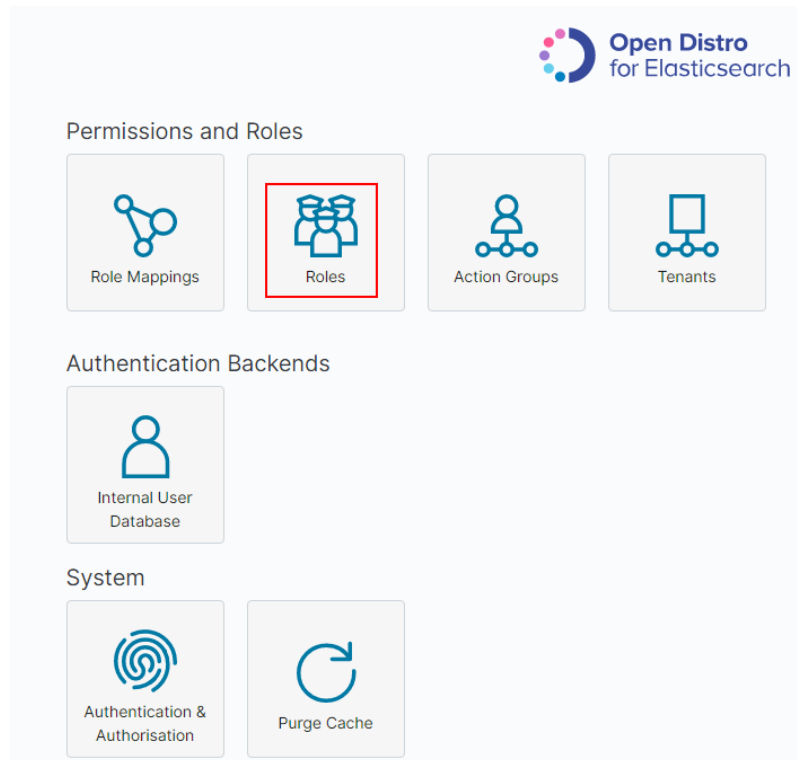
The user will be displayed in the user list.

Creating a Role and Granting Permissions

Create a role and grant permissions to the role.

1. Click **Roles**.

Figure 2-64 Adding a role




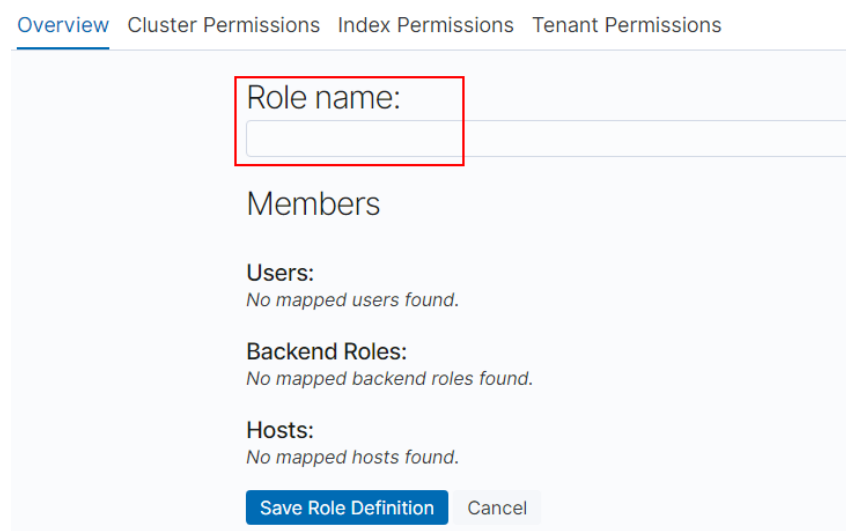
2. On the **Open Distro Security Roles** page, click  .
 - a. On the **Overview** tab page, set the role name.

Figure 2-65 Entering a role name

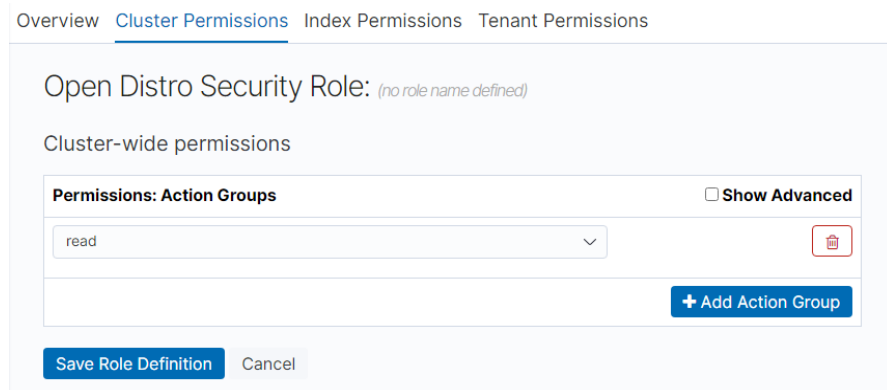


- b. On the **Cluster Permissions** tab page, set CSS cluster permissions. Set cluster permissions based on service requirements. If this parameter is not specified for a role, the role has no cluster-level permissions.
 - **Permissions: Action Groups:** You can click **Add Action Group** to set cluster permissions. For example, if you select the **read** permission

for a cluster, you can only view information such as the cluster status and cluster nodes.

- **Permissions: Single Permissions:** Select **Show Advanced** and click **Add Single Permission** to set more refined permissions for the cluster. For example, if this parameter is set to **indices:data/read**, you can only read specified indexes.

Figure 2-66 Cluster Permissions tab page



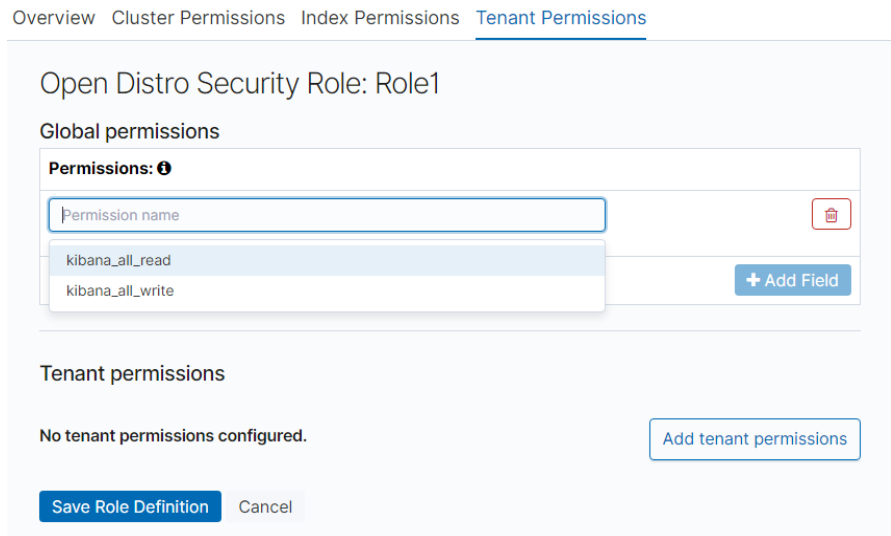
- Configure index permissions on the **Index Permissions** page.
 - **Index patterns:** Set this parameter to the name of the index whose permission needs to be configured. For example, `my_store`.

NOTE

Use different names for the index and the user.

 - **Permissions: Action Groups:** Click **Add Action Group** and set the permission as required. For example, select the read-only permission **Search**.
- On the **Tenant Permissions** page, set role permissions based on service requirements.
 - **Global permissions:** Click **Add Field** to set the kibana read and write permissions of a role, for example, `kibana_all_read` or `kibana_all_write`.
 - **Tenant permissions:** Click **Add tenant pattern** to add a tenant mode and set the `kibana_all_read` or `kibana_all_write` permission for a new tenant mode.

Figure 2-67 Tenant Permissions tab



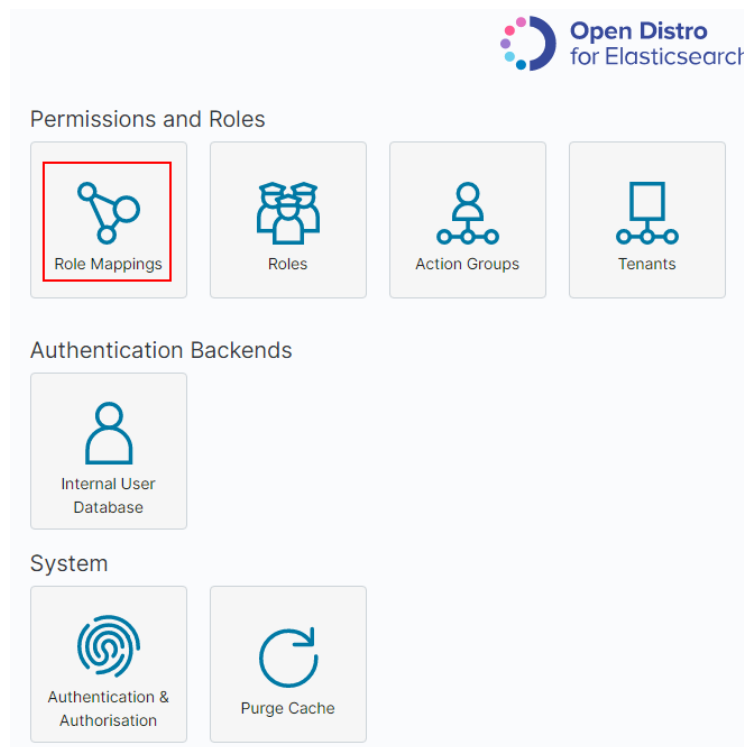
3. Click **Save Role Definition** and you can view the configured role.

Configuring a Role for a User

After creating a role and granting permissions to the role, you need to map the role to a user so that the user can obtain the permissions of the mapped role.

1. Click **Role Mappings**. On the displayed **Role Mappings** page, map the roles.

Figure 2-68 Role mapping



2. On the **Role Mappings** page, click  to select a role and add users.

- **Role:** Select the name of the role to be mapped.
- **Users:** Click **Add User** and enter the name of the user whose role is mapped.

Figure 2-69 Users and roles

The screenshot shows a configuration page with the following elements:

- Role:** A dropdown menu with the text 'Role' and a downward arrow.
- Users:** A section containing a text input field with the placeholder 'username' and a red trash icon to its right. Below this is a blue button labeled '+ Add User'.
- Backend roles:** A section containing a blue button labeled '+ Add Backend Role'.
- Hosts:** A section containing a blue button labeled '+ Add Host'.
- Buttons:** At the bottom left, there are two buttons: 'Submit' (blue) and 'Cancel' (grey).

3. Click **Submit**.
4. Verify that the configuration takes effect in Kibana.

2.16.4 Connecting User-Built Kibana to an Elasticsearch Cluster

To interconnect user-built Kibana with CSS Elasticsearch clusters, the following conditions must be met:

- The local environment must support access from external networks.
- Kibana is built using ECS in the same VPC as Elasticsearch. Kibana can be accessed from the local public network.
- Only Kibana images of the OSS version can be connected to Elasticsearch on CSS.

Example of a Kibana configuration file:

- **Security mode:**

```

elasticsearch.username: "****"
elasticsearch.password: "****"
elasticsearch.ssl.verificationMode: none
server.ssl.enabled: false
server.rewriteBasePath: false
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: https://10.0.0.xxx:9200
elasticsearch.requestHeadersWhitelist: ["securitytenant","Authorization"]
opendistro_security.multitenancy.enabled: true
opendistro_security.multitenancy.tenants.enable_global: true
opendistro_security.multitenancy.tenants.enable_private: true

```

```
opendistro_security.multitenancy.tenants.preferred: ["Private", "Global"]
opendistro_security.multitenancy.enable_filter: false
```

 **NOTE**

- In security mode, the **opendistro_security_kibana** plug-in must be installed. For details, see <https://github.com/opendistro-for-elasticsearch/security-kibana-plugin/tags?after=v1.3.0.0>.
 - The version of the installed plug-in must be the same as that of the cluster. To check the version of the plug-in version, run the **GET _cat/plugins** command.
- **Non-security mode**

```
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx
elasticsearch.hosts: http://10.0.0.xxx:9200
```


3 Logstash

3.1 Creating a Cluster

Procedure

1. Log in to [CSSmanagement console](#).
2. On the **Dashboard** page, click **Create Cluster** in the upper right corner. The **Create** page is displayed.
Alternatively, choose **Clusters > Logstash** in the navigation tree on the left. Click **Create Cluster** in the upper right corner. The **Create** page is displayed.
3. Click **Create Cluster**. The **Create Cluster** page is displayed.
4. Configure **Billing Mode** and **Required Duration**.

Table 3-1 Billing parameters

Parameter	Description
Billing Mode	<p>Select Yearly/Monthly or Pay-per-use.</p> <ul style="list-style-type: none"> • Yearly/Monthly: You pay for the cluster by year or month, in advance. The service duration range is one month to three years. If you plan to use a cluster for more than nine months, you are advised to purchase a yearly package for a better price. A yearly package costs the same as 10 monthly packages. • Pay-per-use: You are billed by actual duration of use, with a billing cycle of one hour. For example, 58 minutes of usage will be rounded up to an hour and billed.
Required Duration	<p>The duration for which the purchased EIP will use. The duration must be specified if the Billing Mode is set to Yearly/Monthly.</p> <p>Configure automatic renewal as required.</p>

- Specify **Region** and **AZ**.

Table 3-2 Region and AZ parameters

Parameter	Description
Region	Select a region for the cluster from the drop-down list on the right.
AZ	Select AZs associated with the cluster region. For details, see What Are Regions and AZs?

- Set basic information about the cluster. Specifically, set **Version** and **Name**.

Table 3-3 Basic parameters




Parameter	Description
Cluster Type	Type of a cluster. Currently, Logstash and Elasticsearch are supported. This section describes how to create a Logstash cluster. For details about how to create an Elasticsearch cluster, see Creating an Elasticsearch Cluster . For details about how to create an Opensearch cluster, see Creating a Cluster .
Version	Version 7.10.0 is supported.
Name	Cluster name, which contains 4 to 32 characters. Only letters, numbers, hyphens (-), and underscores (_) are allowed and the value must start with a letter. NOTE After a cluster is created, you can modify the cluster name as required. Click the name of a cluster to be modified. On the displayed Basic Information page, click  next to the cluster name. After the modification is completed, click  to save the modification. If you want to cancel the modification, click  .

Figure 3-1 Configuring basic information

Version

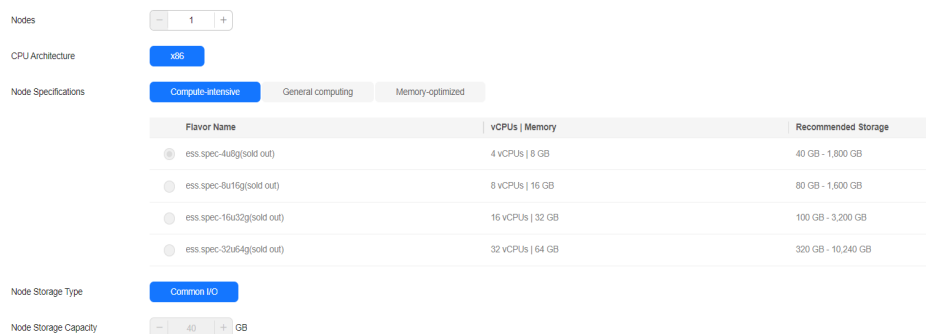
Name

- Set host specifications of the cluster.

Table 3-4 Parameter description

Parameter	Description
Nodes	Number of nodes in a cluster.
CPU Architecture	Currently, Logstash clusters support x86 computing. The supported type is determined by the actual regional environment.
Node Specifications	Specifications of nodes in a cluster. Select a node type and node specification.
Node Storage Type	Extreme SSD, Common I/O, High I/O, and Ultra-high I/O are supported. NOTE If the type of storage in use is not supported, the storage type is not displayed.
Node Storage Capacity	Storage capacity.

Figure 3-2 Configuring host specifications



8. Set the enterprise project.

When creating a CSS cluster, you can bind an enterprise project to the cluster if you have enabled the enterprise project function. You can select an enterprise project created by the current user from the drop-down list on the right or click **View Project Management** to go to the **Enterprise Project Management** console and create a new project or view existing projects.

9. Click **Next: Configure Network**. Configure the cluster network.

Table 3-5 Parameter description

Parameter	Description
VPC	<p>A VPC is a secure, isolated, and logical network environment.</p> <p>Select the target VPC. Click View VPC to enter the VPC management console and view the created or shared VPC names and IDs. If no VPCs are available, create one.</p> <p>NOTE The VPC must contain CIDRs. Otherwise, cluster creation will fail. By default, a VPC will contain CIDRs.</p>
Subnet	<p>A subnet provides dedicated network resources that are isolated from other networks, improving network security.</p> <p>Select the target subnet. You can access the VPC management console to view the names and IDs of the existing subnets in the VPC.</p>
Security Group	<p>A security group implements access control for ECSs that have the same security protection requirements in a VPC. To view more details about the security group, click View Security Group.</p> <p>NOTE Ensure that Port Range/ICMP Type is Any or a port range includes port 9200 for the selected security group.</p>

Figure 3-3 Configuring network specifications



- Click **Next: Configure Advanced Settings**. You can select **Default** or **Custom** for **Advanced Settings**.

If you select **Custom**, you can set the tag and shared directory.

- **Tags:** You can set tags to identify cloud resources.

If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.

- **Shared Directory:** You can configure a shared directory for the cluster. If you want to enable the shared directory, log in to the SFS console to [create an SFS Turbo file system](#) and select a created file system.

Once a shared directory is created for a cluster, it cannot be canceled or stopped. Even if the cluster is deleted, the corresponding shared directory will not be deleted.


11. Click **Next: Confirm**. Check the configuration and click **Next** to create a cluster.
12. Click **Back to Cluster List** to switch to the **Clusters** page. The cluster you created is listed on the displayed page and its status is **Creating**. If the cluster is successfully created, its status will change to **Available**.

If the cluster creation fails, create the cluster again.

3.2 Creating a Cluster in a Shared VPC

A VPC subnet can be shared by multiple Huawei Cloud IAM accounts. You can create CSS clusters in a shared VPC subnet.

Step 1: Creating VPC Share

1. Log in to the [Huawei Cloud management console](#).
2. Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
3. Choose **Shared by Me > Resource Shares**.
4. Click **Create Resource Share** in the upper right corner.
5. On the displayed **Specify Resource Share Details** page, configure basic information and specify the subnet to be shared. Search for **vpc: subnet** and select the target subnet for sharing. Click **Next: Associate Permissions** in the lower right corner.

NOTE

When creating a resource share, you can specify up to 20 resources to share at a time. However, you can update the resource share you created to add more resources. For details, see [Updating a Resource Share](#).

6. On the **Associate Permissions** page, associate a RAM managed permission with each resource type, and then click **Next: Specify Principals** in the lower right corner.

RAM managed permissions available for your selection are system permissions predefined by RAM. Some resource types may have multiple permissions available. You can select as needed. For the details of each permission, see [Viewing the RAM Permissions Library](#).

To create a CSS cluster in a shared VPC, you need to select the **default vpc subnet statement** permission.

7. On the **Grant Access to Principals** page, specify the principals that you want to have access to the resources, and then click **Next: Confirm** in the lower right corner.

In this step, you can select either **Allow sharing with any Huawei Cloud principal** or **Allow sharing only within your organization**. If you select the latter, choose any principals that are within your organization.

You can set **Principal Type** to **Organization** or **Huawei Cloud account ID**. The **Organization** option is available only when the toggle key **Sharing with**

Organizations is turned on. For details, see [Enabling Sharing with Organizations](#).

- Review and confirm the configuration details of your resource share and select **I have read and agree to *Privacy Statement*** on the **Confirm** page. Then, click **Submit** in the lower right corner.

After a resource share is created, RAM sends a sharing invitation to the specified principals. The principals can access and use the shared resources only after they accept the invitation. If the specified principals are within your organization and sharing with Organizations is enabled, the principals can access and use the shared resources without accepting the invitation.

 **NOTE**

Each principal can be shared with a maximum of 100 VPC subnets.

Step 2: Accepting VPC Share


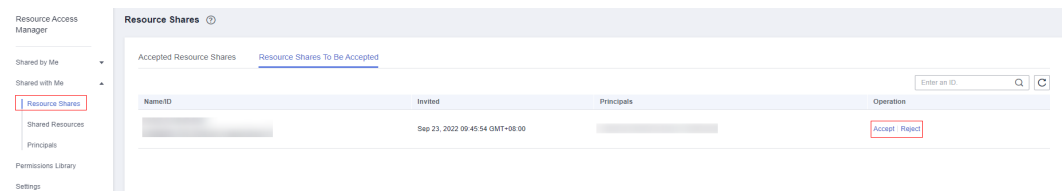
- Log in to the [Huawei Cloud management console](#).
- Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
- Choose **Shared with Me > Resource Shares**.
- Click the **Resource Shares To Be Accepted** tab, and select the resource share for which you are invited. Then, click **Accept** or **Reject** in the **Operation** column.

Figure 3-4 Responding to a resource sharing invitation



- Click **OK** in the displayed dialog box.

After you accept invitations from certain resource shares, you can view them on the **Accepted Resource Shares** page. You can click a resource share name to view its configuration details.

 **NOTE**

Each principal can accept the invitations to resource shares involving a maximum of 100 VPC subnets.

Step 3: Creating a Cluster in the Shared VPC Subnet

- Log in to the CSS console. In the navigation pane on the left, choose **Clusters** and select a cluster type.
For example, log in to the CSS console and choose **Clusters > Elasticsearch** in the navigation pane on the left.
- On the **Clusters** page, click **Create Cluster**.

3. On the **Basic Configuration** page, configure the cluster parameters. For details, see [Creating an Elasticsearch Cluster](#), [Creating a Logstash Cluster](#), and [Creating an OpenSearch Cluster](#).

On the **Network Configuration** page, select the VPC and subnet that are shared with the current account for **VPC** and **Subnet** to create a cluster using the shared VPC.

- **VPC:** Select the name and ID of the VPC that is shared with the current account.
- **Subnet:** Select a subnet for your cluster. You can access the VPC service to view the shared subnet name and ID.

You can create a CSS cluster in the shared VPC subnet.

3.3 Configuring a Cluster

3.3.1 Configuration Center

You can modify configuration files of Logstash clusters in the configuration center, facilitating data migration from different data sources to destinations. Normally, the destinations are Elasticsearch clusters.

Testing Connectivity

Before migrating the data of a Logstash cluster, you can test whether the network connection between the data source and the Logstash cluster is established. You can also enter the IP address or domain name and port of the destination to check the network connectivity between the Logstash cluster and the destination.

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. Click the name of the target cluster. The **Cluster Information** page is displayed. Click **Configuration Center**. Alternatively, click **Configuration Center** in the **Operation** column of the target cluster.
3. On the **Configuration Center** page, click **Test Connectivity**.
4. Enter the IP address and port of the data source and click **Test**.

Figure 3-5 Testing connectivity

The screenshot shows a 'Test Connectivity' form. It has two input fields: 'IP address or domain name' and 'Port'. Below the 'IP address or domain name' field is an 'Add' button and the text 'You can still add 9 more.'. To the right of the 'Port' field is a 'Test' button. At the bottom right, there are 'Cancel' and 'Test' buttons.

NOTE

You can test a maximum of 10 IP addresses or domain names at a time. You can click **Add** to add more IP addresses or domain names and click **Test** on the bottom to test the connectivity of multiple IP addresses or domain names at a time.

Creating a Configuration File

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. Click the name of the target cluster. The **Cluster Information** page is displayed. Click **Configuration Center**. Alternatively, click **Configuration Center** in the **Operation** column of the target cluster.
3. On the **Configuration Center** page, click **Create** in the upper right corner.

You can create a configuration file using a system template or a custom template, or directly create a configuration file.

- To use a system template, click **Apply** in the **Operation** column of the target template, and then configure **Name**, **Configuration File Content**, and **Hidden Content**.

Currently, the following system template types are supported:

- **redis**: You can import data from a Redis database to an Elasticsearch cluster.
- **elasticsearch**: You can migrate data between Elasticsearch clusters.
- **jdbc**: You can import data from a Java Database Connectivity (JDBC) to an Elasticsearch cluster.
- **kafka**: You can import data from Kafka to an Elasticsearch cluster.
- **beats**: You can import data from Beats to an Elasticsearch cluster.
- **dis**: You can import data from DIS to an Elasticsearch cluster.

For details about how to set parameters for each template, see [Parameters for Configuring a System Template](#).

- To directly create a configuration file, enter **Name** and **Configuration File**. The created configuration file cannot exceed 100 KB. A maximum of 50 configuration files can be created.
 - **Hidden Content**: Enter a sensitive string and press Enter to create it. The string will be replaced with asterisks (*) in configurations. (Up to 20 strings are allowed, and each can be up to 512 bytes long.)
4. After the configuration is complete, click **Next** and set parameters.
Configure the file in the pipeline during data migration.

Table 3-6 Parameters

Parameter	Description
pipeline.workers	Number of working threads in the Filters + Outputs phases of parallel pipelines. The default value is the number of CPU cores. The recommended value ranges from 1 to 20.

Parameter	Description
pipeline.batch.size	Maximum number of events that a worker thread collects from inputs before attempting to execute its filters and outputs. A larger value is more effective but increases memory overhead. The default value is 125 .
pipeline.batch.delay	When creating pipeline event batches, the period (in milliseconds) waiting for each event before dispatching an undersized batch to pipeline worker threads. The default value is 50 .
queue.type	An internal queue model for event buffering. memory indicates a memory-based traditional queue, and persisted indicates a disk-based ACKed persistent queue. The default value is memory .
queue.checkpoint.writes	Maximum number of written events before forcing a checkpoint when persistent queues are enabled. The default value is 1024 .
queue.max_bytes	Total capacity of the persistent queue. Make sure the capacity of your disk drive is greater than the value you specify. The default value is 1024 . Unit: MB

5. After the configuration is complete, click **Create**.

On the **Configuration Center** page, you can view the created configuration file. If the status of the configuration file is **Available**, the configuration file is successfully created. You can also edit the created configuration file, add it to a custom template, or delete it.

- To edit a configuration file, click **Edit** in the **Operation** column of a file to modify the file content and parameters.
- You can add a created configuration file to a custom template.
- To delete a configuration file that is not required, click **Delete** in the **Operation** column.

 **NOTE**

You can also click **Operation Record** or **View Running Log** to view the operation records and running logs.

Starting a Configuration File

Created configuration files are displayed on the **Configuration Center** page.

1. Select a configuration file you want to start and click **Start** in the upper left corner.

You can start up 50 configuration files at a time.

2. In the **Start Logstash** dialog box, select **Keepalive** based on service requirements.

The **Keepalive** function is applicable to long-term services. After this function is enabled, a daemon process is configured on each node. When Logstash is faulty, the daemon process automatically starts and rectifies the fault. The **Keepalive** function is not applicable to short-term services. If the function is enabled for a short-term service, the task will fail if no data is available at the source end.

3. Click **OK** to start the configuration file.

You can view the startup configuration file in the pipe list.

 **NOTE**

You can also click **View Operation Record** or **View Running Log** to view the operation records and running logs.

Configuration File Hot Start

When Logstash is running, you can use the hot start function to add a pipe.

 **NOTE**

- Configuration files using the **logstash stdin** plugin cannot use the hot start function.
- If the hot start of a configuration file fails and the Logstash process exits abnormally, the recovery mechanism will be used to restart the original Logstash process. Exercise caution when performing this operation.
- Only one configuration file can be selected for hot start, and the number of configuration items in the **Running** state in the pipe list is less than 20.

1. Select the target configuration file and click **Hot Start** in the upper left corner.

By default, the status of the **Keepalive** function in the dialog box is the same as that in the pipe list.

2. Click **OK** to start the hot start of the configuration file.

You can view the hot start configuration file in the pipe list.

Configuration File Hot Stop

When Logstash is running, you can use the hot stop function to remove a pipeline.

1. Select the target configuration in the pipeline list and click **Hot Stop** above the pipeline list.
2. Click **OK** in the dialog box.

If the hot stop is successful, the target configuration is removed from the pipeline list and the pipeline data migration is interrupted.


Stopping All Configuration Files

To stop data migration of all configuration files in the pipeline list, click **Stop All** above the pipe list.

Click **OK** in the dialog box. If all pipelines are stopped, data migration will be interrupted.

If all pipelines are stopped successfully, data migration of all pipelines is stopped.

Exporting Configuration Files

You can click  in the upper right corner to export the configuration files to the local host in batches.

3.3.2 Example Logstash Configuration File

NOTE

In the following example, the access types of the Elasticsearch clusters on the source and destination ends are the same. This means the source and destination ends both use security clusters or both use non-security clusters with HTTPS not enabled.

If the access types of the Elasticsearch clusters on the source and destination ends are different, you can combine the input and output parts of the following three sample files to obtain the required configuration file.

Non-security Cluster

If the security mode is not enabled for an Elasticsearch cluster, the access example is as follows:

```
input {
  elasticsearch {
    # Source Elasticsearch address
    hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
    # List of indexes to be migrated, separated by commas (,).
    index => "xxx,xxx,xxx"
    # Retain the default values.
    docinfo => true
  }
}

filter {
  # Delete fields added by Logstash.
  mutate {
    remove_field => ["@timestamp", "@version"]
  }
}

output {
  elasticsearch {
    # Destination Elasticsearch cluster address
    hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
    # Index name of the destination cluster. The following configurations must be the same as that of the
    source cluster.
    index => "%{[@metadata][_index]}"
    # ID of the destination data. If you do not need to retain the original ID, delete the following line to
    improve the performance.
    document_id => "%{[@metadata][_id]}"
    # Retain the default values.
    manage_template => false
    ilm_enabled => false
  }
}
```

Security Cluster (HTTPS Access Disabled)

If the security mode is enabled for the created cluster but HTTPS access is disabled, the access example is as follows:

```
input {
  elasticsearch {
    # Username of the source end
```

```
user => "xxx"
# Password of the source end
password => "xxx"
# IP address of the source Elasticsearch
hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
# List of indexes to be migrated, separated by commas (,).
index => "xxx,xxx,xxx"
# Retain the default values.
docinfo => true
}
}

filter {
# Delete fields added by Logstash.
mutate {
remove_field => ["@timestamp", "@version"]
}
}

output {
elasticsearch {
# Username of the destination end
user => "xxx"
# Password of the destination end
password => "xxx"
# Destination Elasticsearch cluster address
hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
# Index name of the target cluster. The following configurations must be the same as that of the
source cluster.
index => "%{[@metadata][_index]}"
# ID of the destination data. If you do not need to retain the original ID, delete the following line to
improve the performance.
document_id => "%{[@metadata][_id]}"
# Retain the default values.
manage_template => false
ilm_enabled => false
}
}
```

Security Cluster (HTTPS Access Enabled)

If the security mode and HTTPS access are enabled for the created cluster, the access example is as follows:

```
input {
elasticsearch {
# Username of the source end
user => "xxx"
# Password of the source end
password => "xxx"
# IP address of the source Elasticsearch
hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
# List of indexes to be migrated, separated by commas (,).
index => "xxx,xxx,xxx"
# Source Elasticsearch certificate. For clusters on the cloud, retain the following information. For user-
built Logstash clusters, download the certificate from the cluster details page. Enter the corresponding
certificate name and path here.
ca_file => "/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"
# Retain the default values.
docinfo => true
ssl => true
}
}

filter {
# Delete fields added by Logstash.
mutate {
remove_field => ["@timestamp", "@version"]
}
}
```

```

}
output {
  elasticsearch {
    # Username of the destination end
    user => "xxx"
    # Password of the destination end
    password => "xxx"
    # Destination Elasticsearch cluster address
    hosts => ["xx.xx.xx.xx:9200", "xx.xx.xx.xx:9200"]
    # Index name of the target cluster. The following configurations must be the same as that of the
    source cluster.
    index => "%{[@metadata][_index]}"
    # ID of the destination data. If you do not need to retain the original ID, delete the following line to
    improve the performance.
    document_id => "%{[@metadata][_id]}"
    # Destination Elasticsearch certificate. For clusters on the cloud, retain the following information. For
    user-built Logstash clusters, download the certificate to the node from the cluster details page. Enter the
    corresponding certificate name and path here.
    cacert => "/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs"
    # Retain the default values.
    manage_template => false
    ilm_enabled => false
    ssl => true
    ssl_certificate_verification => false
  }
}

```

3.3.3 Parameters for Configuring a System Template

- **redis**: You can import data from a Redis database to an Elasticsearch cluster. For details, see <https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-redis.html>.

Table 3-7 Parameter description

Parameter	Mandatory	Description
data_type	Yes	Data source type. The options are list , channel , and pattern_channel . <ul style="list-style-type: none"> • If the parameter is set to list, the BLPOP key is used. • If the parameter is set to channel, the SUBSCRIBE key is used. • If the parameter is set to pattern_channel, the PSUBSCRIBE key is used.
key	Yes	Redis list or channel name
host	Yes	IP address of the Redis server
port	No	Number of the port to be connected. Default value: 6379 .
hosts	Yes	IP address of the node in the Elasticsearch cluster

Parameter	Mandatory	Description
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.

- elasticsearch:** You can migrate data between Elasticsearch clusters.
 For details, see <https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-elasticsearch.html>.

Table 3-8 Configuration items

Configuration Item	Mandatory	Description
hosts	Yes	IP address of the node in the Elasticsearch cluster to which data is imported.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index from which data is to be migrated.
docinfo	No	Document information. Value: true or false . If this parameter is specified, add Elasticsearch document information in the event, such as the index, type, and ID.

Configuration Item	Mandatory	Description
ca_file	No	The default value is /rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs . For Logstash clusters on the cloud, retain the default value or enter the user-defined certificate path if a user-defined certificate is used. For self-built Logstash clusters, you can download the certificate file on the details page of the Elasticsearch cluster with SSL enabled and enter the path here.
ssl	No	Set this parameter to true to enable SSL for the source ES cluster.
hosts	Yes	IP address of the node in the Elasticsearch cluster that outputs data.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.
document_type	No	This parameter is valid only when docinfo is set to true . If docinfo is set to false , delete this parameter from the configuration file.
document_id	No	This parameter is valid only when docinfo is set to true . If docinfo is set to false , delete this parameter from the configuration file.
cacert	No	The default value is /rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/certs . For Logstash clusters on the cloud, retain the default value or enter the user-defined certificate path if a user-defined certificate is used. For self-built Logstash clusters, you can download the certificate file on the details page of the Elasticsearch cluster with SSL enabled and enter the path here.

Configuration Item	Mandatory	Description
ssl	No	Set this parameter to true to enable SSL for the destination ES cluster.
ssl_certificate_verification	No	Set this parameter to false to enable SSL and ignore the server certificate verification.

- **jdbc**: You can import data from a Java Database Connectivity (JDBC) to an Elasticsearch cluster.

For details, see <https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-jdbc.html>.

Table 3-9 Parameter description

Parameter	Mandatory	Description
jdbc_driver_library	Yes	Path of the JDBC driver library. <ul style="list-style-type: none"> • For version 7.10.0, set this field to jdbc_driver_library => "/rds/datastore/logstash/v7.10.0/package/logstash-7.10.0/extend/jars/mariadb-java-client-2.7.0.jar". • For version 5.6.16, set this field to jdbc_driver_library => "/rds/datastore/logstash/v5.6.16/package/logstash-5.6.16/extend/jars/mariadb-java-client-2.7.0.jar". Currently, only these existing drivers are supported. User-defined upload is not supported.
jdbc_driver_class	Yes	JDBC driver class to be loaded, for example, "org.mariadb.jdbc.Driver".
jdbc_connection_string	Yes	JDBC connection string
jdbc_user	Yes	JDBC username
jdbc_password	Yes	JDBC password
statement	Yes	SQL statement of the input data.
hosts	Yes	IP address of the node in the Elasticsearch cluster that outputs data.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.

Parameter	Mandatory	Description
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.

- kafka:** You can import data from Kafka to an Elasticsearch cluster.
<https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-kafka.html>

Table 3-10 Parameter description

Parameter	Mandatory	Description
bootstrap_servers	Yes	IP address and port number of the Kafka instance
topics	Yes	List of topics to be subscribed to
group_id	Yes	Identifier of the group to which the consumer belongs.
auto_offset_reset	Yes	Initial offset in Kafka. <ul style="list-style-type: none"> earliest: Automatically reset the offset to the earliest offset. latest: Automatically reset the offset to the latest offset. none: Report an exception to the consumer if the previous offset for the consumer group is not found. anything else: The system throws an exception to the consumer.
hosts	Yes	IP address of the node in the Elasticsearch cluster that outputs data.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.

Parameter	Mandatory	Description
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.

- **dis:** You can import data from DIS to an Elasticsearch cluster.

Table 3-11 Parameter description

Parameter	Mandatory	Description
streams	Yes	Name of the DIS stream. The entered DIS stream name must be the same as the stream name specified when you are creating a DIS stream on the DIS console.
endpoint	Yes	Data API address of the region where DIS resides.
ak	Yes	User's access key (AK). For details, see Checking Authentication Information .
sk	Yes	User's secret key (SK). For details, see Checking Authentication Information .
region	Yes	Region where DIS is supported.
project_id	Yes	Project ID of the region. For details, see Checking Authentication Information .
group_id	Yes	DIS App name, used to identify a consumer group. The value can be any character string.

Parameter	Mandatory	Description
client_id	No	Client ID, which identifies a consumer in a consumer group. If multiple pipelines or Logstash instances are started for consumption, set this parameter to different values. For example, the value of instance 1 is client1 , and the value of instance 2 is client2 .
auto_offset_reset	No	Position where data starts to be consumed from the stream. The options are as follows: <ul style="list-style-type: none"> • earliest: Data is consumed from the earliest one. • latest: Data is consumed from the latest one.
hosts	Yes	IP address of the node in the Elasticsearch cluster that outputs data.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.

- **beats**: You can import data from Beats to an Elasticsearch cluster.
<https://www.elastic.co/guide/en/logstash/7.10/plugins-inputs-beats.html>

Table 3-12 Parameter description

Parameter	Mandatory	Description
port	Yes	Listening port number. The port number 5044 is used for connecting and indexing Elasticsearch through Beats.

Parameter	Mandatory	Description
hosts	Yes	IP address of the node in the Elasticsearch cluster that outputs data.
user	No	Username for logging in to the Elasticsearch cluster. Generally, the value is admin . This parameter is mandatory for a security cluster.
password	No	Password for logging in to the Elasticsearch cluster. The password is set when the cluster is created. This parameter is mandatory for a security cluster.
index	Yes	Index to which the data is to be migrated. Only one index can be configured.

3.4 Changing Cluster Configurations

3.4.1 Scaling Out a Cluster

If the workloads on the data plane of a cluster change, you can add nodes to scale out the cluster. Services are not interrupted during cluster scale-out.

Prerequisites

- The target cluster is available and has no tasks in progress.
- The target cluster has sufficient quotas available.

Constraints

- Node specifications cannot be modified during scale-out.
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- The number of nodes and node storage capacity cannot be expanded at the same time for a yearly/monthly cluster.
- The quota of nodes in different types varies. For details, see [Table 3-13](#).

Table 3-13 Number of nodes in different types

Node Type	Number
ess	ess: 1-32

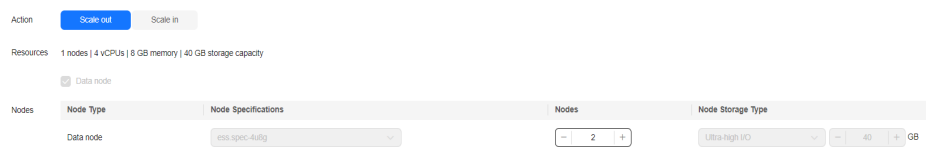
Node Type	Number
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale out** to set parameters.

- **Action:** Select **Scale out**.
- **Resources:** The changed amount of resources.
- **Nodes:** The node number and node storage capacity of the default data node.
 - **Nodes:** For details, see [Table 3-13](#).
 - The value range of **Node Storage Type** depends on the **Node Specifications**. The value must be a multiple of 20.

Figure 3-6 Scaling out a cluster



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling out**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled out.

3.4.2 Scaling in a Cluster

If a cluster can process existing data without fully using its resources, you can scale in the cluster to reduce costs. You are advised to scale in clusters during off-peak hours.

Prerequisites

The cluster is in the **Available** state and has no ongoing task.

Constraints

- Only the number of nodes can be modified during cluster scale-in. The node specifications and node storage capacity cannot be modified.
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- Ensure that the disk usage after scale-in is less than 80% and each AZ of each node type has at least one node.
- When scaling in a cluster, the data in the node to be deleted is migrated to other nodes. The timeout threshold for data migration is five hours. If data migration is not complete within 5 hours, the cluster scale-in fails. You are advised to perform scale-in for multiple times when the cluster has huge amounts of data.
- The quota of nodes in different types varies. For details, see [Table 3-14](#).

Table 3-14 Number of nodes in different types

Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

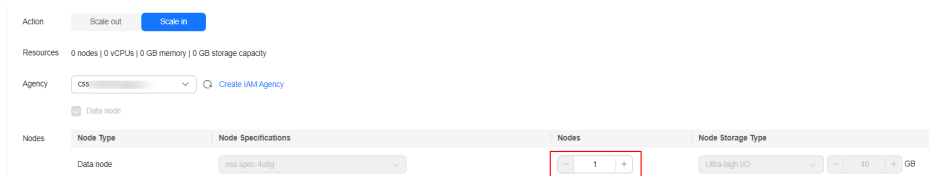
1. Log in to the CSS management console.
2. In the navigation tree on the left, choose **Clusters > Logstash**. The cluster management page is displayed.

3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale in** to set parameters.
 - **Action:** Select **Scale in**.
 - **Resources:** The changed amount of resources.
 - **Agency:** Select an IAM agency to grant the current account the permission to switch AZs.
If no agencies are available, click **Create IAM Agency** to go to the IAM console and create an agency.

NOTE

- The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.
- **Nodes:** The number of the default data nodes. For details about the value range that can be changed, see [Table 3-14](#).

Figure 3-7 Scaling in a cluster




5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling in**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled in.

3.5 Viewing Basic Cluster Information

On the basic information page of a Logstash cluster, you can view the private network address, version, and node of the Logstash cluster.

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. The cluster list page is displayed.
3. Click a cluster name to go to the **Cluster Information** page and view the basic information about the cluster.

Table 3-15 Basic information

Type	Parameter	Description
Cluster Information	Name	Cluster name. The name can be customized. You can click  on the right to change the cluster name.

Type	Parameter	Description
	ID	Unique ID of a cluster, which is automatically generated by the system. Each cluster in the same region has a unique ID.
	Version	Cluster version information.
	Cluster Status	Current status of a cluster
	Task Status	Current task status of a cluster. If no task is in progress, -- is displayed.
	Created	Time when a cluster was created
	Cluster Storage Capacity (GB)	Storage capacity of a cluster
	Used Cluster Storage (GB)	Used storage capacity of a cluster
Configuration	Region	Region where a cluster is located
	AZ	AZ where a cluster is located
	VPC	VPC to which the cluster belongs
	Subnet	Subnet to which the cluster belongs
	Security Group	Security group to which a cluster belongs. To change the security group of a cluster, click Change Security Group on the right. NOTICE Before changing the security group, ensure that the port 9200 required for service access has been enabled. Incorrect security group configuration may cause service access failures. Exercise caution when performing this operation.
	Cluster Routing	Click Modify on the right of the parameter to modify the cluster route information. For details, see Modifying the Return Route .
	Enterprise Project	Enterprise project to which a cluster belongs. You can click the project name to view the basic information about the enterprise project.

Type	Parameter	Description
	Private Network Address	Private IP address and port number of a cluster, which can be used to access the cluster. If the cluster has only one node, the IP address and port number of only one node are displayed, for example, 10.62.179.32:9200 . If the cluster has multiple nodes, the IP addresses and port numbers of all nodes are displayed, for example, 10.62.179.32:9200,10.62.179.33:9200 .
Node	Node Specifications	Specifications of nodes in a cluster
	Node Storage Type	Storage capacity and storage type of nodes in a cluster
	Nodes	Number of nodes in a cluster

Modifying the Return Route

The Logstash service is a hosting service. The subnet selected during cluster creation is not the primary NIC of the Logstash node. You need to add a return route pointing to the target IP address (server where source data is stored) for each node in the cluster.

1. Click the name of a cluster to view its details.
2. Click **Modify** next to the **Cluster Routing**.

Figure 3-8 Configuring cluster routing

Configuration

Region	<input type="text"/>
AZ	<input type="text"/>
VPC	vpc- <input type="text"/>
Subnet	subnet- <input type="text"/>
Security Group	dws- <input type="text"/> Change Security Group
Cluster Routing	Modify View
Enterprise Project	default
IPv4 Access Address	192. <input type="text"/>

3. Modifying the cluster route information

Figure 3-9 Modifying the cluster routing

Modify Cluster Routing

IP Address	<input type="text"/>
Subnet Mask	<input type="text"/>
Modification Type	<input type="button" value="Add"/> <input type="button" value="Delete"/>
<input type="button" value="Cancel"/> <input type="button" value="OK"/>	

- **IP Address:** Enter the first 16 or 24 bits of the server IP address where the source data is stored. For example, if the source IP address is 192.168.1.1, enter 192.168.0.0 to the text box.
- **Subnet Mask:** If the **IP Address** contains 16 bits of the server IP address, set the subnet mask to **255.255.0.0**. If the **IP Address** contains 24 bits of the server IP address, set the subnet mask to **255.255.255.0**.

 NOTE

The subnet mask must cover the IP network segment. That is, after the subnet mask and IP address are converted into binary values, the number of 0 at the end of the IP address must be greater than the number of 0 at the end of the subnet mask.

- **Modification Type:** Select **Add** or **Delete**.
4. Click **OK**.

3.6 Managing Tags

Tags are cluster identifiers. Adding tags to clusters can help you identify and manage your cluster resources.

If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.

Adding Tags to a Cluster

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. The Logstash clusters page is displayed.
3. Find the target cluster and click its name. The **Basic Information** page is displayed.
4. Click the **Tags** tab and add tags to the cluster.

You can select a predefined tag and set **Tag value** for the tag. You can click **View Predefined Tag** to switch to the TMS management console and view existing tags.

You can also create new tags by specifying **Tag key** and **Tag value**.

You can add a maximum of 10 tags for a CSS cluster. If the entered tag is incorrect, you can click **Delete** on the right of the tag to delete the tag. If you do not want to add tags, leave this parameter blank.

Table 3-16 Naming rules for a tag key and value

Parameter	Description
Tag key	Must be unique in a cluster. The value cannot contain more than 64 characters. It can contain only numbers, letters, Chinese characters, and the following special characters: <code>_.:=-@</code> The value cannot start or end with a space. Cannot be left blank.
Tag value	The value cannot contain more than 64 characters. It can contain only numbers, letters, Chinese characters, and the following special characters: <code>_.:=-@</code> The value cannot start or end with a space. Cannot be left blank.

Searching for Clusters by Tag

1. Log in to the CSS management console.
2. On the **Clusters** page, click **Search by Tag** in the upper right corner of the cluster list.
3. Enter the target tag key and value.
You can select a tag key and value from the drop-down lists. The system returns a list of clusters that exactly match the tag key and value. If you select multiple tags, only clusters including all of the selected tags will be returned.
You can select a maximum of 10 tags at one time.
4. Click **Search**.
The system searches for clusters based on the key and value combinations you selected.

3.7 Binding an Enterprise Project

You can create enterprise projects based on your organizational structure. Then you can manage resources across different regions by enterprise project, add users and user groups to enterprise projects, and grant different permissions to the users and user groups. This section describes how to bind and modify an enterprise project for a Logstash cluster in CSS.

Prerequisites

To use the enterprise project function, you need to assign permissions to the corresponding account. You can [submit a service ticket](#) to apply for the permissions.

Before binding an enterprise project, you have [created an enterprise project](#).

Binding an Enterprise Project

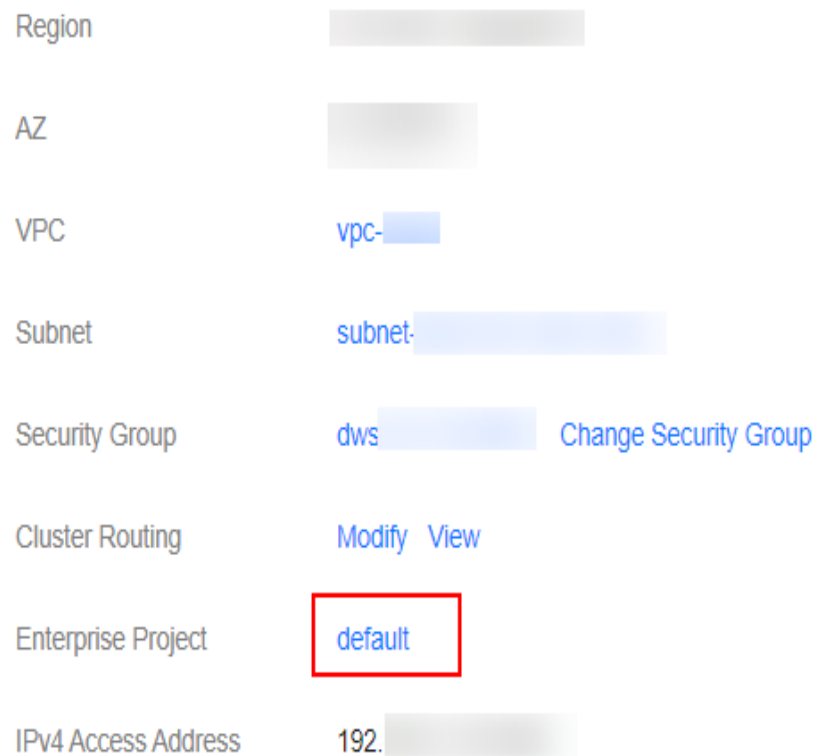
When creating a cluster, you can bind an existing enterprise project to the cluster, or click **View Enterprise Project** to go to the enterprise project management console and create a new project or view existing projects.

Modifying an Enterprise Project

For a cluster that has been created, you can modify its enterprise project based on the site requirements.

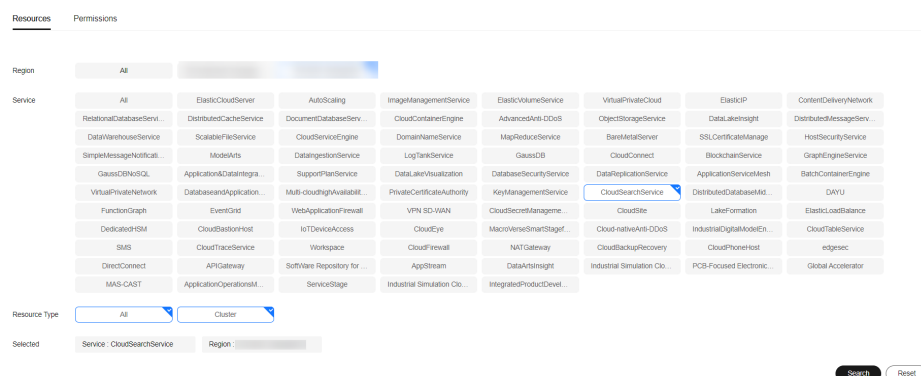
1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters > Logstash**. The cluster management page is displayed.
3. In the cluster list on the displayed page, click the target cluster name to switch to the **Cluster Information** page.
4. On the **Cluster Information** page, click the enterprise project name on the right of **Enterprise Project**. The project management page is displayed.

Figure 3-10 Enterprise project
Configuration



5. On the **Resources** tab page, select the region of the current cluster, and select **CSS** for **Service**. In this case, the corresponding CSS cluster is displayed in the resource list.

Figure 3-11 Filtering CSS clusters



6. Select the cluster whose enterprise project you want to modify and click **Remove**.
7. On the **Remove Resource** page, specify **Mode** and select **Destination Enterprise Project**, and click **OK**.

8. After the resource is removed, you can view the modified enterprise project information on the **Clusters** page.

3.8 Forcibly Restarting VMs in a Cluster

If a Logstash cluster is faulty due to long-term running or other unknown reasons, you can forcibly restart the cluster to restore it. A cluster is unavailable during the restart. Exercise caution when performing this operation.

You can perform the following steps to forcibly restart a cluster:

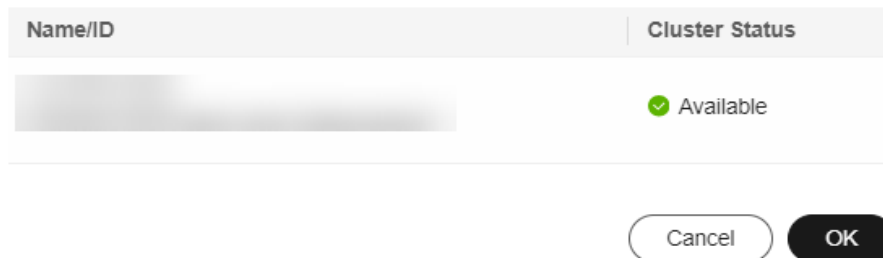
1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. In the **Operation** column of a Logstash cluster, click **More > Force Restart**.
3. In the **Force Restart Cluster VM** dialog box, read the precautions and click **OK**.

Figure 3-12 Forcibly restarting a cluster

Force Restart Cluster VM

The cluster is not available during the restart process, please operate with caution.

The restart process will proactively stop the logstash process. If the 'Keep Resident' value in the pipeline list is set to 'No', it will set the status of all running pipelines to 'Stopped'. If the 'Keep Resident' value is set to 'Yes', it will trigger the logstash process recovery mechanism, setting the pipeline status in operation to 'Recovery'. If the logstash process pipeline status is restored to 'Working' within ten minutes, it will be set to 'Failure'.



During the cluster restart, the cluster status is **Processing** and the task status is **Restarting**. If the cluster status changes to **Available**, the cluster has been restarted successfully.

3.9 Deleting a Cluster

You can delete clusters you no longer require.

Procedure


1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**. On the displayed page, locate the row that contains the target cluster and click **More > Delete** in the **Operation** column.

3. In the displayed dialog box, enter the name of the cluster to be deleted again and click **OK**.

3.10 Managing Logs

CSS provides log backup and search functions to help you locate faults. You can back up cluster logs to OBS buckets and download required log files to analyze and locate faults.

Enabling Log Management

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash**, click the name of the target cluster. The **Basic Information** page is displayed.
3. Click the **Logs** tab and enable **Log Management**.
4. (Optional) If log management is enabled, CSS automatically creates an OBS bucket, backup path, and IAM agency for you to back up logs. The automatically created OBS bucket, backup path, and IAM agency are displayed on the page. If you want to change the OBS bucket, backup path, and IAM agency, click  on the right of **Log Backup Configuration**.

In the displayed **Edit Log Backup Configuration** dialog box, you can either select an existing OBS bucket and IAM agency or create a new bucket and agency. To create a new bucket, click **Create Bucket**. To create a new agency, click **Create IAM Agency**. For details, see [Creating a Bucket](#) and [Creating an Agency](#).

Table 3-17 Parameter description

Parameter	Description	Remarks
OBS Bucket	Name of the OBS bucket used for storing logs	The OBS bucket must be in the same region as the cluster.
Backup Path	Storage path of logs in the OBS bucket	<p>The backup path configuration rules are as follows:</p> <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.

Parameter	Description	Remarks
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in OBS	<p>The following requirements must be met for existing IAM agencies or those that will be created:</p> <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Cloud Service must be Elasticsearch. • The agency must have the OBS Administrator permission for the OBS project in Global service.

5. Back up logs.

a. Automatically backing up logs

Click the icon on the right of **Auto Backup** to enable the auto backup function.

 indicates that the auto backup function is enabled, and  indicates that it is disabled.

After enabling the auto backup function, set the backup start time on the **Edit Auto Backup Policy** page. When the scheduled time arrives, the system will back up logs automatically.

b. Manually backing up logs

On the **Log Backup** tab page, click **Back Up**. On the displayed page, click **Yes** to start backup.

If **Task Status** in the log backup list is **Successful**, the backup is successful.


 **NOTE**

All logs in the cluster are copied to a specified OBS path. You can view or download log files from the OBS path.

6. Search for logs.

You can search for logs by cluster node. A maximum of 1,000 latest logs are displayed.

On the **Log Search** page, select the target node, log type, and log level, and

click  . The search results are displayed.

Viewing Logs

After backing up logs, you can click **Backup Path** to go to the OBS console and view the logs.

Deprecation logs and running logs of CSS are backed up. [Table 3-18](#) lists the logs stored in the OBS bucket.

Table 3-18 Log types

Log Name	Description
logstash-deprecation.log	Deprecation log
logstash-plain.log	Logstash running log

3.11 Managing Certificates

CSS provides a default CA certificate, which is used to trust an HTTPS server when the client connects to the server. You can also upload a customized certificate file based on your service requirements.

Viewing the Default Certificate

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash** and click the name of a Logstash cluster. The basic cluster information page is displayed.
3. Click **Certificate Management**. On the **Default Certificates** tab page, view the default CA certificate of the CSS Elasticsearch cluster, including certificate name, certificate path, certificate status, and description.

Uploading a Custom Certificate

1. Prepare a custom certificate and upload it to an OBS bucket. For details, see section [Uploading a File](#).

If no OBS bucket is available, go to the OBS console to create one. For details, see [Creating a Bucket](#).

The certificate name contains 4 to 32 characters, must start with a letter, and end with .cer/.crt/.rsa/.jks/.pem/.p10/.pfx/.p12/.csr/.der/.keystore. The value can contain letters, digits, hyphens (-), underscores (_), and dots (.). Other special characters are not allowed.

Up to 50 certificates can be uploaded.

The certificate file size cannot exceed 1 MB.

2. Log in to the CSS management console.
3. Choose **Clusters > Logstash** and click the name of a Logstash cluster. The basic cluster information page is displayed.
4. Choose **Certificate Management** and click the **Custom Certificates** tab.
5. Click **Upload Certificate** in the upper left corner. In the displayed dialog box, select **OBS Bucket** and **Certificate Object**.
 - **OBS Bucket:** Select an OBS bucket to store the certificates.
If no OBS bucket is available, click **Create Bucket** on the right to go to OBS console and create a bucket. For details, see [Creating a Bucket](#).
 - **Certificate Object:** Click **Select**. In the **Select Certificate Object** dialog box, select the custom certificate that has been uploaded to the OBS bucket in advance and click **OK**.

6. Click **OK** to upload the custom certificate.

Deleting a Custom Certificate

1. Log in to the CSS management console.
2. Choose **Clusters > Logstash** and click the name of a Logstash cluster. The basic cluster information page is displayed.
3. Choose **Certificate Management** and click the **Custom Certificates** tab.
4. Select the target certificate and click **Delete** in the **Operation** column. In the confirmation dialog box, click **OK**.

4 OpenSearch

4.1 Creating a Cluster

This section describes how to create an OpenSearch cluster.

 **NOTE**

Public IP address access and Kibana public access can be used only after security mode is enabled.

Context

- If you choose the pay-per-use or yearly/monthly billing mode, you can directly create a cluster.
- When creating a cluster, the number of nodes that can be added varies according to the node type. For details, see [Table 4-1](#).

Table 4-1 Number of nodes in different types

Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32

Node Type	Number
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure




1. Log in to the CSS management console.
2. In the upper right corner of the page, click **Create Cluster**. The **Create** page is displayed.
3. Specify **Region** and **AZ**.

Table 4-2 Parameter description for Region and AZ

Parameter	Description
Region	Select a region for the cluster from the drop-down list on the right.
AZ	Select AZs associated with the cluster region. You can select a maximum of three AZs. For details, see Deploying a Cross-AZ Cluster .

4. Set basic information about the cluster. Specifically, set **Version** and **Name**.

Table 4-3 Basic parameters

Parameter	Description
Type	Select OpenSearch
Version	You can select 1.3.6 .
Name	<p>Cluster name, which contains 4 to 32 characters. Only letters, numbers, hyphens (-), and underscores (_) are allowed and the value must start with a letter.</p> <p>NOTE After a cluster is created, you can modify the cluster name as required. Click the name of a cluster to be modified. On the displayed Basic Information page, click  next to the cluster name. After the modification is completed, click  to save the modification. If you want to cancel the modification, click .</p>

5. Set host specifications of the cluster.

Table 4-4 Specification parameters

Parameter	Description
Nodes	<p>Number of nodes in a cluster. Select a number from 1 to 32. You are advised to configure three or more nodes to ensure high availability of the cluster.</p> <ul style="list-style-type: none"> • If neither a master node nor client node is enabled, the nodes specified by this parameter are used to serve as both the master node and client node. Nodes provide the cluster management, data storage, cluster access, and data analysis functions. To ensure data stability in a cluster, you are advised to set this parameter to a value no less than 3. • If only the master node function is enabled, nodes specified by this parameter are used to store data and provide functions of client nodes. • If both the master and client node functions are enabled, the nodes specified by this parameter are only used for storing data. • If only the client node function is enabled, nodes specified by this parameter are used to store data and provide functions of the master node.
CPU Architecture	Support x86 and Kunpeng . The supported type is determined by the actual regional environment.

Parameter	Description
Node Specifications	Specifications of nodes in a cluster. You can select a specification as required. Each cluster supports only one specification. For details, see ECS Types .
Node Storage Type	In the current version, the following options are available: Common I/O , High I/O , and Ultra-high I/O . NOTE If the type of storage in use is not supported, the storage type is not displayed.
Node Storage Capacity	Storage space. Its value varies with node specifications. The node storage capacity must be a multiple of 20.
Master node	The master node manages all nodes in a cluster. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the master node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes. After enabling the master node, specify Node Specifications , Nodes , and Node Storage Type . The value of Nodes must be an odd number greater than or equal to 3. Up to nine nodes are supported. The value of Node Storage Capacity is fixed. You can select a storage type as required.
Client node	The client node allows clients to access clusters and analyze data. If more than 20 nodes are required to store and analyze a large amount of data, you are advised to enable the client node to ensure cluster stability. Otherwise, you are advised to set only the Nodes parameter and use the nodes as both master and client nodes. After enabling the client node, specify Node Specifications , Nodes and Node Storage Type . The value of Nodes ranges from 1 to 32. The value of Node Storage Capacity is fixed. You can select a storage type as required.
Cold data node	The cold data node is used to store historical data, for which query responses can be returned in minutes. If you do not require a quick query response, store historical data on cold data nodes to reduce costs. After enabling cold data node, configure Node Specifications , Nodes , Node Storage Type , and Node Storage Capacity . The value of Nodes ranges from 1 to 32. Select Node Storage Type and Node Storage Capacity as required. After the cold data node is enabled, CSS automatically adds cold and hot tags to related nodes.

6. Set the enterprise project.

When creating a CSS cluster, you can bind an enterprise project to the cluster if you have enabled the enterprise project function. You can select an enterprise project created by the current user from the drop-down list on the right or click **View Project Management** to go to the **Enterprise Project Management** console and create a new project or view existing projects.

7. Click **Next: Configure Network**. Configure the cluster network.**Table 4-5** Parameter description

Parameter	Description
VPC	<p>A VPC is a secure, isolated, and logical network environment.</p> <p>Select the target VPC. Click View VPC to enter the VPC management console and view the created or shared VPC names and IDs. If no VPCs are available, create one.</p> <p>NOTE The VPC must contain CIDRs. Otherwise, cluster creation will fail. By default, a VPC will contain CIDRs.</p>
Subnet	<p>A subnet provides dedicated network resources that are isolated from other networks, improving network security.</p> <p>Select the target subnet. You can access the VPC management console to view the names and IDs of the existing subnets in the VPC.</p>
Security Group	<p>A security group implements access control for ECSs that have the same security protection requirements in a VPC. To view more details about the security group, click View Security Group.</p> <p>NOTE Ensure that Port Range/ICMP Type is Any or a port range includes port 9200 for the selected security group.</p>
Security Mode	<p>After the security mode is enabled, communication will be encrypted and authentication required for the cluster.</p> <ul style="list-style-type: none">• The default administrator account is admin.• Set and confirm the Administrator Password. This password will be required when you access this cluster.

Parameter	Description
HTTPS Access	<p>HTTPS access can be enabled only after the security mode of the cluster is enabled. After HTTPS access is enabled, communication is encrypted when you access the cluster.</p> <p>NOTE Security clusters use HTTPS for communication, which is much slower than non-security clusters that use HTTP for communication. If you want fast read performance and the permission provided by the security mode to isolate resources (such as indexes, documents, and fields), you can disable the HTTPS Access function. After HTTPS Access is disabled, HTTP protocol is used for cluster communication. In this case, data security cannot be ensured and public IP address cannot be used.</p>
Public IP Address	<p>If HTTPS Access is enabled, you can configure Public Network Access and obtain an IP address for public network access. This IP address can be used to access this security cluster through the public network. For details, see Accessing a Cluster from a Public Network.</p>

8. Click **Next: Configure Advanced Settings**. Configure the automatic snapshot creation and other functions.
 - a. Configure **Cluster Snapshot**. Set basic configuration and snapshot configuration.

The cluster snapshot function is enabled by default. You can also disable this function as required. To store automatic snapshots in OBS, an agency will be created to access OBS. Additional cost will be incurred if snapshots are stored in standard storage.

Table 4-6 Cluster snapshot parameter

Parameter	Description
OBS bucket	<p>Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket.</p> <p>The created or existing OBS bucket must meet the following requirements:</p> <ul style="list-style-type: none"> ● Storage Class is Standard. ● Region must be the same as that of the created cluster.

Parameter	Description
Backup Path	<p>Storage path of the snapshot in the OBS bucket.</p> <p>The backup path configuration rules are as follows:</p> <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (.). • The backup path cannot contain more than 1,023 characters.
IAM Agency	<p>IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency.</p> <p>The created or existing IAM agency must meet the following requirements:</p> <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

Table 4-7 Automatic snapshot creation parameter

Parameter	Description
Snapshot Name Prefix	<p>The snapshot name prefix contains 1 to 32 characters and must start with a lowercase letter. Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. A snapshot name consists of a snapshot name prefix and a timestamp, for example, snapshot-1566921603720.</p>
Time Zone	<p>Time zone for the backup time, which cannot be changed. Specify Backup Started Time based on the time zone.</p>
Backup Start Time	<p>The time when the backup starts automatically every day. You can specify this parameter only in full hours, for example, 00:00 or 01:00. The value ranges from 00:00 to 23:00. Select a time from the drop-down list.</p>

Parameter	Description
Retention Period (days)	<p>The number of days that snapshots are retained in the OBS bucket. The value ranges from 1 to 90. You can specify this parameter as required. The system automatically deletes expired snapshots every hour at half past the hour.</p> <p>For example, if you set the automatic snapshot creation policy as shown in Figure 2-1, the system, at 00:30 35 days later, will automatically delete the automated snapshots that were created at 00:00.</p>

Figure 4-1 Setting parameters for automatic snapshot creation

Automatic Snapshot Creation

Snapshot Name Prefix ×

Time Zone GMT+08:00

Backup Start Time ▾ ▾

Retained Snapshots ?

- b. Configure advanced settings for the cluster.
 - **Default:** The **VPC Endpoint Service**, **Kibana Public Access**, and **Tag** functions are disabled by default. You can manually enable these functions after the cluster is created.
 - **Custom:** You can enable the **VPC Endpoint Service**, **Kibana Public Access**, and **Tag** functions as required.

Table 4-8 Parameters for advanced settings

Parameter	Description
VPC Endpoint Service	<p>After enabling this function, you can obtain a private domain name for accessing the cluster in the same VPC. For details, see Accessing a Cluster Using a VPC Endpoint.</p> <p>NOTE The VPC endpoint service cannot be enabled for a shared VPC.</p>

Parameter	Description
Kibana Public Access	You can configure this parameter only when security mode is enabled for a cluster. After enabling this function, you can obtain a public IP address for accessing Kibana. For details, see Accessing a Cluster from a Kibana Public Network .
Tag	Adding tags to clusters can help you identify and manage your cluster resources. You can customize tags or use tags predefined by Tag Management Service (TMS). For details, see Managing Tags . If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.


9. Click **Next: Confirm**. Check the configuration and click **Next** to create a cluster.
10. Click **Back to Cluster List** to switch to the **Clusters** page. The cluster you created is listed on the displayed page and its status is **Creating**. If the cluster is successfully created, its status will change to **Available**.

If the cluster creation fails, create the cluster again.

4.2 Creating a Cluster in a Shared VPC

A VPC subnet can be shared by multiple Huawei Cloud IAM accounts. You can create CSS clusters in a shared VPC subnet.

Step 1: Creating VPC Share

1. Log in to the [Huawei Cloud management console](#).
2. Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
3. Choose **Shared by Me > Resource Shares**.
4. Click **Create Resource Share** in the upper right corner.
5. On the displayed **Specify Resource Share Details** page, configure basic information and specify the subnet to be shared. Search for **vpc: subnet** and select the target subnet for sharing. Click **Next: Associate Permissions** in the lower right corner.

NOTE

When creating a resource share, you can specify up to 20 resources to share at a time. However, you can update the resource share you created to add more resources. For details, see [Updating a Resource Share](#).

6. On the **Associate Permissions** page, associate a RAM managed permission with each resource type, and then click **Next: Specify Principals** in the lower right corner.

RAM managed permissions available for your selection are system permissions predefined by RAM. Some resource types may have multiple permissions available. You can select as needed. For the details of each permission, see [Viewing the RAM Permissions Library](#).

To create a CSS cluster in a shared VPC, you need to select the **default vpc subnet statement** permission.

7. On the **Grant Access to Principals** page, specify the principals that you want to have access to the resources, and then click **Next: Confirm** in the lower right corner.

In this step, you can select either **Allow sharing with any Huawei Cloud principal** or **Allow sharing only within your organization**. If you select the latter, choose any principals that are within your organization.

You can set **Principal Type** to **Organization** or **Huawei Cloud account ID**. The **Organization** option is available only when the toggle key **Sharing with Organizations** is turned on. For details, see [Enabling Sharing with Organizations](#).

8. Review and confirm the configuration details of your resource share and select **I have read and agree to Privacy Statement** on the **Confirm** page. Then, click **Submit** in the lower right corner.

After a resource share is created, RAM sends a sharing invitation to the specified principals. The principals can access and use the shared resources only after they accept the invitation. If the specified principals are within your organization and sharing with Organizations is enabled, the principals can access and use the shared resources without accepting the invitation.

 **NOTE**

Each principal can be shared with a maximum of 100 VPC subnets.

Step 2: Accepting VPC Share


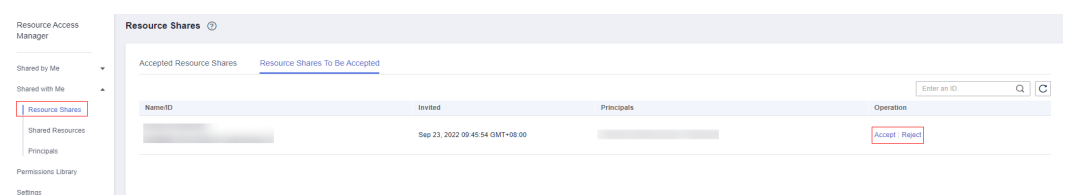
1. Log in to the [Huawei Cloud management console](#).
2. Click  in the upper left corner and choose **Management & Governance > Resource Access Manager**. The **Resource Access Manager** page is displayed.
3. Choose **Shared with Me > Resource Shares**.
4. Click the **Resource Shares To Be Accepted** tab, and select the resource share for which you are invited. Then, click **Accept** or **Reject** in the **Operation** column.

Figure 4-2 Responding to a resource sharing invitation



5. Click **OK** in the displayed dialog box.

After you accept invitations from certain resource shares, you can view them on the **Accepted Resource Shares** page. You can click a resource share name to view its configuration details.

 **NOTE**

Each principal can accept the invitations to resource shares involving a maximum of 100 VPC subnets.

Step 3: Creating a Cluster in the Shared VPC Subnet

1. Log in to the CSS console. In the navigation pane on the left, choose **Clusters** and select a cluster type.
For example, log in to the CSS console and choose **Clusters > Elasticsearch** in the navigation pane on the left.
2. On the **Clusters** page, click **Create Cluster**.
3. On the **Basic Configuration** page, configure the cluster parameters. For details, see [Creating an Elasticsearch Cluster](#), [Creating a Logstash Cluster](#), and [Creating an OpenSearch Cluster](#).

On the **Network Configuration** page, select the VPC and subnet that are shared with the current account for **VPC** and **Subnet** to create a cluster using the shared VPC.

- **VPC**: Select the name and ID of the VPC that is shared with the current account.
- **Subnet**: Select a subnet for your cluster. You can access the VPC service to view the shared subnet name and ID.

You can create a CSS cluster in the shared VPC subnet.

4.3 Accessing a Cluster

4.3.1 Quickly Accessing an OpenSearch Cluster

OpenSearch clusters have built-in Kibana and Cerebro components. You can quickly access an OpenSearch cluster through Kibana and Cerebro.

Accessing a Cluster Through Kibana

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column to go to the OpenSearch Dashboards login page.
 - Non-security cluster: The OpenSearch Dashboards console is displayed.
 - Security cluster: Enter the username and password on the login page and click **Log In** to go to the Kibana console. The default username is **admin** and the password is the one specified during cluster creation.
4. After the login is successful, access the cluster and perform related operations on the OpenSearch Dashboards.

Accessing a Cluster Through Cerebro

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page, locate the target cluster and click **More > Cerebro** in the **Operation** column to go to the Cerebro login page.
 - Non-security cluster: Click the cluster name on the Cerebro login page to go to the Cerebro console.
 - Security cluster: Click the cluster name on the Cerebro login page, enter the username and password, and click **Authenticate** to go to the Cerebro console. The default username is **admin** and the password is the one specified during cluster creation.
4. After the login is successful, you can access clusters through Cerebro.

4.3.2 Accessing a Cluster from a Public Network

You can access a security cluster that has the HTTPS access enabled through the public IP address provided by the system.

By default, CSS uses a shared load balancer for public network access. You can use a dedicated load balancer to improve performance. For details about its configuration, see [Connecting to a Dedicated Load Balancer](#).

NOTE

If public network access is enabled for CSS, then EIP and bandwidth resources will be used and billed.

Configuring Public Network Access

1. Log in to the CSS management console.
2. On the **Create Cluster** page, enable **Security Mode**. Set the administrator password and enable HTTPS access.
3. Select **Automatically assign** for **Public IP Address** and set related parameters.

Figure 4-3 Configuring public network access

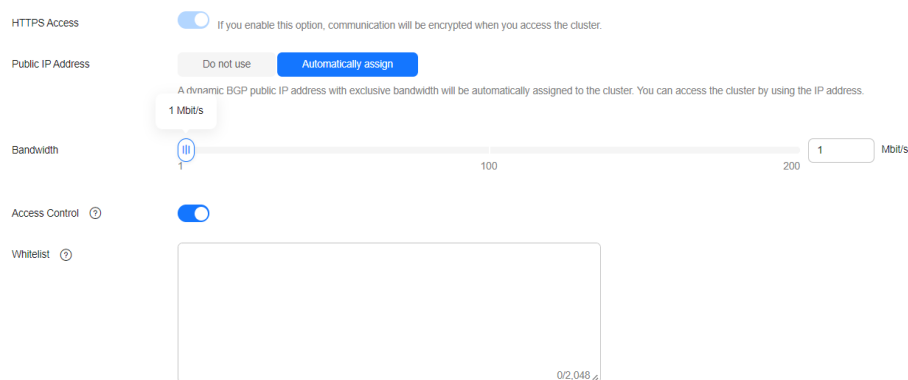


Table 4-9 Public network access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address
Access Control	If you disable this function, all IP addresses can access the cluster through the public IP address. If you enable access control, only IP addresses in the whitelist can access the cluster through the public IP address.
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled.

Managing Public Network Access

You can configure, modify, view the public network access of, or disassociate the public IP address from a cluster.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page, click the name of the target cluster. On the **Basic Information** page that is displayed, manage the public network access configurations.

Figure 4-4 Modifying public network access configurations

Configuration	
Region	[Redacted]
AZ	[Redacted]
VPC	vpc- [Redacted]
Subnet	subne [Redacted]
Security Group	dws [Redacted] Change Security Group
Security Mode	Enabled
Reset Password	Reset
Enterprise Project	default
Public IP Address	100 [Redacted] Disassociate
Access Control	Disabled Set
Bandwidth	5 Mbit/s Edit
HTTPS Access	Enabled Download Certificate
IPv4 Access Address	192 [Redacted]

- Configure public network access
If you did not configure the public network access during cluster creation, you can configure it on the cluster details page after configuring the cluster.
Click **Associate** next to **Public IP Address**, set the access bandwidth, and click **OK**.
If the association fails, wait for several minutes and try again.
- Modify public network access
For a cluster for which you have configured public network access, you can click **Edit** next to **Bandwidth** to modify the bandwidth, or you can click **Set** next to **Access Control** to set the access control function and the whitelist for access.
- View the associated public IP address
On the basic information page of a cluster, you can view the public IP address associated with the cluster.

- Disassociate a public IP address from a cluster
To disassociate the public IP address, click **Disassociate** next to **Public IP Address**.

Accessing a Cluster Through the Public IP Address

After configuring the public IP address, you can use it to access the cluster.

For example, run the following cURL commands to view the index information in the cluster. In this example, the public access IP address of one node in the cluster is **10.62.179.32** and the port number is **9200**.

- If the cluster you access does not have the security mode enabled, run the following command:

```
curl 'http://10.62.179.32:9200/_cat/indices'
```
- If the cluster you access has the security mode enabled, access the cluster using HTTPS and add the username, password and **-u** to the cURL command.

```
curl -u username:password -k 'https://10.62.179.32:9200/_cat/indices'
```

4.3.3 Accessing a Cluster Using a VPC Endpoint

If the VPC endpoint service is enabled, you can use a private domain name or node IP address generated by the endpoint to access the cluster. When the VPC endpoint service is enabled, a VPC endpoint will be created by default. You can select **Private Domain Name Creation** as required. VPC endpoint creation requires specific permissions. For details, see [VPCEP Permissions](#).

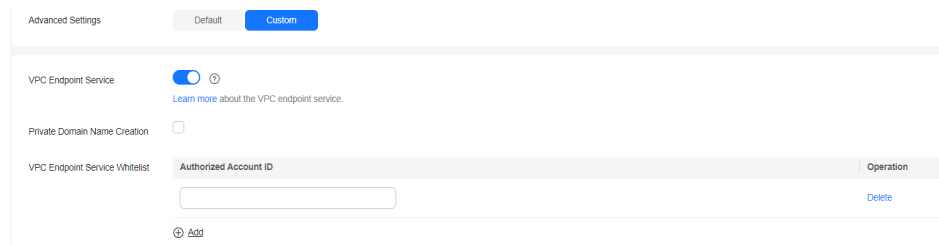
VPC Endpoint uses a shared load balancer for intranet access. If your workloads require quick access, you are advised to connect a dedicated load balancer to the cluster. For details, see [Connecting to a Dedicated Load Balancer](#).

CAUTION

The public IP address access and VPC endpoint service share a load balancer. If you have configured a public access whitelist, public and private IP addresses that access the cluster through VPCEP are restricted because the public IP address access shares the load balancer with the VPC endpoint service. In this case, you need to add IP address **198.19.128.0/17** to the public access whitelist to allow traffic through VPCEP.

Enabling the VPC Endpoint Service

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner.
3. On the **Create Cluster** page, set **Advanced Settings** to **Custom**. Enable the VPC endpoint service.

Figure 4-5 Enabling the VPC endpoint service

- **Private Domain Name Creation:** If you enable this function, the system automatically creates a private domain name for you, which you can use to access the cluster.
- **VPC Endpoint Service Whitelist:** You can add an authorized account ID to the VPC endpoint service whitelist. Then you can access the cluster using the private domain name or the node IP address.
- You can click **Add** to add multiple accounts.
- Click **Delete** in the **Operation** column to delete the accounts that are not allowed to access the cluster.

NOTE

- If the authorized account ID is set to *, all users are allowed to access the cluster.
- You can view authorized account IDs on the **My Credentials** page.
- After the VPC endpoint service is enabled for a cluster, you will be billed per use. For more information, see [Billing Modes](#).

Managing VPC Endpoint Service

You can enable the VPC endpoint service while creating a cluster, and also enable it by performing the following steps after cluster creation.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
4. Click the **VPC Endpoint Service** tab, and turn on the button next to **VPC Endpoint Service**.

In the displayed dialog box, you can determine whether to enable the private domain name. Click **Yes** to enable the VPC endpoint service.

NOTE

- If the VPC endpoint service is enabled, you can use a private domain name or node IP address generated by the VPC endpoint to access the cluster. For details, see [Accessing the Cluster Using the Private Domain Name or Node IP Address](#).
 - If you disable the VPC endpoint service, none of the users can access the cluster using the private domain name.
5. (Optional) Click **Modify** next to **VPC Endpoint Service Whitelist** to update the existing whitelist.
 6. Manage VPC endpoints.

The **VPC Endpoint Service** page displays all VPC endpoints connected to the current VPC endpoint service.

Figure 4-6 Managing VPC endpoints

VPC Endpoint ID	Status	Max. Connections	Owner	Created	Operation
[Redacted]	Accepted	3000	[Redacted]	Mar 20, 2024 10:41:30 GMT+08:00	Accept Reject

Click **Accept** or **Reject** in the **Operation** column to change the node status. If you reject the connection with a VPC endpoint, you cannot access the cluster through the private domain name generated by that VPC endpoint.

Accessing the Cluster Using the Private Domain Name or Node IP Address

1. Obtain the private domain name or node IP address.

Log in to the CSS console, click the target cluster name and go to the **Cluster Information** page. Click the **VPC Endpoint Service** tab and view the private domain name.

Figure 4-7 Viewing the node IP address and private domain name

Cluster Information

Name	cn-[Redacted]
Private IP Address	192.[Redacted]
Private Domain Name	vpcep-[Redacted]
VPC Endpoint Service Whitelist	-- Modify

2. Run the **cURL** command to execute the API or call the API by using a program before accessing the cluster. For details about Elasticsearch operations and APIs, see the [Elasticsearch Reference](#).

The ECS must meet the following requirements:

- Sufficient disk space is allocated for the ECS.
- The ECS and the cluster must be in the same VPC. After enabling the VPC endpoint service, you can access the cluster from the ECS even when the cluster is not in the same VPC as the ECS.
- The security group of the ECS must be the same as that of the cluster.
If this requirement is not met, modify the ECS security group or configure the inbound and outbound rules of the ECS security group to allow the ECS security group to be accessed by all security groups of the cluster. For details, see [Configuring Security Group Rules](#).
- Configure security group rule settings of the target CSS cluster. Set TCP protocol and port 9200 or a port range including port 9200 for both the outbound and inbound directions.

For example, run the following cURL command to view the index information in the cluster. In this example, the private network address is **vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com** and port **9200** is used to access the cluster.

- If the cluster you access does not have the security mode enabled, run the following command:

```
curl 'http://vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com:9200/_cat/indices'
```

- If the cluster you access has the security mode enabled, access the cluster using HTTPS and add the username, password, and **-u** to the cURL command.

```
curl -u username:password -k 'https://vpcep-7439f7f6-2c66-47d4-b5f3-790db4204b8d.region01.huaweicloud.com:9200/_cat/indices'
```

4.3.4 (Optional) Interconnecting with a Dedicated Load Balancer

4.3.4.1 Scenario Description

CSS integrates shared load balancers and allows you to bind public network access and enable the VPC Endpoint service. Dedicated load balancers provide more functions and higher performance than shared load balancers. This section describes how to connect a cluster to a dedicated load balancer.

Advantages of connecting a cluster to a dedicated load balancer:

- A non-security cluster can also use capabilities of the Elastic Load Balance (ELB) service.
- You can use customized certificates for HTTPS bidirectional authentication.
- Seven-layer traffic monitoring and alarm configuration are supported, allowing you to view the cluster status at any time.

There are eight service forms for clusters in different security modes to connect to dedicated load balancers. [Table 4-10](#) describes the ELB capabilities for the eight service forms. [Table 4-11](#) describes the configurations for the eight service forms.

NOTICE

You are not advised connecting a load balancer that has been bound to a public IP address to a non-security cluster. Access from the public network using such a load balancer may bring security risks because non-security clusters can be accessed over HTTP without security authentication.

Table 4-10 ELB capabilities for different clusters

Security Mode	Service Form Provided by ELB for External Systems	ELB Load Balancing	ELB Traffic Monitoring	ELB Two-way Authentication
Non-security	No authentication	Supported	Supported	Not supported
	One-way authentication Two-way authentication	Supported	Supported	Supported

Security Mode	Service Form Provided by ELB for External Systems	ELB Load Balancing	ELB Traffic Monitoring	ELB Two-way Authentication
Security mode + HTTP	Password authentication	Supported	Supported	Not supported
	One-way authentication + Password authentication	Supported	Supported	Supported
	Two-way authentication + Password authentication			
Security mode + HTTPS	One-way authentication + Password authentication Two-way authentication + Password authentication	Supported	Supported	Supported

Table 4-11 Configuration for interconnecting different clusters with ELB

Security Mode	Service Form Provided by ELB for External Systems	ELB Listener			Backend Server Group		
		Frontend Protocol	Port	SSL Parsing Mode	Backend Protocol	Health Check Port	Health Check Path
Non-security	No authentication	HTTP	9200	No authentication	HTTP	9200	/
	One-way authentication	HTTPS	9200	One-way authentication	HTTP	9200	
	Two-way authentication	HTTPS	9200	Two-way authentication	HTTP	9200	

Security mode + HTTP	Password authentication	HTTP	9200	No authentication	HTTP	9200	/_opendistro/_security/health
	One-way authentication + Password authentication	HTTPS	9200	One-way authentication	HTTP	9200	
	Two-way authentication + Password authentication	HTTPS	9200	Two-way authentication	HTTP	9200	
Security mode + HTTPS	One-way authentication + Password authentication	HTTPS	9200	One-way authentication	HTTPS	9200	
	Two-way authentication + Password authentication	HTTPS	9200	Two-way authentication	HTTPS	9200	

4.3.4.2 Connecting to a Dedicated Load Balancer

This section describes how to connect a CSS cluster to a dedicated load balancer.

(Optional) Preparing a Self-signed Certificate

If the target ELB listener uses the HTTP protocol, skip this step.

Prepare and upload a self-signed certificate.

NOTE

You are advised to use a certificate purchased in Cloud Certificate Manager (CCM) or issued by an authoritative organization.

1. Log in to a Linux client where the OpenSSL tool and JDK are installed.
2. Run the following commands to create a self-signed certificate:

```
mkdir ca
mkdir server
mkdir client

#Use OpenSSL to create a CA certificate.
cd ca
#Create the OpenSSL configuration file ca_cert.conf for the CA certificate.
cat >ca_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt = no
```

```
[ req_distinguished_name ]
O           = ELB
EOF
#Create private key file ca.key for the CA certificate.
openssl genrsa -out ca.key 2048
#Create the CSR file ca.csr for the CA certificate.
openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf
#Create a self-signed CA certificate ca.crt.
openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key
#Convert the CA certificate format to p12.
openssl pkcs12 -export -clcerts -in ca.crt -inkey ca.key -out ca.p12
#Convert the CA certificate format to JKS.
keytool -importkeystore -srckeystore ca.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
ca.jks

#Use the CA certificate to issue a server certificate.
cd ../server
#Create the OpenSSL configuration file server_cert.conf for the server certificate. Change the CN
field to the domain name or IP address of the server as required.
cat >server_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt             = no

[ req_distinguished_name ]
O           = ELB
CN          = 127.0.0.1
EOF
#Create the private key file server.key for the server certificate.
openssl genrsa -out server.key 2048
#Create the CSR request file server.csr for the server certificate.
openssl req -out server.csr -key server.key -new -config ./server_cert.conf
#Use the CA certificate to issue the server certificate server.crt.
openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -
CAkey ../ca/ca.key
#Convert the server certificate format to p12.
openssl pkcs12 -export -clcerts -in server.crt -inkey server.key -out server.p12
#Convert the service certificate format to JKS.
keytool -importkeystore -srckeystore server.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
server.jks

#Use the CA certificate to issue a client certificate.
cd ../client
#Create the OpenSSL configuration file client_cert.conf for the client certificate. Change the CN field
to the domain name or IP address of the server as required.
cat >client_cert.conf <<EOF
[ req ]
distinguished_name = req_distinguished_name
prompt             = no

[ req_distinguished_name ]
O           = ELB
CN          = 127.0.0.1
EOF
#Create private key client.key for the client certificate.
openssl genrsa -out client.key 2048
#Create the CSR file client.csr for the client certificate.
openssl req -out client.csr -key client.key -new -config ./client_cert.conf
#Use the CA certificate to issue the client certificate client.crt.
openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -
CAkey ../ca/ca.key
#Convert the client certificate to a p12 file that can be identified by the browser.
openssl pkcs12 -export -clcerts -in client.crt -inkey client.key -out client.p12
#Convert the client certificate format to JKS.
keytool -importkeystore -srckeystore client.p12 -srcstoretype PKCS12 -deststoretype JKS -destkeystore
client.jks
```


3. Upload the self-signed certificate. For details, see [Configuring the Server Certificate and Private Key](#).

Creating a Dedicated Load Balancer

1. Log in to the ELB management console.
2. Create a dedicated load balancer. For details, see [Creating a Dedicated Load Balancer](#). [Table 4-12](#) describes the parameters required for connecting a CSS cluster with a dedicated load balancer.

Table 4-12 Parameters for interconnecting a CSS cluster with a dedicated load balancer

Parameter	Description	Example
Type	Load balancer type. Select Dedicated .	Dedicated
Billing Mode	Billing mode of the dedicated load balancer.	Pay-per-use
Region	Region where the CSS cluster is located.	-
IP as Backend Servers	A CSS cluster can be connected only after the cross-VPC backend is enabled.	Enabled
Network Type	Type of the network used by the load balancer to provide services for external systems.	Private IPv4 network
VPC	VPC where the load balancer works. This parameter is mandatory no matter which network type is selected. Select the VPC of the CSS cluster.	-
Subnet	Subnet where the load balancer is to be created. This parameter is mandatory no matter which network type is selected. Select the subnet of the CSS cluster.	-

Parameter	Description	Example
Specifications	You are advised to select Application load balancing (HTTP/HTTPS) , which provides better functions and performance.	Application load balancing (HTTP/HTTPS) Small I

Interconnecting with a Load Balancer

NOTE

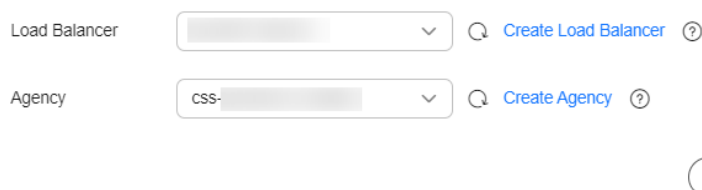
A cluster in security mode with HTTPS access enabled does not support HTTP protocol authentication. If you need to enable HTTP protocol authentication, disable the security mode of the cluster.



Before changing the security mode, disable load balancing. After the security mode is changed, enable load balancing.



1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page, select the cluster you want to connect to the load balancer and click the cluster name. The **Cluster Information** page is displayed.
4. In the navigation pane, choose **Load Balancing**. Enable load balancing and configure basic load balancing information.
 - **Load Balancer:** Select a created load balancer. You can also click **Create Load Balancer** to create one.
 - **Agency:** Select an agency name. If no agency is available, click **Create Agency** to create one. The selected agency must have the **ELB Administrator** and **ELB FullAccess** permissions.

Figure 4-8 Enabling load balancing

Basic Configuration

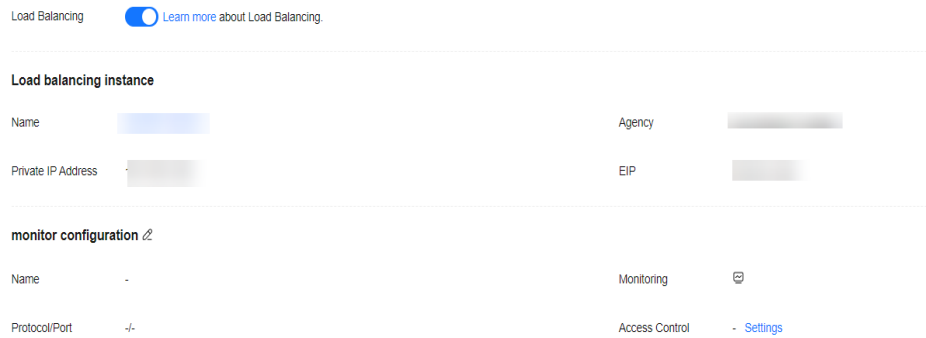


Load Balancer  [Create Load Balancer](#) 

Agency  [Create Agency](#) 

5. Click **OK**. The listener configuration page is displayed.

Figure 4-9 Creating a listener




- In the **Listener Configuration** area, click  to configure listener information.

Figure 4-10 Configuring a listener

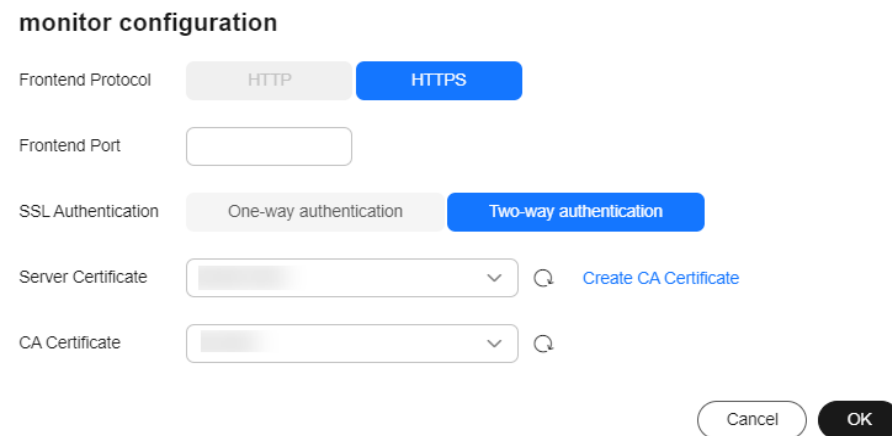


Table 4-13 Listener configuration information

Parameter	Description
Frontend Protocol	The protocol used by the client and listener to distribute traffic. Select a protocol as required.
Frontend Port	The port used by the client and listener to distribute traffic. For example, 9200. You need to specify this parameter as required.
SSL Authentication	Authentication mode for the client to access the server. Select a parsing mode as required.

Parameter	Description
Server Certificate	The server certificate is used for SSL handshake negotiation. The certificate content and private key must be provided. When SSL Authentication is set to Two-way authentication , this parameter is mandatory.
CA Certificate	Also called client CA public key certificate. It is used to verify the issuer of a client certificate. When the HTTPS two-way authentication is enabled, an HTTPS connection can be established only when the client can provide the certificate issued by a specified CA. This parameter is mandatory only when the Frontend Protocol is set to HTTPS .

- (Optional) In the **Connection Mode** area, you can click **Settings** next to **Access Control** to configure the IP addresses or network segments that are allowed to access the system. If you do not set the IP addresses or network segments, all IP addresses are allowed to access the system by default.

In the **Health Check** area, you can view the health check result of each node IP address. The following table describes the health check results.

Health Check Result	Description
Normal	The IP address of the node is properly connected.
Abnormal	The node IP address is connected and unavailable.

Accessing a Cluster Using the Curl Command

Run the following commands to check whether the dedicated load balancer can be connected to a cluster.

Table 4-14 Commands for accessing different clusters

Security Mode	Service Form Provided by ELB for External Systems	Curl Command for Accessing a Cluster
Non-security	No authentication	<code>curl http://IP:9200</code>

Security Mode	Service Form Provided by ELB for External Systems	Curl Command for Accessing a Cluster
	One-way authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200</code>
	Two-way authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200</code>
Security mode + HTTP	Password authentication	<code>curl http://IP:9200 -u user:pwd</code>
	One-way authentication + Password authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
	Two-way authentication + Password authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
Security mode + HTTPS	One-way authentication + Password authentication	<code>curl -k --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>
	Two-way authentication + Password authentication	<code>curl --cacert ./ca.crt --cert ./client.crt --key ./client.key https://IP:9200 -u user:pwd</code>

Table 4-15 Variables

Variable	Description
IP	ELB IP address
user	Username for accessing the CSS cluster
pwd	Password of the user

If the Elasticsearch cluster information is returned, the connection is successful. For example, if a security cluster using the HTTPS protocol is connected to a load balancer using two-way authentication, the information shown in [Figure 4-11](#) is returned.

Figure 4-11 Accessing a cluster

```
root@~# curl --cacert ./ca/ca.crt --cert ./client.crt --key ./client.key https://:9200 -u admin:
{"name": "css-test1-ess-asn-1-1",
 "cluster_name": "css-test1",
 "cluster_uuid": "nX81L1jT_2CMBFe1bgnA",
 "version": {
  "number": "7.10.2",
  "build_flavor": "oss",
  "build_type": "tar",
  "build_hash": "unknown",
  "build_date": "unknown",
  "build_snapshot": true,
  "lucene_version": "8.7.0",
  "minimum_wire_compatibility_version": "6.7.0",
  "minimum_index_compatibility_version": "6.0.0-beta1"
 },
 "tagline": "You Know, for Search"
}
```

4.3.4.3 Sample Code for Two-Way Authentication During the Access to a Cluster

This section provides the sample code for two-way authentication during the access to a cluster from a Java client.

ESSecuredClientWithCerDemo Code

```
import org.apache.commons.io.IOUtils;
import org.apache.http.auth.AuthScope;
import org.apache.http.auth.UsernamePasswordCredentials;
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.client.BasicCredentialsProvider;
import org.apache.http.HttpHost;
import org.apache.http.nio.conn.ssl.SSLIOSessionStrategy;
import org.elasticsearch.action.search.SearchRequest;
import org.elasticsearch.action.search.SearchResponse;
import org.elasticsearch.client.RequestOptions;
import org.elasticsearch.client.RestClient;
import org.elasticsearch.client.RestClientBuilder;
import org.elasticsearch.client.RestHighLevelClient;
import org.elasticsearch.index.query.QueryBuilders;
import org.elasticsearch.search.SearchHit;
import org.elasticsearch.search.SearchHits;
import org.elasticsearch.search.builder.SearchSourceBuilder;
import java.io.FileInputStream;
import java.io.IOException;
import java.security.KeyStore;
import java.security.SecureRandom;
import javax.net.ssl.HostnameVerifier;
import javax.net.ssl.KeyManagerFactory;
import javax.net.ssl.SSLContext;
import javax.net.ssl.SSLSession;
import javax.net.ssl.TrustManagerFactory;
public class ESSecuredClientWithCerDemo {
    private static final String KEY_STORE_PWD = "";
    private static final String TRUST_KEY_STORE_PWD = "";
    private static final String CA_JKS_PATH = "ca.jks";
    private static final String CLIENT_JKS_PATH = "client.jks";
    private static final String ELB_ADDRESS = "127.0.0.1";
    private static final int ELB_PORT = 9200;
    private static final String CSS_USERNAME = "user";
    private static final String CSS_PWD = "";
    public static void main(String[] args) {
        // Create a client.
        RestHighLevelClient client = initESClient(ELB_ADDRESS, CSS_USERNAME, CSS_PWD);
        try {
            // Search match_all, which is equivalent to {"query":{"match_all":{"}}}.

```

```
SearchRequest searchRequest = new SearchRequest();
SearchSourceBuilder searchSourceBuilder = new SearchSourceBuilder();
searchSourceBuilder.query(QueryBuilders.matchAllQuery());
searchRequest.source(searchSourceBuilder);
// query
SearchResponse searchResponse = client.search(searchRequest, RequestOptions.DEFAULT);
System.out.println("query result: " + searchResponse.toString());
SearchHits hits = searchResponse.getHits();
for (SearchHit hit : hits) {
    System.out.println(hit.getSourceAsString());
}
System.out.println("query success");
Thread.sleep(2000L);
} catch (InterruptedException | IOException e) {
    e.printStackTrace();
} finally {
    IOUtils.closeQuietly(client);
}
}
private static RestHighLevelClient initESClient(String clusterAddress, String userName, String password) {
    final CredentialsProvider credentialsProvider = new BasicCredentialsProvider();
    credentialsProvider.setCredentials(AuthScope.ANY, new UsernamePasswordCredentials(userName,
password));
    SSLContext ctx = null;
    try {
        KeyStore ks = getKeyStore(CLIENT_JKS_PATH, KEY_STORE_PWD, "JKS");
        KeyManagerFactory kmf = KeyManagerFactory.getInstance("SunX509");
        kmf.init(ks, KEY_STORE_PWD.toCharArray());
        KeyStore tks = getKeyStore(CA_JKS_PATH, TRUST_KEY_STORE_PWD, "JKS");
        TrustManagerFactory tmf = TrustManagerFactory.getInstance("SunX509");
        tmf.init(tks);
        ctx = SSLContext.getInstance("SSL", "SunJSSE");
        ctx.init(kmf.getKeyManagers(), tmf.getTrustManagers(), new SecureRandom());
    } catch (Exception e) {
        e.printStackTrace();
    }
    SSLIOSessionStrategy sessionStrategy = new SSLIOSessionStrategy(ctx, new HostnameVerifier() {
        @Override
        public boolean verify(String arg0, SSLSession arg1) {
            return true;
        }
    });
    SecuredHttpClientConfigCallback httpClientConfigCallback = new
SecuredHttpClientConfigCallback(sessionStrategy,
credentialsProvider);
    RestClientBuilder builder = RestClient.builder(new HttpHost(clusterAddress, ELB_PORT, "https"))
.setHttpClientConfigCallback(httpClientConfigCallback);
    RestHighLevelClient client = new RestHighLevelClient(builder);
    return client;
}
private static KeyStore getKeyStore(String path, String pwd, String type) {
    KeyStore keyStore = null;
    FileInputStream is = null;
    try {
        is = new FileInputStream(path);
        keyStore = KeyStore.getInstance(type);
        keyStore.load(is, pwd.toCharArray());
    } catch (Exception e) {
        e.printStackTrace();
    } finally {
        IOUtils.closeQuietly(is);
    }
    return keyStore;
}
}
```

SecuredHttpClientConfigCallback Code

```
import org.apache.http.client.CredentialsProvider;
import org.apache.http.impl.nio.client.HttpAsyncClientBuilder;
import org.apache.http.nio.conn.ssl.SSLIOStrategy;
import org.elasticsearch.client.RestClientBuilder;
import org.elasticsearch.common.Nullable;
import java.util.Objects;

class SecuredHttpClientConfigCallback implements RestClientBuilder.HttpClientConfigCallback {
    @Nullable
    private final CredentialsProvider credentialsProvider;
    /**
     * The {@link SSLIOStrategy} for all requests to enable SSL / TLS encryption.
     */
    private final SSLIOStrategy sslStrategy;
    /**
     * Create a new {@link SecuredHttpClientConfigCallback}.
     *
     * @param credentialsProvider The credential provider, if a username/password have been supplied
     * @param sslStrategy The SSL strategy, if SSL / TLS have been supplied
     * @throws NullPointerException if {@code sslStrategy} is {@code null}
     */
    SecuredHttpClientConfigCallback(final SSLIOStrategy sslStrategy,
        @Nullable final CredentialsProvider credentialsProvider) {
        this.sslStrategy = Objects.requireNonNull(sslStrategy);
        this.credentialsProvider = credentialsProvider;
    }
    /**
     * Get the {@link CredentialsProvider} that will be added to the HTTP client.
     *
     * @return Can be {@code null}.
     */
    @Nullable
    CredentialsProvider getCredentialsProvider() {
        return credentialsProvider;
    }
    /**
     * Get the {@link SSLIOStrategy} that will be added to the HTTP client.
     *
     * @return Never {@code null}.
     */
    SSLIOStrategy getSSLStrategy() {
        return sslStrategy;
    }
    /**
     * Sets the {@linkplain HttpAsyncClientBuilder#setDefaultCredentialsProvider(CredentialsProvider)
    credential provider},
     *
     * @param httpClientBuilder The client to configure.
     * @return Always {@code httpClientBuilder}.
     */
    @Override
    public HttpAsyncClientBuilder customizeHttpClient(final HttpAsyncClientBuilder httpClientBuilder) {
        // enable SSL / TLS
        httpClientBuilder.setSSLStrategy(sslStrategy);
        // enable user authentication
        if (credentialsProvider != null) {
            httpClientBuilder.setDefaultCredentialsProvider(credentialsProvider);
        }
        return httpClientBuilder;
    }
}
```

pom.xml Code

```
<?xml version="1.0" encoding="UTF-8"?>
<project xmlns="http://maven.apache.org/POM/4.0.0"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/
```



```
maven-4.0.0.xsd">
  <modelVersion>4.0.0</modelVersion>
  <groupId>1</groupId>
  <artifactId>ESClient</artifactId>
  <version>1.0-SNAPSHOT</version>
  <name>ESClient</name>

  <properties>
    <maven.compiler.source>8</maven.compiler.source>
    <maven.compiler.target>8</maven.compiler.target>
    <project.build.sourceEncoding>UTF-8</project.build.sourceEncoding>
    <elasticsearch.version>7.10.2</elasticsearch.version>
  </properties>
  <dependencies>
    <dependency>
      <groupId>org.elasticsearch.client</groupId>
      <artifactId>transport</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>org.elasticsearch</groupId>
      <artifactId>elasticsearch</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>org.elasticsearch.client</groupId>
      <artifactId>elasticsearch-rest-high-level-client</artifactId>
      <version>${elasticsearch.version}</version>
    </dependency>
    <dependency>
      <groupId>commons-io</groupId>
      <artifactId>commons-io</artifactId>
      <version>2.11.0</version>
    </dependency>
  </dependencies>
</project>
```

4.4 Index Backup and Restoration

4.4.1 Backup and Restoration Overview

You can back up index data in clusters. If data loss occurs or you want to retrieve data of a specified duration, you can restore the index data. Index backup is implemented by creating cluster snapshots. When creating a backup for the first time, you are advised to back up data of all indexes.

- **Managing Automatic Snapshot Creation:** Snapshots are automatically created at a specified time each day according to the rules you create. You can enable or disable the automatic snapshot creation function and set the automatic snapshot creation policy.
- **Manually Creating a Snapshot:** You can manually create a snapshot at any time to back up all data or data of specified indexes.
- **Restoring Data:** You can use existing snapshots to restore the backup index data to a specified cluster.
- **Deleting a Snapshot:** Delete unnecessary snapshots and release resources.

4.4.2 Managing Automatic Snapshot Creation

Snapshots are automatically created at a specified time each day according to the rules you create. You can enable or disable the automatic snapshot creation function and set the automatic snapshot creation policy.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- When creating a backup for the first time, you are advised to back up data of all indexes.
- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.
- Before creating a snapshot, you need to perform basic configurations, including configuring the OBS bucket for storing snapshots, the snapshot backup path, and IAM agency used for security authentication.
- If there are available snapshots in the snapshot list when you configure the OBS bucket for storing cluster snapshots, you cannot change the bucket for snapshots that are created later. Exercise caution when you configure the OBS bucket.
- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, enable the snapshot function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.
- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- The first snapshot of a cluster is a full snapshot, and subsequent snapshots are incremental snapshots. CSS snapshot files depend on each other.

Managing Automatic Snapshot Creation

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page that is displayed, click the name of the target cluster. In the navigation pane on the left, choose **Cluster Snapshots**.

Alternatively, on the **Clusters** page, locate the row that contains the target cluster and click **More > Back Up and Restore** in the **Operation** column to switch to the **Cluster Snapshots** page.



4. On the displayed **Cluster Snapshots** page, click the icon to the right of **Cluster Snapshot** to enable the cluster snapshot function.
5. Enable the cluster snapshot function. OBS buckets and IAM agencies are automatically created by CSS to store snapshots. The automatically created OBS bucket and IAM agency are displayed on the page. You can also click  on the right of **Basic Configuration** to edit the configuration.

Table 4-16 Cluster snapshot parameter

Parameter	Description
OBS bucket	Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket . The created or existing OBS bucket must meet the following requirements: <ul style="list-style-type: none"> • Storage Class is Standard. • Region must be the same as that of the created cluster.
Backup Path	Storage path of the snapshot in the OBS bucket. The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \\:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency . The created or existing IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

6. Enable the automatic snapshot creation function. The **Configure Automatic Snapshot Creation** dialog box is displayed. If the automatic snapshot creation function is enabled, you can click  on the right of **Automatic Snapshot Creation** to modify the snapshot policy.
 - **Snapshot Name Prefix:** Enter a maximum of 32 characters starting with a lowercase letter. Only lowercase letters, digits, hyphens (-), and

underscores (_) are allowed. A snapshot name consists of a snapshot name prefix and a timestamp, for example, **snapshot-2018022405925**.

- **Time Zone:** indicates the time zone for the backup time. Specify backup start time based on the time zone.
- **Index:** Enter an index name. You can select an index for backup. Use commas (,) to separate multiple indexes. Uppercase letters, spaces, and the following special characters are not allowed: "\<|>/? If you do not specify this parameter, data of all indexes in the cluster is backed up by default. You can use the asterisk (*) to back up data of certain indexes. For example, if you enter **index***, then data of indices with the name prefix of **index** will be backed up.
Run the **GET /_cat/indices** command in Kibana to query the names of all indexes in the cluster.
- **Backup Started:** indicates the time when the backup starts automatically every day. You can specify this parameter only in hours and not minutes, for example, **00:00** or **01:00**. The value ranges from **00:00** to **23:00**. Select the backup time from the drop-down list box.
- **Retention Period (days):** indicates the duration when snapshots are retained in the OBS bucket, in days. The value ranges from **1** to **90**. You can specify this parameter as required. The system automatically deletes snapshots that are retained over the specified retention period on the half hour. For example, if you set the snapshot policy as shown in [Figure 4-12](#), the system will automatically delete in 35 days at 00:30 the automated snapshots that were created 35 days earlier at 00:00.

Figure 4-12 Automatically creating a snapshot

Automatic Snapshot Creation

Snapshot Name Prefix X

Time Zone

Backup Start Time v v

? - | 7 | +

7. Click **OK** to save the snapshot policy.
Snapshots that are automatically created according to the snapshot policy are displayed in the snapshot list, along with manually created snapshots. You can distinguish them by the **Snapshot Type** setting. In the upper right corner of the snapshot list, enter the keyword of the snapshot name or snapshot ID to search for the desired snapshots.
8. (Optional) Disable the automatic snapshot creation function.
After you disable the automatic snapshot creation function, the system stops automatic creation of snapshots. If the system is creating a snapshot based on the automatic snapshot creation policy and the snapshot is not yet displayed in the snapshot list, you cannot disable the automatic snapshot creation function. In this case, if you click the button next to **Automatic Snapshot Creation**, a message is displayed, indicating that you cannot disable the function. You are advised to disable the function after the system completes automatic creation of the snapshot, and the created snapshot is displayed in the snapshot list.
When disabling the automatic snapshot creation function, you can choose whether to delete the snapshots that have been automatically created by selecting **Delete automated snapshots** in the displayed dialog box. By default, automatically created snapshots are not deleted.
 - If you do not select **Delete automated snapshots**, automatically created snapshots are not deleted when you disable the automatic snapshot creation function. You can manually delete them later. For details, see [Deleting a Snapshot](#). If you retain the automatically created snapshots and enable automatic snapshot creation again, then all snapshots whose **Snapshot Type** is **Automated** can only be automatically deleted by the system. Specifically, the system automatically deletes snapshots based on the snapshot policy configured when you enable the automatic snapshot creation function again. For example, if you set **Retention Period (days)** to **10**, the system will automatically delete the snapshots that have been retained for more than 10 days.
 - If you select **Delete automated snapshots**, all snapshots with **Snapshot Type** set to **Automated** in the snapshot list will be deleted when you disable the automatic snapshot creation function.

4.4.3 Manually Creating a Snapshot

You can manually create a snapshot at any time to back up all data or data of specified indexes.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- When creating a backup for the first time, you are advised to back up data of all indexes.
- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.
- Before creating a snapshot, you need to perform basic configurations, including configuring the OBS bucket for storing snapshots, the snapshot backup path, and IAM agency used for security authentication.
- If there are available snapshots in the snapshot list when you configure the OBS bucket for storing cluster snapshots, you cannot change the bucket for snapshots that are created later. Exercise caution when you configure the OBS bucket.
- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, then enable the snapshot function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.
- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- The first snapshot of a cluster is a full snapshot, and subsequent snapshots are incremental snapshots. CSS snapshot files depend on each other.

Manually Creating a Snapshot


1. In the CSS navigation pane on the left, click **Clusters**.
2. On the **Clusters** page that is displayed, click the name of the target cluster. In the navigation pane on the left, choose **Cluster Snapshots**.
Alternatively, on the **Clusters** page, locate the row that contains the target cluster and click **More > Back Up and Restore** in the **Operation** column to switch to the **Cluster Snapshots** page.
3. On the displayed **Cluster Snapshots** page, click the icon to the right of **Cluster Snapshot** to enable the cluster snapshot function.
4. Enable the cluster snapshot function. OBS buckets and IAM agencies are automatically created to store snapshots. The automatically created OBS bucket and IAM agency are displayed on the page. You can also click  on the right of **Basic Configuration** to edit the configuration.

Table 4-17 Cluster snapshot parameter

Parameter	Description
OBS bucket	Select an OBS bucket for storing snapshots from the drop-down list box. You can also click Create Bucket on the right to create an OBS bucket. For details, see Creating a Bucket . The created or existing OBS bucket must meet the following requirements: <ul style="list-style-type: none"> • Storage Class is Standard.
Backup Path	Storage path of the snapshot in the OBS bucket. The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*\?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The backup path cannot contain more than 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in OBS You can also click Create IAM Agency on the right to create an IAM agency. For details, see Creating an Agency . The created or existing IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • The agency must have the OBS Administrator permission for the OBS project in Global service.

5. After basic configurations are completed, click **Create**.
 - **Snapshot Name** indicates the name of the manually created snapshot, which can contain 4 to 64 characters and must start with a lowercase letter. Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. For snapshots you create manually, you can specify the snapshot name. The system will not automatically add the time information to the snapshot name.
 - **Index**: Enter an index name. You can select an index for backup. Use commas (,) to separate multiple indexes. Uppercase letters, spaces, and the following special characters are not allowed: "\<|>/? If you do not specify this parameter, data of all indexes in the cluster is backed up by default. You can use the asterisk (*) to back up data of certain indices. For example, if you enter **index***, then data of indices with the name prefix of **index** will be backed up.
Run the **GET /_cat/indices** command in Kibana to query the names of all indexes in the cluster.

- **Description:** indicates the description of the created snapshot. The value contains 0 to 256 characters, and certain special characters (<>) are not allowed.

Figure 4-13 Manually creating a snapshot

Create Snapshot

The screenshot shows a 'Create Snapshot' dialog box with the following fields and values:

- Snapshot Name:** A text input field containing 'snapshot-2125'.
- Index:** A text input field that is empty, with a character count of '0/1,024' in the bottom right corner.
- Description:** A text input field that is empty, with a character count of '0/256' in the bottom right corner.

At the bottom right of the dialog box, there are two buttons: 'Cancel' and 'OK'.

6. Click **OK**.

After the snapshot is created, it will be displayed in the snapshot list. The status **Available** indicates that the snapshot is created successfully, along with manually created snapshots. You can distinguish them by the **Snapshot Type** setting. In the upper right corner of the snapshot list, enter the keyword of the snapshot name or snapshot ID to search for the desired snapshots.

4.4.4 Restoring Data

You can use existing snapshots to restore the backup index data to a specified cluster.

Prerequisites

To use the function of creating or restoring snapshots, the account or IAM user logging in to the CSS management console must have both of the following permissions:

- **OBS Administrator** for project **OBS** in region **Global service**
- **Elasticsearch Administrator** in the current region

Precautions

- Cluster snapshots will increase the CPU usage and disk I/O. You are advised to take cluster snapshots during off-peak hours.

- If snapshots have been stored in the OBS bucket, the OBS bucket cannot be changed. You can disable the snapshot function, then enable the snapshot function, and specify a new OBS bucket. After you disable the snapshot function, you cannot use previously created snapshots to restore the cluster.
- If a cluster is in the **Unavailable** status, you can use the cluster snapshot function only to restore clusters and view existing snapshot information.
- During backup and restoration of a cluster, you can perform only certain operations, including scaling out, accessing Kibana, viewing metric, and deleting other snapshots of clusters. However, you cannot perform the following operations: restarting or deleting the cluster, deleting a snapshot that is in the **Creating** or **Restoring** status, and creating or restoring another snapshot. If a snapshot is being created or restored for a cluster, any automatic snapshot creation task initiated for the cluster will be canceled.
- Cluster data cannot be queried during snapshot restoration.
- If you restore a CSS cluster snapshot to another cluster, indexes with the same name in the destination cluster will be overwritten. If the snapshot and the destination cluster use different shards, the indexes with the same name will not be overwritten.
- The version of the destination cluster used for restoration must be the same as or higher than that of the source cluster.

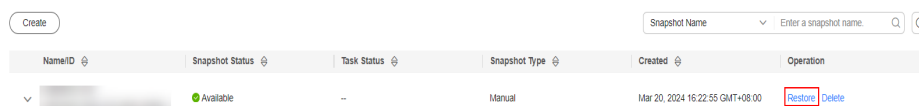
Restoring Data

You can use snapshots whose **Snapshot Status** is **Available** to restore cluster data. The stored snapshot data can be restored to other clusters.

Restoring data will overwrite current data in clusters. Therefore, exercise caution when restoring data.

1. In the **Snapshots** area, locate the row that contains the snapshot you want to restore and click **Restore** in the **Operation** column.

Figure 4-14 Selecting a snapshot



2. On the **Restore** page, set restoration parameters.

Index: Enter the name of the index you want to restore. If you do not specify any index name, data of all indexes will be restored. The value can contain 0 to 1,024 characters. Uppercase letters, spaces, and certain special characters (including "\<|>/?") are not allowed. You can use the asterisk (*) to match multiple indexes. For example, **index*** indicates that all indexes with the prefix **index** in snapshots are restored.

Rename Pattern: Enter a regular expression. Indexes that match the regular expression are restored. The default value **index_(.+)** indicates restoring data of all indexes. The value contains 0 to 1,024 characters. Uppercase letters, spaces, and certain special characters (including "\<|>/?,) are not allowed.

Rename Replacement: Enter the index renaming rule. The default value **restored_index_\$1** indicates that **restored_** is added in front of the names of all restored indexes. The value contains 0 to 1,024 characters. Uppercase

letters, spaces, and certain special characters (including "\<|>/?," are not allowed.

 **NOTE**

The **Rename Pattern** and **Rename Replacement** take effect only when they are configured at the same time.

Cluster: Select the name of the cluster to be restored. You can select the cluster of the current version. However, you can only restore the snapshot to clusters whose status is **Available**. If the status of the current cluster is **Unavailable**, you cannot restore the snapshot to the current cluster. If the target cluster you selected has an index with the same name as the original cluster, data in the index will be overwritten after the restoration. Exercise caution when performing this operation.

Figure 4-15 Restoring a snapshot

Restore

Index ?

Rename Pattern ?

Rename Replacement ?

★ Cluster ?

3. Click **OK**. If restoration succeeds, **Task Status** of the snapshot in the snapshot list will change to **Restoration succeeded**, and the index data is generated again according to the snapshot information.

Figure 4-16 Successful restoration

NameID	Snapshot Status	Task Status	Snapshot Type	Created	Operation
<input type="checkbox"/>	Available	Restoration succeeded	Manual	Mar 20, 2024 12:07:06 GMT+08:00	Restore Delete

4.4.5 Deleting a Snapshot

If you no longer need a snapshot, delete it to release storage resources. If the automatic snapshot creation function is enabled, snapshots that are automatically created cannot be deleted manually, and the system automatically deletes these snapshots on the half hour after the time specified by **Retention Period (days)**. If you disable the automatic snapshot creation function while retaining the automated snapshots, then you can manually delete them later. If you do not manually delete the automatically created snapshots and enable the automatic

snapshot creation function again, then all snapshots with **Snapshot Type** set to **Automated** in the snapshot list of the cluster can only be automatically deleted by the system.

NOTE

After a snapshot is deleted, its data cannot be restored. Exercise caution when deleting a snapshot.

1. In the snapshot list, locate the snapshot that you want to delete.
2. Click **Delete** in the **Operation** column. In the dialog box that is displayed, confirm the snapshot information and click **OK**.

4.5 Scaling In/Out a Cluster

4.5.1 Overview

You can scale in or out a cluster and change cluster specifications. In this way, you can improve cluster efficiency and reduce O&M costs.

Scaling Out a Cluster

- If a data node (ess) processes many data writing and querying requests and responds slowly, you can expand its storage capacity to improve its efficiency. If some nodes turn unavailable due to the excessive data volume or misoperations, you can add new nodes to ensure the cluster availability.
- **Cold data nodes** (ess-cold) are used to share the workload of data nodes. To prevent cold data loss, you can expand the storage capacity of the cold data node or add new ones.

Changing Specifications

- If the allocation of new indexes or shards takes too long or the node coordination and scheduling are inefficient, you can change the master node (ess-master) specifications.
- If too many tasks need to be distributed or too many results have been aggregated, you can change the client node (ess-client) specifications.
- If the speed of data writing and query decreases suddenly, you can change the data node (ess) specifications.
- If cold data query becomes slow, you can change the cold node (ess-cold) specifications.

Scaling in a Cluster

- If a cluster can process existing data without fully using its resources, you can scale in the cluster to reduce costs.

Removing Specified Nodes

- If a cluster can process existing data without fully using its nodes, you can remove one or more specified nodes from the cluster to reduce costs.

Replacing a Specified Node

- If a node in the cluster is faulty, you can create a new node with the same specifications to replace it.

Adding Master/Client Nodes

- If the workloads on the data plane of a cluster increase, you can dynamically scale the cluster by adding master/client nodes.

Changing the Security Mode

After a cluster is created, its security mode can be changed in the following methods:

- Change a non-security cluster to a security cluster that uses HTTP or HTTPS protocol.
- Change a security cluster that uses HTTP or HTTPS protocol to a non-security cluster.
- Change the protocol of a security cluster.

Changing AZs

You can **Add AZ** or **Migrate AZ**.

- **Add AZ:** Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster to improve cluster availability.
- **Migrate AZ:** Completely migrate data from the current AZ to another AZ that has sufficient resources.

4.5.2 Scaling Out a Cluster

If the workloads on the data plane of a cluster change, you can scale out the cluster by increasing the number or capacity of its nodes. Services are not interrupted during cluster scale-out.

Prerequisites

- The target cluster is available and has no tasks in progress.
- The target cluster has sufficient quotas available.

Constraints

- The **Node Specifications** cannot be modified during scale-out. You can modify **Node Specifications** by referring to [Changing Specifications](#).
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- The number of nodes and node storage capacity cannot be expanded at the same time for a yearly/monthly cluster.
- The quota of nodes in different types varies. For details, see [Table 4-18](#).

Table 4-18 Number of nodes in different types

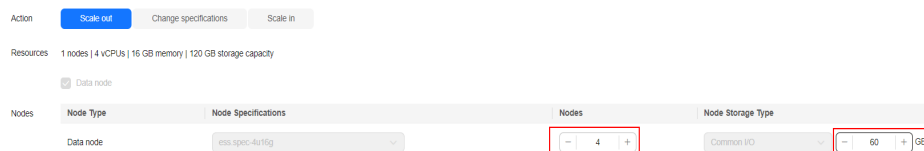
Node Type	Number
ess	ess: 1-32
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9

Node Type	Number
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale out** to set parameters.
 - **Action**: Select **Scale out**.
 - **Resource**: The changed amount of resources.

- **Nodes:** The number of nodes and node storage capacity of the default data node.
 - **Nodes:** For details, see [Table 4-18](#).
 - The value range of **Node Storage Type** depends on the **Node Specifications**. The value must be a multiple of 20.

Figure 4-17 Scaling out a cluster

5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling out**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled out.

4.5.3 Changing Specifications

If the workloads on the data plane of a cluster change, you can change its node specifications as needed.

Prerequisites

- The target cluster is available and has no tasks in progress.
- The target cluster has sufficient quotas available.
- When changing the node specifications, ensure that all service data has copies so the services will not be interrupted.

Run the **GET _cat/indices?v** command in Kibana. If the returned **rep** value is greater than **0**, the data has copies. If the returned **rep** value is **0**, the data has no copies. In this case, create snapshot for the cluster by referring to [Manually Creating a Snapshot](#).

- If the data volume is large, it may take long to modify the node specifications. You are advised to modify specifications during off-peak hours.

Constraints

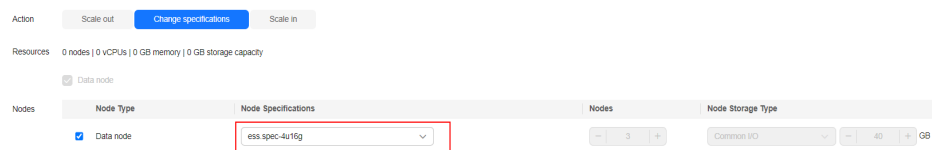
- The number of nodes and the capacity of node storage cannot be changed. You can add nodes and increase the node storage capacity by referring to [Scaling Out a Cluster](#). For details about how to reduce the number of nodes, see [Scaling in a Cluster](#).
- After decreasing cluster specifications, the cluster performance will deteriorate and service capabilities will be affected. Exercise caution when performing this operation.
- If a cluster has multiple node types, you can change the specifications of only one type at a time. After the change, nodes in other types still maintain their original specifications.

- Kibana is unavailable during specification change.
- During the specification modification, the nodes are stopped and restarted in sequence. It is a rolling process.

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Change Specifications** to set parameters.
 - **Action:** select **Change specifications**.
 - **Resources:** The changed amount of resources.
 - **Nodes:** Specifications of the default data nodes. Select the required specifications from the **Node Specifications** drop-down list and select the node that you want to change the specifications.
 - If a cluster has master nodes, client nodes, or cold data nodes, you can change their specifications.

Figure 4-18 Changing cluster specifications



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. In the dialog box that is displayed, confirm whether to select **Verify index copies** and **Cluster status check** and click **OK** to start the specifications change.

Index copy verification:

By default, CSS checks for indexes that do not have copies. You can skip this step, but the lack of index copies may affect services during a cluster specifications change.

- If you selected **Verify index copies** and the cluster has no master node, indexes must have at least one copy and the cluster must have at least three nodes.
- If you selected **Verify index copies** and the cluster has no master node, indexes must have at least one copy.

Cluster status check:

The cluster status is checked before the specifications change by default. The specifications of nodes are changed one by one to ensure success and data security. If a cluster is overloaded and services are faulty, the request for a specifications change will not be delivered. In this case, you can disable cluster status check. If you ignore the cluster status check before the specifications

change, the cluster may be faulty and services may be interrupted. Exercise caution when performing this operation.

8. Click **Back to Cluster List** to switch to the **Clusters** page. The **Cluster Status** is **Configuration modified**. When **Cluster Status** changes to **Available**, the cluster specifications have been successfully modified.

4.5.4 Scaling in a Cluster

If a cluster can process existing data without fully using its resources, you can scale in the cluster to reduce costs. You are advised to scale in clusters during off-peak hours.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Only the number of nodes can be modified during cluster scale-in. The node specifications and node storage capacity cannot be modified. You can modify node specifications by referring to [Changing Specifications](#). You can modify node storage capacity by referring to [Scaling Out a Cluster](#).
- If you change the number and storage capacity of a specified type of node, nodes in other types will not be changed.
- Ensure that the disk usage after scale-in is less than 80% and each AZ of each node type has at least one node.
- When scaling in a cluster, the data in the node to be deleted is migrated to other nodes. The timeout threshold for data migration is five hours. If data migration is not complete within 5 hours, the cluster scale-in fails. You are advised to perform scale-in for multiple times when the cluster has huge amounts of data.
- For a cluster without master nodes, the number of remaining data nodes (including cold data nodes and other types of nodes) after scale-in must be greater than half of the original node number, and greater than the maximum number of index replicas.
- For a cluster with master nodes, the number of removed master nodes in a scale-in must be fewer than half of the original master node number. After scale-in, there has to be an odd number of master nodes, and there has to be at least three of them.
- A cluster with two nodes cannot be scaled in. You can create a cluster using a single node and then [Migrating Cluster Data Through Backup and Restoration](#).
- The quota of nodes in different types varies. For details, see [Table 4-19](#).

Table 4-19 Number of nodes in different types

Node Type	Number
ess	ess: 1-32

Node Type	Number
ess, ess-master	ess: 1-200 ess-master: an odd number ranging from 3 to 9
ess, ess-client	ess: 1-32 ess-client: 1-32
ess, ess-cold	ess: 1-32 ess-cold: 1-32
ess, ess-master, ess-client	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32
ess, ess-master, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-cold: 1-32
ess, ess-client, ess-cold	ess: 1-32 ess-client: 1-32 ess-cold: 1-32
ess, ess-master, ess-client, ess-cold	ess: 1-200 ess-master: an odd number ranging from 3 to 9 ess-client: 1-32 ess-cold: 1-32
<p>Details about the four node types:</p> <ul style="list-style-type: none"> ● ess: the default node type that is mandatory for cluster creation. The other three node types are optional. ● ess-master: master node ● ess-client: client node ● ess-cold: cold data node 	

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Scale Cluster** tab and click **Scale in** to set parameters.

- **Action:** Select **Scale in**.
- **Resources:** The changed amount of resources.
- **Agency:** Select an IAM agency to grant the current account the permission to switch AZs.

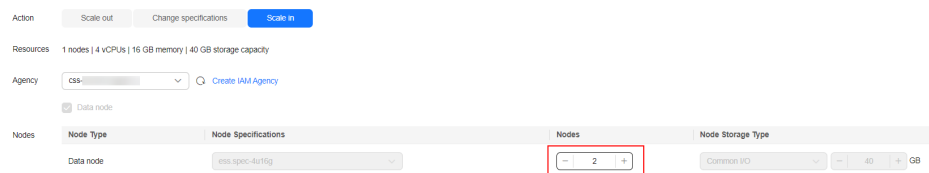
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

NOTE

The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.

- **Nodes:** The number of the default data nodes. For details about the value range that can be changed, see [Table 4-19](#).

Figure 4-19 Scaling in a cluster



5. Click **Next**.
6. Confirm the information and click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling in**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled in.

4.5.5 Removing Specified Nodes

If a cluster can process existing data without fully using its nodes, you can remove one or more specified nodes from the cluster to reduce costs. You are advised to scale in clusters during off-peak hours.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Ensure that the disk usage after scale-in is less than 80% and each AZ of each node type has at least one node.
- In a cross-AZ cluster, the difference between the numbers of the same type nodes in different AZs cannot exceed 1.
- For a cluster without master nodes, the number of removed data nodes and cold data nodes in a scale-in must be fewer than half of the original number of data nodes and cold data nodes, and the number of remaining data nodes and cold data nodes after a scale-in must be greater than the maximum number of index replicas.
- For a cluster with master nodes, the number of removed master nodes in a scale-in must be fewer than half of the original master node number. After

scale-in, there has to be an odd number of master nodes, and there has to be at least three of them.

Procedure

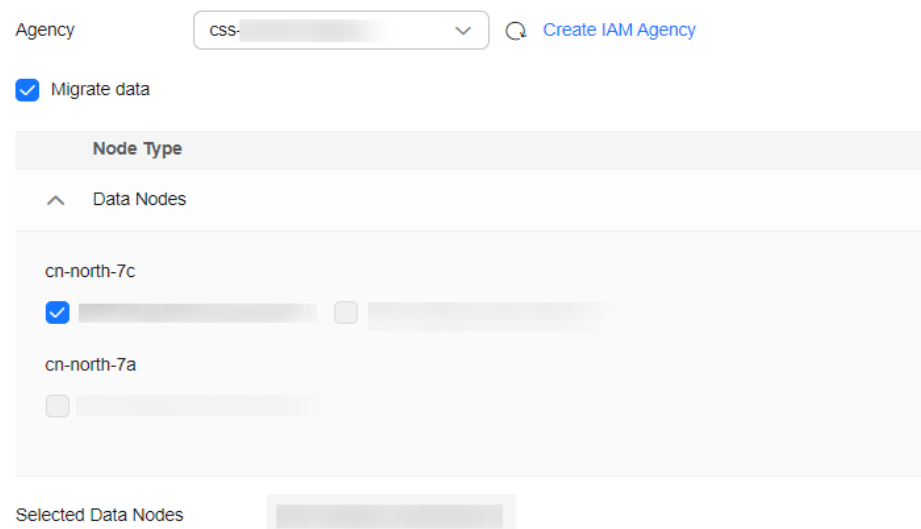
1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, click the **Scale In** tab.
5. On the **Scale In** tab page, set the following parameters:
 - **Agency:** Select an IAM agency to grant the current account the permission to switch AZs.
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

NOTE

The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.

- **Whether to perform data migration:** If this option is selected, data migration is performed. If the target node contains disabled indexes or indexes that have no replicas, this option must be selected.
- In the data node table, select the node to be scaled in.

Figure 4-20 Deleting specified nodes



6. Click **Next**.
7. Confirm the information and click **Submit**.
8. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Scaling in**. When **Cluster Status** changes to **Available**, the cluster has been successfully scaled in.

4.5.6 Replacing a Specified Node

If a node in the cluster is faulty, you can create a new node with the same specifications to replace it.

Prerequisites

The target cluster is available and has no tasks in progress.

Constraints

- Only one node can be replaced at a time.
- The ID, IP address, specifications, and AZ of the new node will be the same as those of the original one.
- The configurations you modified manually will not be retained after node replacement. For example, if you have manually added a return route to the original node, you need to add it to the new node again after the node replacement is complete.
- If the node you want to replace is a data node (ess) or cold data node (ess-cold), pay attention to the following precautions:
 - a. For data node replacement, data from the original node will be migrated to other nodes, and then the node will be rebuilt. Therefore, the total number of replicas and primary shards of each index in the cluster must be less than the total number of data nodes (including ess and ess-cold) in the cluster. The time required for node replacement is closely related to the time required for migrating data to other nodes.
 - b. The AZ of the node to be replaced must have two or more data nodes (including ess and ess-cold).
 - c. If the cluster of the node to be replaced does not have a master node (ess-master), the number of available data nodes (including ess and ess-cold) in the cluster must be greater than or equal to 3.
 - d. The preceding precautions do not apply if you are replacing a master node (ess-master) or client node (ess-client).
 - e. The precautions 1 to 4 do not apply if you are replacing a faulty node, regardless of its type. Faulty nodes are not included in `_cat/nodes`.

Procedure

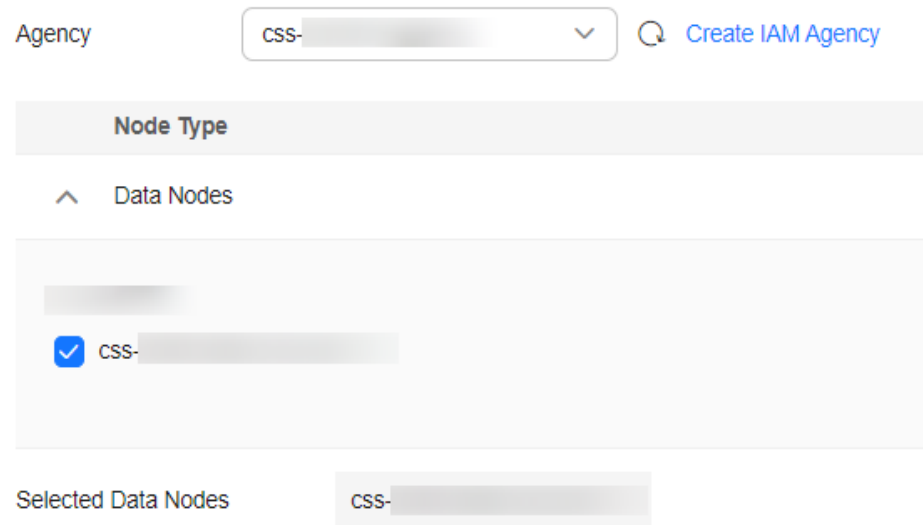
1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, click the **Replace Node** tab.
5. On the **Replace Node** tab page, set the following parameters:
 - **Agency**: Select an IAM agency to grant the current account the permission to switch AZs.
If no agency is available, click **Create IAM Agency** to go to the IAM console and create an agency.

 **NOTE**

The selected agency must be assigned the **Tenant Administrator** or **VPC Administrator** policy.

- **Whether to perform data migration:** If this option is selected, data migration is performed. If the target node has disabled indexes or indexes that have no replicas, this option must be selected.
- Select the node to be replaced in the data node table.

Figure 4-21 Replacing a specified node



6. Click **Submit**.
7. Click **Back to Cluster List** to switch to the **Clusters** page. The **Task Status** is **Upgrading**. When **Cluster Status** changes to **Available**, the node has been successfully replaced.

4.5.7 Adding Master/Client Nodes

If workloads on the data plane of a cluster increase, you can add master or client nodes as needed. Services are not interrupted while they are added.

Prerequisites

The cluster is in the **Available** state and has no ongoing task.

Constraints

- If a cluster already has master and client nodes, the **Add Master/Client Node** tab is not displayed on the **Modify Configuration** page. In this case, you need to add the master or client nodes by referring to [Scaling Out a Cluster](#).
- When you add master or client nodes, the number of nodes that can be configured varies depending on the node type. For details, see [Table 4-20](#).

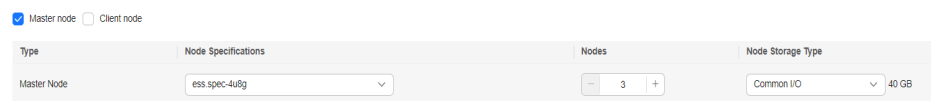
Table 4-20 Number of nodes in different types

Node Type	Number
Master node	An odd number ranging from 3 to 9
Client node	1 to 32

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. On the **Modify Configuration** page, choose the **Add Master/Client Node** tab.
5. Select the target node type and set the node specifications, quantity, and storage.
 - Master and client nodes cannot be added at the same time.
 - If a cluster already has a master or client node, you can only add nodes of the other type.

Figure 4-22 Adding a master or client node



6. Click **Next**.
7. Confirm the information and click **Submit**.
Return to the cluster list page. The **Task Status** of the cluster is **Scaling out**.
 - If you added a master node and **Cluster Status** changed to **Available**, the master node has been successfully added.
 - If you added a client node and **Cluster Status** changed to **Available**, the client node has been added. You can restart data nodes and cold data nodes to shut down Cerebro and Kibana processes on the nodes.

4.5.8 Changing the Security Mode

After a cluster is created, its security mode can be changed in the following methods:

- [Switching from the Non-Security Mode to Security Mode](#)
- [Switching from the Security to Non-Security Mode](#)
- [Switching the Protocol of Security Clusters](#)

Context

You can create clusters in multiple security modes. For details about the differences between security modes, see [Table 4-21](#).

Table 4-21 Cluster security modes

Security Mode	Scenario	Advantage	Disadvantage
Non-Security Mode	Intranet services and test scenarios	Simple. Easy to access.	Poor security. Anyone can access such clusters.
Security Mode + HTTP Protocol	User permissions can be isolated, which is applicable to scenarios sensitive to cluster performance.	Security authentication is required for accessing such clusters, which improves cluster security. Accessing a cluster through HTTP protocol can retain the high performance of the cluster.	Cannot be accessed from the public network.
Security Mode + HTTPS Protocol	Scenarios that require high security and public network access.	Security authentication is required for accessing such clusters, which improves cluster security. HTTPS protocol allows public network to access such clusters.	The performance of clusters using HTTPS is 20% lower than that of using HTTP.

Prerequisites

- You are advised to back up data before changing the cluster security mode.
- The target cluster is available and has no tasks in progress.

Constraints

- A cluster automatically restarts when its security mode is being changed. Services are interrupted during the restart. The authentication mode for calling the cluster will change after the restart, and client configurations need to be adjusted accordingly.
- If a cluster has already opened the Kibana session box, a session error message will be displayed after you change the cluster security mode. In this case, clear the cache and open Kibana again.

Switching from the Non-Security Mode to Security Mode

You can change a non-security cluster to a security cluster that uses HTTP or HTTPS. After a cluster's security mode is enabled, security authentication is required for accessing the cluster.

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. Choose the **Configure Security Mode** tab.
5. Enable the security mode. Enter and confirm the administrator password of the cluster.

Figure 4-23 Changing the security mode

6. Enable or disable **HTTPS Access**.
 - If you enable **HTTPS Access**: The HTTPS protocol is used to encrypt cluster communication and you can configure public networks to access the cluster.
 - If you disable **HTTPS Access**: The HTTP protocol is used and you cannot configure public networks to access the cluster.
7. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

Switching from the Security to Non-Security Mode

You can change a security cluster that uses HTTP or HTTPS to a non-security cluster. After a cluster's security mode is disabled, security authentication is no longer required for accessing the cluster.

NOTICE

- Clusters in non-security mode can be accessed without security authentication, and HTTP protocol is used to transmit data. Ensure the security of the cluster access environment and do not expose the access interface to the public network.
- During the switchover from the security mode to the non-security mode, the indexes of the original security cluster will be deleted. Back up data before disabling the security mode.
- If a security cluster has been bound to a public IP address, unbind it before changing the security mode.
- If a security cluster has enabled Kibana public network access, disable it before changing the security mode.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters**. On the displayed **Clusters** page, locate the target cluster and choose **More > Modify Configuration** in the **Operation** column.
3. Choose the **Configure Security Mode** tab.
4. Disable the security mode.

Figure 4-24 Disabling the security mode

Security Mode After the security mode is disabled, the cluster can be accessed without security authentication and data is transmitted in plaintext using HTTP. Therefore, ensure the security of the access environment and do not expose access interfaces to the public network.

5. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

Switching the Protocol of Security Clusters

You can change the protocol of a security cluster.

NOTICE

If a security cluster has been bound to a public IP address, you need to unbind it before changing HTTPS protocol to HTTP.

1. Log in to the CSS management console.
2. In the navigation pane on the left, choose **Clusters**. On the displayed **Clusters** page, locate the target cluster and choose **More > Modify Configuration** in the **Operation** column.
3. Choose the **Configure Security Mode** tab.
4. Enable or disable **HTTPS Access**.

Figure 4-25 Configuring the protocol

Security Mode	<input checked="" type="checkbox"/> If you enable this option, you will need to verify your identity to access the cluster
Administrator Username	admin
Administrator Password	<input type="password"/>
Confirm Password	<input type="password"/>
HTTPS Access	<input checked="" type="checkbox"/> If you enable this option, communication will be encrypted when you access the cluster.

- If you enable **HTTPS Access**:
HTTPS protocol is used to encrypt cluster communication and you can configure public network access.
- If you disable **HTTPS Access**: An alarm message is displayed. Click **OK** to disable the function.

When the HTTP protocol is used, cluster communication is no longer encrypted and the public network access function cannot be enabled.

5. Click **Submit**. Confirm the information and the cluster list page is displayed. The **Task Status** of the cluster is **The security mode is changing**. When the cluster status changes to **Available**, the security mode has been successfully changed.

4.5.9 Changing AZs

CSS supports cross-AZ deployment. You can add an AZ to obtain more resources or improve cluster availability, and can migrate your current AZ to one with higher specifications. This section describes how to add or migrate your AZs.

Description

You can **Add AZ** or **Migrate AZ**.

- **Add AZ**: Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster to improve cluster availability.
- **Migrate AZ**: Completely migrate data from the current AZ to another AZ that has sufficient resources.

Prerequisites

- Ensure that an AZ with sufficient resources exists.
- The target cluster is available and has no tasks in progress.
- Make sure that no non-standard operations have been performed in the cluster. If you have made non-standard modifications, such as modifying return routes, system parameters, and Kibana configurations, these modifications will be lost after the AZ change and your services may be affected.

Constraints

- To ensure service continuity, the total number of data nodes and cold data nodes in a cluster must be greater than or equal to 3.
- During the change, nodes are brought offline one by one and then new nodes are created. Ensure that the disk capacity of other nodes can store all the data of the node after a single node is brought offline.
- To prevent backup allocation failures after a node is brought offline during the change, ensure that the maximum number of primary and standby index shards of an index can be allocated to the remaining data nodes and cold data nodes. That is, the maximum number of primary and standby shards of an index plus 1 is less than or equal to the total number of data nodes and cold data nodes in the current cluster.
- You are advised to back up data before the change to prevent data loss caused by upgrade faults.
- Before a change completes, some nodes may have been moved to a new AZ. In this case, the AZs before and after the change are both displayed. After the change succeeds, the new AZs and their nodes will be displayed properly.
- When adding AZs, the current AZ must be retained in the change. When adding one or two AZs to a single-AZ cluster, you must change AZs for all nodes at the same time. When adding an AZ to a dual-AZ cluster, you can change AZs for a single type of nodes or all nodes in a cluster at a time. For example, in a cluster using the dual-AZ architecture, you can use the three-AZ architecture for master nodes alone. During HA modification, the nodes with the smallest configurations are modified to rebuild the cluster. After the HA modification is complete, the YML configuration of the nodes that are not modified is also updated. You need to restart the cluster to make the modification take effect.
- When migrating an AZ, you can select only one target AZ. You can migrate AZs for a single type of nodes or all nodes in a cluster at a time. For example, in a cluster with two AZs, you can migrate the AZ of the master node to the other AZ. After adding AZs, you need to restart the cluster to make the modification take effect.

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose a cluster type. The cluster management page is displayed.
3. Choose **More > Modify Configuration** in the **Operation** column of the target cluster. The **Modify Configuration** page is displayed.
4. Click the **Change AZ** tab.
5. On the **Change AZ** page, set parameters.

Table 4-22 Parameters for changing AZs

Parameter	Description
Operation Type	<ul style="list-style-type: none"> • Add AZ: Add one or two AZs to a single-AZ cluster, or add an AZ to a dual-AZ cluster. During HA modification, the nodes with the smallest configurations are modified to rebuild the cluster. After the HA modification is complete, the YML configuration of the nodes that are not modified is also updated. You need to restart the cluster to make the modification take effect. • Migrate AZ: Migrate data from one AZ to another. After adding AZs, you need to restart the cluster to make the modification take effect.
Node Type	Select a type of node or All nodes to change their AZ. NOTE When adding one or two AZs to a single-AZ cluster, you can only select All nodes to change AZs for all nodes at a time.
Current AZ	Current AZ of a cluster
Target AZ	Target AZ. <ul style="list-style-type: none"> • Add AZ: Select up to three AZs, which must include all your current AZs. • Migrate AZ: Select only one target AZ, which cannot be your current AZ.
Agency	Select an IAM agency to grant the current account the permission to change AZs. If no agencies are available, click Create IAM Agency to go to the IAM console and create an agency. NOTE The selected agency must be authorized with the Tenant Administrator or VPC Administrator policy.

6. Click **Submit**. Determine whether to check for the backup of all indexes and click **OK** to start the change.

Figure 4-26 Checking full index snapshots

Check Full Index Snapshot

You are advised to check the full index snapshot before the change.
 If you have backed up data, check whether the backup period exceeds one month. If yes, back up the latest data.
 Currently, only the index name can be matched, and the specific content and backup time cannot be verified.

Check full index snapshot

Cancel **OK**

7. The current AZ change task is displayed in the task list. If the task status is **Running**, expand the task list and click **View Progress** to view the progress details.
 If the task status is **Failed**, you can retry or terminate the task.

- Retry a task: Click **Retry** in the **Operation** column of a task.
- Terminate a task: Click **Terminate** in the **Operation** column of a task.

If the AZ of the original node is not changed after the task is terminated, you can recover the node by referring to [Replacing a Specified Node](#).

 **NOTE**

If the AZ of some nodes have been changed and the AZ form of the cluster has changed, stopping the switchover task may make the deliver of the previous switchover request fail. Exercise caution when stopping the switchover task.


4.6 Managing Clusters

4.6.1 Viewing Basic Information About an Opensearch Cluster

On the basic information page of an Opensearch cluster, you can view the private network address, public network address, version, and node of the cluster.

1. Log in to the CSS management console.
2. Choose **Clusters > Opensearch**. The cluster list page is displayed.
3. Click a cluster name to go to the **Cluster Information** page and view the basic information about the cluster.

Table 4-23 Basic information

Type	Parameter	Description
Cluster Information	Name	Cluster name. The name can be customized. You can click  on the right to change the cluster name.
	ID	Unique ID of a cluster, which is automatically generated by the system. Each cluster in the same region has a unique ID.
	Version	Cluster version information.
	Cluster Status	Current status of a cluster
	Task Status	Current task status of a cluster. If no task is in progress, -- is displayed.
	Created	Time when a cluster was created
	Cluster Storage Capacity (GB)	Storage capacity of a cluster
	Used Cluster Storage (GB)	Used storage capacity of a cluster

Type	Parameter	Description
Configuration	Region	Region where a cluster is located
	AZ	AZ where a cluster is located
	VPC	VPC to which the cluster belongs
	Subnet	Subnet to which the cluster belongs
	Security Group	<p>Security group to which a cluster belongs.</p> <p>To change the security group of a cluster, click Change Security Group on the right.</p> <p>NOTICE</p> <p>Before changing the security group, ensure that the port 9200 required for service access has been enabled. Incorrect security group configuration may cause service access failures. Exercise caution when performing this operation.</p>
	Security Mode	<p>Security mode of a cluster.</p> <ul style="list-style-type: none"> Enabled: The current cluster is a security cluster. Disabled: The current cluster is a non-security cluster.
Reset Password	<p>This parameter is displayed only for security clusters.</p> <p>Click Reset to change the password of the administrator account admin of the security cluster.</p> <p>NOTE</p> <p>Requirements for administrator passwords:</p> <ul style="list-style-type: none"> The password can contain 8 to 32 characters. The password must contain at least three of the following character types: uppercase letters, lowercase letters, digits, and special characters. The following special characters are supported: ~!@#%&^&()*-_=+ [{}];:;<.>/? Do not use the administrator name, or the administrator name spelled backwards. You are advised to change the password periodically. 	

Type	Parameter	Description
	Enterprise Project	Enterprise project to which a cluster belongs. You can click the project name to view the basic information about the enterprise project.
	Public IP Address	Public network access information, which is displayed only for clusters in security mode. <ul style="list-style-type: none"> For a security cluster with public network access enabled, the configured public network address is displayed. You can use this address to access the security cluster from the public network. For a security cluster with public network access disabled, -- is displayed. When using a public IP address to access a cluster, you are advised to enable access control and configure an access whitelist to improve cluster security. For details about how to configure the public network access, see Accessing a Cluster from a Public Network .
	HTTPS Access	Indicates whether to enable the HTTPS access protocol for a cluster. <ul style="list-style-type: none"> Disabled: The HTTP protocol is used for cluster access. Enabled: The HTTPS protocol is used for cluster access. Only security clusters can enable this function. If HTTPS Access is enabled, you can click Download Certificate to obtain the CER security certificate for accessing the security cluster. Currently, the security certificate cannot be used in the public network environment.

Type	Parameter	Description
	Private IPv4 Address	Private IP address and port number of a cluster, which can be used to access the cluster. If the cluster has only one node, the IP address and port number of only one node are displayed, for example, 10.62.179.32:9200 . If the cluster has multiple nodes, the IP addresses and port numbers of all nodes are displayed, for example, 10.62.179.32:9200,10.62.179.33:9200 .
Node	Node Specifications	Specifications of nodes in a cluster
	Node Storage Type	Storage capacity and storage type of nodes in a cluster
	Nodes	Number of nodes in a cluster

4.6.2 Managing Tags

Tags are cluster identifiers. Adding tags to clusters can help you identify and manage your cluster resources.

You can add tags to a cluster when creating the cluster or add them on the details page of the created cluster.

If your organization has enabled tag policies for CSS, you must comply with the tag policy rules when creating clusters, otherwise, clusters may fail to be created. Contact the organization administrator to learn more about tag policies.

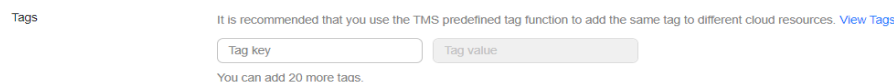
Managing Tags of a New Cluster

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner. The **Create Cluster** page is displayed.
3. On the **Create Cluster** page, set **Advanced Settings** to **Custom**. Add tags for a cluster.

You can select a predefined tag and set **Tag value** for the tag. You can click **View Predefined Tag** to switch to the TMS management console and view existing tags.

You can also create new tags by specifying **Tag key** and **Tag value**.

Figure 4-27 Adding tags during cluster creation



You can add a maximum of 20 tags for a CSS cluster. If the entered tag is incorrect, you can click **Delete** on the right of the tag to delete the tag. If you do not want to add tags, leave this parameter blank.

Table 4-24 Naming rules for a tag key and value

Parameter	Description
Tag key	<p>Must be unique in a cluster.</p> <p>The value cannot contain more than 64 characters.</p> <p>It can contain only numbers, letters, and the following special characters: _:=+@/ The value cannot start or end with a space.</p> <p>Cannot be left blank.</p>
Tag value	<p>The value cannot contain more than 64 characters.</p> <p>It can contain only numbers, letters, and the following special characters: _:=+@/ The value cannot start or end with a space.</p> <p>Cannot be left blank.</p>

Managing Tags of Existing Clusters

You can modify, delete, or add tags for a cluster.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters** > **Opensearch**.
3. On the **Clusters** page, click the name of a cluster for which you want to manage tags.

The **Basic Information** page is displayed.

4. In the navigation pane on the left, choose the **Tags** tab. You can add, modify, or delete tags.

- View

On the **Tags** page, you can view details about tags of the cluster, including the number of tags and the key and value of each tag.

- Add

Click **Add** in the upper left corner. In the displayed **Add Tag** dialog box, enter the key and value of the tag to be added, and click **OK**.

- Modify

You can only change the value of an existing tag.

In the **Operation** column of a tag, click **Edit**. In the displayed **Edit Tag** page, enter a new tag value and click **OK**.

- Delete

In the **Operation** column of a tag, click **Delete**. After confirmation, click **Yes** on the displayed **Delete Tag** page.

Searching for Clusters by Tag

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > Opensearch**.
3. On the **Clusters** page, click **Search by Tag** in the upper right corner of the cluster list.
4. Select or enter the tag key and tag value you want to search for, and click **Add** to add the tag to the search text box.

You can select a tag key or tag value from their drop-down lists. The system returns a list of clusters that exactly match the tag key or tag value. If you enter multiple tags, the cluster that meets requirements of all the tags will be filtered.

You can add a maximum of 10 tags at one time.

5. Click **Search**.
The system searches for the target cluster by tag key and value.

4.6.3 Managing Logs

CSS provides log backup and search functions to help you locate faults. You can back up cluster logs to OBS buckets and download required log files to analyze and locate faults.

If logs are backed up in OBS buckets, extra fees are charged. For details, see the [Billing Modes](#).

Log Query

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. In the navigation pane on the left, choose **Log Management**.
4. Query logs on the log management page.

Select the node, log type, and log level you want to query, and then click



. The query result is displayed.

When you search for logs, the latest 10,000 logs are matched. A maximum of 100 logs are displayed.

Enabling Log Backup

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. Click the **Logs** tab and toggle on the **Log Management** switch.
4. In the **Edit Log Backup Configuration** dialog box, set the parameters.

In the displayed dialog box, **OBS Bucket**, **Backup Path**, and **IAM Agency** are automatically created for log backup. You can change the default value by referring to [Table 4-25](#).


If the **Log Management** function has been enabled for the cluster, you can click  on the right of **Log Backup Configuration** and modify the configuration in the displayed **Edit Log Backup Configuration** dialog box. For details, see [Table 4-25](#).

Table 4-25 Parameters for configuring log backup


Parameter	Description	Remarks
OBS Bucket	Select an OBS bucket from the drop-down list for storing logs. You can also click Create Bucket on the right to create an OBS bucket.	The OBS bucket and the cluster must be in the same region. NOTE To let an IAM user access an OBS bucket, you need to grant the GetBucketStoragePolicy , GetBucketLocation , ListBucket , and ListAllMyBuckets permissions to the user.
Backup Path	Storage path of logs in the OBS bucket	The backup path configuration rules are as follows: <ul style="list-style-type: none"> • The backup path cannot contain the following characters: \:*?"<> • The backup path cannot start with a slash (/). • The backup path cannot start or end with a period (. • The total length of the backup path cannot exceed 1,023 characters.
IAM Agency	IAM agency authorized by the current account for CSS to access or maintain data stored in the OBS bucket. You can also click Create IAM Agency on the right to create an IAM agency.	The IAM agency must meet the following requirements: <ul style="list-style-type: none"> • Agency Type must be Cloud service. • Set Cloud Service to Elasticsearch or CSS. • Mandatory policies: OBS Administrator

5. Back up logs.

- Automatically backing up logs

Click the icon on the right of **Auto Backup** to enable the auto backup function.

After the automatic backup function is enabled, set the backup start time in the **Configure Auto Backup** dialog box. When the scheduled time arrives, the system will back up logs automatically.

After the **Automatic Snapshot Creation** function is enabled, you can click  on the right of the parameter to change the backup start time.

– Manually backing up logs

On the **Log Backup** tab page, click **Back Up**. On the displayed page, click **Yes** to start backup.


If **Task Status** in the log backup list is **Successful**, the backup is successful.

 **NOTE**

All logs in the cluster are copied to a specified OBS path. You can view or download log files from the path of the OBS bucket.

6. Search for logs.

On the **Log Search** page, select the target node, log type, and log level, and

click  . The search results are displayed.

When you search for logs, the latest 10,000 logs are matched. A maximum of 100 logs are displayed.

Viewing Logs

After backing up logs, you can click **Backup Path** to go to the OBS console and view the logs.

Figure 4-28 Accessing OBS

Log Backup Configuration

OBS Bucket



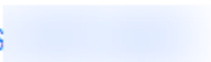
Backup Path

CSS_



IAM Agency

CSS



Backed up logs mainly include deprecation logs, run logs, index slow logs, and search slow logs. **Table 4-26** lists the storage types of the OBS bucket.

Table 4-26 Log types

Log Name	Description
clustername_deprecation.log	Deprecation log
clustername_index_indexing_slowlog.log	Search slow log
clustername_index_search_slowlog.log	Index slow log
clustername.log	Elasticsearch run log
clustername_access.log	Access log

4.6.4 Configuring YML Parameters

You can modify the `elasticsearch.yml` file.

Modifying Parameter Configurations

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster. The cluster information page is displayed.
3. Click **Parameter Configurations** and click **Edit** to modify module parameters as required.

Table 4-27 Module parameters

Module Name	Parameter	Description
Cross-domain Access	http.cors.allow-credentials	Indicates whether to return the Access-Control-Allow-Credentials of the header during cross-domain access. Value: true or false Default value: false
	http.cors.allow-origin	Origin IP address allowed for cross-domain access, for example, 122.122.122.122:9200
	http.cors.max-age	Default browser cache duration. The cache is automatically cleared after the time range you specified. Unit: s Default value: 1728000

Module Name	Parameter	Description
	http.cors.allow-headers	Headers allowed for cross-domain access, including X-Requested-With , Content-Type , and Content-Length . Use commas (,) and spaces to separate headers.
	http.cors.enabled	Indicates whether to allow cross-domain access. Value: true or false Default value: false
	http.cors.allow-methods	Methods allowed for cross-domain access, including OPTIONS , HEAD , GET , POST , PUT , and DELETE . Use commas (,) and spaces to separate methods.
Reindexing	reindex.remote.whitelist	Configure this parameter to migrate data from the current cluster to the target cluster through the reindex API. The example value is 122.122.122.122:9200.
Custom Cache	indices.queries.cache.size	Cache size in the query phase Value range: 1 to 100. Unit: % Default value: 10%
Queue Size in a Thread Pool	thread_pool.force_merge.size	Queue size in the force merge thread pool. The value is an integer. Default value: 1
Customize	You can add parameters based on your needs.	Customized parameters NOTE <ul style="list-style-type: none"> • Enter multiple values in the format of [value1, value2, value3...]. • Separate values by commas (,) and spaces. • Colons (:) are not allowed.

4. After the modification is complete, click **Submit**. In the displayed **Submit Configuration** dialog box, select the box indicating "I understand that the modification will take effect after the cluster is restarted." and click **Yes**.
If the **Status** is **Succeeded** in the parameter modification list, the modification has been saved. Up to 20 modification records can be displayed.
5. Return to the cluster list and choose **More > Restart** in the **Operation** column to restart the cluster and make the modification take effect.
 - You need to restart the cluster after modification, or **Configuration unupdated** will be displayed in the **Task Status** column on the **Clusters** page.

- If you restart the cluster after the modification, and **Task Status** displays **Configuration error**, the parameter configuration file fails to be modified.

4.6.5 Viewing the Default Plugin List

CSS clusters have default plug-ins. You can view the default plugin information on the console or Kibana.

Viewing Plugins on the Console

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters**. Click the target cluster name and go to the **Cluster Information** page of the cluster.
3. Click the **Plugins** tab.
4. On the **Default** tab page, view default plugins supported by the current version.

Viewing Plugins on the Kibana

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters**. Locate the target cluster and click **Access Kibana** in the **Operation** column to log in to OpenSearch Dashboard.
3. Go to **Dev Tools** and run the following command to view the cluster plugin information:

```
GET _cat/plugins?v
```

The following is an example of the response body:

name	component	version
css-3657-ess-esn-1-1	analysis-dynamic-synonym	1.3.6
css-3657-ess-esn-1-1	analysis-icu	1.3.6
css-3657-ess-esn-1-1	analysis-ik	1.3.6
css-3657-ess-esn-1-1	analysis-kuromoji	1.3.6
css-3657-ess-esn-1-1	analysis-logtxt	1.0.0
css-3657-ess-esn-1-1	analysis-nori	1.3.6
css-3657-ess-esn-1-1	analysis-pinyin	1.3.6
css-3657-ess-esn-1-1	analysis-stconvert	1.3.6
css-3657-ess-esn-1-1	hpack	2.0.0
css-3657-ess-esn-1-1	ingest-attachment	1.3.6
css-3657-ess-esn-1-1	obs-store-plugin	1.3.6
css-3657-ess-esn-1-1	opensearch-alerting	1.3.6.0
css-3657-ess-esn-1-1	opensearch-anomaly-detection	1.3.6.0
css-3657-ess-esn-1-1	opensearch-asynchronous-search	1.3.6.0
css-3657-ess-esn-1-1	opensearch-cross-cluster-replication	1.3.6.0
css-3657-ess-esn-1-1	opensearch-index-management	1.3.6.0
css-3657-ess-esn-1-1	opensearch-job-scheduler	1.3.6.0
css-3657-ess-esn-1-1	opensearch-knn	1.3.6.0
css-3657-ess-esn-1-1	opensearch-ml	1.3.6.0
css-3657-ess-esn-1-1	opensearch-observability	1.3.6.0
css-3657-ess-esn-1-1	opensearch-performance-analyzer	1.3.6.0
css-3657-ess-esn-1-1	opensearch-reports-scheduler	1.3.6.0
css-3657-ess-esn-1-1	opensearch-security	1.3.6.0
css-3657-ess-esn-1-1	opensearch-sql	1.3.6.0
css-3657-ess-esn-1-1	repository-obs	1.3.6

name indicates the cluster node name, **component** indicates the plugin name, and **version** indicates the plugin version.

4.6.6 Binding an Enterprise Project

You can create enterprise projects based on your organizational structure. Then you can manage resources across different regions by enterprise project, add users and user groups to enterprise projects, and grant different permissions to the users and user groups. This section describes how to bind an OpenSearch cluster to an enterprise project and how to modify an enterprise project.

Prerequisites

To use the enterprise project function, you need to assign permissions to the corresponding account. You can [submit a service ticket](#) to apply for the permissions.

Before binding an enterprise project, you have [created an enterprise project](#).

Binding an Enterprise Project

When creating a cluster, you can bind an existing enterprise project to the cluster, or click **View Enterprise Project** to go to the enterprise project management console and create a new project or view existing projects.

Modifying an Enterprise Project

For a cluster that has been created, you can modify its enterprise project based on the site requirements.

1. Log in to the CSS management console.
2. In the left navigation pane, choose **Clusters > OpenSearch**. The cluster list page is displayed.
3. In the cluster list on the displayed page, click the target cluster name to switch to the **Cluster Information** page.
4. On the **Cluster Information** page, click the enterprise project name on the right of **Enterprise Project**. The project management page is displayed.

Figure 4-29 Enterprise Project Configuration

Region

AZ

VPC

Subnet

Security Group [Change Security Group](#)

Cluster Routing [Modify](#) [View](#)

Enterprise Project

IPv4 Access Address

5. On the **Resources** tab page, select the region of the current cluster, and select **CSS** for **Service**. In this case, the corresponding CSS cluster is displayed in the resource list.

Figure 4-30 Filtering CSS clusters

Resources Permissions

Region

Service	All	ElasticCloudServer	AutoScaling	ImageManagementService	ElasticVolumeService	VirtualPrivateCloud	ElasticIP	ContentDeliveryNetwork
RelationalDatabaseServ...	DistributedCacheService	DocumentDatabaseServ...	CloudContainerEngine	AdvancedAnti-DDoS	ObjectStorageService	DataLakeInsight	DistributedMessageServ...	
DatasyncService	ScalableFileService	CloudServiceEngine	DomainNameService	MapReduceService	BareMetalServer	SSLCertificateManage	HostSecurityService	
SimpleMessageNotificat...	ModelArts	DataIngestionService	LogTankService	GaussDB	CloudConnect	BlockchainService	GraphEngineService	
GaussDBforSQL	Application&DataIntegr...	SupportPlanService	DataLakeVisualization	DatabaseSecurityService	DataReplicationService	ApplicationServiceMesh	BatchContainerEngine	
VirtualPrivateNetwork	DatabaseandApplication...	Multi-cloudHighAvailabi...	PrivateCertificateAuthority	KeyManagementService	CloudSearchService	DistributedDatabaseMid...	DAU	
FunctionGraph	EventGrid	WebApplicationFirewall	VPN SD-WAN	CloudSecretManagement	CloudSite	LateFormation	ElasticLoadBalance	
Dedicated-EBM	CloudElasticHost	IoTDeviceAccess	CloudEye	MacroVerseSmartStage	Cloud-nativeAnti-DDoS	IndustrialDigitalModelEn...	CloudTableService	
SMS	CloudTraceService	Workspace	CloudFirewall	NATGateway	CloudBackupRecovery	CloudPhoneHost	edgesec	
DirectConnect	APIGateway	Software Repository for...	AppStream	DataArmsInsight	Industrial Simulation Co...	PCB-Focused Electronic...	Global Accelerator	
MAS-CAST	ApplicationOperationalM...	ServiceStage	Industrial Simulation Co...	IntegratedProductDevel...				

Resource Type

Selected Service: CloudSearchService Region:

6. Select the cluster whose enterprise project you want to modify and click **Remove**.
7. On the **Remove Resource** page, specify **Mode** and select **Destination Enterprise Project**, and click **OK**.
8. After the resource is removed, you can view the modified enterprise project information on the **Clusters** page.

4.6.7 Restarting a Cluster

If a cluster becomes faulty, you can restart it to check if it can run normally.

Prerequisites

- The target cluster is not frozen and has no task in progress.
- If a cluster is available, ensure that it has stopped processing service requests (such as importing data and searching for data). Otherwise, data may be lost when the cluster is restarted. You are advised to perform this operation during off-peak hours.

Context

CSS supports quick restart and rolling restart.

Quick Restart

- All clusters support this function.
- If you select a node type for quick restart, all nodes of the selected type will be restarted together.
- If you select a node name for quick restart, only the specified node will be restarted.
- The cluster is unavailable during quick restart.

Rolling Restart

- Rolling restart is supported only when a cluster has at least three nodes (including master nodes, client nodes, and cold data nodes).
- Rolling restart can be performed only by specifying node types. If you select a node type for rolling restart, the nodes of the selected type will be restarted in sequence.
- During the rolling restart, only the nodes that are being restarted are unavailable and other nodes can run normally.
- When the data volume is large, rolling restart will take a long time.

Quick Restart

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. In the **Operation** column of the target cluster, choose **More > Restart**.
4. On the **Restart Cluster** page, select **Quick Restart**.

You can quick restart nodes by **Node type** or **Node name**. If you select **Node type**, then you can select multiple node types and perform quick restart at the time. If you select **Node name**, you can perform quick restart only on one node at a time.

5. Refresh the page and check the cluster status. During the restart, the cluster status is **Processing**, and the task status is **Restarting**. If the cluster status changes to **Available**, the cluster has been restarted successfully.

Rolling Restart

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. In the **Operation** column of the target cluster, choose **More > Restart**.
4. On the **Restart Cluster** page, select **Rolling Restart**.
You can perform rolling restart by **Node type**. Select specific node types for restart.
5. Refresh the page and check the cluster status. During the restart, the cluster status is **Processing**, and the task status is **Restarting**. If the cluster status changes to **Available**, the cluster has been restarted successfully.

4.6.8 Deleting a Cluster

You can delete clusters that you no longer need.

NOTE

- If you delete a cluster, the cluster service data will be cleared. Exercise caution when performing this operation.
- The snapshots of a cluster stored in OBS are not deleted with the cluster. You can restore a deleted cluster using its snapshots stored in the OBS bucket. For details, see [Can I Restore a Deleted Cluster?](#)

Procedure

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Locate the target cluster and click **More > Delete** in the **Operation** column.
4. In the displayed dialog box, enter the name of the cluster to be deleted and click **OK**.

4.7 Customizing Word Dictionaries

4.7.1 Managing Word Dictionaries

You can configure the custom word dictionary to identify the segments of specified words. For example, you can search for the keyword of company names, such as, Huawei, and network buzzwords.

NOTE

- Hot update is supported. The updated custom word dictionary can take effect without cluster restart.
- Custom word dictionaries are generally used for Chinese word segmentation. They can also be used to segment English words based on special characters except **#&+-.@_**

Context

Custom word dictionary uses the IK and synonym analyzer.

The IK analyzer has a main word dictionary and a stop word dictionary. The synonym analyzer has a synonym word dictionary. Before configuring a custom word dictionary, upload the prepared word dictionary file to OBS. For details, see [Uploading the Word Dictionary File to OBS](#).

The IK analyzer uses the ik_max_word and ik_smart word segmentation policies. The synonym analyzer uses the ik_synonym word segmentation policy.

- ik_max_word: splits the text at a fine granularity.
- ik_smart: splits the text at a coarse granularity.

Prerequisites

- To use the custom word dictionary, the account or IAM user used for logging in to the CSS management console must have both of the following permissions:
 - **OBS Administrator** for project **OBS** in region **Global service**
 - **Elasticsearch Administrator** in the current region
- Prepare the word dictionary file on the local PC as required by referring to [Uploading the Word Dictionary File to OBS](#).

Uploading the Word Dictionary File to OBS

Before configuring a custom word dictionary, upload the word dictionary to an OBS bucket.

1. Prepare the word dictionary file according to [Table 4-28](#).

Table 4-28 Dictionary description

Word Dictionary Type	Introduction	Requirement
Main Word Dictionary	Main words are the words on which users want to perform word segmentation. The main word dictionary is a collection of main words.	The main word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. Letters must be in lowercase. The maximum size of a main word dictionary file is 100 MB.
Stop Word Dictionary	Stop words are the words which users can ignore. A stop word dictionary is a collection of stop words.	The stop word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. The maximum size of a stop word dictionary file is 80 MB.

Word Dictionary Type	Introduction	Requirement
Synonym Dictionary	Synonyms are words with the same meaning. A synonym dictionary is a collection of synonyms.	The synonym dictionary file must be a text file encoded using UTF-8 without BOM, with a pair of comma-separated synonyms per line. The maximum size of a synonym dictionary file is 80 MB.

2. Upload the word dictionary file to an OBS bucket. For details, see [Uploading an Object](#). The OBS bucket to which data is uploaded must be in the same region as the cluster.

Managing Word Dictionaries

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. On the **Clusters** page, click the name of the target cluster.
4. Click the **Word Dictionaries** tab.
5. On the displayed **Word Dictionaries** page, set the switch to enable or disable the custom word library function.
 - **OBS Bucket:** indicates the OBS bucket where the main word dictionary file, stop word dictionary file, and synonym dictionary file are stored. If no OBS bucket is available, create one by referring to [Creating a Bucket](#). The OBS bucket must be in the same region as the cluster.
 - **Main word dictionary object:** The main word dictionary file must be a text file encoded using UTF-8 without BOM. One subword occupies a line. Letters must be in lowercase. The maximum size of a main word dictionary file is 100 MB.
 - **Stop word dictionary object:** The stop word dictionary file must be a text file encoded using UTF-8 without BOM, with one subword per line. The maximum size of a stop word dictionary file is 80 MB.
 - **Synonym word dictionary object:** The synonym dictionary file must be a text file encoded using UTF-8 without BOM. One pair of comma-separated synonyms occupies a line. The maximum size of a synonym dictionary file is 80 MB.

Figure 4-31 Configuring a custom word dictionary

Word Dictionaries

OBS Bucket [Create Bucket](#) ?

Main Word Dictionary Stop Word Dictionary Synonym Dictionary

Main Word Dictionary ?

Stop Word Dictionary ?

Synonym Dictionary ?

6. Click **Save**. In the displayed **Confirm** dialog box, click **OK**. The word dictionary information is displayed in the lower part of the page. The word dictionary status is **Updating**. Wait for about one minute. After the word dictionary configuration is complete, the word dictionary status will change to **Succeeded**, indicating that the configured word dictionary has taken effect in the cluster.

Modifying a Word Dictionary

You can separately update the main word dictionary, the stop word dictionary, and the synonym dictionary.

On the **Word Dictionaries** page, modify **OBS Bucket**, **Main Word Dictionary**, **Stop Word Dictionary**, or **Synonym Dictionary**, and click **Save**. In the displayed dialog box, click **OK**. When the word dictionary status changes from **Updating** to **Successful**, the custom word dictionary is modified.

Figure 4-32 Configuring a custom word dictionary

Word Dictionaries

OBS Bucket [Create Bucket](#) ?

Main Word Dictionary Stop Word Dictionary Synonym Dictionary

Main Word Dictionary ?

Stop Word Dictionary ?

Synonym Dictionary ?

Disabling a Word Dictionary

You can disable your word dictionary when it is no longer in need.

On the **Word Dictionaries** page, disable the function and click **OK** in the displayed dialog box. After the word dictionary is disabled, the word dictionary configuration information will not be displayed.

4.7.2 Example

Application Scenarios

Configure a custom word dictionary for the cluster, set main words, stop words, and synonyms. Search for the target text by keyword and synonym and view the search results.

Step 1: Configure a Custom Word Dictionary

1. Prepare a word dictionary file (a text file encoded using UTF-8 without BOM) and upload it to the target OBS path.

Set the main word dictionary file, stop word dictionary file, and synonym word dictionary file.

NOTE

The default word dictionary contains common stop words such as **are** and **the**. You do not need to upload such stop words.

2. Log in to the CSS management console.
3. In the navigation pane, choose **Clusters > OpenSearch**.
4. On the **Clusters** page, click the name of the target cluster.
5. Click the **Word Dictionaries** tab. Configure the word dictionary file for the step 1 by referring to [Managing Word Dictionaries](#).
6. After the word dictionary takes effect, return to the cluster list. Locate the target cluster and click **Kibana** in the **Operation** column to access the cluster.
7. On the Kibana page, click **Dev Tools** in the navigation tree on the left. The operation page is displayed.
8. Run the following commands to check the performance of different word segmentation policies.
 - Use the `ik_smart` word segmentation policy to split the target text.

Example code:

```
POST /_analyze
{
  "analyzer":"ik_smart",
  "text":"Text used for word segmentation"
}
```

After the operation is completed, view the word segmentation result.

```
{
  "tokens": [
    {
      "token": "word-1",
      "start_offset": 0,
      "end_offset": 4,
      "type": "CN_WORD",
      "position": 0
    }
  ]
}
```

```

},
{
  "token": "word-2",
  "start_offset": 5,
  "end_offset": 8,
  "type": "CN_WORD",
  "position": 1
}
]
}

```

- Use the ik_max_word word segmentation policy to split the target text.

Example code:

```

POST /_analyze
{
  "analyzer": "ik_max_word",
  "text": "Text used for word segmentation"
}

```

After the operation is completed, view the word segmentation result.

```

{
  "tokens" : [
    {
      "token": "word-1",
      "start_offset" : 0,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 0
    },
    {
      "token": "word-3",
      "start_offset" : 0,
      "end_offset" : 2,
      "type" : "CN_WORD",
      "position" : 1
    },
    {
      "token": "word-4",
      "start_offset" : 0,
      "end_offset" : 1,
      "type" : "CN_WORD",
      "position" : 2
    },
    {
      "token": "word-5",
      "start_offset" : 1,
      "end_offset" : 3,
      "type" : "CN_WORD",
      "position" : 3
    },
    {
      "token": "word-6",
      "start_offset" : 2,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 4
    },
    {
      "token": "word-7",
      "start_offset" : 3,
      "end_offset" : 4,
      "type" : "CN_WORD",
      "position" : 5
    },
    {
      "token": "word-2",
      "start_offset" : 5,
      "end_offset" : 8,
      "type" : "CN_WORD",

```



```
    "position" : 6
  },
  {
    "token" : "word-8",
    "start_offset" : 5,
    "end_offset" : 7,
    "type" : "CN_WORD",
    "position" : 7
  },
  {
    "token" : "word-9",
    "start_offset" : 6,
    "end_offset" : 8,
    "type" : "CN_WORD",
    "position" : 8
  },
  {
    "token" : "word-10",
    "start_offset" : 7,
    "end_offset" : 8,
    "type" : "CN_WORD",
    "position" : 9
  }
]
}
```

Step 2: Use Keywords for Search

1. Create the **book** index and configure the word segmentation policy.
In this example, both **analyzer** and **search_analyzer** are set to **ik_max_word**. You can also use **ik_smart**.

```
PUT /book
{
  "settings": {
    "number_of_shards": 2,
    "number_of_replicas": 1
  },
  "mappings": {
    "properties": {
      "content": {
        "type": "text",
        "analyzer": "ik_max_word",
        "search_analyzer": "ik_max_word"
      }
    }
  }
}
```

2. Import the text information to the **book** index.

```
PUT /book/_doc/1
{
  "content": "Imported text"
}
```

3. Use a keyword to search for the text and view the search results.

```
GET /book/_doc/_search
{
  "query": {
    "match": {
      "content": "Keyword"
    }
  }
}
```

Search result

```
{
  "took" : 16,
  "timed_out" : false,
```

```

"_shards" : {
  "total" : 2,
  "successful" : 2,
  "skipped" : 0,
  "failed" : 0
},
"hits" : {
  "total" : {
    "value" : 1,
    "relation" : "eq"
  },
  "max_score" : 1.7260926,
  "hits" : [
    {
      "_index" : "book",
      "_type" : "_doc",
      "_id" : "1",
      "_score" : 1.7260926,
      "_source" : {
        "content" : "Imported text"
      }
    }
  ]
}

```

Step 3: Use Synonyms for Search

1. Create the **myindex** index and configure the word segmentation policy.

```

PUT myindex
{
  "settings": {
    "analysis": {
      "filter": {
        "my_synonym": {
          "type": "dynamic_synonym"
        }
      },
      "analyzer": {
        "ik_synonym": {
          "filter": [
            "my_synonym"
          ],
          "type": "custom",
          "tokenizer": "ik_smart"
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "desc": {
        "type": "text",
        "analyzer": "ik_synonym"
      }
    }
  }
}

```

2. Import the text information to the **myindex** index.

```

PUT /myindex/_doc/1
{
  "desc": "Imported text"
}

```

3. Conduct search based on the synonym and view the search results.

```

GET /myindex/_search
{
  "query": {

```

```
"match": {
  "desc": "Keyword"
}
}
```

Search result

```
{
  "took" : 1,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 0.1519955,
    "hits" : [
      {
        "_index" : "myindex",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 0.1519955,
        "_source" : {
          "desc" : "Imported text"
        }
      }
    ]
  }
}
```

4.8 Converting Between Simplified and Traditional Chinese (Using the Conversion Plugin)

By default, a simplified-traditional Chinese conversion plugin is installed in CSS. The plugin implements conversion between simplified and traditional Chinese. With this plugin, you can search index data containing the corresponding simplified Chinese based on the traditional Chinese keyword, and vice versa.

The simplified-traditional Chinese conversion plugin can be used as the analyzer, tokenizer, token-filter, or char-filter.

The simplified-traditional Chinese conversion plugin provides the following two conversion types:

- s2t: converts simplified Chinese to traditional Chinese.
- t2s: converts traditional Chinese to simplified Chinese.

Example

1. Log in to the CSS management console.
2. In the navigation pane on the left, click **Clusters** to switch to the **Clusters** page.
3. In the cluster list, locate the row containing the cluster and click **Access Kibana** in the **Operation** column.

If the target cluster has the security mode enabled, enter the username and password you set when you created the cluster.

4. In the Kibana navigation pane on the left, choose **Dev Tools**.
5. On the **Console** page, run the following command to create index **stconvert** and specify a user-defined mapping to define the data type:

```
PUT /stconvert
{
  "settings": {
    "number_of_shards": 1,
    "number_of_replicas": 0,
    "analysis": {
      "analyzer": {
        "ts_ik": {
          "tokenizer": "ik_smart",
          "char_filter": [
            "tsconvert",
            "stconvert"
          ]
        }
      },
      "char_filter": {
        "tsconvert": {
          "type": "stconvert",
          "convert_type": "t2s"
        },
        "stconvert": {
          "type": "stconvert",
          "convert_type": "s2t"
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "desc": {
        "type": "text",
        "analyzer": "ts_ik"
      }
    }
  }
}
```

The command output is similar to the following:

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : "stconvert"
}
```

6. On the **Console** page, run the following command to import data to index **stconvert**:

```
POST /stconvert/_doc/1
{
  "desc": "Text in traditional Chinese"
}
```

If the value of **failed** in the command output is **0**, the data is imported successfully.

7. On the **Console** page, run the following command to search for the keyword and view the search result:

```
GET /stconvert/_search
{
  "query": {
    "match": {
      "desc": "Keyword"
    }
  }
}
```

```
}  
}
```

The command output is similar to the following:

```
{  
  "took" : 15,  
  "timed_out" : false,  
  "_shards" : {  
    "total" : 1,  
    "successful" : 1,  
    "skipped" : 0,  
    "failed" : 0  
  },  
  "hits" : {  
    "total" : 1,  
    "max_score" : 0.5753642,  
    "hits" : [  
      {  
        "_index" : "stconvert",  
        "_type" : "type",  
        "_id" : "1",  
        "_score" : 0.5753642,  
        "_source" : {  
          "desc": "Text in traditional Chinese"  
        }  
      }  
    ]  
  }  
}
```

4.9 Configuring SMN Alarms

Scenarios

By default, CSS has installed the open-source alert plugin **opensearch-alerting** for OpenSearch clusters to send notifications when data meets specific conditions. This plugin consists of three components: **Alerts**, **Monitors**, and **Destinations**. CSS integrates the SMN service in the **Destinations** component and can send alarm messages only through the SMN service as the destination.

This section describes how to configure the SMN alarm function for OpenSearch clusters on **OpenSearch Dashboards**.

NOTE

For details about the official guide of the plug-in **OpenSearch Alerting**, visit [Alerting - OpenSearch Documentation](#).

Constraints and Limitations

By default, the open-source alert plug-in **opensearch-alerting** is installed for OpenSearch clusters of version 1.3.6.

Prerequisites

- The SMN service has been authorized. For details, see [\(Optional\) Authorizing CSS to Use SMN](#).
- You have created a topic on the SMN console. For details, see [Creating a Topic](#).

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters > OpenSearch**, select the target cluster and click **Access Kibana** in the **Operation** column.
3. On the **OpenSearch Dashboards** page, choose **OpenSearch Plugins > Alerting** in the navigation tree on the left.
4. Create an SMN destination to send alert messages.
 - a. On the **Alerting** page, click the **Destinations** tab and click **Add destination** to configure destination information.

Table 4-29 Destinations parameters

Parameter	Description
Name	User-defined destination name
Type	Retain the default value SMN .
Topic	Select the SMN topic you have created for sending alarm messages. For details, see Creating a Topic .

Figure 4-33 Add destination

- b. Click **Create** to return to the destination list. The created SMN destination is displayed in the list.

Figure 4-34 Destination list

- 5. Create a monitoring task and configure the alarm triggering condition and monitoring frequency.
 - a. Click the **Monitors** tab on the **Alerting** page and click **Create monitor** to configure monitoring information.

Table 4-30 Monitor parameters

Parameter	Description
Monitor details	
Monitor name	User-defined monitor name

Parameter	Description
Monitor type	Monitor type. The value can be Per query monitor (common monitoring), Per bucket monitor (aggregation bucket monitoring), and Per cluster metrics monitor (cluster metric monitoring).
Monitor defining method	Monitor defining method. Extraction query editor is recommended. <ul style="list-style-type: none"> • Visual editor • Extraction query editor • Anomaly detector The options of Monitor defining method are determined by the Monitor type you selected.
Detector	If Monitor defining method is set to Anomaly detector , select an exception detection task.
Frequency	Select the monitoring frequency and set the monitoring interval. The options include: <ul style="list-style-type: none"> • By interval • Daily • Weekly • Monthly • Custom cron expression
Data source	
Index	When Monitor defining method is set to Visual editor or Extraction query editor , you need to specify the index to be monitored.
Time field	When Monitor defining method is set to Visual editor , you need to specify the time field to define counting parameters such as count .
Query	
Metrics	When Monitor defining method is set to Visual editor , you need to set the metrics range for extracting statistics.
Time range for the last	When Monitor defining method is set to Visual editor , you need to set the monitoring time range for plug-ins.
Data filter	When Monitor defining method is set to Visual editor , you need to set filters for data search.
Group by	When Monitor defining method is set to Visual editor , you need to specify a field so that each value of the field triggers an alarm.

Parameter	Description
Define extraction query	When Monitor defining method is set to Extraction query editor , you need to enter the query statement to define the monitoring.
Request type	When Monitor type is set to Per cluster metrics monitor , you need to specify the request type to monitor cluster metrics, such as the running status and CPU usage.

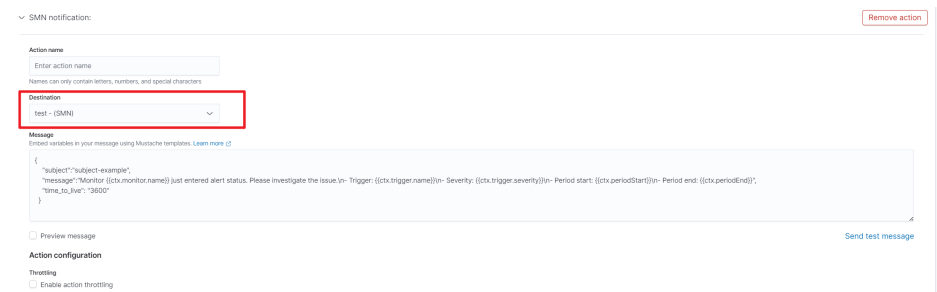
- b. Click **Add trigger** to add triggers and specify the alarm triggering conditions and actions to be triggered when an alarm is reported.
- c. On the **Triggers** page, set the alarm triggering sensitivity and message release on the destination end.

Table 4-31 Trigger parameters

Parameter	Description
Trigger name	User-defined trigger name
Severity level	Sensitivity of a trigger, that is, the number of alarms that are triggered before an alarm message is sent. 1 indicates the highest sensitivity.
Trigger condition	Trigger condition. An alarm is triggered when the trigger condition is hit.
Action name	Trigger action name
Destination	Select the SMN destination created in section 4 .
Message	Alarm message body By default, the subject and body are defined when the destination is an email. For details, see Message Publishing .
Perform action	When Monitor type is set to Per bucket monitor , you need to set whether to send alarms in combination. The value can be: <ul style="list-style-type: none"> • Per execution: A combination alarm is sent when multiple alarm triggering conditions are hit. • Per alert: Alarms are sent separately when multiple alarm triggering conditions are hit.

Parameter	Description
Actionable alerts	<p>When Monitor type is set to Per bucket monitor, set this parameter to Per alert. You need to set the alarms that can be executed after alarm triggering conditions are hit.</p> <ul style="list-style-type: none"> ● De-duplicated: Alarms that have been triggered. OpenSearch retains the existing alarms to prevent the plugin from creating duplicate alarms. ● New: Newly created alarms. ● Completed: Alarms that are no longer ongoing.
Throttling	<p>Message sending frequency. It limits the number of notification messages can be received in a specified period.</p> <p>For example, if this parameter is set to 10 minutes, SMN sends only one alarm notification in the next 10 minutes even if the trigger condition is hit for multiple times. After 10 minutes, SMN sends another alarm notification if the alarm condition is met.</p>

Figure 4-35 Setting the destination of a trigger action



- d. Click **Send test message**. If a subscriber receives an email, as shown in **Figure 4-37**, the trigger is configured successfully.

Figure 4-36 Sending a test message

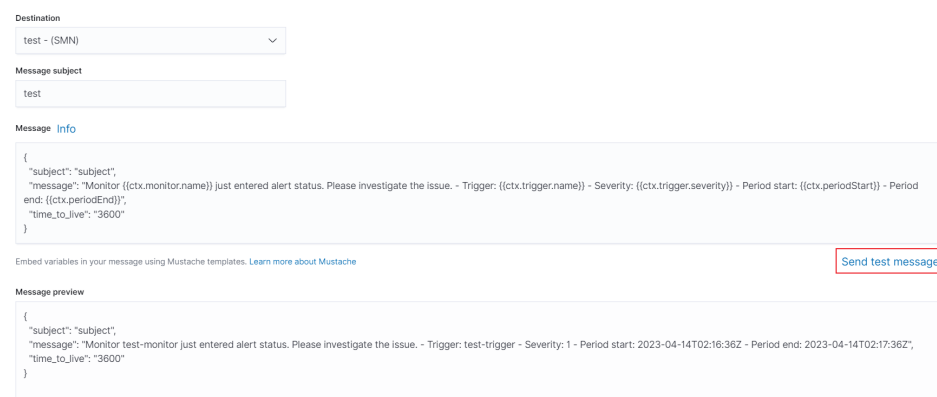
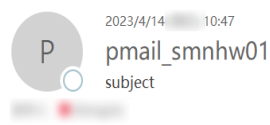


Figure 4-37 Email notification



Monitor test-monitor just entered alert status. Please investigate the issue. - Trigger: test-trigger - Severity: 1 - Period start: 2023-04-14T02:46:09.977Z - Period end: 2023-04-14T02:47:09.977Z

- e. Click **Create** to return to the monitor details page. The detector is successfully created.

4.10 Switching Hot and Cold Data

CSS provides you with cold data nodes. You can store data that requires query response in seconds on hot data nodes with high performance and store historical data that requires query response in minutes on cold data nodes with large capacity and low specifications.

NOTE

- When creating a cluster, you need to configure data nodes. After cold data nodes are selected, the original data nodes become hot data nodes.
- You can enable the cold data node, master node, and client node functions at the same time.
- You can increase nodes and expand storage capacity of cold data nodes. The maximum storage capacity is determined by the node specifications. Local disks do not support storage capacity expansion.

Switching Between Hot and Cold Data

If you enable cold data nodes when creating a cluster, the cold data nodes are labeled with **cold**. Other data nodes become hot nodes and are labeled with **hot**. You can specify indexes to allocate data to cold or hot nodes.

You can configure a template to store indexes on the specified cold or hot node.

Log in to the **Kibana Console** page of the cluster, store the indexes starting with **myindex** on the cold node. In this way, you can use a template to store the **myindex*** data on the cold data node.

Run the following command to create a template:

```
PUT _template/test
{
  "order": 1,
  "index_patterns": "myindex*",
  "settings": {
    "refresh_interval": "30s",
    "number_of_shards": "3",
    "number_of_replicas": "1",
    "routing.allocation.require.box_type": "cold"
  }
}
```

You can perform operations on the created index.

```
PUT myindex/_settings
{
  "index.routing.allocation.require.box_type": "cold"
}
```

You can cancel the configurations of hot and cold data nodes.

```
PUT myindex/_settings
{
  "index.routing.allocation.require.box_type": null
}
```

4.11 Managing Indexes

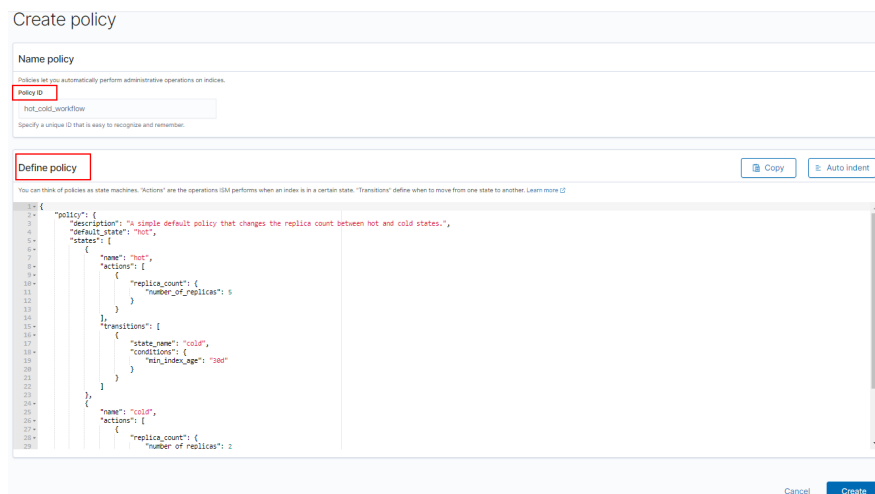
4.11.1 Creating and Managing Index Policies

You can manage the indexes of OpenSearch clusters. ISM is a plugin that allows you to automate periodic and administrative operations based on changes on the index age, index size, or number of documents. When using the ISM plug-in, you can define policies that automatically handle index rollovers or deletions based on your needs.

Creating an Index Policy

1. Log in to Kibana and choose or **Index Management** on the left. The index management page is displayed.
2. Click **Create policy** to create an index policy.
3. In the **Configuration method** dialog box, select **JSON editor** and click **Continue**. The page for creating an index policy is displayed.
4. Enter a policy ID in the **Policy ID** text box and enter your policy in the **Define policy** text box.

Figure 4-38 Configuring a policy



5. Click **Create**.

Attaching a Policy to an Index

You can attach a policy to one or more indexes and add the policy ID to an index template. When you create indexes using that index template pattern, the policy will be attached to all created indexes.

- **Method 1: OpenSearch Dashboard CLI**

On the **Dev Tools** page of the **OpenSearch Dashboards**, run the following command to associate the policy ID with the index template:

```
PUT _template/<template_name>
{
  "index_patterns": ["index_name-*"],
  "settings": {
    "opendistro.index_state_management.policy_id": "policy_id"
  }
}
```

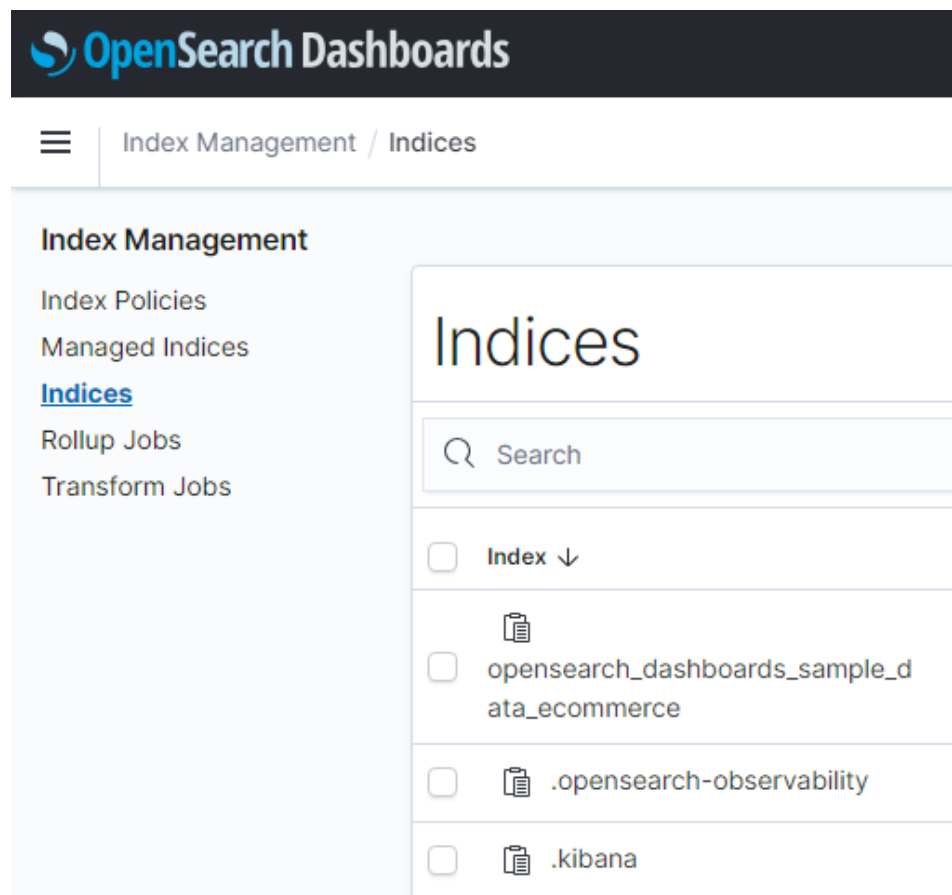
- **<template_name>**: Replace it with the name of a created index template.
- **policy_id**: Replace it with a custom policy ID.

For details about how to create an index template, see [Index Templates](#).

- **Method 2: OpenSearch Dashboards Console**

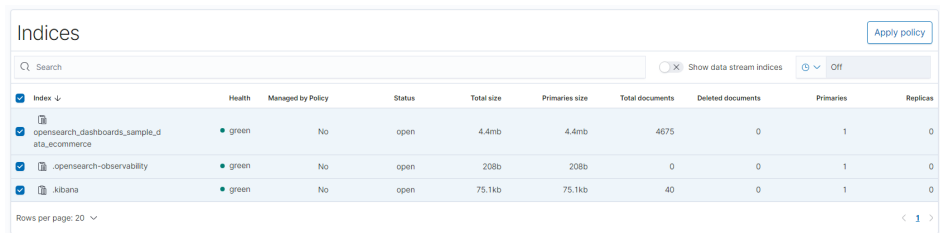
- On the **Index Management** page of the **OpenSearch Dashboards**, choose **Indices**.

Figure 4-39 Choosing Indices



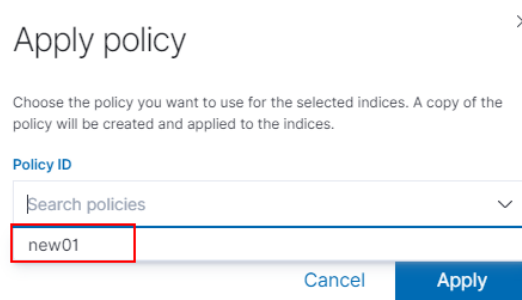
- In the **Indices** list, select the target index to which you want to attach a policy.
- Click **Apply policy** in the upper right corner.

Figure 4-40 Adding a policy



- d. Select the policy you created from the **Policy ID** drop-down list.

Figure 4-41 Selecting an index policy

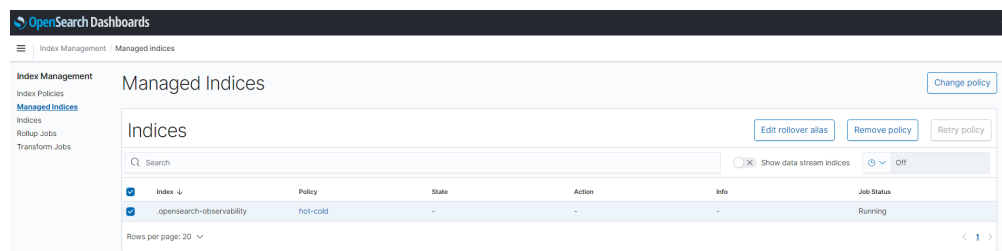


- e. Click **Apply**.
After you attach a policy to an index, ISM creates a job that runs every 5 minutes by default, to execute the policy, check conditions, and convert the index to different statuses.

Managing Index Policies

1. On the **Index Management** page of the **OpenSearch Dashboards**, choose **Managed Indices**.
2. If you want to change the policy, click **Change policy**. For details, see [Changing Policies](#).

Figure 4-42 Changing policies



3. To delete a policy, select your policy, and click **Remove policy**.
4. To retry a policy, select your policy, and click **Retry policy**.

For details, see [Index State Management](#).

4.11.2 Changing an Index Policy

You can change any managed index policy. ISM has constraints to ensure that policy changes do not break indexes.

If an index is stuck in its current status and you want to update its policy immediately, make sure that the new policy includes the same status (same name, action, and order) as the old policy. In this case, ISM applies the new policy even if the old policy is being executed.

If the new policy you use does not include the same status as the old policy, ISM updates the policies only after all actions in the current status are completed. Alternatively, you can select a specific status in the old policy and make the new policy take effect.

Perform the following steps to change a policy in the **OpenSearch Dashboards**:

1. On the **Index Management** page of the **OpenSearch Dashboards**, select the index policy you want to change.
2. Click **Change policy** in the upper right corner. In the **Choose managed indices** and **Choose new policy** areas, select information about the new policy.

Figure 4-43 Changing an index policy

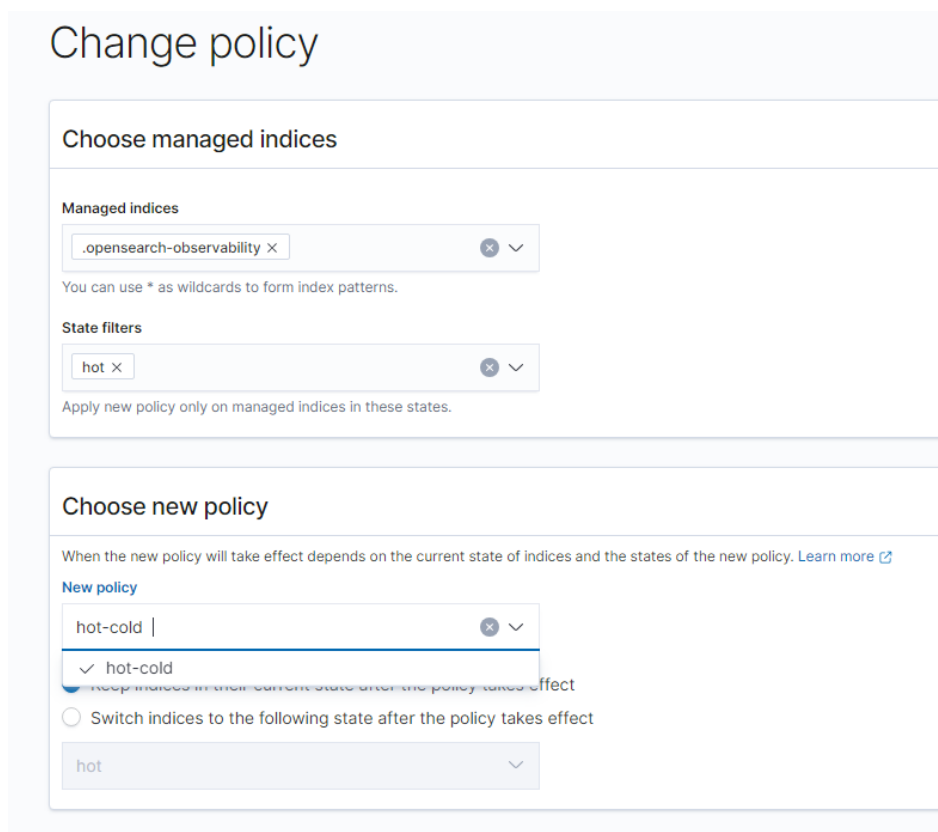


Table 4-32 Parameters required for changing a policy

Parameter	Description
Managed indices	Select the indexes to which you want to attach the new policy. Multiple indexes can be selected.

Parameter	Description
State filters	Select an index status. When a status is selected, the new policy is attached to an index in this status.
New policy	Select a new policy.

3. After configuration is complete, click **Change**.

4.12 Intelligent O&M

4.12.1 Overview of Intelligent O&M

CSS provides intelligent O&M to detect potential cluster risks and provide risk handling suggestions.

Currently, the Elasticsearch and Opensearch clusters support intelligent O&M.

Intelligent O&M supports the following functions:

- **Creating a Scan Task**
Before using the intelligent O&M function, you need to create a scan task.
- **Viewing Cluster Risk Items**
After a scan task is started, you can view details about cluster risk items in the intelligent O&M list.
- **Deleting a Scan Task**
After processing all risk items found in a scan task, you can delete the scan task.

4.12.2 Creating a Scan Task

If the intelligent O&M function is enabled for CSS, you need to start a scan task.

Prerequisites

A CSS cluster has been created. For details, see [Creating an OpenSearch Cluster in Security Mode](#).

Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click the name of the cluster for which you want to perform intelligent O&M. The basic information page of the cluster is displayed.
3. Choose **Intelligent O&M** from the navigation pane.
4. On the Intelligent O&M page, click **Scan** in the upper left corner.
5. In the dialog box, enter the basic information about the scan task and click **OK**.

Table 4-33 Detection task information

Parameter	Description
Name	Name of a scan task.
Description	Brief description of a scan task.
SMN Topic	This parameter is available if you select Send SMN notification upon task completion . If no SMN topic has been created, go to the SMN console to create one.
Notification Level	This parameter is available if you select Send SMN notification upon task completion . If the scan result contains a risk at this level or higher, SMN will send an alarm notification that lists all the risk items in the result.

After a scan task is created, you can view it in the intelligent O&M list.

Follow-up Operations

View cluster risks and diagnose the cluster health status. For details, see [Viewing Cluster Risk Items](#).

4.12.3 Viewing Cluster Risk Items

After a scan task is started, you can view details about cluster risk items in the intelligent O&M list.

Prerequisites

A scan task has been started. For details, see [Creating a Scan Task](#).



Check Items

The following items will be checked and the detected risks will be displayed in the intelligent O&M list:

- Check the current health status of the cluster. Red: Some primary shards are not allocated. Yellow: Some secondary shards are not allocated. Green: that all shards are allocated.
- Check the number of nodes in the cluster and the number of AZs to evaluate the high availability status of the distributed Elasticsearch cluster.
- Check whether index replicas are enabled. If replicas are not enabled and a fault occurs, an index may be unavailable, and the data in a cluster using local disks may be lost.
- Check for Kibana index conflicts in clusters.


- Check disk usage. If the disk usage of a node is too high, new index shards may fail to be allocated to the node and the cluster performance may be affected.
- Check whether the storage usage of cluster data nodes or cold data nodes is balanced. Unbalanced storage distribution may result in unbalanced cluster loads and increase read/write latency.
- Check whether any node in the current cluster is disconnected or unavailable for 5 consecutive minutes.
- Check for nodes with too many shards. A large number of shards will consume too many node resources, increasing read/write latency and slowing down metadata update.
- Check the size of all shards. A large shard may affect performance deterioration, occupy too much node memory, and slow down shard restoration during scaling or fault recovery.
- Check whether the current cluster has an available new version.
- Check for snapshot creation failures and snapshot records in the cluster in the last seven days.


Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click a cluster name to go to the basic information page of the cluster.
3. Choose **Intelligent O&M** from the navigation pane.
4. On the intelligent O&M list page, select a started scan task. Click  on the left of the task name to view its creation time, abstract, ID, and risk items. Click  on the left of a risk item to view its details, including the check item, risk description, and risk suggestion.

You can handle cluster risks in a timely manner based on the suggestions.

Figure 4-44 Risk items

Name/ID	Description	Task Status	Risk Item	Created	SMN Notification	Operation
		● Completed	High-risk items: 1 Medium-risk items: 1 Warnings: 0	Mar 20, 2024 14:41:54 GMT+...	● Sent	Delete Export Risk

Created	Mar 20, 2024 14:41:54 GMT+08:00	ID	
Summary	Analysis completed. High-risk items: 1; medium-risk items: 1; warnings: 0		
<div style="border: 1px solid #ccc; padding: 5px;"> <p>High-Risk Items(1)</p> <p>Check Item Check the number of nodes in the cluster and the number of AZs(Available Zones) to evaluate the high availability status of the distributed Elasticsearch cluster.</p> <p>Risk Description The current cluster has one or two nodes. If a node is faulty, the entire cluster may become unavailable. The service availability risk is high.</p> <p>Suggestion You are advised to change the cluster to a multi-AZ cluster. Procedure: On the CSS cluster console, choose Clusters > Elasticsearch. In the Operation column of a cluster, choose More > Modify Configuration. Click the Change AZ tab and add AZs. Click the Scale Cluster tab and change the number of nodes.</p> </div>			
<div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;"> <p>Medium-Risk Items(1)</p> </div>			
<div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;"> <p>Warnings(0)</p> </div>			

4.12.4 Deleting a Scan Task

After processing all risk items found in a scan task, you can delete the scan task. After a scan task is deleted, the system deletes all diagnosis information corresponding to the scan task.

Prerequisites

A scan task has been started. For details, see [Creating a Scan Task](#).

Procedure

1. Log in to the CSS management console.
2. On the cluster management page, click a cluster name to go to the basic information page of the cluster.
3. Choose **Intelligent O&M** from the navigation pane.
4. Locate a scan task you want to delete and click **Delete** in the **Operation** column.
5. In the dialog box, click **OK**.

4.13 OpenSearch Dashboards

4.13.1 Logging In to the OpenSearch Dashboards

Restrictions

- You can customize the username, role name, and tenant name in the **OpenSearch Dashboards**. Chinese characters are not allowed.
- **OpenSearch Dashboards** does not support Chinese characters.

Prerequisites

An OpenSearch cluster has been created.

Procedure

- Logging in to the console
 - a. Log in to the CSS management console.
 - b. In the navigation pane, choose **Clusters > OpenSearch**.
 - c. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column to go to the OpenSearch login page.
 - Non-security cluster: The OpenSearch Dashboards console is displayed.
 - Security cluster: Enter the username and password on the login page and click **Log In** to go to the OpenSearch console. The default username is **admin** and the password is the one specified during cluster creation.

Figure 4-45 Logging in to OpenSearch



- d. After the login is successful, access the cluster and perform related operations on the OpenSearch Dashboards.
- Logging in using an EIP
If you have enabled Kibana public access during cluster creation, you can use the Kibana public IP address to log in to the cluster. For details, see [Kibana Public Access](#).

4.13.2 Accessing a Cluster from a Kibana Public Network

For CSS clusters that have security mode enabled, you can enable Kibana public access. After the configuration is complete, an IP address will be provided to access Kibana of this cluster over the Internet.

You can configure Kibana public access during cluster creation, or after a cluster in security mode is created.

NOTE

The whitelist for Kibana public network access depends on the ELB whitelist. After you updated the whitelist, the new settings take effect immediately for new connections. For existing persistent connections using the IP addresses that have been removed from the whitelist, the new settings take effect about 1 minute after these connections are stopped.

Configuring Kibana Public Access When Creating a Cluster

1. Log in to the CSS management console.
2. Click **Create Cluster** in the upper right corner. The **Create Cluster** page is displayed.
3. On the **Create Cluster** page, enable **Security Mode**.
4. Set **Advanced Settings** to **Custom**, enable **Kibana Public Access**, and set parameters.

Table 4-34 Kibana public access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address Value range: 1 to 100 Unit: Mbit/s
Access Control	If you disable this function, all IP addresses can access Kibana through the public IP address. If you enable this function, only IP addresses or IP address in the whitelist can access Kibana through the public IP address.
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled. You are advised to enable this function.

After the cluster is created, click the cluster name to go to the **Basic Information** page. On the **Kibana Public Access** page, you can view the Kibana public IP address.

Configuring Kibana Public Access for an Existing Cluster

You can enable, disable, modify, and view Kibana public access for an existing cluster that has security mode enabled.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
4. Click the **Kibana Public Access** tab. Turn on the **Kibana Public Access** switch to enable the Kibana public access function.
5. On the displayed page, set parameters.

Table 4-35 Kibana public access parameters

Parameter	Description
Bandwidth	Bandwidth for accessing Kibana with the public IP address Value range: 1 to 100 Unit: Mbit/s
Access Control	If you disable this function, all IP addresses can access Kibana through the public IP address. If you enable this function, only IP addresses or IP address in the whitelist can access Kibana through the public IP address.

Parameter	Description
Whitelist	IP address or IP address range allowed to access a cluster. Use commas (,) to separate multiple addresses. This parameter can be configured only when Access Control is enabled. You are advised to enable this function.

6. After you set the parameters, click **OK**.

Modifying Kibana Public Access

For clusters configured Kibana public access, you can modify its bandwidth and access control or disable this function.

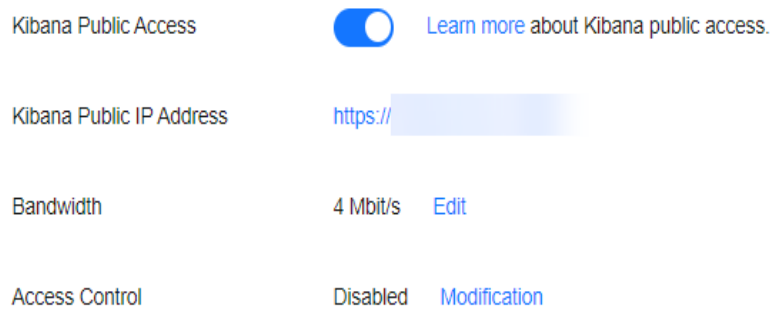
1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
4. Click the **Kibana Public Access** tab to modify the Kibana public access configuration.
 - Modifying bandwidth
Click **Modify** on the right of **Bandwidth**. On the **Modify Bandwidth** page, modify the bandwidth and click **OK**.
 - Modifying access control
Click **Modify** on the right of **Access Control**. On the **Modify Access Control** page, set **Access Control** and **Whitelist**, and click **OK**.
 - Disabling Kibana public access
Toggle off the **Kibana Public Access** switch.

Accessing OpenSearch Dashboard with the Public IP Address

After configuring Kibana public access, you will obtain a public IP address that you can use to access OpenSearch Dashboard of this cluster.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Choose **Clusters** in the navigation pane. On the **Clusters** page, click the name of the target cluster.
4. Click the **Kibana Public Access** tab to obtain the Kibana public IP address.

Figure 4-46 Obtaining the Kibana public IP address



5. Use this IP address to access OpenSearch Dashboard of this cluster through the Internet.

4.13.3 Creating and Authorizing a User on the OpenSearch Dashboards

Prerequisites

The security mode has been enabled for the OpenSearch cluster.

Parameters

Table 4-36 Parameters for creating and authorizing a user on Kibana

Parameter	Description
Permission	Single permission, for example, creating an index (for example, indices:admin/create)
Action group	A group of permissions. For example, the predefined SEARCH action group grants roles permissions to use _search and _msearchAPI .
Role	A role is a combination of permissions and action groups, including operation permissions on clusters, indexes, documents, or fields.
Backend role	(Optional) Other external roles from the backend such as LDAP/Active Directory
User	A user can send operation requests to Elasticsearch clusters. The user has credentials such as username and password, and zero or multiple backend roles and custom attributes.

Parameter	Description
Role mapping	A user will be assigned a role after successful authentication. Role mapping is to map a role to a user (or a backend role). For example, the mapping from kibana_user (role) to jdoe (user) means that John Doe obtains all permissions of kibana_user after being authenticated by kibana_user . Similarly, the mapping from all_access (role) to admin (backend role) means that any user with the backend role admin (from the LDAP/Active Directory server) has all the permissions of role all_access after being authenticated. You can map each role to multiple users or backend roles.

 **NOTE**

You can customize the username, role name, and tenant name in the **OpenSearch Dashboards**. Chinese characters are not allowed.

Procedure

Step 1 Log in to the OpenSearch Dashboards.

1. Log in to the CSS management console.
2. In the navigation pane, choose **Clusters > OpenSearch**.
3. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
4. Enter the administrator username and password to log in to the OpenSearch Dashboards.
 - Username: **admin** (default administrator account name)
 - Password: Enter the administrator password you set when creating the cluster in security mode.

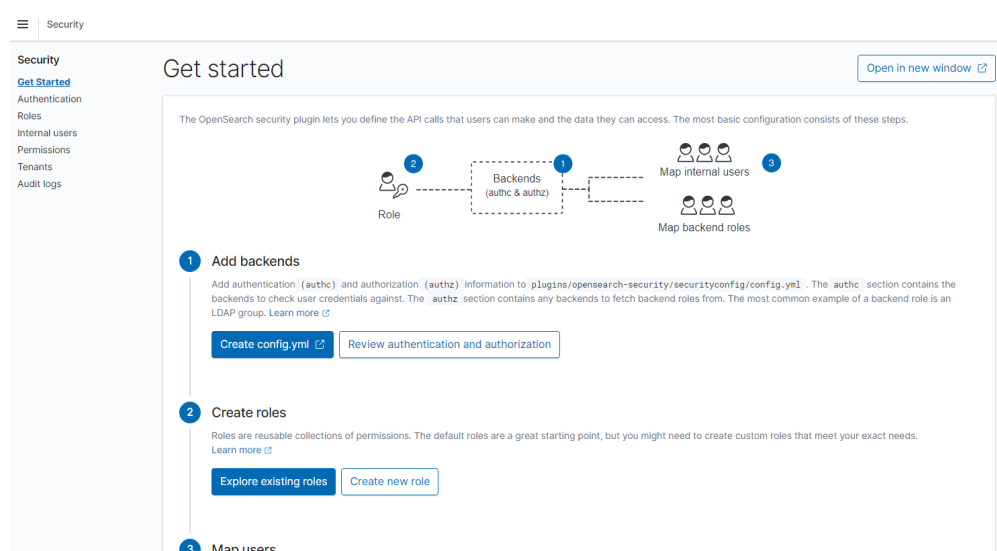
Figure 4-47 Logging in to OpenSearch



Step 2 Creating a user.

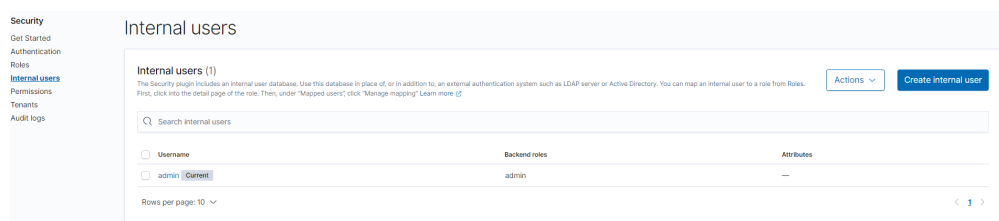
1. On the **OpenSearch Dashboards** page, choose **Security**. The **Security** page is displayed.

Figure 4-48 Going to the Security page



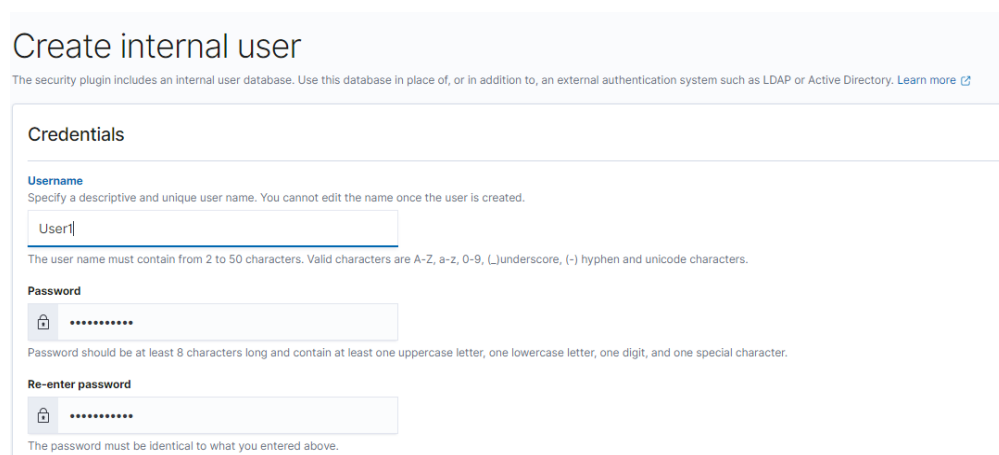
2. Choose **Internal users** on the left. The user creation page is displayed.

Figure 4-49 Creating a user



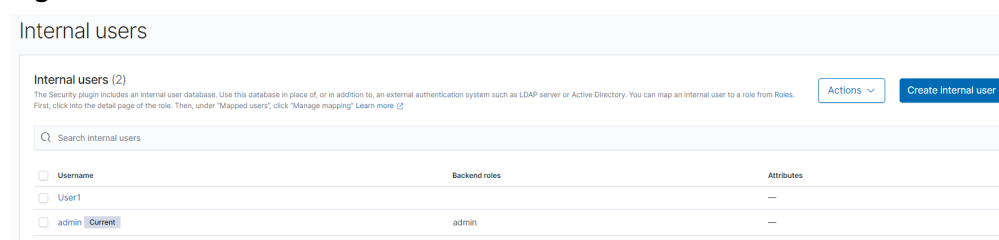
3. Click **Create internal user**. The user information configuration page is displayed.
4. In the **Credentials** area, enter the username and password.

Figure 4-50 Entering the username and password



5. Click **Create**. After the user is created, it is displayed in the user list.

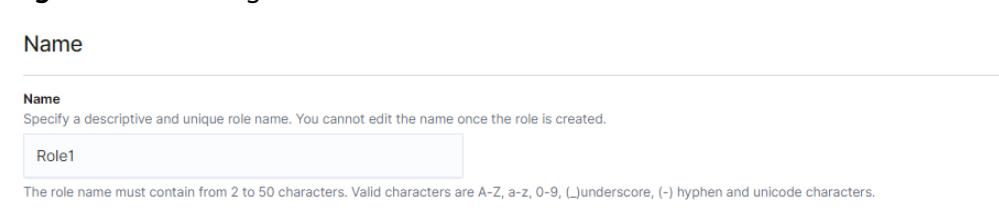
Figure 4-51 User information



Step 3 Create a role and grant permissions to the role.

1. Select **Roles** from the **Security** drop-down list box.
2. On the **Roles** page, click **Create role**. The role creation page is displayed.
3. In the **Name** area, set the role name.

Figure 4-52 Setting a role name



4. On the **Cluster Permissions** page, set the cluster permission. Set cluster permissions based on service requirements. If this parameter is not specified for a role, the role has no cluster-level permissions.

Figure 4-53 Assigning cluster-level permissions

Cluster permissions

Specify how users in this role can access the cluster. By default, no cluster permission is granted. [Learn more](#)

Cluster Permissions

Specify permissions using either action groups or single permissions. An action group is a list of single permissions. You can often

✕ ▼
Create new permission group

5. In the **Index Permissions** area, set the index permission.

Figure 4-54 Setting index permissions

Index permissions

Index permissions allow you to specify how users in this role can access the specific indices. By default, no index permission is granted. [Learn more](#)

▼ my_store

Index

✕

Specify index pattern using *

Index permissions

You can specify permissions using both action groups or single permissions. A permission group is a list of single permissions. You can often achieve your

✕ ▼
Create new permission group

6. On the **Tenant Permissions** page, set role permissions.

Figure 4-55 Role permissions

Tenant permissions

Tenants are useful for safely sharing your work with other OpenSearch Dashboards users. You can control which roles have access to a tenant and

Tenant

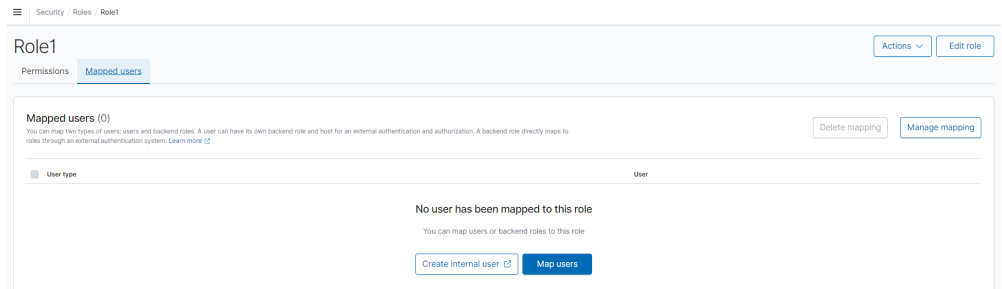
✕ ▼
Read and Write ▼
Remove

Add another tenant permission

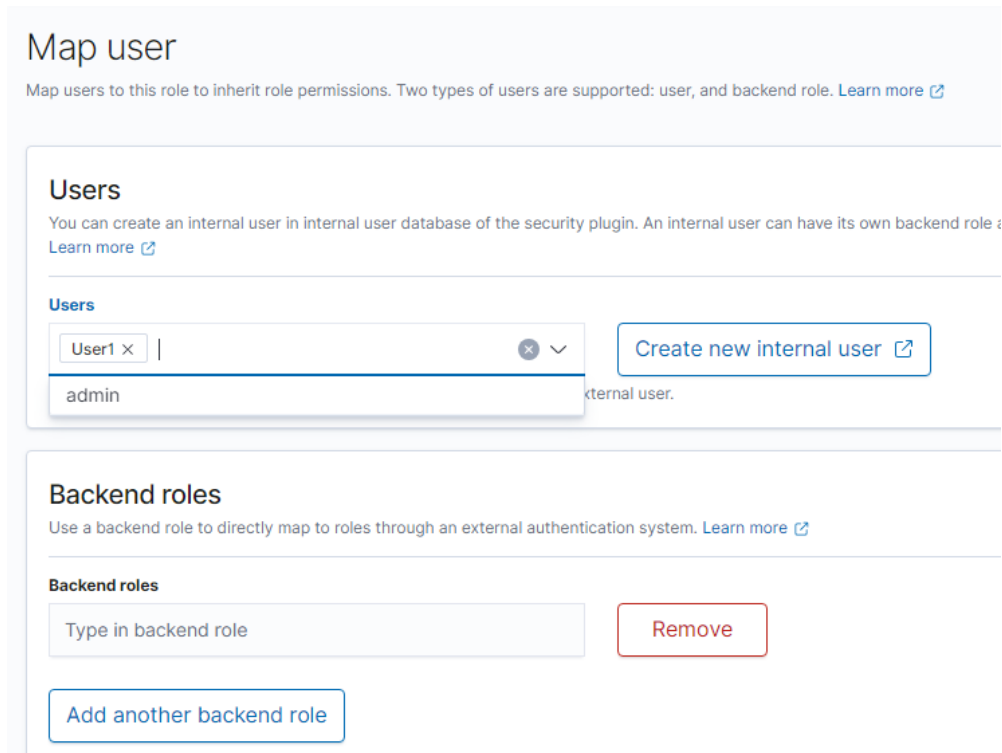
After the setting is complete, you can view the created role on the **Roles** page.

Step 4 Map a user with a role to bind them.

1. Select **Roles** from the **Security** drop-down list box.
2. On the **Roles** page, select the role to be mapped. The role mapping page is displayed.



3. On the **Mapped users** tab page, click **Map users** and select the user to be mapped from the **users** drop-down list box.



4. Click **Map**.
5. After the configuration is complete, you can check whether the configuration takes effect in OpenSearch Dashboards.

----End

5 Viewing the Cluster Runtime Status and Storage Capacity Status

On the **Dashboard** page of the CSS management console, you can view information about the status and storage capacity of existing clusters.

Table 5-1 Cluster status description

Status	Description
Available	The cluster is running properly and is providing services.
Abnormal	The cluster creation failed or the cluster is unavailable. If a cluster is in the unavailable status, you can delete the cluster or use snapshots created when the cluster is available to restore data to other clusters. However, operations such as expanding cluster capacity, accessing Kibana, creating snapshots, and restoring snapshots to the cluster are not allowed. When a cluster is in the unavailable status, data importing is not recommended to avoid data loss. You can view the cluster metrics or restart the cluster. However, the operations may fail. If the operations fail, contact technical support in a timely manner.
Processing	The cluster is being restarted, scaled, backed up, or recovered.
Creating	The cluster is being created.

Table 5-2 Cluster storage capacity status description

Status	Description
Normal	The storage capacity usage of all nodes in a cluster is less than 50%.

Status	Description
Warning	The storage capacity usage of any node in a cluster is greater than or equal to 50% and less than 80%.
Danger	The storage capacity usage of any node in a cluster is greater than or equal to 80%. You are advised to increase the storage space of the cluster to achieve normal data search or analysis.
Abnormal	The cluster storage capacity usage is unknown. For example, if the status of a cluster is Abnormal due to faults, the storage space status of the cluster will be Abnormal .

6 Enhanced Cluster Features

6.1 Vector Retrieval

6.1.1 Description

Image recognition and retrieval, video search, and personalized recommendation impose high requirements on the latency and accuracy of high-dimensional space vector retrieval. To facilitate large-scale vector search, CSS integrates the vector search feature powered by Huawei's vector search engine and the Elasticsearch plug-in mechanism.

Principles

Vector search works in a way similar to traditional search. To improve vector search performance, we need to:

- **Narrow down the matched scope**

Similar to traditional text search, vector search use indexes to accelerate the search instead of going through all data. Traditional text search uses inverted indexes to filter out irrelevant documents, whereas vector search creates indexes for vectors to bypass irrelevant vectors, narrowing down the search scope.
- **Reduce the complexity of calculating a single vector**

The vector search method can quantize and approximate high dimensional vectors first. By doing this, you can acquire a smaller and more relevant data set. Then more sophisticated algorithms are applied to this smaller data set to perform computation and sorting. This way, complex computation is performed on only part of the vectors, and efficiency is improved.

Vector search means to retrieve the k-nearest neighbors (KNN) to the query vector in a given vector data set by using a specific measurement method. Generally, CSS only focuses on Approximate Nearest Neighbor (ANN), because a KNN search requires excessive computational resources.

Functions

The engine developed by Huawei Cloud integrates a variety of vector indexes, such as brute-force search, Hierarchical Navigable Small World (HNSW) graphs, product quantization, and IVF-HNSW. It also supports multiple similarity calculation methods, such as Euclidean, inner product, cosine, and Hamming. The recall rate and retrieval performance of the engine are better than those of open-source engines. It can meet the requirements for high performance, high precision, low costs, and multi-modal computation.

The search engine also supports all the capabilities of the native Elasticsearch, including distribution, multi-replica, error recovery, snapshot, and permission control. The engine is compatible with the native Elasticsearch ecosystem, including the cluster monitoring tool Cerebro, the visualization tool Kibana, and the real-time data ingestion tool Logstash. Several client languages, such as Python, Java, Go, and C++, are supported.

Constraints

- Only Elasticsearch clusters of versions 7.6.2 and 7.10.2 and OpenSearch clusters of version 1.3.6 support vector search.
- The vector search plug-in performs in-memory computing and requires more memory than common indexes do. It is recommended that the memory of the cluster node be greater than or equal to 8 GB and the cluster computing specifications be memory-optimized.

6.1.2 Cluster Planning for Vector Retrieval

Off-heap memory is used for index construction and query in vector retrieval. Therefore, the required cluster capacity is related to the index type and off-heap memory size. You can estimate the off-heap memory required by full indexing to select proper cluster specifications. The memory usage of vector search is high, CSS disables the vector search plug-in by default for clusters whose memory is 8 GB or less.

There are different methods for estimating the size of off-heap memory required by different types of indexes. The calculation formulas are as follows:

- **GRAPH Index**

$$mem_needs = (dim \times dim_size + neighbors \times 4) \times num + delta$$

NOTE

If you need to update indexes in real time, consider the off-heap memory overhead required for vector index construction and automatic merge. The actual size of required **mem_needs** is at least 1.5 to 2 times of the original estimation.

- **PQ Index**

$$mem_needs = frag_num \times frag_size \times num + delta$$

- **FALT and IVF Indexes**

$$mem_needs = dim \times dim_size \times num + delta$$

Table 6-1 Parameter description

Parameter	Description
dim	Vector dimensions
neighbors	Number of neighbors of a graph node. The default value is 64 .
dim_size	Number of bytes required by each dimension. The default value is four bytes in the float type.
num	Total number of vectors
delta	Metadata size. This parameter can be left blank.
frag_num	Number of vector segments during quantization and coding. If this parameter is not specified when an index is created, the value is determined by the vector dimension dim . if dim <= 256: frag_num = dim / 4 elif dim <= 512: frag_num = dim / 8 else : frag_num = 64
frag_size	Size of the center point during quantization and coding. The default value is 1 . If the value of frag_num is greater than 256 , the value of frag_size is 2 .

These calculation methods can estimate the size of off-heap memory required by a complete vector index. To determine cluster specifications, you also need to consider the heap memory overhead of each node.

Heap memory allocation policy: The size of the heap memory of each node is half of the node physical memory, and the maximum size is **31 GB**.

For example, if you create a Graph index for the SIFT10M dataset, set **dim** to **128**, **dim_size** to **4**, **neighbors** to default value **64**, and **num** to **10 million**, the off-heap memory required by the Graph index is as follows:

$$mem_needs = (128 \times 4 + 64 \times 4) \times 10000000 \approx 7.5GB$$

Considering the overhead of heap memory, a single server with **8 vCPUs** and **16 GB memory** is recommended. If real-time write or update is required, you need to apply for larger memory.

6.1.3 Creating a Vector Index

Prerequisites

- You have created a cluster by referring to [Cluster Planning for Vector Retrieval](#). The cluster must be an Elasticsearch cluster of version 7.6.2 or 7.10.2, or an OpenSearch cluster of version 1.3.6.

- Cluster advanced settings have been configured as required by referring to [Advanced Cluster Configurations](#).

Creating a Vector Index

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. Click **Dev Tools** in the navigation tree on the left and run the following command to create a vector index.

Create an index named **my_index** that contains a vector field **my_vector** and a text field **my_label**. The vector field creates the graph index and uses Euclidean distance to measure similarity.

```
PUT my_index
{
  "settings": {
    "index": {
      "vector": true
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "dimension": 2,
        "indexing": true,
        "algorithm": "GRAPH",
        "metric": "euclidean"
      },
      "my_label": {
        "type": "text"
      }
    }
  }
}
```

Table 6-2 Parameters for creating an index

Type	Parameter	Description
Index settings parameters	vector	To use a vector index, set this parameter to true .
Field mappings parameters	type	Field type, for example, vector .
	dimension	Vector dimension. Value range: [1, 4096]

Type	Parameter	Description
	indexing	<p>Whether to enable vector index acceleration. The value can be:</p> <ul style="list-style-type: none"> • false: disables vector index acceleration. If this parameter is set to false, vector data is written only to docvalues, and only ScriptScore and Rescore can be used for vector query. • true: enables vector index acceleration. If this parameter is set to true, an extra vector index is created. The index algorithm is specified by the algorithm field and VectorQuery can be used for data query. <p>Default value: false</p>

Type	Parameter	Description
	algorithm	<p>Index algorithm. This parameter is valid only when indexing is set to true.</p> <p>The value can be:</p> <ul style="list-style-type: none"> • FLAT: brute-force algorithm that calculates the distance between the target vector and all vectors in sequence. The algorithm relies on sheer computing power and its recall rate reaches 100%. You can use this algorithm if you require high recall accuracy. • GRAPH: Hierarchical Navigable Small Worlds (HNSW) algorithm for graph indexes. This algorithm is mainly used in scenarios where high performance and precision are required and the data records of a single shard is fewer than 10 million. • GRAPH_PQ: combination of the HNSW algorithm and the PQ algorithm. The PQ algorithm reduces the storage overhead of original vectors, so that HNSW can easily search for data among hundreds of millions of records. • IVF_GRAPH: combination of IVF and HNSW. The entire space is divided into multiple cluster centroids, which makes search much faster but slightly inaccurate. You can use this algorithm if you require high performance when searching for data among hundreds of millions of records. • IVF_GRAPH_PQ: combination of the PQ algorithm with the IVF or HNSW algorithm to further improve the system capacity and reduce the system overhead. This algorithm is applicable to scenarios where there are more than 1 billion files in shards and high retrieval performance is required. • PV_GRAPH: Improved Hierarchical Navigable Small Worlds (HNSW) algorithm for graph index. This algorithm is applicable to scenarios where there are fewer than 10 million files in a single shard, available memory resources are sufficient, and high performance and precision are required. This algorithm supports the vector and scalar joint filtering. Currently, the sub_fields parameter supports only the keyword type. Compared with post-filtering and Boolean

Type	Parameter	Description
		<p>query, it greatly improves the filling rate of returned results and the search performance. Only Elasticsearch cluster 7.10.2 supports the PV_GRAPH index.</p> <p>Default value: GRAPH</p> <p>NOTE If IVF_GRAPH or IVF_GRAPH_PQ is specified, you need to pre-build and register a central point index. For details, see (Optional) Pre-Building and Registering a Center Point Vector.</p>
	Table 6-3	If Indexing is set to true , CSS provides optional parameters for vector search to achieve higher query performance or precision.
	metric	<p>Method of calculating the distance between vectors.</p> <p>The value can be:</p> <ul style="list-style-type: none"> • euclidean: Euclidean distance • inner_product: inner product distance • cosine: cosine distance • hamming: Hamming distance, which can be used only when dim_type is set to binary. <p>Default value: euclidean</p>
	dim_type	<p>Type of the vector dimension value.</p> <p>The value can be binary and float (default).</p>
	sub_fields	<p>Define the auxiliary scalar field of vectors. Only the keyword type is supported. This parameter must be specified if you need to use the vector and scalar joint filtering feature. It takes effect only when algorithm is set to PV_GRAPH.</p>

Table 6-3 Optional parameters

Type	Parameter	Description
Graph index configuration parameters	neighbors	<p>Number of neighbors of each vector in a graph index. The default value is 64. A larger value indicates higher query precision. A larger index results in a slower build and query speed.</p> <p>Value range: [10, 255]</p>

Type	Parameter	Description
	shrink	Cropping coefficient during HNSW build. The default value is 1.0f . Value range: (0.1, 10)
	scaling	Scaling ratio of the upper-layer graph nodes during HNSW build. The default value is 50 . Value range: (0, 128]
	efc	Queue size of the neighboring node during HNSW build. The default value is 200 . A larger value indicates a higher precision and slower build speed. Value range: (0, 100000]
	max_scan_num	Maximum number of nodes that can be scanned. The default value is 10000 . A larger value indicates a higher precision and slower indexing speed. Value range: (0, 1000000]
PQ index configuration parameters	centroid_num	Number of cluster centroids of each fragment. The default value is 255 . Value range: (0, 65535]
	fragment_num	Number of fragments. The default value is 0 . The plug-in automatically sets the number of fragments based on the vector length. Value range: [0, 4096]

Importing Vector Data

Run the following command to import vector data. When writing vector data to the **my_index** index, you need to specify the vector field name and vector data.

- If the input vector data is an array of floating-point numbers separated by commas (,):

```
POST my_index/_doc
{
  "my_vector": [1.0, 2.0]
}
```

- If the input vector data is a Base64 string encoded using little endian:

When writing binary vectors or high dimensional vectors that have a large number of valid bits, the Base64 encoding format is efficient for data transmission and parsing.

```
POST my_index/_doc
{
  "my_vector": "AACAPwAAAEA="
}
```

- To write a large amount of data, bulk operations are recommended.

```
POST my_index/_bulk
{"index": {}}
{"my_vector": [1.0, 2.0], "my_label": "red"}
{"index": {}}
{"my_vector": [2.0, 2.0], "my_label": "green"}
{"index": {}}
{"my_vector": [2.0, 3.0], "my_label": "red"}
```

Advanced Cluster Configurations

- When importing data offline, you are advised to set **refresh_interval** of indexes to **-1** to disable automatic index refreshing and improve batch write performance.
- You are advised to set **number_of_replicas** to **0**. After the offline data import is complete, you can modify the parameter value as needed.
- The parameters of other advanced functions as follows:

Table 6-4 Cluster parameters

Parameter	Description
native.cache.circuit_breaker.enabled	Whether to enable the circuit breaker for off-heap memory. Default value: true
native.cache.circuit_breaker.cpu.limit	Upper limit of off-heap memory usage of the vector index. For example, if the overall memory of a host is 128 GB and the heap memory occupies 31 GB, the default upper limit of the off-heap memory usage is 43.65 GB, that is, (128 - 31) x 45%. If the off-heap memory usage exceeds its upper limit, the circuit breaker will be triggered. Default value: 45%
native.cache.expire.enabled	Whether to enable the cache expiration policy. If this parameter is set to true , some cache items that have not been accessed for a long time will be cleared. Value: true or false Default value: false
native.cache.expire.time	Expiration time. Default value: 24h
native.vector.index_threads	Number of threads used for creating underlying indexes. Each shard uses multiple threads. Set a relatively small value to avoid resource preemption caused by the build queries of too many threads. Default value: 4

6.1.4 Querying Vectors

Standard Query

Standard vector query syntax is provided for vector fields with vector indexes. The following command will return n (specified by **size/topk**) data records that are most close to the query vector.

```
POST my_index/_search
{
  "size":2,
  "_source": false,
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1, 1],
        "topk":2
      }
    }
  }
}
```

Table 6-5 Parameters for standard query

Parameter	Description
vector (the first one)	Indicates that the query type is VectorQuery .
my_vector	Indicates the name of the vector field you want to query.
vector (the second one)	Indicates the vector value you want to query, which can be an array or a Base64 string
topk	Same as the value of size generally.
Table 6-6	Indicates optional query parameters. You can adjust the vector index parameters to achieve higher query performance or precision.

Table 6-6 Optional query parameters

Type	Parameter	Description
Graph index configuration parameters	ef	Queue size of the neighboring node during the query. A larger value indicates a higher query precision and slower query speed. The default value is 200 . Value range: (0, 100000]
	max_scan_num	Maximum number of scanned nodes. A larger value indicates a higher query precision and slower query speed. The default value is 10000 . Value range: (0, 1000000]

Type	Parameter	Description
IVF index configuration parameters	nprobe	Number of center points. A larger value indicates a higher query precision and slower query speed. The default value is 100 . Value range: (0, 100000]

Compound Query

Vector search can be used together with other Elasticsearch subqueries, such as Boolean query and post-filtering, for compound query.

In the following two examples, top 10 (**topk**) results closest to the query vector are queried first. **filter** retains only the results whose **my_label** field is **red**.

- Example of a Boolean query

```
POST my_index/_search
{
  "size": 10,
  "query": {
    "bool": {
      "must": {
        "vector": {
          "my_vector": {
            "vector": [1, 2],
            "topk": 10
          }
        }
      }
    },
    "filter": {
      "term": { "my_label": "red" }
    }
  }
}
```

- Example of post-filtering

```
GET my_index/_search
{
  "size": 10,
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1, 2],
        "topk": 10
      }
    }
  },
  "post_filter": {
    "term": { "my_label": "red" }
  }
}
```

ScriptScore Query

You can use **script_score** to perform Nearest Neighbor Search (NSS) on vectors. The query syntax is provided below.

The pre-filtering condition can be any query. **script_score** traverses only the pre-filtered results, calculates the vector similarity, and sorts and returns the results.

The performance of this query depends on the size of the intermediate result set after the pre-filtering. If the pre-filtering condition is set to **match_all**, brute-force search is performed on all data.

```
POST my_index/_search
{
  "size":2,
  "query": {
    "script_score": {
      "query": {
        "match_all": {}
      },
      "script": {
        "source": "vector_score",
        "lang": "vector",
        "params": {
          "field": "my_vector",
          "vector": [1.0, 2.0],
          "metric": "euclidean"
        }
      }
    }
  }
}
```

Table 6-7 script_score parameters

Parameter	Description
source	Script description. Its value is vector_score if the vector similarity is used for scoring.
lang	Script syntax description. Its value is vector .
field	Vector field name
vector	Vector data to be queried
metric	Measurement method, which can be euclidean , inner_product , cosine , and hamming . Default value: euclidean

Re-Score Query

If the **GRAPH_PQ** or **IVF_GRAPH_PQ** index is used, the query results are sorted based on the asymmetric distance calculated by PQ. CSS supports re-scoring and ranking of query results to improve the recall rate.

Assuming that **my_index** is a PQ index, an example of re-scoring the query results is as follows:

```
GET my_index/_search
{
  "size": 10,
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1.0, 2.0],
        "topk": 100
      }
    }
  }
}
```

```

},
"rescore": {
  "window_size": 100,
  "vector_rescore": {
    "field": "my_vector",
    "vector": [1.0, 2.0],
    "metric": "euclidean"
  }
}
}
}

```

Table 6-8 Rescore parameter description

Parameter	Description
window_size	Vector retrieval returns <i>topk</i> search results and ranks the first <i>window_size</i> results.
field	Vector field name
vector	Vector data to be queried
metric	Measurement method, which can be euclidean , inner_product , cosine , and hamming . Default value: euclidean

Painless Syntax Extension

CSS extension supports multiple vector distance calculation functions, which can be directly used in customized painless scripts to build flexible re-score formulas.

The following is an example:

```

POST my_index/_search
{
  "size": 10,
  "query": {
    "script_score": {
      "query": {
        "match_all": {}
      },
      "script": {
        "source": "1 / (1 + euclidean(params.vector, doc[params.field]))",
        "params": {
          "field": "my_vector",
          "vector": [1, 2]
        }
      }
    }
  }
}

```

The following table lists the distance calculation functions supported by the CSS.

Function Signature	Description
euclidean(Float[], DocValues)	Euclidean distance function

Function Signature	Description
cosine(Float[], DocValues)	Cosine similarity function
innerproduct(Float[], DocValues)	Inner product function
hamming(String, DocValues)	Hamming distance function Only vectors whose dim_type is binary are supported. The input query vector must be a Base64-encoded character string.

6.1.5 Optimizing the Performance of Vector Retrieval

Optimizing Write Performance

- To reduce the cost of backup, disable the backup function before data import and enable it afterwards.
- Set **refresh_interval** to **120s** or a larger value. Larger segments can reduce the vector index build overhead caused by merging.
- Increase the value of **native.vector.index_threads** (the default value is **4**) to increase the number of threads for vector index build.

```
PUT _cluster/settings
{
  "persistent": {
    "native.vector.index_threads": 8
  }
}
```

Optimizing Query Performance

- After importing data in batches, you can run the **forcemerge** command to improve the query efficiency.
- If the off-heap memory required by the vector index exceeds the circuit breaker limit, index entry swap-in and swap-out occur, which affects the query performance. In this case, you can increase the circuit breaker threshold of off-heap memory.

```
POST index_name/_forcemerge?max_num_segments=1
PUT _cluster/settings
{
  "persistent": {
    "native.cache.circuit_breaker.cpu.limit": "75%"
  }
}
```

- If the fetch overhead is high, you can configure **_source** to reduce the **fdt** file size to reduce the fetch overhead.

```
PUT my_index
{
  "settings": {
    "index": {
      "vector": "true"
    },
    "index.soft_deletes.enabled": false
  },
  "mappings": {
    "_source": {
```

```
"excludes": ["my_vector"]
},
"properties": {
  "my_vector": {
    "type": "vector",
    "dimension": 128,
    "indexing": true,
    "algorithm": "GRAPH",
    "metric": "euclidean"
  }
}
}
```

6.1.6 (Optional) Pre-Building and Registering a Center Point Vector

When you perform operations in [Creating a Vector Index](#), if **IVF_GRAPH** and **IVF_GRAPH_PQ** index algorithms are selected, you need to pre-build and register the center point vector.

Context

The vector index acceleration algorithms **IVF_GRAPH** and **IVF_GRAPH_PQ** are suitable for ultra-large-scale computing. These two algorithms allow you to narrow down the query range by dividing a vector space into subspaces through clustering or random sampling. Before pre-build, you need to obtain all center point vectors by clustering or random sampling.

Then, pre-construct and register the center point vectors to create the **GRAPH** or **GRAPH_PQ** index and register them with the Elasticsearch cluster. All nodes in the cluster can share the index file. Reuse of the center index among shards can effectively reduce the training overhead and the number of center index queries, improving the write and query performance.

Procedure

1. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
2. Click **Dev Tools** in the navigation tree on the left.
3. Create a center point index table.
 - For example, if the created index is named **my_dict**, **number_of_shards** of the index must be set to **1**. Otherwise, the index cannot be registered.
 - If you want to use the **IVF_GRAPH** index, set **algorithm** of the center point index to **GRAPH**.
 - If you want to use the **IVF_GRAPH_PQ** index, set **algorithm** of the center point index to **GRAPH_PQ**.

```
PUT my_dict
{
  "settings": {
    "index": {
      "vector": true
    },
    "number_of_shards": 1,
    "number_of_replicas": 0
  },
  "mappings": {
```

```

"properties": {
  "my_vector": {
    "type": "vector",
    "dimension": 2,
    "indexing": true,
    "algorithm": "GRAPH",
    "metric": "euclidean"
  }
}
}
}

```

4. Write the center point vector to the created index.

Write the center point vector obtained through sampling or clustering into the created **my_dict** index by referring to [Importing Vector Data](#).

5. Call the registration API.

Register the created **my_dict** index with a **Dict** object with a globally unique identifier name (**dict_name**).

```

PUT _vector/register/my_dict
{
  "dict_name": "my_dict"
}

```

6. Create an **IVF_GRAPH** or **IVF_GRAPH_PQ** index.

You do not need to specify the dimension and metric information. Simply specify the registered dictionary name.

```

PUT my_index
{
  "settings": {
    "index": {
      "vector": true
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "indexing": true,
        "algorithm": "IVF_GRAPH",
        "dict_name": "my_dict",
        "offload_ivf": false
      }
    }
  }
}
}

```

Table 6-9 Field mappings parameters

Parameter	Description
dict_name	Specifies the name of the depended central point index. The vector dimension and measurement metric of the index are the same as those of the Dict index.

Parameter	Description
offload_ivf	Unloads the IVF inverted index implemented by the underlying index to Elasticsearch. In this way, the use of non-heap memory and the overhead of write and merge operations are reduced. However, the query performance also deteriorates. You can use the default value. Value: true or false Default value: false

6.1.7 Managing the Vector Index Cache

The vector retrieval engine is developed in C++ and uses off-heap memory. You can use the following APIs to manage the index cache.

- **View cache statistics.**

GET `/_vector/stats`

In the implementation of the vector plug-in, the vector index is the same as other types of Lucene indexes. Each segment constructs and stores an index file. During query, the index file is loaded to the non-heap memory. The plug-in uses the cache mechanism to manage the non-heap memory. You can use this API to query the non-heap memory usage, number of cache hits, and number of loading times.

- **Preload the vector index.**

PUT `/_vector/warmup/{index_name}`

You can use this API to preload the vector index specified by **index_name** to the off-heap memory for query.

- **Clear the cache.**

PUT `/_vector/clear/cache`

PUT `/_vector/clear/cache/index_name`

The caching mechanism limits the non-heap memory usage when vector indexes are used. When the total index size exceeds the cache size limit, index entry swap-in and swap-out occur, which affects the query performance. You can use this API to clear unnecessary index cache to ensure the query performance of hot data indexes.

6.1.8 Sample Code for Vector Search on a Client

Elasticsearch provides standard REST APIs and clients developed using Java, Python, and Go.

Based on the open-source dataset **SIFT1M** (<http://corpus-texmex.irisa.fr/>) and Python Elasticsearch client, this section provides a code snippet for creating a vector index, importing vector data, and querying vector data on the client.

Prerequisites

The Python dependency package has been installed on the client. If it is not installed, run the following commands to install it:

```
pip install numpy
pip install elasticsearch==7.6.0
```

Sample Code

```
import numpy as np
import time
import json

from concurrent.futures import ThreadPoolExecutor, wait
from elasticsearch import Elasticsearch
from elasticsearch import helpers

endpoint = 'http://xxx.xxx.xxx.xxx:9200/'

# Construct an Elasticsearch client object
es = Elasticsearch(endpoint)

# Index mapping information
index_mapping = """
{
  "settings": {
    "index": {
      "vector": "true"
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "dimension": 128,
        "indexing": true,
        "algorithm": "GRAPH",
        "metric": "euclidean"
      }
    }
  }
}
"""

# Create an index.
def create_index(index_name, mapping):
    res = es.indices.create(index=index_name, ignore=400, body=mapping)
    print(res)

# Delete an index.
def delete_index(index_name):
    res = es.indices.delete(index=index_name)
    print(res)

# Refresh indexes.
def refresh_index(index_name):
    res = es.indices.refresh(index=index_name)
    print(res)

# Merge index segments.
def merge_index(index_name, seg_cnt=1):
    start = time.time()
    es.indices.forcemerge(index=index_name, max_num_segments=seg_cnt, request_timeout=36000)
    print(f" Complete the merge within {time.time() - start} seconds")

# Load vector data.
def load_vectors(file_name):
    fv = np.fromfile(file_name, dtype=np.float32)
    dim = fv.view(np.int32)[0]
    vectors = fv.reshape(-1, 1 + dim)[: , 1:]
```



```
return vectors

# Load the ground_truth data.
def load_gts(file_name):
    fv = np.fromfile(file_name, dtype=np.int32)
    dim = fv.view(np.int32)[0]
    gts = fv.reshape(-1, 1 + dim)[:, 1:]
    return gts

def partition(ls, size):
    return [ls[i:i + size] for i in range(0, len(ls), size)]

# Write vector data.
def write_index(index_name, vec_file):
    pool = ThreadPoolExecutor(max_workers=8)
    tasks = []

    vectors = load_vectors(vec_file)
    bulk_size = 1000
    partitions = partition(vectors, bulk_size)

    start = time.time()
    start_id = 0
    for vecs in partitions:
        tasks.append(pool.submit(write_bulk, index_name, vecs, start_id))
        start_id += len(vecs)
    wait(tasks)
    print(f" Complete the writing within {time.time() - start} seconds")

def write_bulk(index_name, vecs, start_id):
    actions = [
        {
            "_index": index_name,
            "my_vector": vecs[j].tolist(),
            "_id": str(j + start_id)
        }
        for j in range(len(vecs))
    ]
    helpers.bulk(es, actions, request_timeout=3600)

# Query an index.
def search_index(index_name, query_file, gt_file, k):
    print("Start query! Index name: " + index_name)

    queries = load_vectors(query_file)
    gt = load_gts(gt_file)

    took = 0
    precision = []
    for idx, query in enumerate(queries):
        hits = set()
        query_json = {
            "size": k,
            "_source": False,
            "query": {
                "vector": {
                    "my_vector": {
                        "vector": query.tolist(),
                        "topk": k
                    }
                }
            }
        }
        res = es.search(index=index_name, body=json.dumps(query_json))
```

```
for hit in res['hits']['hits']:
    hits.add(int(hit['_id']))
precision.append((len(hits.intersection(set(gt[idx, :k]))) / k)
took += res['took']

print("precision: " + str(sum(precision) / len(precision)))
print(f" Complete the retrieval within {took / 1000:.2f} seconds; average took size is {took /
len(queries):.2f} ms")

if __name__ == "__main__":
    vec_file = r"./data/sift/sift_base.fvecs"
    qry_file = r"./data/sift/sift_query.fvecs"
    gt_file = r"./data/sift/sift_groundtruth.ivecs"

    index = "test"
    create_index(index, index_mapping)
    write_index(index, vec_file)
    merge_index(index)
    refresh_index(index)

    search_index(index, qry_file, gt_file, 10)
```

6.1.9 Using PV_GRAPH to Search for Vector Indexes

PV_GRAPH deeply optimizes the HNSW algorithm and supports the vector and scalar joint filtering. When the vector and scalar joint filtering is used, the result filling rate and query performance can be greatly improved compared with post-filtering and Boolean query.

Prerequisites

An Elasticsearch cluster of version 7.10.2 has been created by referring to [Cluster Planning for Vector Retrieval](#).

Creating an Index

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. Click **Dev Tools** in the navigation tree on the left and run the following command to create a vector index.

Create an index named **my_index** that contains a vector field **my_vector** and two sub-fields **country** and **category**.

```
PUT my_index
{
  "settings": {
    "index": {
      "vector": true
    }
  },
  "mappings": {
    "properties": {
      "my_vector": {
        "type": "vector",
        "dimension": 2,
        "indexing": true,
        "algorithm": "PV_GRAPH",
        "metric": "euclidean",
        "sub_fields": ["country", "category"]
      }
    }
  }
}
```

```
}  
}  
}
```

For details about the parameters for creating an index, see [Table 6-2](#).

NOTE

The **metric** parameter of the PV_GRAPH index algorithm can only be set to **euclidean** or **inner_product**.

Importing the Vector and Scalar Data

When algorithm is set to **PV_GRAPH** and **sub_fields** is specified, the following data writing grammars are supported. The **sub_fields** parameter supports only the **keyword** type and you can specify multiple values for it.

```
# Write a single data record.  
POST my_index/_doc  
{  
  "my_vector": {  
    "data": [1.0, 1.0],  
    "country": "cn",  
    "category": ["1", "2"]  
  }  
}  
  
# Write multiple data records in batches.  
POST my_index/_bulk  
{"index": {}}  
{"my_vector": {"data": [1.0, 2.0], "country": "cn", "category": "1"}}  
{"index": {}}  
{"my_vector": {"data": [2.0, 2.0], "country": "cn", "category": ["1", "2"]}}  
{"index": {}}  
{"my_vector": {"data": [2.0, 3.0], "country": "eu", "category": "2"}}
```

Querying a Vector

Based on the existing Elasticsearch APIs, the **filter** parameter is added to **vector** to support vector and scalar joint filtering. The values of **sub_fields** can be used for scalar filtering. Currently, the JSON format is supported. The **should**, **must**, **must_not**, **term**, and **terms** queries are supported. The syntax is the same as that of Elasticsearch query. The restrictions are as follows:

Currently, up to four layers are supported for filtering nesting.

- **must_not** cannot be nested or contain nest layers.
- The first layer can contain only one query keyword (such as **must**).

The fields defined in **sub_fields** during index creation are the scalar fields used in the joint filtering and take effect only when the **algorithm** is set to **PV_GRAPH**. If the specified filtering field does not exist, the filtering request becomes invalid and the query is processed with no filtering conditions.

```
# Example of single-label and single-value matching query  
GET my_index/_search  
{  
  "query": {  
    "vector": {  
      "my_vector": {  
        "vector": [1.0, 1.0],  
        "topk": 10,  
        "filter": {  
          "term": { "country": "cn" }  
        }  
      }  
    }  
  }  
}
```

```

    }
  }
}

# Example of single-label and multi-value matching query
GET my_index/_search
{
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1.0, 1.0],
        "topk": 10,
        "filter": {
          "terms": { "country": ["cn", "eu"] }
        }
      }
    }
  }
}

# Example of multi-label matching query
GET my_index/_search
{
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1.0, 1.0],
        "topk": 10,
        "filter": {
          "must": [
            {
              "term": {"country": "cn"}
            },
            {
              "terms": {"category": ["1", "2"]}
            }
          ]
        }
      }
    }
  }
}

# Example of must_not matching query
GET my_index/_search
{
  "query": {
    "vector": {
      "my_vector": {
        "vector": [1.0, 1.0],
        "topk": 10,
        "filter": {
          "must_not": [
            {
              "term": {"country": "eu"}
            }
          ]
        }
      }
    }
  }
}

```

For details about vector query parameters, see [Table 6-5](#).

6.2 Storage-Compute Decoupling

6.2.1 Context

You can store hot data on SSD to achieve the optimal query performance, and store historical data in OBS to reduce data storage costs.

Application Scenarios

A large volume of data is written to and stored in SSDs. If historical data is no longer updated (is turned into cold data) and its QPS decreases, you can call CSS APIs to dump hot data from SSDs to OBS buckets. This operation freezes indexes, decoupling compute from storage.

Constraints

- Currently, only Elasticsearch clusters of versions 7.6.2 and 7.10.2 and OpenSearch clusters of version 1.3.6 support decoupled storage and computing.
- The storage-compute decoupling feature depends on OBS. Therefore, you must comply with the restrictions on OBS bandwidth and QPS. For details, see [OBS Restrictions](#). If these restrictions are violated, the performance of queries on OBS will deteriorate. For example, the speed of restoring shards and querying data will become slow.

6.2.2 Freezing an Index

Precautions

- Before freezing an index, ensure no data is being written to it. The index will be set to read only before being frozen, and data write will fail.
- After an index is frozen:
 - It becomes read-only.
 - The index data will be dumped to OBS. This process occupies network bandwidth.
 - The query latency of a dumped index will increase. During aggregation, the latency of processing complex queries and reading a large volume of data is long.
 - It cannot be unfrozen. That is, a read-only index cannot be changed to writable.
 - After the freezing is complete, the index data in your local disks will be deleted.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.

- Click **Dev Tools** in the navigation tree on the left.
- Run the following command to freeze a specified index and dump it to OBS:
POST `/${index_name}/_freeze_low_cost`

Table 6-10 Parameter description

Parameter	Description
index_name	Name of the index to be frozen.

Information similar to the following is displayed:

```
{
  "freeze_uuid": "pdsRgUtSTymVDWR_HoTGFw"
}
```

Table 6-11 Response parameter

Parameter	Description
freeze_uuid	After an index freezing request is submitted, an asynchronous job will be started. The request returns the asynchronous job ID, which can be used to query the progress of the asynchronous job.

 **NOTE**

After an index freezing request is submitted, data cannot be written to the index. During the index freezing, query requests are not affected. After the freezing is complete, the index is closed and then opened. During this period, the index cannot be queried, and the cluster may be in the **red** status for a short time. The index is restored after being opened.

- Run the following command to check the freezing task progress:
GET `_freeze_low_cost_progress/${freeze_uuid}`

Table 6-12 Parameter description

Parameter	Description
freeze_uuid	Asynchronous task ID, which is obtained in 4 .

Information similar to the following is displayed:

```
{
  "stage": "STARTED",
  "shards_stats": {
    "INIT": 0,
    "FAILURE": 0,
    "DONE": 0,
    "STARTED": 3,
    "ABORTED": 0
  },
}
```

```

"indices" : {
  "data1" : [
    {
      "uuid" : "7OS-G1-tRke2jHZPlckexg",
      "index" : {
        "name" : "data1",
        "index_id" : "4b5PHXJITLaS6AurImfQ9A",
        "shard" : 2
      },
      "start_ms" : 1611972010852,
      "end_ms" : -1,
      "total_time" : "10.5s",
      "total_time_in_millis" : 10505,
      "stage" : "STARTED",
      "failure" : null,
      "size" : {
        "total_bytes" : 3211446689,
        "finished_bytes" : 222491269,
        "percent" : "6.0%"
      },
      "file" : {
        "total_files" : 271,
        "finished_files" : 12,
        "percent" : "4.0%"
      },
      "rate_limit" : {
        "paused_times" : 1,
        "paused_nanos" : 946460970
      }
    },
    {
      "uuid" : "7OS-G1-tRke2jHZPlckexg",
      "index" : {
        "name" : "data1",
        "index_id" : "4b5PHXJITLaS6AurImfQ9A",
        "shard" : 0
      },
      "start_ms" : 1611972010998,
      "end_ms" : -1,
      "total_time" : "10.3s",
      "total_time_in_millis" : 10359,
      "stage" : "STARTED",
      "failure" : null,
      "size" : {
        "total_bytes" : 3221418186,
        "finished_bytes" : 272347118,
        "percent" : "8.0%"
      },
      "file" : {
        "total_files" : 372,
        "finished_files" : 16,
        "percent" : "4.0%"
      },
      "rate_limit" : {
        "paused_times" : 5,
        "paused_nanos" : 8269016764
      }
    },
    {
      "uuid" : "7OS-G1-tRke2jHZPlckexg",
      "index" : {
        "name" : "data1",
        "index_id" : "4b5PHXJITLaS6AurImfQ9A",
        "shard" : 1
      },
      "start_ms" : 1611972011021,
      "end_ms" : -1,
      "total_time" : "10.3s",
      "total_time_in_millis" : 10336,

```

```

"stage" : "STARTED",
"failure" : null,
"size" : {
  "total_bytes" : 3220787498,
  "finished_bytes" : 305789614,
  "percent" : "9.0%"
},
"file" : {
  "total_files" : 323,
  "finished_files" : 14,
  "percent" : "4.0%"
},
"rate_limit" : {
  "paused_times" : 3,
  "paused_nanos" : 6057933087
}
}
]
}

```

Table 6-13 Response parameters

Parameter	Description
stage	Status. Its value can be: <ul style="list-style-type: none"> ● INIT: The instance has just started or is being initialized. ● FAILURE: failed ● DONE: complete ● STARTED: started ● ABORTED: Canceled. This field is reserved.
shards_stats	Numbers of shards in each state.
indices	Index status details.

Table 6-14 Return values of **indices**

Parameter	Description
uuid	UUID of the freezing operation
index	Index and shard information
start_ms	Start time
end_ms	End time. If no end time is specified, the value -1 is displayed.
total_time	Time spent
total_time_in_millis	Time spent, in milliseconds
stage	Status of the current shard.
failure	Failure cause. If no failure occurs, null is displayed.

Parameter	Description
size.total_bytes	Size of files to be frozen, in bytes
size.finished_bytes	Frozen bytes
size.percent	Percentage of frozen bytes
file.total_bytes	Number of files to be frozen
file.finished_bytes	Number of frozen files
file.percent	Percentage of frozen files
rate_limit.paused_times	Number of times that freezing is suspended due to rate limit
rate_limit.paused_nanos	Duration of freezing task suspension due to rate limit, in nanoseconds

The following parameters are added to a frozen index. For details, see [Table 6-15](#).

Table 6-15 Frozen index parameters

Parameter	Description
index.frozen_low_cost	Indicates whether an index is frozen. The value is true .
index.blocks.write	Indicates whether data writing is denied in a frozen index. The value is true .
index.store.type	Storage type of an index. The value is obs .

- After an index is frozen, its data will be cached. Run the following command to check the current cache status: For details about the cache, see [Configuring Cache](#).

```
GET _frozen_stats
GET _frozen_stats/${node_id}
```

Table 6-16 Parameter description

Parameter	Description
node_id	Node ID, which can be used to obtain the cache status of a node.

Information similar to the following is displayed:

```
{
  "_nodes" : {
    "total" : 3,
    "successful" : 3,
```

```
"failed" : 0
},
"cluster_name" : "css-zzz1",
"nodes" : {
  "7uwKO38RRoaON37YsXhCYw" : {
    "name" : "css-zzz1-ess-esn-2-1",
    "transport_address" : "10.0.0.247:9300",
    "host" : "10.0.0.247",
    "ip" : "10.0.0.247",
    "block_cache" : {
      "default" : {
        "type" : "memory",
        "block_cache_capacity" : 8192,
        "block_cache_blocksize" : 8192,
        "block_cache_size" : 12,
        "block_cache_hit" : 14,
        "block_cache_miss" : 0,
        "block_cache_eviction" : 0,
        "block_cache_store_fail" : 0
      }
    }
  },
  "obs_stats" : {
    "list" : {
      "obs_list_count" : 17,
      "obs_list_ms" : 265,
      "obs_list_avg_ms" : 15
    },
    "get_meta" : {
      "obs_get_meta_count" : 79,
      "obs_get_meta_ms" : 183,
      "obs_get_meta_avg_ms" : 2
    },
    "get_obj" : {
      "obs_get_obj_count" : 12,
      "obs_get_obj_ms" : 123,
      "obs_get_obj_avg_ms" : 10
    },
    "put_obj" : {
      "obs_put_obj_count" : 12,
      "obs_put_obj_ms" : 2451,
      "obs_put_obj_avg_ms" : 204
    },
    "obs_op_total" : {
      "obs_op_total_ms" : 3022,
      "obs_op_total_count" : 120,
      "obs_op_avg_ms" : 25
    }
  },
  "reader_cache" : {
    "hit_count" : 0,
    "miss_count" : 1,
    "load_success_count" : 1,
    "load_exception_count" : 0,
    "total_load_time" : 291194714,
    "eviction_count" : 0
  }
},
"73EDpEqoQES749umJqxOzQ" : {
  "name" : "css-zzz1-ess-esn-3-1",
  "transport_address" : "10.0.0.201:9300",
  "host" : "10.0.0.201",
  "ip" : "10.0.0.201",
  "block_cache" : {
    "default" : {
      "type" : "memory",
      "block_cache_capacity" : 8192,
      "block_cache_blocksize" : 8192,
      "block_cache_size" : 12,
      "block_cache_hit" : 14,
```

```

    "block_cache_miss" : 0,
    "block_cache_eviction" : 0,
    "block_cache_store_fail" : 0
  }
},
"obs_stats" : {
  "list" : {
    "obs_list_count" : 17,
    "obs_list_ms" : 309,
    "obs_list_avg_ms" : 18
  },
  "get_meta" : {
    "obs_get_meta_count" : 79,
    "obs_get_meta_ms" : 216,
    "obs_get_meta_avg_ms" : 2
  },
  "get_obj" : {
    "obs_get_obj_count" : 12,
    "obs_get_obj_ms" : 140,
    "obs_get_obj_avg_ms" : 11
  },
  "put_obj" : {
    "obs_put_obj_count" : 12,
    "obs_put_obj_ms" : 1081,
    "obs_put_obj_avg_ms" : 90
  },
  "obs_op_total" : {
    "obs_op_total_ms" : 1746,
    "obs_op_total_count" : 120,
    "obs_op_avg_ms" : 14
  }
},
"reader_cache" : {
  "hit_count" : 0,
  "miss_count" : 1,
  "load_success_count" : 1,
  "load_exception_count" : 0,
  "total_load_time" : 367179751,
  "eviction_count" : 0
}
},
"EF8WoLCUQbqJl1Pkqo9-OA" : {
  "name" : "css-zzz1-ess-esn-1-1",
  "transport_address" : "10.0.0.18:9300",
  "host" : "10.0.0.18",
  "ip" : "10.0.0.18",
  "block_cache" : {
    "default" : {
      "type" : "memory",
      "block_cache_capacity" : 8192,
      "block_cache_blocksize" : 8192,
      "block_cache_size" : 12,
      "block_cache_hit" : 14,
      "block_cache_miss" : 0,
      "block_cache_eviction" : 0,
      "block_cache_store_fail" : 0
    }
  },
  "obs_stats" : {
    "list" : {
      "obs_list_count" : 17,
      "obs_list_ms" : 220,
      "obs_list_avg_ms" : 12
    },
    "get_meta" : {
      "obs_get_meta_count" : 79,
      "obs_get_meta_ms" : 139,
      "obs_get_meta_avg_ms" : 1
    }
  },

```

```

"get_obj" : {
  "obs_get_obj_count" : 12,
  "obs_get_obj_ms" : 82,
  "obs_get_obj_avg_ms" : 6
},
"put_obj" : {
  "obs_put_obj_count" : 12,
  "obs_put_obj_ms" : 879,
  "obs_put_obj_avg_ms" : 73
},
"obs_op_total" : {
  "obs_op_total_ms" : 1320,
  "obs_op_total_count" : 120,
  "obs_op_avg_ms" : 11
}
},
"reader_cache" : {
  "hit_count" : 0,
  "miss_count" : 1,
  "load_success_count" : 1,
  "load_exception_count" : 0,
  "total_load_time" : 235706838,
  "eviction_count" : 0
}
}
}
}

```

7. Run the following command to reset the cache status:

```
POST _frozen_stats/reset
```

Information similar to the following is displayed:

```

{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "Es-0325-007_01",
  "nodes" : {
    "mqTdk2YRSPyOSXfesREFSg" : {
      "result" : "ok"
    }
  }
}
}

```

 **NOTE**

This command is used to debug performance issues. If you reset the cache status and run this command, you can check the cache command status. You do not need to run this command during service running.

8. Run the following command to check all the frozen indexes:

```
GET _cat/freeze_indices?stage=${STAGE}
```

Table 6-17 Parameter description

Parameter	Description
STAGE	Its value can be: <ul style="list-style-type: none"> • start: List of indexes that are being frozen • done: List of indexes that have been frozen • unfreeze: List of indexes that are not frozen • Empty or other values: List of all indexes that are being frozen or have been frozen

Information similar to the following is displayed:

```
green open data2 0bNtxWDtRbOSkS4JYaUgMQ 3 0 5 0 7.9kb 7.9kb
green open data3 oYMLvw31QnyasqUNuyP6RA 3 0 51 0 23.5kb 23.5kb
```

 **NOTE**

The parameters and return values of this command are the same as those of `_cat/indices` of Elasticsearch.

6.2.3 Configuring Cache

After data is dumped to OBS, some data is cached to reduce access to OBS and improve Elasticsearch query performance. Data that is requested for the first time is obtained from OBS. The obtained data is cached in the memory. In subsequent queries, the system searches for data in the cache first. Data can be cached in memory or files.

Elasticsearch accesses different files in different modes. The cache system supports multi-level cache and uses blocks of different sizes to cache different files. For example, a large number of small blocks are used to cache `.fdx` and `.tip` files, and a small number of large blocks are used to cache `.fdt` files.

Table 6-18 Cache configurations

Parameter	Type	Description
<code>low_cost.obs.blockcache.names</code>	Array	The cache system supports multi-level cache for data of different access granularities. This configuration lists the names of all caches. If this parameter is not set, the system has a cache named default . To customize the configuration, ensure there is a cache named default . Default value: default

Parameter	Type	Description
low_cost.obs.blockcache.<NAME>.type	ENUM	Cache type, which can be memory or file . If it is set to memory , certain memory will be occupied. If it is set to file , cache will be stored in disks. You are advised to use ultra-high I/O disks to improve cache performance. Default value: memory
low_cost.obs.blockcache.<NAME>.blockshift	Integer	Size of each block in the cache. Its value is the number of bytes shifted left. For example, if this parameter is set to 16 , the block size is 2¹⁶ bytes, that is, 65536 bytes (64 KB). Default value: 13 (8 KB)
low_cost.obs.blockcache.<NAME>.bank.count	Integer	Number of cache partitions. Default value: 1
low_cost.obs.blockcache.<NAME>.number.blocks.perbank	Integer	Number of blocks included in each cache partition. Default value: 8192
low_cost.obs.blockcache.<NAME>.exclude.file.types	Array	Extensions of files that are not cached. If the extensions of certain files are neither in the exclude list nor in the include list, they are stored in the default cache.
low_cost.obs.blockcache.<NAME>.file.types	Array	Extensions of cached files. If the extensions of certain files are neither in the exclude list nor in the include list, they are stored in the default cache.

The following is a common cache configuration. It uses two levels of caches, **default** and **large**. The **default** cache uses 64 KB blocks and has a total of 30 x 4096 blocks. It is used to cache files except .fdt files. The **large** cache uses 2 MB blocks and contains 5 x 1000 blocks. It is used to cache .fdx, .dvd, and .tip files.

```
low_cost.obs.blockcache.names: ["default", "large"]
low_cost.obs.blockcache.default.type: file
low_cost.obs.blockcache.default.blockshift: 16
low_cost.obs.blockcache.default.number.blocks.perbank: 4096
low_cost.obs.blockcache.default.bank.count: 30
low_cost.obs.blockcache.default.exclude.file.types: [".fdt"]

low_cost.obs.blockcache.large.type: file
low_cost.obs.blockcache.large.blockshift: 21
low_cost.obs.blockcache.large.number.blocks.perbank: 1000
low_cost.obs.blockcache.large.bank.count: 5
low_cost.obs.blockcache.large.file.types: [".fdx", ".dvd", ".tip"]
```

Table 6-19 Other parameters

Parameter	Type	Description
index.frozen.obs.max_bytes_per_sec	String	Maximum rate of uploading files to OBS during freezing. It takes effect immediately after you complete configuration. Default value: 150MB
low_cost.obs.index.upload.threshold.use.multipart	String	If the file size exceeds the value of this parameter during freezing, the multipart upload function of OBS is used. Default value: 1GB
index.frozen.reader.cache.expire.duration.seconds	Integer	Timeout duration. To reduce the heap memory occupied by frozen indexes, the reader caches data for a period of time after the index shard is started, and stops caching after it times out. Default value: 300s
index.frozen.reader.cache.max.size	Integer	Maximum cache size. Default value: 100

6.2.4 Enhanced Cold Data Query Performance

Context

When you query data on the **Discover** page of Kibana for the first time, all data needs to be obtained from OBS because there is no cache. If a large number of documents are returned, it takes a long time to obtain the corresponding time fields and file metadata from OBS. To accelerate queries the first time they run on the **Discover** page, you can cache data locally.

Prerequisites

This feature is available in Elasticsearch clusters of versions 7.6.2 and 7.10.2 and OpenSearch clusters created after February 2023.

API for Querying Cold Data from Local Cache

This API can be used to query the cold data from local cache.

Example request:

```
GET /_frozen_stats/local_cache
GET /_frozen_stats/local_cache/{nodeId}
```

Response example:

```
{
  "_nodes" : {
```

```

    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "elasticsearch",
  "nodes" : {
    "6by3lPy1R3m55Dcq3liK8Q" : {
      "name" : "node-1",
      "transport_address" : "127.0.0.1:9300",
      "host" : "127.0.0.1",
      "ip" : "127.0.0.1",
      "local_cache" : {
        "get_stats" : {
          "get_total_count" : 562,           //Total number of times data was retrieved from the local
cold data cache.
          "get_hit_count" : 562,           //Total number of hits in the local cold data cache.
          "get_miss_count" : 0,           //Total number of local cold data cache misses.
          "get_total_ns" : 43849200,      //Total duration for retrieving data from the local cold
data cache.
          "get_avg_ns" : 78023            //Average duration for retrieving data from the local cold
data cache.
        },
        "load_stats" : {
          "load_count" : 2,               //Number of times cold data was loaded from the local
cache
          "load_total_ms" : 29,           //Total duration for loading cold data from the local cache
          "load_avg_ms" : 14,            //Average duration for loading cold data from the local
cache
          "load_fail_count" : 0,          //Number of failure times for loading cold data from the
local cache
          "load_overflow_count" : 0       //Number of times the local cold data cache exceeds
the cache pool size.
        },
        "reload_stats" : {
          "reload_count" : 0,            //Number of times the local cold data cache was
regenerated.
          "reload_total_ms" : 0,         //Total duration for regenerating the local cold data
cache.
          "reload_avg_ms" : 0,          //Average duration for regenerating the local cold data
cache.
          "reload_fail_count" : 0        //Number of failures in regenerating the local cold data
cache.
        },
        "init_stats" : {
          "init_count" : 0,              //Number of times the local cold data cache was initialized.
          "init_total_ms" : 0,           //Total duration for initializing the local cold data cache.
          "init_avg_ms" : 0,            //Average duration for initializing the local cold data
cache.
          "init_fail_count" : 0          //Number of failures in initializing the local cold data
cache.
        }
      }
    }
  }
}

```


Configuring Parameters

Configuration Item	Type	Unit	Value Range	Scope	Can Be Dynamically Modified	Description
low_cost.local_cache.max.capacity	Integer	-	The value ranges from 10 to 5000. The default value is 500 .	node	Yes	<p>Maximum number of available cold data caches on a node. Each shard corresponds to a cache object.</p> <p>NOTE</p> <ul style="list-style-type: none"> If the heap memory usage remains high, decrease the value. If the value of load_overflow_count keeps increasing rapidly, increase the value.
index.low_cost.local_cache.threshold	Integer	%	The value ranges from 0 to 100. The default value is 50 .	index	Yes	<p>Threshold for enabling the local cache of cold data.</p> <p>NOTE</p> <ul style="list-style-type: none"> If the percentage of date fields is less than the value of this parameter, the cold data of the date type will be cached locally. Otherwise, this parameter is not used. If the date fields of the current index occupy most of the data volume of the current index, you are not advised using this function.

Configuration Item	Type	Unit	Value Range	Scope	Can Be Dynamically Modified	Description
index.low_cost.local_cache.evict_time	String	Days	The value ranges from 1d to 365d. The default value is 30d .	index	Yes	<p>Wait time before cold data is deleted from local cache. The value is determined based on index.frozen_date (time when the freezing is successful).</p> <p>NOTE</p> <ul style="list-style-type: none"> For indexes that have been frozen in old clusters and do not have index.frozen_date specified, the value of this parameter is determined based on the index creation time. You are advised to adjust the deletion time based on the disk usage to avoid high disk usage.

Modifying Parameters

- Run the following command to modify **low_cost.local_cache.max.capacity**:

```
PUT _cluster/settings
{
  "persistent": {
    "low_cost.local_cache.max.capacity":1000
  }
}
```

- Run the following command to modify **index.low_cost.local_cache.threshold**:

```
PUT es_write_pref2-00000000021/_settings
{
  "index.low_cost.local_cache.threshold":20
}
```

- Run the following command to modify **index.low_cost.local_cache.evict_time**:

```
PUT es_write_pref2-00000000021/_settings
{
  "index.low_cost.local_cache.evict_time":"7d"
}
```

6.2.5 Monitoring OBS Operations

To clearly display the operations of the storage and compute decoupling plugin in OBS, the real-time OBS rate metric is added to CSS and recorded in the system index.

Prerequisite

This feature is available in Elasticsearch clusters of versions 7.6.2 and 7.10.2 and OpenSearch clusters created after March 2023.

Description

- The [GET _frozen_stats/obs_rate](#) API is used to query the real-time rate of OBS operations.
- The system index [.freeze_obs_rate-YYYY.mm.dd](#) is added to store the real-time OBS operation rate and OBS operation data, helping you monitor the OBS operations.
- The [low_cost.obs_rate_index.evict_time](#) parameter is added to control the storage duration of the [.freeze_obs_rate-YYYY.mm.dd](#) index

GET _frozen_stats/obs_rate API

- Calculation method: The average OBS operation rate in the last 5 seconds is calculated every 5 seconds.

- Example request:

```
GET _frozen_stats/obs_rate
GET _frozen_stats/obs_rate/{nodeId}
```

{nodeId} indicates the ID of the node whose OBS operation rate you want to query.

- Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "elasticsearch",
  "nodes" : {
    "df1DvcSwTJ-fkiILT2zE3A" : {
      "name" : "node-1",
      "transport_address" : "127.0.0.1:9300",
      "host" : "127.0.0.1",
      "ip" : "127.0.0.1",
      "update_time" : 1671777600482,           // Time when the current statistics are
      "obs_rate" : {                          updated.
        "list_op_rate" : 0.0,                 // Rate of OBS list operations. Unit: times/s.
        "get_meta_op_rate" : 0.0,            // Rate of OBS get meta operations. Unit: times/s.
        "get_obj_op_rate" : 0.0,            // Rate of OBS get operations. Unit: times/s.
        "put_op_rate" : 0.0,                // Rate of OBS put operations. Unit: times/s.
        "obs_total_op_rate" : 0.0,          // Rate of all OBS operations. The unit is times/s.
        "obs_upload_rate" : "0.0 MB/s",     // Data upload rate of OBS, in MB/s.
        "obs_download_rate" : "0.0 MB/s"   // Data download rate of OBS, in MB/s.
      }
    }
  }
}
```

System Index

- System index name: [.freeze_obs_rate-YYYY.mm.dd](#).
- Example: [.freeze_obs_rate-2023.01.23](#)

 NOTE

The default retention period of indexes is 30 days.

Configuration Item

Configuration Item	Type	Scope	Can Be Dynamically Modified	Description
low_cost.obs_rate_index.evict_time	String	node	Yes	The retention period of the .freeze_obs_rate-YYYY.mm.dd index. <ul style="list-style-type: none"> Value range: 1d to 365d Default value: 30d Unit: day

For example, run the following command to modify the retention period of the **.freeze_obs_rate-YYYY.mm.dd** index:

```
PUT _cluster/settings
{
  "persistent": {
    "low_cost.obs_rate_index.evict_time": "7d"
  }
}
```

6.3 Enhanced Import Performance

6.3.1 Context

Feature Description

CSS provides enhanced data import function. It optimizes bulk route, and speeds up processing through indexes and word segmentation, improving import performance and reduces bulk rejection. This function applies to clusters that contain a large number of index shards and text indexes, and have high import throughput.

Prerequisites

An Elasticsearch cluster of version 7.10.2 or OpenSearch cluster has been created on the CSS console.

Constraints

- Currently, only Elasticsearch clusters of version 7.10.2 and OpenSearch clusters of version 1.3.6 support the import performance enhancement.

- After the local shard preferential bulk routing optimization and bulk routing optimization are enabled, data writing is not routed based on IDs, and routing-related functions are restricted. For example, ID-based GET requests may fail. The optimization of local shard preferential bulk routing depends on the random distribution of client bulk requests and the balanced distribution of primary shards.
- The prerequisite for enabling **index.native_analyzer** is that the **index.native_speed_up** has been enabled.
- The **index.native_speed_up** function cannot be enabled for indexes that contain the **nested** field.

6.3.2 Instructions

6.3.2.1 Bulk Route Optimization

According to the default routing rule of Elasticsearch, data in a bulk request is routed to different shards. When massive data is written and a large number of index shards exist, excessive internal requests forwarding may trigger bulk rejection. In a large-scale cluster, the long tail effect causes a high bulk request latency.

You can specify the **index.bulk_routing** configuration item to enable bulk route optimization. This function reduces the requests that need to be internally forwarded. For clusters containing a large number of shards, this function can improve write performance and reduce bulk rejection.

Procedure

1. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
2. In the navigation tree on the left, choose **Dev Tools**.
3. On the **Dev Tools** page, run the following command:

```
PUT my_index
{
  "settings": {
    "index.bulk_routing": "local_pack"
  }
}
```

Table 6-20 Values of **index.bulk_routing**

Value	Description
default	The default routing mechanism of Elasticsearch is used. Records in a bulk request are split and routed independently.
pack	Data of a single bulk request is randomly routed to the same shard.

Value	Description
local_pac k	The data of a single bulk request is routed to the local shard of the data node that receives the bulk request. If the node does not contain the corresponding index shard, the data is randomly routed to another node that contains the index shard.

6.3.2.2 Bulk Aggregation Optimization

You can specify the **index.aggr_perf_batch_size** configuration item to enable or disable batch import optimization. After the batch import function is enabled, documents in bulk requests are written in batches. This function reduces the overhead of memory application, application lock, and other calls, improving data import performance.

NOTE

The value range of **index.aggr_perf_batch_size** is [1, Integer.MAX_VALUE]. The default value is **1**, indicating that the batch import function is disabled. If the value is greater than 1, the batch import function is enabled and the value of **MIN(bulk_doc_size, aggr_perf_batch_size)** indicates the batch size.

Procedure

1. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
2. In the navigation tree on the left, choose **Dev Tools**.
3. On the **Dev Tools** page, run the following command:

```
PUT my_index
{
  "settings": {
    "index.aggr_perf_batch_size": "128"
  }
}
```

6.3.2.3 Text Index Acceleration

- You can configure **index.native_speed_up** to enable or disable text index acceleration. This function optimizes the index process and memory usage to accelerate index building for text fields (text and keyword).
- You can configure **index.native_analyzer** to enable or disable word segmentation acceleration. For texts that require common word segmentation, you can use the analyzer to accelerate word segmentation.

Procedure

1. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
2. In the navigation tree on the left, choose **Dev Tools**.
3. On the **Dev Tools** page, run the following command:

```
PUT my_index
{
```

```
"settings": {
  "index.native_speed_up": true,
  "index.native_analyzer": true
},
"mappings": {
  "properties": {
    "my_field": {
      "type": "text"
    }
  }
}
}
```

6.3.2.4 Optimization of Other Parameters

After the import performance is enhanced, the number of index merge tasks increases accordingly. You can adjust the following configuration to reduce the impact of merge task overhead on the import performance:

You can increase the value of `index.merge.scheduler.max_thread_count` to increase the number of shard merge threads and reduce the traffic limit on data import. The default value is **4** and you are advised to set it to **8**.

Procedure

1. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
2. In the navigation tree on the left, choose **Dev Tools**.
3. On the **Dev Tools** page, run the following command:

```
PUT my_index
{
  "settings": {
    "index.merge.scheduler.max_thread_count": 8
  }
}
```

6.3.3 Performance Data

- Test environment
 - Cluster: 3 Huawei Cloud M6 ECSs (8 vCPUs | 64 GB memory)
 - Data: [open-source web server access logs](#) and internal service dataset (dns_logs)
 - Configuration: 120 shards, no replicas, and all the enhanced features enabled
- Test result

Type	Performance (Before)	Performance (After)	Improved By
Open-source dataset	85 Mbit/s	131 Mbit/s	54%
Service dataset	124 Mbit/s	218 Mbit/s	76%

6.4 Flow Control 2.0

6.4.1 Context

Feature Description

CSS can control traffic at the node level. You can configure the blacklist and whitelist, the maximum concurrent HTTP connections, and the maximum HTTP connections for a node. You can also configure backpressure based on client traffic in the node memory and block access in one click. CSS can also collect statistics on node access IP addresses and URIs. Each function has an independent control switch, which is disabled by default. To restore default values of parameters, set them to **null**.

After the client write traffic backpressure and control is enabled, large requests will be rejected when too much node heap memory has been occupied. This function prevents nodes from being suspended and reduces the risk of node unavailability.

- **HTTP/HTTPS flow control:**
 - You can control client IP address access by setting IP addresses and subnets in HTTP/HTTPS blacklist or whitelist. If an IP address is in the blacklist, the client is disconnected and all its request are rejected. Whitelist rules take precedence over blacklist rules. If a client IP address exists in both the blacklist and whitelist, the client request will not be rejected.
 - HTTP/HTTPS concurrent connection flow control limits the total number of HTTP connections to a node per second.
 - HTTP/HTTPS new connection flow control limits the number of new connections to a node.
- **Memory flow control**

Memory flow control limits the write traffic based on the node heap memory. You can back pressure requests to the client, trigger resource recycling as much as possible, and then accept requests based on the available heap memory.
- **Request sampling**

Request sampling can record the access of client IP addresses and the type of requests from the client. Based on the statistics, you can identify the access traffic of client IP addresses and analyze the client write and query requests.
- **One-click traffic blocking**

One-click access blocking can block all the access traffic of a node, excluding the traffic from Kibana and CSS O&M and monitoring APIs.
- **Flow control**

Flow control provides an independent API for viewing traffic statistics and records the number of current client connections and client backpressure connections. You can evaluate the flow control threshold and analyze the cluster loads based on the statistics.

- **Access logs**

Access logs record the URLs and bodies of HTTP/HTTPS requests received by nodes within a period of time. You can analyze the current traffic pressure based on the access logs.

Constraints

- Currently, only Elasticsearch clusters of versions 7.6.2 and 7.10.2 support the traffic control feature.
- Elasticsearch clusters of versions 7.6.2 and 7.10.2 created after February 2023 support only traffic control 2.0. Clusters created before February 2023 support only traffic control 1.0. For details, see [Flow Control 1.0](#).

6.4.2 HTTP/HTTPS Flow Control

You can run commands in Kibana to enable or disable HTTP/HTTPS flow control for your cluster.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable HTTP/HTTPS flow control.

- **Enabling HTTP/HTTPS flow control for a node**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": true,
    "flowcontrol.http.allow": ["192.168.0.1/24", "192.168.2.1/24"],
    "flowcontrol.http.deny": "192.168.1.1/24",
    "flowcontrol.http.concurrent": 1000,
    "flowcontrol.http.newconnect": 1000,
    "flowcontrol.http.warmup_period": 0
  }
}
```

 **NOTE**

If all parameters are set to **null**, they will be restored to default values.

- **Disabling HTTP/HTTPS flow control for a node**

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": false
  }
}
```

For details about the parameters in this command, see [Table 6-21](#).

Table 6-21 HTTP/HTTPS flow control parameters

Parameter	Type	Description
flowcontrol.http.enabled	Boolean	Whether to enable HTTP/HTTPS flow control. This function is disabled by default. Enabling it may affect node access performance. Value: true or false Default value: false
flowcontrol.http.allow	List<String> >	IP address whitelist. It can contain multiple IP addresses and masks, or an IP address list. Use commas (,) to separate multiple values. Example: <i>xx.xx.xx.xx/24,xx.xx.xx.xx/24</i> , or <i>xx.xx.xx.xx.xx,xx.xx.xx</i> . The default value is null.
flowcontrol.http.deny	List<String> >	IP address blacklist. Multiple IP addresses and masks or an IP address list can be configured. Use commas (,) to separate multiple IP addresses and masks. The default value is null.
flowcontrol.http.concurrent	Integer	Maximum concurrent HTTP/HTTPS connections. Default value: Number of available cores on a node x 400
flowcontrol.http.newconnect	Integer	Maximum new connections that can be created for HTTP/HTTPS requests per second. Default value: Number of available cores on a node x 200
flowcontrol.http.warmup_period	Integer	Time required for the HTTP/HTTPS connection setup speed to reach the maximum. If flowcontrol.http.newconnect is set to 100 and flowcontrol.http.warmup_period is set to 5000ms , it indicates the system can set up 100 connections per second in 5 seconds. Value range: 0-10000 Unit: ms Default value: 0

6.4.3 Memory Flow Control

Context

Elasticsearch provides a circuit breaker, which will terminate requests or return the error code **429** if the memory usage exceeds its threshold. However, the circuit breaker rejects a request only after the node reads the entire request, which occupies heap memory. To prevent a request from being fully received by a node before the request is rejected, you can control the client traffic based on the real-time status of the node heap memory.

Parameter Description

The following table describes memory flow control parameters.

Table 6-22 Memory flow control parameters

Parameter	Type	Description
flowcontrol.memory.enabled	Boolean	Whether to enable memory flow control. After this function is enabled, the memory usage is continuously monitored. The value can be: <ul style="list-style-type: none"> true false (default value)
flowcontrol.memory.heap_limit	String	Maximum global heap memory usage of a node. If the value of this parameter is exceeded, traffic backpressure is performed. Value range: 10%–100% Default value: 90%
flowcontrol.holding.inflight_factor	Float	Backpressure release factor. The principle is similar to that of the circuit breaker parameter network.breaker.inflight_requests.overhead . When the memory usage reaches the limit, a larger value indicates stronger backpressure. The write traffic will be limited. Value range: ≥ 0.5 Default value: 1.0

Parameter	Type	Description
flowcontrol.holding.max	TimeValue	<p>Maximum delay of each request. If the delay exceeds the value of this parameter, you can disconnect the request backpressure or disconnect the request link. For details, see the configuration of flowcontrol.holding.max_strategy.</p> <p>Value range: $\geq 15s$ Default value: 60s</p>
flowcontrol.holding.max_strategy	String	<p>Policy after the maximum delay time is exceeded. The value can be:</p> <ul style="list-style-type: none"> • keep (default value): If the heap memory is still high, continue the backpressure. The server determines when to execute the request based on the real-time memory. • soft: The requests will be executed even if the heap memory is still high. The inFlight circuit breaker will determine whether to execute or reject the requests. • hard: If the heap memory is still high, requests will be discarded and the client connection of the requests will be disconnected.
flowcontrol.memory.ounce_free_max	String	<p>Maximum memory that can be opened at a time for a suspended request queue. This parameter is used to prevent a cluster from being entirely suspended due to temporary low memory under high pressure.</p> <p>Value range: 1 to 50 Default value: 10%</p>
flowcontrol.memory.ounces_gc	Boolean	<p>Whether to trigger garbage collection to ensure write stability when the write pressure is too high. (The backpressure connection pool is checked every second. The write pressure is regarded high if all the existing connections are blocked and new write requests cannot be released.) The value can be:</p> <ul style="list-style-type: none"> • true (default value) • false

 NOTE

- **flowcontrol.memory.enabled** and **flowcontrol.memory.heap_limit** are the most important parameters. *enabled* indicates the memory flow control switch, and *heap_limit* indicates the heap memory threshold of a node.
- The default value 90% of **flowcontrol.memory.heap_limit** is a conservative threshold. When the heap memory usage is greater than 90%, the system stops reading large requests that exceed 64 KB from the client until the heap memory decreases. If the heap memory decreases to 85%, the maximum client data that can be read is 5% of the maximum heap memory. If the heap memory usage has been higher than 90% for a long time, client connection requests cannot be read. In this case, the GC algorithm is triggered to perform garbage collection until the heap memory usage is lower than the threshold.
- Generally, you can set the **flowcontrol.memory.heap_limit** threshold to 80% or less to ensure that the node has certain heap memory for operations besides data writing, such as Elasticsearch query and segment merge.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable memory flow control.

- Enable memory flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory.heap_limit": "80%"
  }
}
```

- Disable cluster memory flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": false
  }
}
```

6.4.4 Request Sampling

Context

Request sampling can record the access of client IP addresses and the type of requests from the client. Based on the statistics, you can identify the access traffic of client IP addresses and analyze the client write and query requests.

Table 6-23 Request statistics parameters

Parameter	Type	Description
flowcontrol.log.access.enabled	Boolean	Whether to collect statistics on the IP addresses of clients that accessed the ES cluster recently and the number of requests. The value can be: <ul style="list-style-type: none"> • true • false (default value)
flowcontrol.log.access.count	Integer	Number of client IP addresses that accessed a cluster recently. Value range: 0–100 Default value: 10

 **NOTE**

- IP address statistics switches control whether to collect request type statistics and whether to enable logging.
- **flowcontrol.log.access.enabled** controls whether to collect statistics on client requests, including the number of bulk write, search, and msearch requests.

6.4.5 One-click Traffic Blocking

You can block all connections in one click, except the connections that passes through O&M APIs, to handle unexpected traffic burst and quickly recover your cluster.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable one-click traffic blocking.

- Enable one-click traffic blocking

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.break.enabled": true
  }
}
```

- Disable one-click traffic blocking

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.break.enabled": false
  }
}
```

6.4.6 Access Statistics and Traffic Control Information Query

Flow control can be implemented via an independent API.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run the commands to query traffic control information.

- Check the traffic control status of all nodes.

```
GET /_nodes/stats/filter/v2
```

- View traffic control details of all nodes.

```
GET /_nodes/stats/filter/v2?detail
```

- View the traffic control status of a specific node.

```
GET /_nodes/{nodeId}/stats/filter/v2
```

{nodeId} indicates the ID of the node you want to check.

Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-xxxx",
  "nodes" : {
    "d3qnVlpPTtSoadkV0LQEkA" : {
      "name" : "css-xxxx-ess-esn-1-1",
      "host" : "192.168.x.x",
      "timestamp" : 1672236425112,
      "flow_control" : {
        "http" : {
          "current_connect" : 52,
          "rejected_concurrent" : 0,
          "rejected_rate" : 0,
          "rejected_black" : 0,
          "rejected_breaker" : 0
        }
      },
      "access_items" : [
        {
          "remote_address" : "10.0.0.x",
          "search_count" : 0,
          "bulk_count" : 0,
          "other_count" : 4
        }
      ],
      "holding_requests" : 0
    }
  }
}
```

Table 6-24 Response parameters

Parameter	Description
current_connect	Number of HTTP connections of a node, which is recorded even if flow control is disabled. This value is equal to the current_open value of GET /_nodes/stats/http API. It includes the current client connections of nodes.
rejected_concurrent	Number of concurrent connections rejected during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
rejected_rate	Number of new connections rejected during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
rejected_black	Number of requests rejected based on the blacklist during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
rejected_breaker	Number of rejected new connections after one-click traffic blocking is enabled.
remote_address	IP addresses and the number of requests.
search_count	Number of times that a client accessed a database using _search and _msearch .
bulk_count	Number of times that a client accessed a database using _bulk .
other_count	Number of times that a client accessed a database using other requests.

6.4.7 Temporary Access Statistics Logs

Context

You can check access logs in either of the following ways:

- Enable and check access logs via an independent API. Configure the API parameters to record the access log time and size. The access log content is returned through a REST API.
- Print access logs. Your access logs are printed as files in backend logs. This section describes how to temporarily access logs in this mode.

When the access log function is enabled or disabled, the parameters involved in the command are as follows:

Table 6-25 Access log parameters

Parameter	Type	Description
duration_limit	String	Duration recorded in an access log. Value range: 10 to 120 Unit: s Default value: 30
capacity_limit	String	Size of an access log. After access logging is enabled, the size of recorded requests is checked. If the size exceeds the value of this parameter, the access logging stops. Value range: 1 to 5 Unit: MB Default value: 1

 **NOTE**

Access logging stops if either **duration_limit** or **capacity_limit** reaches the threshold.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable access logs.
 - Enable access logs for all nodes in a cluster.
PUT /_access_log?duration_limit=30s&capacity_limit=1mb
 - Enable access logs for a node in a cluster.
PUT /_access_log/{nodeId}?duration_limit=30s&capacity_limit=1mb
{nodeId} indicates the ID of the node where you want to enable access logs.
4. View access logs.
 - Check the access logs of all nodes in a cluster.
GET /_access_log
 - Check the access logs of a node in a cluster.
GET /_access_log/{nodeId}

Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-flowcontroller",
  "nodes" : {
```

```
"8x-ZHu-wTemBQwpcGivFKg" : {
  "name" : "css-flowcontroller-ess-esn-1-1",
  "host" : "10.0.0.98",
  "count" : 2,
  "access" : [
    {
      "time" : "2021-02-23 02:09:50",
      "remote_address" : "/10.0.0.98:28191",
      "url" : "/_access/security/log?pretty",
      "method" : "GET",
      "content" : ""
    },
    {
      "time" : "2021-02-23 02:09:52",
      "remote_address" : "/10.0.0.98:28193",
      "url" : "/_access/security/log?pretty",
      "method" : "GET",
      "content" : ""
    }
  ]
}
}
```

Table 6-26 Response parameters

Parameter	Description
name	Node name
host	Node IP address
count	Number of node access requests in a statistical period
access	Details about node access requests in a statistical period. For details, see Table 6-27 .

Table 6-27 access

Parameter	Description
time	Request time
remote_address	Source IP address and port number of the request
url	Original URL of the request
method	Method corresponding to the request path
content	Request content

5. Run the following commands to delete access logs.

- Delete access logs of all nodes in a cluster.

```
DELETE /_access_log
```

- Delete access logs of a specified node in a cluster.

```
DELETE /_access_log/{nodeId}
```

{nodeId} indicates the ID of the node where you want to enable access logs.

6.4.8 Recording Access Logs in Files

The traffic control function can record cluster access logs and write the logs to background log files. You can back up the logs to OBS for viewing. You can run the following command to enable the function of recording access logs to files:

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.log.file.enabled": true
  }
}
```

Table 6-28 Parameters

Parameter	Type	Description
flowcontrol.log.file.enabled	Boolean	Indicates whether to record the log details of each request to the background log file. The value can be: <ul style="list-style-type: none"> true false (default value)

 **NOTE**

- After the function of recording access logs to files is enabled, access from a client to a cluster node is recorded in the **{Cluster name}_access_log.log** file. You can use the log backup function to view detailed access logs.
- After the fault is located, you are advised to disable this function.

6.5 Flow Control 1.0

6.5.1 Context

Feature Description

CSS can control traffic at the node level. You can configure the blacklist and whitelist, the maximum concurrent HTTP connections, and the maximum HTTP connections for a node. You can also configure the maximum heap memory used by specific request paths, the maximum CPU usage, and block access in one click, and collect statistics on node access IP addresses and URIs. Each function has an independent control switch, which is disabled by default. To restore default values of parameters, set them to **null**.

If flow control is enabled, requests will be blocked at the entry, which relieves the cluster pressure in high-concurrency scenario and avoids unavailability issues.

- **HTTP/HTTPS Flow Control**
 - You can control client IP address access by setting IP addresses and subnets in HTTP/HTTPS blacklist or whitelist. If an IP address is in the blacklist, the client is disconnected and all its request are rejected. Whitelist rules take precedence over blacklist rules. If a client IP address

- exists in both the blacklist and whitelist, the client request will not be rejected.
- HTTP/HTTPS concurrent connection flow control limits the total number of HTTP connections to a node per second.
 - HTTP/HTTPS new connection flow control limits the number of new connections to a node.
- **Memory Flow Control**

Memory flow control limits request paths based on the node heap memory. You can configure memory flow control whitelist, global memory usage threshold, and heap memory threshold for a single path. Global memory flow control threshold takes precedence over the memory threshold of a single path. Paths in the whitelist will not be blocked in memory flow control.
 - **Global Path Whitelist for Flow Control**

You can configure the global path whitelist for flow control as required when you need to use custom plug-ins.
 - **Request Sampling**

Request sampling can record the number of access requests from client IP addresses and the request paths of sampled users. Based on the statistics, you can identify the access traffic of client IP addresses and analyze the access traffic of request paths.
 - **Flow Control**

Flow control provides an independent API for viewing traffic statistics and records the number of times the API is triggered. You can evaluate the flow control threshold and analyze the cluster load based on the statistics.
 - **Access Logs**

Access logs record the URLs and bodies of HTTP/HTTPS requests received by nodes within a period of time. You can analyze the current traffic pressure based on the access logs.
 - **CPU Flow Control**

You can configure the node CPU usage threshold to limit the accessed traffic on a single node.
 - **One-click Traffic Blocking**

One-click access blocking can block all the access traffic of a node, excluding the traffic from Kibana and Elasticsearch monitor APIs.

Constraints

- Currently, only versions 7.6.2 and 7.10.2 support the flow control feature.
- Flow control may affect the performance of some nodes.
- If flow control is enabled, user requests that exceed the flow control threshold will be rejected.
- Memory flow control and CPU flow control are based on request paths. The length and number of paths cannot be too large, or the cluster performance will be affected.

6.5.2 HTTP/HTTPS Flow Control

Context

You can run commands in Kibana to enable or disable HTTP/HTTPS flow control for your cluster. The command parameters are as follows.

Table 6-29 HTTP/HTTPS flow control parameters

Parameter	Type	Description
flowcontrol.http.enabled	Boolean	Whether to enable HTTP/HTTPS flow control. This function is disabled by default. Enabling it may affect node access performance. Value: true or false Default value: false
flowcontrol.http.allow	List<String>	IP address whitelist. It can contain multiple IP addresses and masks, or an IP address list. Use commas (,) to separate multiple values. Example: <i>xx.xx.xx.xx/24,xx.xx.xx.xx/24</i> , or <i>xx.xx.xx.xx.xx,xx.xx.xx</i> . The default value is null.
flowcontrol.http.deny	List<String>	IP address blacklist. Multiple IP addresses and masks or an IP address list can be configured. Use commas (,) to separate multiple IP addresses and masks. The default value is null.
flowcontrol.http.concurrent	Integer	Maximum concurrent HTTP/HTTPS connections. Default value: Number of available cores on a node x 400
flowcontrol.http.newconnect	Integer	Maximum new connections that can be created for HTTP/HTTPS requests per second. Default value: Number of available cores on a node x 200

Parameter	Type	Description
flowcontrol.http.warmup_period	Integer	Time required for the HTTP/HTTPS connection setup speed to reach the maximum. If flowcontrol.http.newconnect is set to 100 and flowcontrol.http.warmup_period is set to 5000ms , it indicates the system can set up 100 connections per second in 5 seconds. Value range: 0-10000 Unit: ms Default value: 0

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable HTTP/HTTPS flow control.

- Enabling HTTP/HTTPS flow control for a node

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": true,
    "flowcontrol.http.allow": ["192.168.0.1/24", "192.168.2.1/24"],
    "flowcontrol.http.deny": "192.168.1.1/24",
    "flowcontrol.http.concurrent": 1000,
    "flowcontrol.http.newconnect": 1000,
    "flowcontrol.http.warmup_period": 0
  }
}
```

NOTE

If all parameters are set to **null**, they will be restored to default values.

- Disabling HTTP/HTTPS flow control for a node

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.http.enabled": false
  }
}
```

6.5.3 Memory Flow Control

Context

Elasticsearch provides a circuit breaker, which will terminate requests if the memory usage exceeds its threshold. However, Elasticsearch does not check the heap memory usage when an API is called, and does not allow users to configure the threshold for a single request. In this case, memory usage can only be

calculated during request processing, which may lead to frequent circuit breaking and cannot avoid heap memory waste. To solve this problem, CSS checks the heap memory usage when receiving REST requests, blocking excess API requests and protecting nodes. You can configure global memory flow control, or configure the request path and heap memory threshold for a specific request path. Before a request is processed, the system checks the configured heap memory threshold. If the threshold is exceeded, the request path will be blocked.

 **NOTE**

- Memory flow control may affect request processing performance.
- If the memory flow control is enabled, some Kibana search requests may fail.
- If memory flow control is enabled in Elasticsearch 5.5.1, `_mget` requests will be blocked and Kibana access will be abnormal. You can add `_mget` requests to the request whitelist to avoid this problem.

The following table describes memory flow control parameters.

Table 6-30 Memory flow control parameters

Parameter	Type	Description
<code>flowcontrol.memory.enabled</code>	Boolean	Whether to enable memory flow control. This function is disabled by default. Enabling memory flow control may slightly affect node performance. Value: true or false Default value: false

Parameter	Type	Description
flowcontrol.memory.allow_path	List<String>	<p>Request path whitelist for memory flow control. Whitelisted paths are blocked in memory flow control. Wildcard characters are supported. By default, query APIs controlled by the cluster are not blocked in memory flow control. This prevents the failure to query cluster information when the memory usage reaches the threshold.</p> <p>Example:</p> <ul style="list-style-type: none"> • "flowcontrol.memory.allow_path": "/index/_search", • "flowcontrol.memory.allow_path": "/index*/_search", • "flowcontrol.memory.allow_path": ["/index/_search", "/index1/_bulk"], <p>A maximum of 10 paths can be configured. A path can contain up to 32 characters.</p> <p>The default value is null.</p>
flowcontrol.memory.heap_limit	String	<p>Maximum global heap memory usage of a node. The value cannot be less than 10% of the heap memory.</p> <p>Value range: 10%–100%</p> <p>Default value: 90%</p>

Parameter	Type	Description
flowcontrol.memory.*.filter_path	String	<p>Paths under memory flow control.</p> <p>The default value is **, indicating all paths. If flowcontrol.memory.heap_limit is configured and flowcontrol.memory.*.filter_path is not, it indicates that all the paths, except those in the whitelist, are under control. The whitelist takes precedence over the single-path rule. If a path is specified in both flowcontrol.memory.allow_path and flowcontrol.memory.*.filter_path, the requests from the path will be allowed.</p> <p>For example, if flowcontrol.memory.allow_path and flowcontrol.memory.*.filter_path are both set to abc/_search, then abc/_search will not be under flow control.</p> <p>Maximum length: 32 characters</p>
flowcontrol.memory.*.heap_limit	String	<p>Heap memory usage threshold of request paths. If the heap memory usage exceeds the threshold, flow control will be triggered.</p> <p>Value range: 0-100%</p> <p>Default value: 90%</p>

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.

3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable memory flow control.

- Enabling memory flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory.allow_path": "/index/_search",
    "flowcontrol.memory.heap_limit": "85%"
  }
}
```

- Enabling memory flow control for a request path

Configure the heap memory usage threshold for a request path. You can configure the priorities of such threshold rules.

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory": {
      "flowcontrol_search": {
        "filter_path": "index1/_search",
        "heap_limit": "50%"
      },
      "flowcontrol_bulk": {
        "filter_path": "index*/_bulk",
        "heap_limit": "50%"
      }
    }
  }
}
```

- Deleting the memory flow control configuration of a request path

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": true,
    "flowcontrol.memory": {
      "flowcontrol_search": {
        "filter_path": null,
        "heap_limit": null
      }
    }
  }
}
```

- Disabling cluster memory flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.memory.enabled": false
  }
}
```

6.5.4 Global Path Whitelist for Flow Control

Context

The following table describes the global path whitelist parameters for flow control.

Table 6-31 Global path whitelist parameters for flow control

Parameter	Type	Description
flowcontrol.path.white_list	List<String>	<p>Paths that are not under flow control. These paths are not affected by memory flow control, CPU flow control, or one-click blocking; but are under IP address-based flow control.</p> <p>A maximum of 10 paths can be configured. A path can contain up to 32 characters.</p> <p>This parameter is left blank by default.</p> <p>NOTE You are advised not to configure this parameter, unless required by plug-ins.</p>

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation tree on the left, choose **Dev Tools**. Run the following command to configure the global path whitelist for flow control:

```
PUT _cluster/settings
{
  "persistent": {
    "flowcontrol.path.white_list": "xxxx"
  }
}
```

6.5.5 Request Sampling

Context

Request sampling can record the access IP addresses, the number of accessed nodes, request paths, request URLs, and request bodies, which can be used to obtain the IP addresses and paths of clients that have sent a large number of access requests.

The following table describes request sampling parameters.

Table 6-32 Request sampling parameters

Parameter	Type	Description
flowcontrol.statics.enabled	Boolean	<p>Whether to enable request sampling. Request sampling may affect node performance.</p> <p>Value: true or false</p> <p>Default value: false</p>

Parameter	Type	Description
flowcontrol.statics.threshold	Integer	Number of recent access requests whose statistics are collected. The value 100 indicates that statistics will be collected on the 100 IP addresses and 100 URLs that are most frequently accessed. Minimum value: 10 Maximum value: 1000 Default value: 100
flowcontrol.statics.sample_frequency	Integer	Path sampling frequency. If this parameter is set to 100 , samples are collected from every 100 requests. Minimum value: 50 Default value: 100

 **NOTE**

- The IP address statistics and URL sampling statistics are cached based on their access time. If the cache space reaches the threshold (**flowcontrol.statics.threshold**), the records of the earliest access will be deleted.
- In URL sampling, an access path is uniquely identified by its URL hash.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable sampling.

– Enabling sampling

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.statics.enabled": true,
    "flowcontrol.statics.threshold": 100,
    "flowcontrol.statics.sample_frequency": 50
  }
}
```

– Disabling sampling

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.statics.enabled": false
  }
}
```

6.5.6 Flow Control

Flow control can be implemented via an independent API.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run the commands to query traffic control information.
 - Check the traffic control status of all nodes.
GET `/_nodes/stats/filter`
 - View the traffic control status of a specific node.
GET `/_nodes/{nodeId}/stats/filter`

{nodeId} indicates the ID of the node you want to check.

Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-flowcontroller",
  "nodes" : {
    "ELBRNCMbTj6L1C-Wke-Dnw" : {
      "name" : "css-flowcontroller-ess-esn-1-1",
      "host" : "10.0.0.133",
      "timestamp" : 1613979513747,
      "flow_control" : {
        "transport" : {
          "concurrent_req" : 0,
          "rejected_concurrent" : 0,
          "rejected_new" : 0,
          "rejected_deny" : 0
        },
        "http" : {
          "concurrent_req" : 0,
          "rejected_concurrent" : 0,
          "rejected_new" : 0,
          "rejected_deny" : 0
        },
        "memory" : {
          "memory_allow" : 41,
          "memory_rejected" : 0
        },
        "cpu" : {
          "rejected_cpu" : 0
        }
      },
      "ip_address" : [
        {
          "ip" : "/10.0.0.198",
          "count" : 453
        },
        {
          "ip" : "/198.19.49.1",
          "count" : 42
        }
      ],
      "url_sample" : [
        {
          "url" : "/*/_search?pretty=true",
          "method" : "GET",
          "remote_address" : "/10.0.0.198:16763",
          "count" : 1
        }
      ]
    }
  }
}
```

In the response, the information of each node is separated. The **http** field records the numbers of concurrent connections and new connections. The **memory** records memory flow control statistics. The **ip_address** field records the recent client IP addresses that are accessed most recently. The **url_sample** field records the recent URLs that are requested most frequently. The **cpu** field records CPU flow control statistics.

Table 6-33 Response parameters

Parameter	Description
concurrent_req	Number of TCP connections of a node, which is recorded no matter whether flow control is enabled. This value is similar to the value of current_open of the GET /_nodes/stats/http API but is smaller, because whitelisted IP addresses and internal node IP addresses are not counted.
rejected_concurrent	Number of concurrent connections rejected during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
rejected_new	Number of new connections rejected during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
rejected_deny	Number of requests rejected based on the blacklist during HTTP flow control. This value is not cleared when HTTP flow control is disabled.
memory_allow	Number of allowed requests during memory flow control. This parameter takes effect when memory flow control is enabled, and its value is not cleared after memory flow control is disabled. The requests from the paths in the allow_path whitelist are not recorded. If allow_path is set to **, no requests are recorded.
memory_rejected	Number of rejected requests during memory flow control. This parameter takes effect when memory flow control is enabled, and its value is not cleared after memory flow control is disabled. The requests from the paths in the allow_path whitelist are not recorded. If allow_path is set to **, no requests are recorded.
rejected_cpu	Number of requests rejected when the CPU flow control threshold is exceeded. This parameter takes effect when CPU flow control is enabled, and its value is not cleared after CPU flow control is disabled.
ip_address	IP addresses and the number of requests. For details, see Table 6-34 .

Parameter	Description
url_sample	Request path sampling. The number of URLs of a request are collected based on the configured time and sampling interval. For details, see Table 6-35 .

Table 6-34 ip_address

Parameter	Description
ip	Source IP address for accessing the node.
method	Number of access requests from an IP address.

Table 6-35 url_sample

Parameter	Description
url	Request URL
method	Method corresponding to the request path
remote_address	Source IP address and port number of the request
count	How many times a path is sampled

6.5.7 Access Logs

Context

You can check access logs in either of the following ways:

- Enable and check access logs via an independent API. Configure the API parameters to record the access log time and size. The access log content is returned through a REST API.
- Print access logs. Your access logs are printed as files in backend logs.

Enabling the access log function may affect cluster performance.

The following table describes access log parameters.

Table 6-36 Access log parameters

Parameter	Type	Description
duration_limit	String	Duration recorded in an access log. Value range: 10 to 120 Unit: s Default value: 30

Parameter	Type	Description
capacity_limit	String	Size of an access log. After access logging is enabled, the size of recorded requests is checked. If the size exceeds the value of this parameter, the access logging stops. Value range: 1 to 5 Unit: MB Default value: 1

 **NOTE**

Access logging stops if either **duration_limit** or **capacity_limit** reaches the threshold.

Procedure

- Log in to the CSS management console.
- Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
- In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable access logs.
 - Enabling access logs for all nodes in a cluster
PUT /_access_log?duration_limit=30s&capacity_limit=1mb
 - Enabling access logs for a node in a cluster
PUT /_access_log/{nodeId}?duration_limit=30s&capacity_limit=1mb
{nodeId} indicates the ID of the node where you want to enable access logs.
- Use APIs to check access logs.
 - API for checking the access logs of all nodes in a cluster
GET /_access_log
 - API for checking the access logs of a node in a cluster
GET /_access_log/{nodeId}
{nodeId} indicates the ID of the node where you want to enable access logs.

Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "css-flowcontroller",
  "nodes" : {
    "8x-ZHu-wTemBQwpcGivFKg" : {
      "name" : "css-flowcontroller-ess-esn-1-1",
      "host" : "10.0.0.98",
      "count" : 2,
      "access" : [
        {
          "time" : "2021-02-23 02:09:50",
          "remote_address" : "/10.0.0.98:28191",
          "url" : "/_access/security/log?pretty",
```



```
"persistent": {
  "flowcontrol.accesslog.enabled": false
}
```

6.5.8 CPU Flow Control

Context

CPU flow control can be implemented based on the CPU usage of a node.

You can configure the CPU usage threshold of a node to prevent the node from breaking down due to heavy traffic. You can determine the CPU usage threshold based on the traffic threshold. If the CPU usage of a node exceeds the configured threshold, CPU flow control discards excess node requests to protect the cluster. Traffic within the node or passing through Elasticsearch monitoring APIs are not affected.

The following table describes CPU flow control parameters.

Table 6-39 CPU flow control parameters

Parameter	Type	Description
flowcontrol.cpu.enabled	Boolean	Whether to enable CPU flow control. If this function is enabled, the node access performance may be affected. Value: true or false Default value: false
flowcontrol.cpu.percent_limit	Integer	Maximum CPU usage of a node. Value range: 0-100 Default value: 90
flowcontrol.cpu.allow_path	List	Path whitelist for CPU flow control. The paths specified in the allow_path whitelist are not under CPU flow control. The default value is null. A path can contain up to 32 characters. A maximum of 10 request paths can be configured. Wildcard characters are supported. For example, if this parameter is set to auto_*/_search , all the search requests of the indexes prefixed with auto_ are not under the flow control.

Parameter	Type	Description
flowcontrol.cpu.*.filter_path	String	<p>Paths under CPU flow control. Maximum length: 32 characters Example: "flowcontrol.cpu.search.filter_path": "/index/_search", "flowcontrol.cpu.search.limit": 60, The default value is **, indicating all paths. If limit is configured and filter_path is not, it indicates that all the paths, except those in the whitelist, are under control. The whitelist takes precedence over the single-path rule. If a path is specified in both allow_path and filter_path, the requests from the path will be allowed. For example, if both filter_path and allow_path both set to abc/_search, then abc/_search will not be under flow control.</p>
flowcontrol.cpu.*.limit	Integer	<p>CPU threshold of request paths. If the CPU usage exceeds the threshold, flow control will be triggered. Value range: 0-100 Default value: 90</p>

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable memory flow control.

- Enabling CPU flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.cpu.enabled": true,
    "flowcontrol.cpu.percent_limit": 80,
    "flowcontrol.cpu.allow_path": ["index/_search"]
  }
}
```

- Disabling CPU flow control

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.cpu.enabled": false
  }
}
```

6.5.9 One-click Traffic Blocking

You can block all traffic in one click, except the traffic that passes through O&M APIs, to handle unexpected traffic burst and quickly recover your cluster.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation pane on the left, choose **Dev Tools** and run commands to enable or disable one-click traffic blocking.

- Enabling one-click traffic blocking

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.break.enabled": true
  }
}
```

- Disabling one-click traffic blocking

```
PUT /_cluster/settings
{
  "persistent": {
    "flowcontrol.break.enabled": false
  }
}
```

6.6 Large Query Isolation

6.6.1 Context

The large query isolation feature allows you to separately manage large queries. You can isolate query requests that consume a large amount of memory or take a long period of time. If the heap memory usage of a node is too high, the interrupt control program will be triggered. The program will interrupt a large query based on the policies you configured and cancel the running query tasks of the query.

You can also configure a global query timeout duration. Long queries will be intercepted.

 **NOTE**

Currently, only versions 7.6.2 and 7.10.2 support large query isolation.

6.6.2 Procedure

The large query isolation and global timeout features are disabled by default. If you enable them, the configuration will take effect immediately. Perform the following steps to configure the features:

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
3. In the navigation pane of Kibana on the left, choose **Dev Tools**. Run the following command to enable large query isolation and global timeout features:

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.enabled": true,
    "search.isolator.time.enabled": true
  }
}
```

The two features each has an independent switch and the following parameters.

Table 6-40 Parameters for large query isolation and global timeout duration

Switch	Parameter	Description
search.isolator.enabled	search.isolator.memory.task.limit search.isolator.time.management	Thresholds of a shard query task. A query task exceeding one of these thresholds is regarded as a large query task.
	search.isolator.memory.pool.limit search.isolator.memory.heap.limit search.isolator.count.limit	Resource usage thresholds in the isolation pool. If the resource usage of a query task exceeds one of these thresholds, the task will be intercepted. NOTE search.isolator.memory.heap.limit defines the limit on the heap memory consumed by write, query, and other operations of a node. If the limit is exceeded, large query tasks in the isolation pool will be interrupted.
	search.isolator.strategy search.isolator.strategy.ratio	Policy for selecting a query task in the isolation pool.
search.isolator.time.enabled	search.isolator.time.limit	Global timeout interval of query tasks.

4. Configure the large query isolation and global timeout duration separately.
 - Configure the thresholds of a shard query task. A query task exceeding one of these thresholds is regarded as a large query task.

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.memory.task.limit": "50MB",
    "search.isolator.time.management": "10s"
  }
}
```

Table 6-41 Parameter description

Parameter	Data Type	Description
search.isolator.memory.task.limit	String	<p>Threshold of the memory requested by a query task to perform aggregation or other operations. If the requested memory exceeds the threshold, the task will be isolated and observed.</p> <p>Value range: 0b to the maximum heap memory of a node</p> <p>Default value: 50MB</p> <p>NOTE You can run the following command to query the current heap memory and the maximum heap memory of a cluster:</p> <p>GET _cat/nodes?&h=id,ip,port,r,ramPercent,ramCurrent,heapMax,heapCurrent</p>
search.isolator.time.management	String	<p>Threshold of the duration of a query. (started when cluster resources are used for query). If the duration of a query exceeds the threshold, it will be isolated and observed.</p> <p>Value range: \geq 0ms</p> <p>Default value: 10s</p>

- Configure the resource usage thresholds in the isolation pool. If the resource usage of a query task exceeds one of these thresholds, the task will be intercepted.

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.memory.pool.limit": "50%",
    "search.isolator.memory.heap.limit": "90%",
    "search.isolator.count.limit": 1000
  }
}
```

Table 6-42 Parameter description

Parameter	Data Type	Description
search.isolator.memory.pool.limit	String	Threshold of the heap memory percentage of the current node. If the total memory requested by large query tasks in the isolation pool exceeds the threshold, the interrupt control program will be triggered to cancel one of the tasks. Value range: 0.0 to 100.0% Default value: 50%
search.isolator.memory.heap.limit	String	Heap memory threshold of the current node. If the heap memory of the node exceeds the threshold, the interrupt control program will be triggered to cancel a large query task in the isolation pool. Value range: 0.0 to 100.0% Default value: 90%
search.isolator.count.limit	Integer	Threshold of the number of large query tasks in the current node isolation pool. If the number of observed query tasks exceeds the threshold, the interrupt control program will be triggered to stop accepting new large queries. New large query requests will be directly canceled. Value range: 10–50000 Default value: 1000

 **NOTE**

In addition to **search.isolator.memory.pool.limit** and **search.isolator.count.limit** parameters, you can configure **search.isolator.memory.task.limit** and **search.isolator.time.management** to control the number of query tasks that enter the isolation pool.

- Policy for selecting a query task in the isolation pool.

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.strategy": "fair",
    "search.isolator.strategy.ratio": "0.5%"
  }
}
```

Parameter	Data Type	Description
search.isolator.strategy	String	<p>Policy for selecting large queries when the interrupt control program is triggered. The selected query will be interrupted.</p> <p>NOTE The large query isolation pool is checked every second until the heap memory is within the safe range.</p> <p>Values: fair, mem-first, or time-first</p> <ul style="list-style-type: none"> • mem-first: The query task that uses the most heap memory in the isolation pool is interrupted. • time-first: The query task that has been running for the longest time in the isolation pool is interrupted. • fair: If the difference between the heap memory of shard queries is smaller than <i>Maximum_heap_memory</i> x search.isolator.strategy.ratio, the query that takes the longest time should be interrupted. Otherwise, the query that uses the most heap memory is interrupted. <p>Default value: fair</p>
search.isolator.strategy.ratio	String	<p>Threshold of the fair policy. This parameter takes effect only if search.isolator.strategy is set to fair. If the difference between the memory usage of large query tasks does not exceed the threshold, the query that takes the longest time should be interrupted. If the difference between the memory usage of large query tasks exceeds the threshold, the query that uses the most memory is interrupted.</p> <p>Value range: 0.0 to 100.0% Default value: 1%</p>

- Configure the global timeout duration of query tasks.

```
PUT _cluster/settings
{
  "persistent": {
    "search.isolator.time.limit": "120s"
  }
}
```


Parameter	Data Type	Description
search.isolator.time.limit	String	Global query timeout duration. If this function is enabled, all the query tasks that exceed the specified duration will be canceled. Value range: $\geq 0\text{ms}$ Default value: 120s

6.7 Index Monitoring

6.7.1 Context

CSS monitors various metrics of the running status and change trend of cluster indexes to measure service usage and handle potential risks in a timely manner, ensuring that clusters can run stably.

During index monitoring, the **stats** information about indexes is collected and saved to the monitoring index (**monitoring-eye-css-[yyyy-mm-dd]**) of the cluster, and retained for one week by default.

Currently, only the Elasticsearch clusters of the versions 7.6.2 and 7.10.2 support the index monitoring.

6.7.2 Enabling Index Monitoring

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. Choose **Dev Tools** in the navigation pane on the left and run the following command to enable index monitoring:

```
PUT _cluster/settings
{
  "persistent": {
    "css.monitoring.index.enabled": "true"
  }
}
```

4. (Optional) To monitor a specific index, run the following command on the **Dev Tools** page of Kibana:

```
PUT _cluster/settings
{
  "persistent": {
    "css.monitoring.index.enabled": "true",
    "css.monitoring.index.interval": "30s",
    "css.monitoring.index.indices": ["index_name"],
    "css.monitoring.history.duration": "3d"
  }
}
```

Table 6-43 Parameter description

Parameter	Data Type	Description
css.monitoring.index.enabled	Boolean	Whether to enable index monitoring. If this parameter is set to true , the monitoring will be enabled. Default value: false
css.monitoring.index.interval	Time	Interval for collecting index monitoring data. Minimum value: 1s Default value: 10s
css.monitoring.index.indices	String	Name of an index to be monitored. By default, all indexes are monitored. You can configure specific indexes or a type of indexes to monitor. Example: <ul style="list-style-type: none"> "css.monitoring.index.indices": ["index_name"] indicates only <i>index_name</i> is monitored. "css.monitoring.index.indices": ["log_*"] indicates that only indexes starting with log_ are monitored. "css.monitoring.index.indices": ["index1", "index2"] indicates that index1 and index2 are monitored. Default value: * (indicating that all indexes are monitored)
css.monitoring.history.duration	Time	Retention period of monitoring data storage. The default period is a week. Minimum value: 1d Default value: 7d

NOTICE

Indexes starting with **monitoring-eye-css-*** are regarded as monitoring indexes and will not be monitored.

6.7.3 Checking the Index Read and Write Traffic

You can call an API to query the index read and write traffic within a period of time.

Prerequisites

A cluster has been created and **index monitoring** has been enabled.

Procedure

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
3. Choose **Dev Tools** in the navigation pane on the left and run the following commands to query the index read and write traffic:
 - Check read and write traffic of all the indexes.
GET /_cat/monitoring
 - Check read and write traffic of a specific index.
GET /_cat/monitoring/{indexName}

{indexName} indicates the name of the index whose read and write traffic you want to check.
 - Check the read and write traffic of indexes for different periods.
GET _cat/monitoring?begin=1650099461000
GET _cat/monitoring?begin=2022-04-16T08:57:41
GET _cat/monitoring?begin=2022-04-16T08:57:41&end=2022-04-17T08:57:41

Table 6-44 Parameter description

Parameter	Mandatory	Description
begin	No	Start time (UTC time) of the monitoring data you want to view. Time format: strict_date_optional_time epoch_millis The default start time is five minutes before the current time.
end	No	End time (UTC time) of the monitoring data you want to view. Time format: strict_date_optional_time epoch_millis The default end time is the current time.

NOTE

These parameters cannot be used for system indexes, whose names start with a dot (.).

Information similar to the following is displayed:

```
index  begin          end          status pri rep init unassign docs.count docs.deleted store.size
pri.store.size delete.rate indexing.rate search.rate
test 2022-03-25T09:46:53.765Z 2022-03-25T09:51:43.767Z yellow 1 1 0 1 9 0
5.9kb 5.9kb 0/s 0/s 0/s
```

Table 6-45 Parameters in the returned information

Parameter	Description
index	Index name

Parameter	Description
begin	Start time of the monitoring data you queried.
end	End time of the monitoring data you queried.
status	Index status within the queried monitoring interval.
pri	The number of index shards within the queried monitoring interval.
rep	The number of index replicas within the queried monitoring interval.
init	The number of initialized indexes within the queried monitoring interval.
unassign	The number of unallocated indexes within the queried monitoring interval.
docs.count	The number of documents within the queried monitoring interval.
docs.deleted	The number of deleted documents within the queried monitoring interval.
store.size	Index storage size within the queried monitoring interval.
pri.store.size	Size of the primary index shard within the queried monitoring interval.
delete.rate	Number of indexes deleted per second within the queried monitoring interval.
indexing.rate	Number of indexes wrote per second within the queried monitoring interval.
search.rate	Number of indexes queried per second within the queried monitoring interval.

6.7.4 Checking Index Monitoring Information

You can check preconfigured index monitoring visualizations on the **Dashboard** and **Visualizations** pages of Kibana. You can also customize tables and charts.

Prerequisites

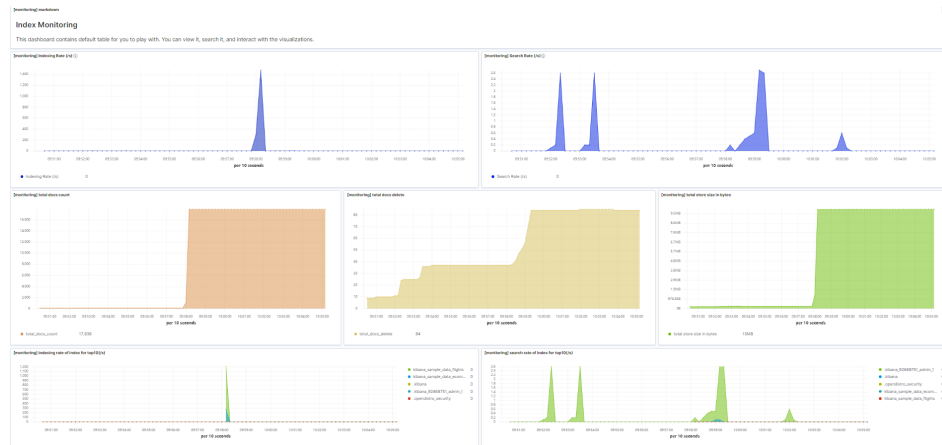
A cluster has been created and **index monitoring** has been enabled.

Checking Dashboard Charts

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
3. In the navigation tree on the left, click **Dashboard**.

4. Click **[Monitoring] Index Monitoring Dashboard** to view the preconfigured dashboard.

Figure 6-1 Preconfigured dashboard charts



The preconfigured dashboard displays the number of read and write operations per second in the cluster and the top 10 indexes with the most read and write operations per second.

Table 6-46 Preconfigured charts

Chart Name	Description
[monitoring] markdown	Markdown chart, which briefly describes the dashboard content.
[monitoring] Indexing Rate (/s)	Number of documents written to a cluster per second.
[monitoring] Search Rate (/s)	Average number of queries per second in a cluster.
[monitoring] indexing rate of index for top10	Top 10 indexes with the most documents written per second.
[monitoring] search rate of index for top10	Top 10 indexes with the most queries per second.
[monitoring] total docs count	Total number of documents in a cluster.
[monitoring] total docs delete	Total number of deleted documents in a cluster.
[monitoring] total store size in bytes	Total storage occupied by documents in a cluster.
[monitoring] indices store_size for top10	Top 10 indexes that occupy the largest storage space.

Chart Name	Description
[monitoring] indices docs_count for top10	Top 10 indexes with the largest number of documents.
[monitoring] indexing time in millis of index for top10(ms)	Top 10 indexes with the longest document write latency in a unit time (ms).
[monitoring] search query time in millis of index for top10(ms)	Top 10 indexes with the longest index query time in a unit time (ms).
[monitoring] segment count of index for top10	Top 10 indexes with the largest number of index segments.
[monitoring] segment memory in bytes of index for top10	Top 10 indexes with the largest heap memory usage of index segments.

NOTICE

The index pattern of **monitoring-eye-css-*** cannot be deleted during index monitoring. Otherwise, the monitoring chart will be abnormal.

Customizing Visualizations Charts

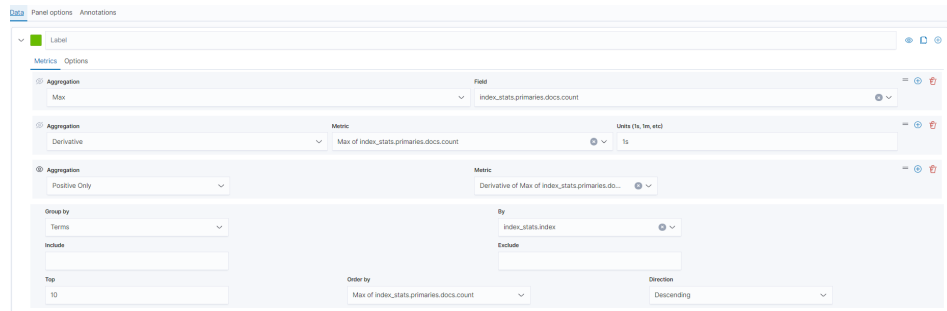
The index monitoring module periodically stores the index/stats information in the **monitoring-eyes-css** index. You can use the Kibana chart function to draw customized charts.

The following procedure describes how to check the trend of the document quantity in a chart as an example.

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. Choose **Visualize**.
4. Click **Create visualization** and select **TSVB**.
5. Set chart parameters and view the visualizations.

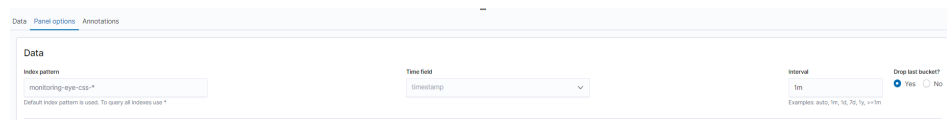
On the **Data** tab page, **index_stats primaries.docs.count** indicates the number of documents in the primary shard. **Derivative** indicates the difference between aggregation buckets. Set **Unit** to **1s**, visualizing network rates as "per second". Select **Positive only** to prevent negative numbers after resetting. To sort statistics by index, set **Group by** to **Terms** and **By** to **index_stats.index**. Statistics will be grouped by index name.

Figure 6-2 TSVB page



To view data in different time segments, set the aggregation interval, or the displayed data will be incomplete. On the **Panel options** tab page, set **Interval** to **1m** or **30m** to adjust the interval of **timestamp**.

Figure 6-3 Setting the interval



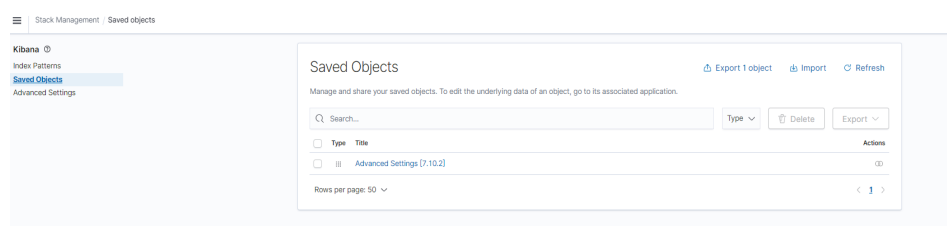
Importing Index Monitoring Charts

You can import or export charts on Kibana. If the index monitoring charts are not displayed, you can import the charts to Kibana again to load the monitoring view.

The following describes how to import a chart to Kibana:

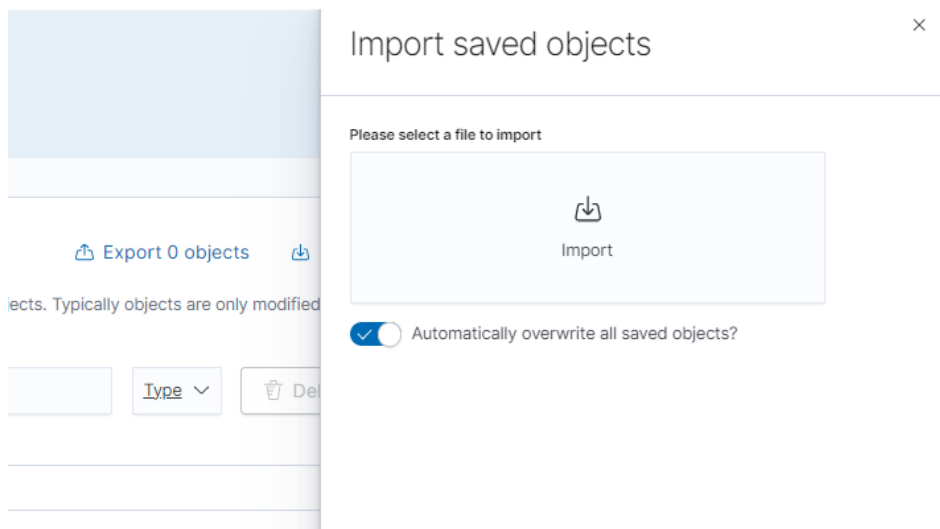
1. Create the **monitoring-kibana.ndjson** file by referring to [kibana-monitor](#).
2. Log in to Kibana and choose **Management > Stack Management > Saved objects**.

Figure 6-4 Selecting saved objects



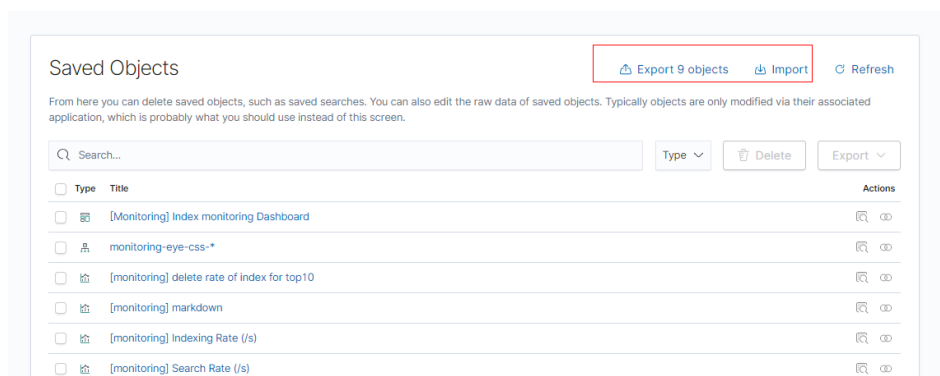
3. Click **Import** and upload the **monitoring-kibana.ndjson** file created in step 1.

Figure 6-5 Uploading a file



4. After the upload is complete, click **Done**. The index monitoring chart is successfully imported.

Figure 6-6 Successfully importing index monitoring charts



6.7.5 kibana-monitor

The configuration file content of **kibana-monitor** is as follows. You are advised to save the file as **monitoring-kibana.ndjson**.

```
{
  "attributes": {
    "description": "",
    "kibanaSavedObjectMeta": {
      "searchSourceJSON": {}
    },
    "title": "[monitoring] segment memory in bytes of index for top10",
    "uiStateJSON": {},
    "version": 1,
    "visState": {
      "title": "[monitoring] segment memory in bytes of index for top10",
      "type": "timeseries",
      "aggs": [],
      "params": {
        "id": "61ca57f0-469d-11e7-af02-69e470af7417",
        "type": "timeseries",
        "series": [
          {
            "id": "61ca57f1-469d-11e7-af02-69e470af7417",
            "color": "#68BC00",
            "split_mode": "terms",
            "split_color_mode": "kibana",
            "metrics": [
              {
                "id": "61ca57f2-469d-11e7-af02-69e470af7417",
                "type": "max",
                "field": "index_stats.total.segments.memory_in_bytes",
                "separate_axis": 0,
                "axis_position": "right",
                "formatter": "bytes",
                "chart_type": "line",
                "line_width": 1,
                "point_size": 1,
                "fill": 0.5,
                "stacked": "none",
                "label": "segments memory in bytes",
                "type": "timeseries",
                "terms_field": "index_stats.index",
                "terms_order_by": "61ca57f2-469d-11e7-af02-69e470af7417",
                "time_field": "timestamp",
                "index_pattern": "monitoring-eye-css-*",
                "interval": "",
                "axis_position": "left",
                "axis_formatter": "number",
                "axis_scale": "normal",
                "show_legend": 1,
                "show_grid": 1,
                "tooltip_mode": "show_all",
                "default_index_pattern": "monitoring-eye-css-*",
                "default_timefield": "timestamp",
                "isModelInvalid": false
              }
            ],
            "id": "3ae5d820-6628-11ed-8cd7-973626cf6f70",
            "references": []
          }
        ],
        "type": "visualization",
        "updated_at": "2022-12-01T12:41:01.165Z",
        "version": "WzlwNiwYXQ=="
      }
    },
    "attributes": {
      "description": "",
      "kibanaSavedObjectMeta": {
        "searchSourceJSON": {}
      },
      "title": "[monitoring] segment count of index for top10",
      "uiStateJSON": {},
      "version": 1,
      "visState": {
        "aggs": [],
        "params": {
          "axis_formatter": "number",
          "axis_position": "left",
          "axis_scale": "normal",
          "default_index_pattern": "monitoring-eye-css-*",
          "default_timefield": "timestamp",
          "filter": {
            "language": "kuery",
            "query": ""
          }
        }
      }
    }
  }
}
```



```
\"id\":\"61ca57f0-469d-11e7-af02-69e470af7417\", \"index_pattern\":\"monitoring-eye-css-  
\"interval\":\"\", \"isModelInvalid\":false, \"series\":[ { \"axis_position\":\"right\", \"chart_type\":\"line\", \"color  
\"rgb(231,102,76,1)\", \"fill\":\"0.5\", \"formatter\":\"number\", \"id\":\"61ca57f1-469d-11e7-  
af02-69e470af7417\", \"label\":\"segment count of index for top10\", \"line_width\":1, \"metrics\":[ { \"field  
\"indices_stats.total.segments.count\", \"id\":\"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\":\"max  
\"}], \"point_size\":1, \"separate_axis\":0, \"split_color_mode\":\"kibana\", \"split_mode\":\"terms\", \"stacked  
\"none\", \"terms_field\":\"index_stats.index\", \"terms_order_by\":\"61ca57f2-469d-11e7-  
af02-69e470af7417\", \"type\":\"timeseries\"}], \"show_grid\":1, \"show_legend\":1, \"time_field\":\"timestamp  
\", \"tooltip_mode\":\"show_all\", \"type\":\"timeseries\"}, { \"title\":\"[monitoring] segment count of index for  
top10\", \"type\":\"metrics\"}], \"id\":\"45d571c0-6626-11ed-8cd7-973626cf6f70\", \"references\":  
[ ], \"type\":\"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlwNywyXQ==\" }  
{ \"attributes\": { \"description\": \"\", \"kibanaSavedObjectMeta\": { \"searchSourceJSON\": { } }, \"title\": \"[monitoring]  
markdown\", \"uiStateJSON\": { }, \"version\": 1, \"visState\": { \"title\": \"[monitoring] markdown\", \"type  
\": \"markdown\", \"params\": { \"fontSize\": 12, \"openLinksInNewTab\": false, \"markdown\": \"### Index  
Monitoring \\nThis dashboard contains default table for you to play with. You can view it, search it, and  
interact with the visualizations.\" }, \"aggs\": [ ] }, \"id\": \"b2811c70-a5f1-11ec-9a68-ada9d754c566\", \"references\":  
[ ], \"type\": \"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlwOCwyXQ==\" }  
{ \"attributes\": { \"description\": \"number of document being indexing for primary and replica  
shards\", \"kibanaSavedObjectMeta\": { \"searchSourceJSON\": { } }, \"title\": \"[monitoring] Indexing Rate (/s)  
\", \"uiStateJSON\": { }, \"version\": 1, \"visState\": { \"title\": \"[monitoring] Indexing Rate (/s)\", \"type\": \"metrics  
\", \"params\": { \"id\": \"61ca57f0-469d-11e7-af02-69e470af7417\", \"type\": \"timeseries\", \"series\": [ { \"id  
\": \"61ca57f1-469d-11e7-af02-69e470af7417\", \"color\": \"rgb(0,32,188,1)\", \"split_mode\": \"everything  
\", \"metrics\": [ { \"id\": \"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\": \"max\", \"field  
\": \"indices_stats_all.total.indexing.index_total\", \"unit\": \"1s\", \"id\": \"fed72db0-a5f8-11ec-  
aa10-992297d21a2e\", \"type\": \"derivative\", \"field\": \"61ca57f2-469d-11e7-af02-69e470af7417\", \"unit  
\": \"\", \"id\": \"14b66420-a5f9-11ec-aa10-992297d21a2e\", \"type\": \"positive_only\", \"field\": \"fed72db0-  
a5f8-11ec-aa10-992297d21a2e\" }, { \"separate_axis\": 0, \"axis_position\": \"right\", \"formatter\": \"number  
\", \"chart_type\": \"line\", \"line_width\": 1, \"point_size\": 1, \"fill\": \"0.5\", \"stacked\": \"none\", \"label\": \"Indexing  
Rate (/s)\", \"type\": \"timeseries\", \"split_color_mode\": \"rainbow\", \"hidden\": false }, { \"time_field  
\": \"timestamp\", \"index_pattern\": \"monitoring-eye-css-*, \"interval\": \"\", \"axis_position\": \"left  
\", \"axis_formatter\": \"number\", \"axis_scale\": \"normal\", \"show_legend\": 1, \"show_grid  
\": 1, \"default_index_pattern\": \"monitoring-eye-css-*, \"default_timefield\": \"timestamp\", \"isModelInvalid  
\": false, \"legend_position\": \"bottom\" }, \"aggs\": [ ] }, \"id\": \"de4f8ab0-a5f8-11ec-9a68-  
ada9d754c566\", \"references\":  
[ ], \"type\": \"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlwOSwyXQ==\" }  
{ \"attributes\": { \"description\": \"number of search request being executed in primary and replica  
shards\", \"kibanaSavedObjectMeta\": { \"searchSourceJSON\": { } }, \"title\": \"[monitoring] Search Rate (/s)  
\", \"uiStateJSON\": { }, \"version\": 1, \"visState\": { \"title\": \"[monitoring] Search Rate (/s)\", \"type\": \"metrics  
\", \"params\": { \"id\": \"61ca57f0-469d-11e7-af02-69e470af7417\", \"type\": \"timeseries\", \"series\": [ { \"id  
\": \"61ca57f1-469d-11e7-af02-69e470af7417\", \"color\": \"rgb(0,33,224,1)\", \"split_mode\": \"everything  
\", \"metrics\": [ { \"id\": \"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\": \"max\", \"field  
\": \"indices_stats_all.total.search.query_total\", \"unit\": \"1s\", \"id\": \"b1093ac0-a5f7-11ec-  
aa10-992297d21a2e\", \"type\": \"derivative\", \"field\": \"61ca57f2-469d-11e7-af02-69e470af7417\", \"unit  
\": \"\", \"id\": \"c17db930-a5f7-11ec-aa10-992297d21a2e\", \"type\": \"positive_only\", \"field\": \"b1093ac0-  
a5f7-11ec-aa10-992297d21a2e\" }, { \"separate_axis\": 0, \"axis_position\": \"right\", \"formatter\": \"number  
\", \"chart_type\": \"line\", \"line_width\": 1, \"point_size\": 1, \"fill\": \"0.5\", \"stacked\": \"none\", \"split_color_mode  
\": \"rainbow\", \"label\": \"Search Rate (/s)\", \"type\": \"timeseries\", \"filter\": { \"query\": \"\", \"language  
\": \"kuery\" } }, { \"time_field\": \"timestamp\", \"index_pattern\": \"monitoring-eye-css-*, \"interval  
\": \"\", \"axis_position\": \"left\", \"axis_formatter\": \"number\", \"axis_scale\": \"normal\", \"show_legend  
\": 1, \"show_grid\": 1, \"default_index_pattern\": \"monitoring-eye-css-*, \"default_timefield\": \"timestamp  
\", \"isModelInvalid\": false, \"legend_position\": \"bottom\" }, \"aggs\": [ ] }, \"id\": \"811df7a0-a5f8-11ec-9a68-  
ada9d754c566\", \"references\":  
[ ], \"type\": \"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxMCwyXQ==\" }  
{ \"attributes\": { \"description\": \"\", \"kibanaSavedObjectMeta\": { \"searchSourceJSON\": { } }, \"title\": \"[monitoring]  
total docs count\", \"uiStateJSON\": { }, \"version\": 1, \"visState\": { \"title\": \"[monitoring] total docs count\", \"type  
\": \"metrics\", \"aggs\": [ ], \"params\": { \"id\": \"61ca57f0-469d-11e7-af02-69e470af7417\", \"type\": \"timeseries  
\", \"series\": [ { \"id\": \"61ca57f1-469d-11e7-af02-69e470af7417\", \"color  
\": \"rgb(218,139,69,1)\", \"split_mode\": \"everything\", \"split_color_mode\": \"kibana\", \"metrics\": [ { \"unit  
\": \"\", \"id\": \"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\": \"max\", \"field  
\": \"indices_stats_all.total.docs.count\" }, { \"separate_axis\": 0, \"axis_position\": \"right\", \"formatter  
\": \"number\", \"chart_type\": \"line\", \"line_width\": 1, \"point_size\": 1, \"fill\": \"0.5\", \"stacked\": \"none\", \"label  
\": \"total_docs_count\", \"type\": \"timeseries\" } ], \"time_field\": \"timestamp\", \"index_pattern\": \"monitoring-  
eye-css-*, \"interval\": \"\", \"axis_position\": \"left\", \"axis_formatter\": \"number\", \"axis_scale\": \"normal  
\", \"show_legend\": 1, \"show_grid\": 1, \"default_index_pattern\": \"monitoring-eye-css-*, \"default_timefield\": \"timestamp  
\", \"isModelInvalid\": false, \"legend_position\": \"bottom  
\" } } ], \"id\": \"eea89780-664b-11ed-8cd7-973626cf6f70\", \"references\":  
[ ], \"type\": \"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxMSwyXQ==\" }  
{ \"attributes\": { \"description\": \"\", \"kibanaSavedObjectMeta\": { \"searchSourceJSON\": { } }, \"title\": \"[monitoring]  
total docs delete\", \"uiStateJSON\": { }, \"version\": 1, \"visState\": { \"title\": \"[monitoring] total docs delete\", \"type
```

```
\"metrics\".\"agg\".\"params\".:{\"id\":\"61ca57f0-469d-11e7-af02-69e470af7417\", \"type\":\"timeseries  
\".\"series\".:[{\"id\":\"61ca57f1-469d-11e7-af02-69e470af7417\", \"color  
\".\"rgba(214,191,87,1)\".\"split_mode\":\"everything\".\"split_color_mode\":\"kibana\", \"metrics\".:[{\"id  
\".\"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\":\"max\", \"field  
\".\"indices_stats_all.total.docs.deleted\"}], \"separate_axis\":0, \"axis_position\":\"right\", \"formatter  
\".\"number\", \"chart_type\":\"line\", \"line_width\":1, \"point_size\":1, \"fill\":0.5, \"stacked\":\"none\", \"label  
\".\"total_docs_delete\", \"type\":\"timeseries\", \"hidden\":false}, {\"time_field\":\"timestamp\", \"index_pattern  
\".\"monitoring-eye-css-*, \"interval\":\"\", \"axis_position\":\"left\", \"axis_formatter\":\"number  
\".\"axis_scale\":\"normal\", \"show_legend\":1, \"show_grid\":1, \"tooltip_mode\":\"show_all  
\".\"default_index_pattern\": \"monitoring-eye-css-*, \"default_timefield\":\"timestamp\", \"isModelInvalid  
\".false, \"drop_last_bucket\":1, \"legend_position\":\"bottom  
\"}], \"id\":\"c7f72ae0-664f-11e7-8cd7-973626cf670\", \"references\":  
[], \"type\":\"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxMiwYXQ==\"}  
{\"attributes\":{\"description\":\"\", \"kibanaSavedObjectMeta\":{\"searchSourceJSON\":\"\"}, \"title\":\"[monitoring]  
total store size in bytes\", \"uiStateJSON\":\"\", \"version\":1, \"visState\":{\"\"title\":\"[monitoring] total store size in  
bytes\", \"type\":\"metrics\", \"agg\": \"\", \"params\": {\"id\":\"61ca57f0-469d-11e7-af02-69e470af7417\", \"type  
\".\"timeseries\", \"series\": [{\"id\":\"61ca57f1-469d-11e7-af02-69e470af7417\", \"color  
\".\"#68BC00\", \"split_mode\":\"everything\", \"split_color_mode\":\"kibana\", \"metrics\": [{\"id  
\".\"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\":\"max\", \"field  
\".\"indices_stats_all.total.store.size.in.bytes\"}], \"separate_axis\":0, \"axis_position\":\"right\", \"formatter  
\".\"bytes\", \"chart_type\":\"line\", \"line_width\":1, \"point_size\":1, \"fill\":0.5, \"stacked\":\"none\", \"label  
\".\"total store size in bytes\", \"type\":\"timeseries\"}], \"time_field\":\"timestamp\", \"index_pattern  
\".\"monitoring-eye-css-*, \"interval\":\"\", \"axis_position\":\"left\", \"axis_formatter\":\"number  
\".\"axis_scale\":\"normal\", \"show_legend\":1, \"show_grid\":1, \"tooltip_mode\":\"show_all  
\".\"default_index_pattern\": \"monitoring-eye-css-*, \"default_timefield\":\"timestamp\", \"isModelInvalid  
\".false, \"legend_position\":\"bottom\", \"background_color_rules\": [{\"id  
\".\"7712e550-664f-11e7-8b5d-8db37e5b4cc4\"}], \"bar_color_rules\": [{\"id  
\".\"77680a30-664f-11e7-8b5d-8db37e5b4cc4\"}], \"id\":\"c7f72ae0-664e-11e7-8cd7-973626cf670\", \"refere  
nces\": [], \"type\":\"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxMywyXQ==\"}  
{\"attributes\":{\"description\":\"\", \"kibanaSavedObjectMeta\":{\"searchSourceJSON\":\"\"}, \"title\":\"[monitoring]  
indexing rate of index for top10(/s)\", \"uiStateJSON\":\"\", \"version\":1, \"visState\":{\"\"title\":\"[monitoring]  
indexing rate of index for top10(/s)\", \"type\":\"metrics\", \"agg\": \"\", \"params\": {\"id  
\".\"61ca57f0-469d-11e7-af02-69e470af7417\", \"type\":\"timeseries\", \"series\": [{\"id  
\".\"61ca57f1-469d-11e7-af02-69e470af7417\", \"color\": \"#68BC00\", \"split_mode\":\"terms\", \"metrics\":  
[ {\"id\":\"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\":\"max\", \"field  
\".\"index_stats.total.indexing.index_total\"}, {\"unit\":\"1s\", \"id\":\"541ed8f0-a5ee-11ec-aa10-992297d21a2e  
\".\"type\":\"derivative\", \"field\":\"61ca57f2-469d-11e7-af02-69e470af7417\", \"unit\":\"\", \"id  
\".\"67ec1f50-a5ee-11ec-aa10-992297d21a2e\", \"type\":\"positive_only\", \"field\":\"541ed8f0-a5ee-11ec-  
aa10-992297d21a2e\"}], \"separate_axis\":0, \"axis_position\":\"right\", \"formatter\":\"number\", \"chart_type  
\".\"line\", \"line_width\":1, \"point_size\":1, \"fill\":0.5, \"stacked\":\"none\", \"label\":\"indexing_rate\", \"type  
\".\"timeseries\", \"split_filters\": [{\"color\":\"#68BC00\", \"id\":\"81004200-a5ee-11ec-aa10-992297d21a2e  
\".\"filter\":{\"query\":\"\", \"language\":\"kuery\"}], \"filter\":{\"query\":\"\", \"language\":\"kuery  
\"}], \"terms_field\":\"index_stats.index\", \"terms_order_by\":\"61ca57f2-469d-11e7-  
af02-69e470af7417\", \"terms_size\":\"10\", \"terms_direction\":\"desc\", \"split_color_mode\":\"rainbow  
\"}], \"time_field\":\"timestamp\", \"index_pattern\":\"monitoring-eye-css-*, \"interval\":\"\", \"axis_position  
\".\"left\", \"axis_formatter\":\"number\", \"axis_scale\":\"normal\", \"show_legend\":1, \"show_grid  
\".\"1\", \"default_index_pattern\":\"monitoring-eye-css-*, \"default_timefield\":\"timestamp\", \"isModelInvalid  
\".false, \"tooltip_mode\":\"show_all\"}], \"id\":\"943b3e00-a5ef-11ec-9a68-ada9d754c566\", \"references\":  
[], \"type\":\"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxNCwyXQ==\"}  
{\"attributes\":{\"description\":\"\", \"kibanaSavedObjectMeta\":{\"searchSourceJSON\":\"\"}, \"title\":\"[monitoring]  
search rate of index for top10(/s)\", \"uiStateJSON\":\"\", \"version\":1, \"visState\":{\"\"title\":\"[monitoring] search  
rate of index for top10(/s)\", \"type\":\"metrics\", \"agg\": \"\", \"params\": {\"id\":\"61ca57f0-469d-11e7-  
af02-69e470af7417\", \"type\":\"timeseries\", \"series\": [{\"id\":\"61ca57f1-469d-11e7-  
af02-69e470af7417\", \"color\":\"rgba(99,157,12,1)\".\"split_mode\":\"terms\", \"metrics\": [{\"id  
\".\"61ca57f2-469d-11e7-af02-69e470af7417\", \"type\":\"max\", \"field  
\".\"index_stats.total.search.query_total\"}, {\"unit\":\"1s\", \"id\":\"fdfdad0-a5ef-11ec-aa10-992297d21a2e  
\".\"type\":\"derivative\", \"field\":\"61ca57f2-469d-11e7-af02-69e470af7417\", \"unit\":\"\", \"id  
\".\"0aaa26a0-a5f0-11ec-aa10-992297d21a2e\", \"type\":\"positive_only\", \"field\":\"fdfdad0-a5ef-11ec-  
aa10-992297d21a2e\"}], \"separate_axis\":0, \"axis_position\":\"right\", \"formatter\":\"number\", \"chart_type  
\".\"line\", \"line_width\":1, \"point_size\":1, \"fill\":0.5, \"stacked\":\"none\", \"label\":\"search rate\", \"type  
\".\"timeseries\", \"terms_field\":\"index_stats.index\", \"terms_order_by\":\"61ca57f2-469d-11e7-  
af02-69e470af7417\", \"split_color_mode\":\"rainbow\"}], \"time_field\":\"timestamp\", \"index_pattern  
\".\"monitoring-eye-css-*, \"interval\":\"\", \"axis_position\":\"left\", \"axis_formatter\":\"number  
\".\"axis_scale\":\"normal\", \"show_legend\":1, \"show_grid\":1, \"default_index_pattern\":\"monitoring-eye-  
css-*, \"default_timefield\":\"timestamp\", \"isModelInvalid\":false, \"tooltip_mode\":\"show_all  
\"}], \"id\":\"ab503550-a5ef-11ec-9a68-ada9d754c566\", \"references\":  
[], \"type\":\"visualization\", \"updated_at\":\"2022-12-01T12:41:01.165Z\", \"version\":\"WzlxNSwyXQ==\"}  
{\"attributes\":{\"description\":\"\", \"kibanaSavedObjectMeta\":{\"searchSourceJSON\":\"\"}, \"title\":\"[monitoring]  
indices store_size for top10\", \"uiStateJSON\":\"\", \"version\":1, \"visState\":{\"\"title\":\"[monitoring] indices
```

```
store_size for top10\","type":"metrics\","aggs":{},"params":{"id":"61ca57f0-469d-11e7-af02-69e470af7417"},"type":"timeseries"},"series":[{"id":"38474c50-a5f5-11ec-aa10-992297d21a2e","color":"#68BC00","split_mode":"terms","metrics":{"id":"38474c51-a5f5-11ec-aa10-992297d21a2e"},"type":"max","field":"index_stats.total.store.size_in_bytes"},"separate_axis":0,"axis_position":"right","formatter":"bytes","chart_type":"line","line_width":1,"point_size":1,"fill":0.5,"stacked":"none","label":"store_size for index","type":"timeseries"},"terms_field":"index_stats.index","terms_order_by":"38474c51-a5f5-11ec-aa10-992297d21a2e"},"filter":{"query":"","language":"kuery"},"split_color_mode":"rainbow"},"time_field":"timestamp","index_pattern":"monitoring-eye-css-*"},"interval":"","axis_position":"left","axis_formatter":"number","axis_scale":"normal","show_legend":1,"show_grid":1,"default_index_pattern":"monitoring-eye-css-*"},"default_timefield":"timestamp"},"isModelInvalid":false,"filter":{"query":"","language":"kuery"},"bar_color_rules":{"id":"7d9d3cb0-a5f5-11ec-aa10-992297d21a2e"},"tooltip_mode":"show_all"},"id":"c78119a0-a5f5-11ec-9a68-ada9d754c566","references":{},"type":"visualization","updated_at":"2022-12-01T12:41:01.165Z","version":"WzlxNiwYXQ=="},"attributes":{"description":"","kibanaSavedObjectMeta":{"searchSourceJSON":"","title":"[monitoring] search query time in millis of index for top10(ms)","uiStateJSON":"","version":1,"visState":{"title":"[monitoring] search query time in millis of index for top10(ms)","type":"metrics\","aggs":{},"params":{"axis_formatter":"number","axis_max":"","axis_min":"","axis_position":"left","axis_scale":"normal","default_index_pattern":"monitoring-eye-css-*"},"default_timefield":"timestamp"},"id":"61ca57f0-469d-11e7-af02-69e470af7417"},"index_pattern":"monitoring-eye-css-*"},"interval":"","isModelInvalid":false,"series":[{"axis_position":"right","chart_type":"line","color":"#68BC00","fill":0.5,"formatter":"number","id":"61ca57f1-469d-11e7-af02-69e470af7417"},"label":"index_query_time_in_millis","line_width":1,"metrics":{"field":"index_stats.total.search.query_time_in_millis"},"id":"61ca57f2-469d-11e7-af02-69e470af7417"},"type":"max"},"unit":"1s"},"id":"42c92b10-6645-11ed-925a-6de90846447d"},"type":"derivative"},"field":"61ca57f2-469d-11e7-af02-69e470af7417"},"point_size":1,"separate_axis":0,"split_color_mode":"kibana"},"split_mode":"terms","stacked":"none","terms_field":"index_stats.index","terms_order_by":"61ca57f2-469d-11e7-af02-69e470af7417"},"type":"timeseries"},"show_grid":1,"show_legend":1,"time_field":"timestamp"},"tooltip_mode":"show_all"},"type":"timeseries"},"background_color":null,"filter":{"query":"","language":"kuery"},"legend_position":"right"},"id":"c8109100-6627-11ed-8cd7-973626cf6f70","references":{},"type":"visualization","updated_at":"2022-12-01T12:41:01.165Z","version":"WzlxNywyXQ=="},"attributes":{"description":"","hits":0,"kibanaSavedObjectMeta":{"searchSourceJSON":{"query":{"language":"kuery"},"query":"","filter":{"optionsJSON":{"hidePanelTitles":false,"useMargins":true},"panelsJSON":[{"gridData":{"x":0,"y":0,"w":48,"h":5,"i":"971ed6c6-81b9-491b-9f08-e3ae9c382abd"},"panelIndex":"971ed6c6-81b9-491b-9f08-e3ae9c382abd"},"embeddableConfig":{"panelRefName":"panel_0"},"gridData":{"x":0,"y":5,"w":24,"h":15,"i":"5a6982e7-0c6c-4733-8a2d-e4c57cdf7397"},"panelIndex":"5a6982e7-0c6c-4733-8a2d-e4c57cdf7397"},"embeddableConfig":{"panelRefName":"panel_1"},"gridData":{"x":24,"y":5,"w":24,"h":15,"i":"662476f4-739c-4a05-858c-2ee8230cf410"},"panelIndex":"662476f4-739c-4a05-858c-2ee8230cf410"},"embeddableConfig":{"panelRefName":"panel_2"},"gridData":{"x":0,"y":20,"w":16,"h":15,"i":"d89c38e2-33f3-4592-b503-20460a6a7a57"},"panelIndex":"d89c38e2-33f3-4592-b503-20460a6a7a57"},"embeddableConfig":{"panelRefName":"panel_3"},"gridData":{"x":16,"y":20,"w":16,"h":15,"i":"1f693b49-79fa-4807-94e8-0c12f51e54f8"},"panelIndex":"1f693b49-79fa-4807-94e8-0c12f51e54f8"},"embeddableConfig":{"panelRefName":"panel_4"},"gridData":{"x":32,"y":20,"w":16,"h":15,"i":"616b143d-74e9-4dac-98ba-5849536f0fba"},"panelIndex":"616b143d-74e9-4dac-98ba-5849536f0fba"},"embeddableConfig":{"panelRefName":"panel_5"},"gridData":{"x":0,"y":35,"w":24,"h":11,"i":"cfa82f27-1b8d-49ba-a7b9-d8809d3b258c"},"panelIndex":"cfa82f27-1b8d-49ba-a7b9-d8809d3b258c"},"embeddableConfig":{"panelRefName":"panel_6"},"gridData":{"x":24,"y":35,"w":24,"h":11,"i":"135d13eb-aab6-43ca-9029-7d26e91d90e3"},"panelIndex":"135d13eb-aab6-43ca-9029-7d26e91d90e3"},"embeddableConfig":{"panelRefName":"panel_7"},"gridData":{"x":0,"y":46,"w":24,"h":11,"i":"28a77de1-9110-49e8-b273-724f880b1653"},"panelIndex":"28a77de1-9110-49e8-b273-724f880b1653"},"embeddableConfig":{"panelRefName":"panel_8"},"gridData":{"x":24,"y":46,"w":24,"h":11,"i":"80ece867-cf23-4935-bfbc-430afa51bcca"},"panelIndex":"80ece867-cf23-4935-bfbc-430afa51bcca"},"embeddableConfig":{"panelRefName":"panel_9"},"gridData":{"x":0,"y":57,"w":24,"h":11,"i":"2ba970aa-c9c4-491b-bdd3-c1b1ee9bc8d3"},"panelIndex":"2ba970aa-c9c4-491b-bdd3-c1b1ee9bc8d3"},"embeddableConfig":{"panelRefName":"panel_10"},"gridData":{"x":24,"y":57,"w":24,"h":11,"i":"f2e1b6ab-ddf7-492e-aaca-9460f11aa4aa"},"panelIndex":"f2e1b6ab-ddf7-492e-aaca-9460f11aa4aa"},"embeddableConfig":{"panelRefName":"panel_11"},"gridData":{"x":0,"y":68,"w":24,"h":11,"i":"dd14182d-d8b9-47f2-bf36-6c3a3b09586c"},"panelIndex":"dd14182d-d8b9-47f2-bf36-6c3a3b09586c"},"embeddableConfig":{"panelRefName":"panel_12"},"gridData":{"x":24,"y":68,"w":24,"h":11,"i":"a47f9333-52b7-49b7-8cac-f470cf405131"},"panelIndex":"a47f9333-52b7-49b7-8cac-f470cf405131"},"embeddableConfig":{"panelRefName":"panel_13"},"timeRestore":false,"title":"[Monitoring] Index monitoring Dashboard","version":1,"id":"524eb000-a5f2-11ec-9a68-ada9d754c566","references":{"id":"b2811c70-a5f1-11ec-9a68-ada9d754c566","name":"panel_0","type":"visualization"},"id":"de4f8ab0-a5f8-11ec-9a68-
```

```
ada9d754c566", "name": "panel_1", "type": "visualization"}, {"id": "811df7a0-a5f8-11ec-9a68-ada9d754c566", "name": "panel_2", "type": "visualization"}, {"id": "eea89780-664b-11ed-8cd7-973626cf6f70", "name": "panel_3", "type": "visualization"}, {"id": "cfbb4e20-664c-11ed-8cd7-973626cf6f70", "name": "panel_4", "type": "visualization"}, {"id": "c7f72ae0-664e-11ed-8cd7-973626cf6f70", "name": "panel_5", "type": "visualization"}, {"id": "943b3e00-a5ef-11ec-9a68-ada9d754c566", "name": "panel_6", "type": "visualization"}, {"id": "ab503550-a5ef-11ec-9a68-ada9d754c566", "name": "panel_7", "type": "visualization"}, {"id": "c78119a0-a5f5-11ec-9a68-ada9d754c566", "name": "panel_8", "type": "visualization"}, {"id": "225f6020-a5f1-11ec-9a68-ada9d754c566", "name": "panel_9", "type": "visualization"}, {"id": "17d49220-662a-11ed-8cd7-973626cf6f70", "name": "panel_10", "type": "visualization"}, {"id": "c8109100-6627-11ed-8cd7-973626cf6f70", "name": "panel_11", "type": "visualization"}, {"id": "45d571c0-6626-11ed-8cd7-973626cf6f70", "name": "panel_12", "type": "visualization"}, {"id": "3ae5d820-6628-11ed-8cd7-973626cf6f70", "name": "panel_13", "type": "visualization"}], "type": "dashboard", "updated_at": "2022-12-01T12:41:01.165Z", "version": "WzlxOCwyXQ==", "exportedCount": 16, "missingRefCount": 0, "missingReferences": []}
```

6.8 Enhanced Cluster Monitoring

6.8.1 P99 Latency Monitoring

Context

The Elasticsearch community only discusses how to monitor the average latency of search requests, which cannot reflect the actual search performance of a cluster. To enhance monitoring, CSS allows you to monitor the P99 latency of search requests in clusters.

Prerequisites

Currently, only clusters of version 7.6.2 and 7.10.2 support P99 latency monitoring.

Obtaining Monitoring Information

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation tree on the left, choose **Dev Tools** and run the following command to check the P99 latency of the current cluster:

```
GET /search/stats/percentile
```

Example response:

```
{
  "overall" : {
    "1.0" : 2.0,
    "5.0" : 2.0,
    "25.0" : 6.5,
    "50.0" : 19.5,
    "75.0" : 111.0,
    "95.0" : 169.0,
    "99.0" : 169.0,
    "max" : 169.0,
    "min" : 2.0
  },
  "last_one_day" : {
    "1.0" : 2.0,
    "5.0" : 2.0,
    "25.0" : 6.5,
    "50.0" : 19.5,
    "75.0" : 111.0,
```

```
"95.0" : 169.0,  
"99.0" : 169.0,  
"max" : 169.0,  
"min" : 2.0  
},  
"latest" : {  
  "1.0" : 26.0,  
  "5.0" : 26.0,  
  "25.0" : 26.0,  
  "50.0" : 26.0,  
  "75.0" : 26.0,  
  "95.0" : 26.0,  
  "99.0" : 26.0,  
  "max" : 26.0,  
  "min" : 26.0  
}  
}
```

NOTE

- In the response, **overall** indicates all the statistics that have been collected since the cluster startup, **last_one_day** indicates the statistics collected in the last day, and **latest** indicates the statistics that have been collected since the last reset.
- The calculated P99 latency is an estimation. It is more precise than the P50 latency.
- The P99 latency of a cluster is cleared and recalculated if the cluster is restarted.

Other Operations

- Define percentage.

You can run the following command to specify the percentage:

```
GET /search/stats/percentile  
{  
  "percents": [1, 50, 90]  
}
```

- Reset the **latest** statistics.

You can run the following command to reset the **latest** statistics:

```
POST /search/stats/reset
```

Example response:

```
{  
  "nodes" : {  
    "css-c9c8-ess-esn-1-1" : "ok"  
  }  
}
```

6.8.2 HTTP Status Code Monitoring

Context

When an external system accesses Elasticsearch through the HTTP protocol, a response and the corresponding status code are returned. The open-source Elasticsearch server does not collect the status code, so users cannot monitor Elasticsearch APIs status or cluster request status. CSS allows you to monitor the HTTP status codes of clusters.

Prerequisites

Currently, only clusters of versions 7.6.2 and 7.10.2 support HTTP status code monitoring.

Obtaining Status Codes

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster and click **Access Kibana** in the **Operation** column.
3. In the navigation tree on the left, choose **Dev Tools**.
4. On the console page of **Dev Tools**, run commands based on the cluster version.

- For clusters of version 7.6.2, run the following command to obtain the status code statistics:

```
GET /_nodes/http_stats
```

Example response:

```
{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0 },
  "cluster_name" : "css-8362",
  "nodes" : {
    "F9IFdQPAPRaOJI7oL7HOxtQ" : {
      "http_code" : {
        "200" : 114,
        "201" : 5,
        "429" : 0,
        "400" : 7,
        "404" : 0,
        "405" : 0
      }
    }
  }
}
```

- For clusters of version 7.10.2, run the following command to obtain the status code statistics:

```
GET _nodes/stats/http
```

Example response:

```
{
// ...
  "cluster_name" : "css-2985",
  "nodes" : {
// ...
    "omvR9_W-TsGApraMApREjA" : {
// ...
      "http" : {
        "current_open" : 4,
        "total_opened" : 37,
        "http_code" : {
          "200" : 25,
          "201" : 7,
          "429" : 0,
          "400" : 3,
          "404" : 0,
          "405" : 0
        }
      }
    }
  }
}
```

6.9 Enhanced Aggregation

6.9.1 Features

The enhanced aggregation is an optimization feature for service awareness. With this feature, you can optimize the aggregation analysis capability of observable services.

Currently, the enhanced aggregation is supported by only clusters of version 7.10.2.

Working Principles

In large-scale dataset aggregation and analysis scenarios, data grouping and aggregation takes a lot of time. Improving the grouping aggregation capability depends on the following key features:

- **Sorting key:** Data is stored in sequence based on the sorting key.
- **Clustering key:** It is contained in the sorting key. Data is clustered based on the clustering key.

In the case of data clustering, enhanced aggregation uses the vectorization technology to process data in batches, improving aggregation performance.

Table 6-47 Feature parameters

Parameter	Description
index.search.turbo.enabled	Indicates whether to enable the feature. The default value is true .
index.sort.field	Sorting key
index.cluter.field	Clustering key

Features

Based on different service requirements, enhanced aggregation can be used in the following three scenarios:

- [Grouping and Aggregation of Low-cardinality Fields](#)
- [High-cardinality Field Histogram Aggregation](#)
- [Low-cardinality and High-cardinality Field Mixing](#)

6.9.2 Grouping and Aggregation of Low-cardinality Fields

Low-cardinality fields have high data clustering performance when being sorted, which facilitates vectorized optimization. Assume that the following query statement exists:

```
POST testindex/_search
{
  "size": 0,
  "aggs": {
    "groupby_region": {
      "terms": {
        "field": "region"
      },
      "aggs": {
        "groupby_host": {
          "terms": {
            "field": "host"
          },
          "aggs": {
            "avg_cpu_usage": {
              "avg": {
                "field": "cpu_usage"
              }
            }
          }
        }
      }
    }
  }
}
```

Assume that the **region** and **host** are low-cardinality fields. To use the enhanced aggregation, set the parameters as follows:

NOTE

The clustering key must be a prefix subset of the sorting key.

```
// Configure an index
"settings": {
  "index": {
    "search": {
      "turbo": {
        "enabled": "true" // Enable optimization
      }
    },
    "sort": { // Specify a sorting key
      "field": [
        "region",
        "host",
        "other"
      ]
    },
    "cluster": {
      "field": [ // Specify a clustering key
        "region",
        "host"
      ]
    }
  }
}
```

6.9.3 High-cardinality Field Histogram Aggregation

High-cardinality fields are usually used for histogram grouping and aggregation instead of single-point grouping and aggregation. For example, collecting the statistics of logs at a certain period. Assume that the following query statement exists:

```
POST testindex/_search?pretty
{
  "size": 0,
  "aggs": {
```



```
"avg_score": {
  "avg": {
    "field": "score"
  },
  "aggs": {
    "groupbytime": {
      "date_histogram": {
        "field": "timestamp",
        "calendar_interval": "day"
      }
    }
  }
}
```

This query groups the field **timestamp** using a histogram and calculates the average score. **timestamp** is a typical high-cardinality field. To use the enhanced aggregation for the preceding query, set parameters as follows:

```
// Configure an index
"settings": {
  "index": {
    "search": {
      "turbo": {
        "enabled": "true" // Enable optimization
      }
    },
    "sort": { // Specify a sorting key
      "field": [
        "timestamp"
      ]
    }
  }
}
```

6.9.4 Low-cardinality and High-cardinality Field Mixing

In the scenario where low-cardinality and high-cardinality fields are mixed, assume that the following query statement exists:

```
POST testindex/_search
{
  "size": 0,
  "aggs": {
    "groupby_region": {
      "terms": {
        "field": "region"
      },
      "aggs": {
        "groupby_host": {
          "terms": {
            "field": "host"
          },
          "aggs": {
            "groupby_timestamp": {
              "date_histogram": {
                "field": "timestamp",
                "interval": "day"
              },
              "aggs": {
                "avg_score": {
                  "avg": {
                    "field": "score"
                  }
                }
              }
            }
          }
        }
      }
    }
  }
}
```

```
}  
}  
}  
}  
}
```

Group the low-cardinality fields and create a histogram using the high-cardinality fields. To use the enhanced aggregation for the preceding query, set the parameters as follows:

NOTE

- A clustering key is the prefix subset of a sorting key.
- High-cardinality fields must be in the sorting key, and high-cardinality fields must follow the last low-cardinality field.

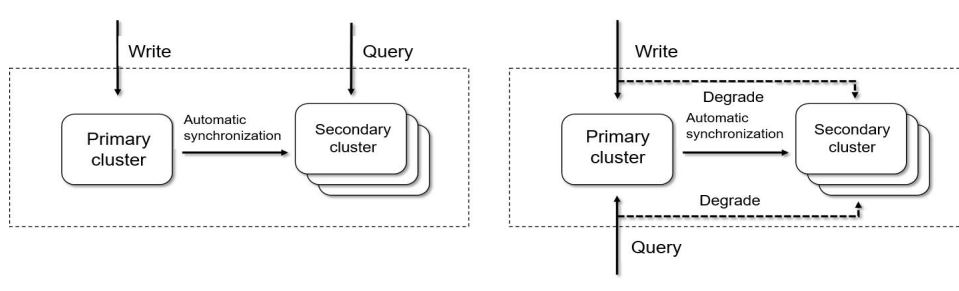
```
// Configure an index  
"settings" : {  
  "index" : {  
    "search" : {  
      "turbo" : {  
        "enabled" : "true" // Enable optimization  
      }  
    },  
    "sort" : { // Specify a sorting key  
      "field" : [  
        "region",  
        "host",  
        "timestamp",  
        "other"  
      ]  
    },  
    "cluster" : {  
      "field" : [ // Specify a clustering key  
        "region",  
        "host"  
      ]  
    }  
  }  
}
```

6.10 Read/Write Splitting

6.10.1 Features

CSS supports read/write splitting. Data written to the primary cluster (**Leader**) can be automatically synchronized to the secondary cluster (**Follower**). In this way, data is written to the primary cluster and queried in the secondary cluster. The read and write can be separated to improve the query performance (as shown in the left part of [Figure 6-7](#)). When the primary cluster is unavailable, the secondary cluster can provide data write and query services (as shown in the right part of [Figure 6-7](#)).

Figure 6-7 Two application scenarios of read/write splitting



Currently, only clusters of versions 7.6.2 and 7.10.2 support read/write isolation. The versions of the primary and secondary clusters must be the same.

6.10.2 Instructions

6.10.2.1 Basic Settings

1. Log in to the CSS management console.
2. Choose **Clusters** in the navigation pane. On the **Clusters** page, locate the target cluster, and click **Access Kibana** in the **Operation** column.
3. Click **Dev Tools** in the navigation tree on the left and perform the following operations:

Configure the primary cluster information.

PUT /_cluster/settings

```
{
  "persistent" : {
    "cluster" : {
      "remote.rest" : {
        "leader1" : {
          "seeds" : [
            "http://10.0.0.1:9200",
            "http://10.0.0.2:9200",
            "http://10.0.0.3:9200"
          ],
          "username": "elastic",
          "password": "*****"
        }
      }
    }
  }
}
```

NOTE

- Secondary clusters must be able to access the REST API (default port: 9200) of the primary cluster.
- The primary cluster name is **leader1** and can be changed.
- The value of **seeds** is the REST address of the primary cluster. Multiple values are supported. When HTTPS access is enabled, the URI schema must be changed to HTTPS.
- **username** and **password** are required only when the security mode is enabled for the primary cluster.
- After the configuration is complete, you can use the **GET _remote/rest/info** API to obtain the connection status with the primary cluster.

6.10.2.2 Synchronizing Specified Indexes

Synchronize a single index.

The request URL and request body parameters are as follows:

```
PUT start_remote_sync
```

Table 6-48 Request body parameters

Parameter	Description
remote_cluster	Name of the primary cluster. The default name is leader1 . You can change the name by configuring the primary cluster information.
remote_index	Name of the index to be synchronized in the primary cluster
local_index	Name of the index being synchronized to the secondary cluster
settings	Index settings of the index being synchronized

After the synchronization function is enabled, indexes in the secondary cluster become read-only and are periodically synchronized with indexes in the primary cluster.

The following are two examples:

1. Synchronize a single index from the primary cluster to the secondary cluster.

```
PUT start_remote_sync
{
  "remote_cluster": "leader1",
  "remote_index": "data1_leader",
  "local_index": "data1_follower"
}
```

2. Synchronize a single index from the primary cluster to the secondary cluster and modify the index configurations.

```
PUT start_remote_sync
{
  "remote_cluster": "leader1",
  "remote_index": "data1_leader",
  "local_index": "data1_follower",
  "settings": {
    "number_of_replicas": 4
  }
}
```

 **NOTE**

The following index configurations cannot be modified:

- number_of_shards
- version.created
- uuid
- creation_date
- soft_deletes.enabled

6.10.2.3 Matching Index Synchronization

The request URL and request body parameters are as follows:

```
PUT auto_sync/pattern/{pattern_name}
```

Table 6-49 Request body parameters

Parameter	Description
remote_cluster	Name of the primary cluster. The default name is leader1 . You can change the name by configuring the primary cluster information.
remote_index_patterns	Mode of the index to be synchronized in the primary cluster. The wildcard (*) is supported.
local_index_pattern	Mode of the index to be synchronized in the secondary cluster. The template can be replaced. For example, if this parameter is set to {{remote_index}}-sync , the index log1 change to log1-sync after synchronization.
apply_exist_index	Whether to synchronize existing indexes in the primary cluster. The default value is true .
settings	Index settings of the index being synchronized

The following are two examples:

- Synchronize a single index from the primary cluster to the secondary cluster.

```
PUT auto_sync/pattern/pattern1
{
  "remote_cluster": "leader1",
  "remote_index_patterns": "log*",
  "local_index_pattern": "{{remote_index}}-sync",
  "apply_exist_index": true
}
```
- Synchronize a single index from the primary cluster to the secondary cluster and modify the index configurations.

```
PUT auto_sync/pattern/pattern1
{
  "remote_cluster": "leader1",
  "remote_index_patterns": "log*",
  "local_index_pattern": "{{remote_index}}-sync",
  "apply_exist_index": true,
  "settings": {
    "number_of_replicas": 4
  }
}
```

NOTE

The following index configurations cannot be modified:

- number_of_shards
- version.created
- uuid
- creation_date
- soft_deletes.enabled

6.10.2.4 Stopping Index Synchronization

You can specify multiple indexes or use wildcard to match the target indexes and terminate their synchronization tasks. Subsequent modifications to the indexes in the primary cluster will not be synchronized to the secondary cluster. The read-only state of the indexes in the secondary cluster is cancelled, and new data can be written to the secondary cluster.

An example request is as follows:

```
PUT log*/stop_remote_sync
```

6.10.2.5 Other Management APIs

- **Querying the created patterns.**

This API is used to query the pattern list and query a specified pattern by name.

An example request is as follows:

```
GET auto_sync/pattern
GET auto_sync/pattern/{pattern_name}
```

The following is an example of the response:

```
{
  "patterns" : [
    {
      "name" : "pattern1",
      "pattern" : {
        "remote_cluster" : "leader",
        "remote_index_patterns" : [
          "log*"
        ],
        "local_index_pattern" : "{{remote_index}}-sync",
        "settings" : { }
      }
    }
  ]
}
```

- **Deleting a created schema.**

This API is used to delete a specified pattern.

An example request is as follows:

```
DELETE auto_sync/pattern/{pattern_name}
```

- **Obtaining the automatic synchronization status.**

This API is used to obtain the synchronization status of matched indexes.

An example request is as follows:

```
GET auto_sync/stats
```

The following is an example of the response:

```
{
  "success_count" : 3,
  "failed_count" : 0,
  "failed_remote_cluster_state_requests_count" : 0,
  "last_fail_exception" : { },
  "last_fail_remote_cluster_requests_exception" : { }
}
```

- **Obtaining the synchronization status of the index that is being synchronized.**

An example request is as follows:

```
GET {index_name}/sync_stats
```

The following is an example of the response:

```
{
  "indices" : {
    "data1_follower" : {
      "shards" : {
        "0" : [
          {
            "primary" : false,
            "total_synced_times" : 27,
            "total_empty_times" : 25,
            "total_synced_files" : 4,
            "total_synced_bytes" : 3580,
            "total_paused_nanos" : 0,
            "total_paused_times" : 0,
            "current" : {
              "files_count" : 0,
              "finished_files_count" : 0,
              "bytes" : 0,
              "finished_bytes" : 0
            }
          }
        ],
      },
    },
    {
      "primary" : true,
      "total_synced_times" : 28,
      "total_empty_times" : 26,
      "total_synced_files" : 20,
      "total_synced_bytes" : 17547,
      "total_paused_nanos" : 0,
      "total_paused_times" : 0,
      "current" : {
        "files_count" : 0,
        "finished_files_count" : 0,
        "bytes" : 0,
        "finished_bytes" : 0
      }
    }
  ]
}
}
```

- **Changing the synchronization period.**

The synchronization period is 30 seconds by default and can be modified.

An example request is as follows (change the synchronization period to 2 seconds):

```
PUT {index_name}/_settings
{
  "index.remote_sync.sync_interval": "2s"
}
```

- **Enabling forcible synchronization**

By default, the plug-in determines whether to synchronize metadata based on whether the number of documents in the index of the primary cluster changes. If the primary cluster only updates documents and the number of documents remains unchanged, the plug-in does not synchronize the updates to the secondary cluster. The configuration can be modified. After this function is enabled, the index metadata of the primary cluster is forcibly synchronized to the secondary cluster in each synchronization period.

The following is an example of enabling forcible synchronization:

```
PUT _cluster/settings
{
  "persistent": {
    "remote_sync.force_synchronize": true
  }
}
```

6.10.3 Best Practices

This section describes how to switch from the primary cluster to the secondary cluster when the primary cluster is faulty.

1. If the synchronization of specified indexes has been configured between the primary and secondary clusters.
 - (1) Call the API for stopping index synchronization in the secondary cluster. In this case, the read and write traffic can be switched to the secondary cluster.
 - (2) After the primary cluster recovers, call the index synchronization API to synchronize data from the secondary cluster to the primary cluster.
2. If the matching pattern for index synchronization has been established between the primary and secondary clusters.
 - (1) Call the API for deleting the created matching pattern for index synchronization in the secondary cluster.
 - (2) Call the API for stopping index synchronization on the secondary cluster (using * for matching). In this case, the read and write traffic can be switched to the secondary cluster.
 - (3) After the primary cluster recovers, call the index synchronization API to synchronize data from the secondary cluster to the primary cluster.

7 Monitoring

7.1 Monitoring Metrics of Elasticsearch & OpenSearch Clusters

Function

This section describes CSS metrics that can be monitored by Cloud Eye as well as their namespaces and dimensions. You can use the management console or [APIs](#) provided by Cloud Eye to view the monitoring metrics and alarms generated for CSS.

Namespace

SYS.ES

Monitoring Metrics

- [Table 7-1](#) describes the monitoring metrics of CSS clusters.
- Monitored object: CSS clusters This section describes monitoring metrics of Elasticsearch clusters. For details about the monitoring metrics supported by Logstash clusters, see [Monitoring Metrics of Logstash Clusters](#).
- Monitoring period (original metric): 1 minute

NOTE

Accumulated value: The value is accumulated from the time when a node is started. After the node is restarted, the value is reset to zero and accumulated again.

Table 7-1 CSS metrics

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
status	Cluster Health Status	Health status of the monitored object	0,1,2,3 0: The cluster is 100% available. 1: The data is complete while some replicas are missing. Exceptions may occur because the high availability is compromised. This is a warning that should prompt investigation. 2: Data is missing and the cluster fails to work. 3: The cluster status is not obtained.	CSS cluster	1 minute
disk_util	Disk Usage	Disk usage of the monitored object. Unit: %	0-100%	CSS cluster	1 minute
max_jvm_heap_usage	Max. JVM Heap Usage	Maximum JVM heap usage of nodes in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
max_jvm_young_gc_time	Max. JVM Young GC Duration	Maximum accumulated JVM Young GC duration of nodes in a CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
max_jvm_young_gc_count	Max. JVM Young GC Count	Maximum accumulated JVM Young GC count of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_jvm_old_gc_time	Max. JVM Old GC Duration	Maximum accumulated JVM Old GC duration of nodes in a CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute
max_jvm_old_gc_count	Max. JVM Old GC Count	Maximum accumulated JVM Old GC count of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
total_fs_size	Total Size of File Systems	Total size of file systems in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
free_fs_size	Available Size of File Systems	Available size of file systems in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
max_cpu_usage	Max. CPU Usage	Maximum node CPU usage in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
max_cpu_time_of_jvm_process	Max. CPU Time of JVM Process	Maximum accumulated CPU usage duration of node JVM processes in a CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute
max_virtual_memory_size_of_jvm_process	Max. Virtual Memory Size of JVM Process	Maximum virtual memory size of node JVM processes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
max_current_opened_http_count	Current Max. Opened HTTP Connections	Maximum number of HTTP connections that are currently open for nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_total_opened_http_count	Total Max. Opened HTTP Connections	Maximum number of HTTP connections that were open for nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
indices_count	Indexes	Number of indexes in a CSS cluster	≥ 0	CSS cluster	1 minute
total_shards_count	Shards	Number of shards in a CSS cluster	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
primary_shards_count	Primary Shards	Number of primary shards in a CSS cluster	≥ 0	CSS cluster	1 minute
docs_count	Documents	Number of documents in a CSS cluster	≥ 0	CSS cluster	1 minute
docs_deleted_count	Deleted Documents	Number of documents deleted in a CSS cluster	≥ 0	CSS cluster	1 minute
nodes_count	Nodes	Number of nodes in a CSS cluster	≥ 0	CSS cluster	1 minute
data_nodes_count	Data Nodes	Number of data nodes in a CSS cluster	≥ 0	CSS cluster	1 minute
coordinating_nodes_count	Coordinating Nodes	Number of coordinating nodes in a CSS cluster	≥ 0	CSS cluster	1 minute
master_nodes_count	Master Nodes	Number of master nodes in a CSS cluster	≥ 0	CSS cluster	1 minute
ingest_nodes_count	Client Nodes	Number of client nodes in a CSS cluster	≥ 0	CSS cluster	1 minute
max_load_average	Max. Node Load	Maximum number of average queuing tasks per minute on nodes in a cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_cpu_usage	Avg. CPU Usage	Average node CPU usage in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
avg_load_average	Avg. Node Load	Average number of queuing tasks per minute on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_jvm_heap_usage	Avg. JVM Heap Usage	Average node JVM heap usage in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
max_open_file_descriptors	Max. Open File Descriptors	Maximum number of node file descriptors that are currently open in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_open_file_descriptors	Avg. Open File Descriptors	Average number of node file descriptors that are currently open in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_max_file_descriptors	Max. Allowed File Descriptors	Maximum number of allowed node file descriptors in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_open_file_descriptors	Open File Descriptors	Number of node file descriptors that are currently open in a cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_write_queue	Tasks in Write Queue	Number of tasks in queue for the write thread pool.	≥ 0	CSS cluster	1 minute
sum_thread_pool_search_queue	Tasks in Search Queue	Total number of tasks in queue for the search thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_force_merge_queue	Tasks in ForceMerge Queue	Total number of tasks in queue for the force merge thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_write_rejected	Rejected Tasks in Write Queue	Total number of tasks rejected by the write thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_search_rejected	Rejected Tasks in Search Queue	Total number of tasks rejected by the search thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_force_merge_rejected	Rejected Tasks in ForceMerge Queue	Total number of tasks rejected by the force merge thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_search_queue	Max. Tasks in Search Queue	Maximum number of tasks in queue for the search thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_force_merge_queue	Max. Tasks in ForceMerge Queue	Maximum number of tasks in queue for the force merge thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_write_threads	Size of Write Thread Pool	Total size of the write thread pools on nodes in the CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_search_threads	Size of Search Thread Pool	Total size of the search thread pools on nodes in the CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_force_merge_threads	Size of ForceMerge Thread Pool	Total size of the force merge thread pools on nodes in the CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_write_queue	Avg. Tasks in Write Queue	Average number of tasks in queue for the write thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_search_queue	Avg. Tasks in Search Queue	Average number of tasks in queue for the search thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_force_merge_queue	Avg. Tasks in ForceMerge Queue	Average number of tasks in queue for the force merge thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_thread_pool_search_threads	Avg. Size of Search Thread Pool	Average size of the search thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_write_threads	Avg. Size of Write Thread Pool	Average size of the write thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_force_merge_threads	Avg. Size of ForceMerge Thread Pool	Average size of the force merge thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_write_rejected	Avg. Rejected Tasks in Write Queue	Average number of tasks rejected by the write thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_flush_queue	Tasks in Flush Queue	Total number of tasks in queue for the flush thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_flush_rejected	Rejected Tasks in Flush Queue	Total number of tasks rejected by the flush thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_flush_queue	Max. Tasks in Flush Queue	Maximum number of tasks in queue for the flush thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_flush_threads	Size of Flush Thread Pool	Total size of the flush thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_flush_queue	Avg. Tasks in Flush Queue	Average number of tasks in queue for the flush thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_flush_threads	Avg. Size of Flush Thread Pool	Average number of tasks in queue for the flush thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_generic_queue	Tasks in Generic Queue	Total number of tasks in queue for the generic thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_generic_rejecte	Rejected Tasks in Generic Queue	Total number of tasks rejected by the generic thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_generic_queue	Max. Tasks in Generic Queue	Maximum number of tasks in queue for the generic thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_generic_threads	Size of Generic Thread Pool	Total size of the generic thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_generic_queue	Avg. Tasks in Generic Queue	Average number of tasks in queue for the generic thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_thread_pool_generic_threads	Avg. Size of Generic Thread Pool	Average number of tasks in queue for the generic thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_management_queue	Tasks in Management Queue	Total number of tasks in queue for the management thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_management_rejected	Rejected Tasks in Management Queue	Total number of tasks rejected by the management thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_management_queue	Max. Tasks in Management Queue	Maximum number of tasks in queue for the management thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_management_threads	Size of Management Thread Pool	Total size of the management thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_management_queue	Avg. Tasks in Management Queue	Average number of tasks in queue for the management thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_management_threads	Avg. Size of Management Thread Pool	Average number of tasks in queue for the management thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_refresh_queue	Tasks in Refresh Queue	Total number of tasks in queue for the refresh thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_refresh_rejected	Rejected Tasks in Refresh Queue	Total number of tasks rejected by the refresh thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
max_thread_pool_refresh_queue	Max. Tasks in Refresh Queue	Maximum number of tasks in queue for the refresh thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_refresh_threads	Size of Refresh Thread Pool	Total size of the refresh thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_refresh_queue	Avg. Tasks in Refresh Queue	Average number of tasks in queue for the refresh thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_refresh_threads	Avg. Size of Refresh Thread Pool	Average number of tasks in queue for the refresh thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_searcher_queue	Tasks in OBS Searcher Queue	Total number of tasks in queue for the OBS Searcher thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_obs_searcher_rejected	Rejected Tasks in OBS Searcher Queue	Total number of tasks rejected by the OBS Searcher thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_obs_searcher_queue	Max. Tasks in OBS Searcher Queue	Maximum number of tasks in queue for the OBS Searcher thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_searcher_threads	Size of OBS Searcher Thread Pool	Total size of the OBS Searcher thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_obs_searcher_queue	Avg. Tasks in OBS Searcher Queue	Average number of tasks in queue for the OBS Searcher thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_thread_pool_obs_searcher_threads	Avg. Size of OBS Searcher Thread Pool	Average number of tasks in queue for the OBS Searcher thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_queue	Tasks in OBS Queue	Total number of tasks in queue for the OBS thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_rejected	Rejected Tasks in OBS Queue	Total number of tasks rejected by the OBS thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_obs_queue	Max. Tasks in OBS Queue	Maximum number of tasks in queue for the OBS thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_threads	Size of OBS Thread Pool	Total size of the OBS thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_thread_pool_obs_queue	Avg. Tasks in OBS Queue	Average number of tasks in queue for the OBS thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_obs_threads	Avg. Size of OBS Thread Pool	Average number of tasks in queue for the OBS thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_upload_queue	Tasks in OBS Upload Queue	Total number of tasks in queue for the OBS Upload thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_upload_rejected	Rejected Tasks in OBS Upload Queue	Total number of tasks rejected by the OBS Upload thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
max_thread_pool_obs_upload_queue	Max. Tasks in OBS Upload Queue	Maximum number of tasks in queue for the OBS Upload thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_upload_threads	Size of OBS Upload Thread Pool	Total size of the OBS Upload thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_obs_upload_queue	Avg. Tasks in OBS Upload Queue	Average number of tasks in queue for the OBS Upload thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_obs_upload_threads	Avg. Size of OBS Upload Thread Pool	Average number of tasks in queue for the OBS Upload thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_thread_pool_obs_download_queue	Tasks in OBS Download Queue	Total number of tasks in queue for the OBS Download thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_download_rejected	Rejected Tasks in OBS Download Queue	Total number of tasks rejected by the OBS Download thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
max_thread_pool_obs_download_queue	Max. Tasks in OBS Download Queue	Maximum number of tasks in queue for the OBS Download thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_thread_pool_obs_download_threads	Size of OBS Download Thread Pool	Total size of the OBS Download thread pools on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_thread_pool_obs_download_queue	Avg. Tasks in OBS Download Queue	Average number of tasks in queue for the OBS Download thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_thread_pool_obs_download_threads	Avg. Size of OBS Download Thread Pool	Average number of tasks in queue for the OBS Download thread pool on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
min_free_fs_size	Min. Available Storage Space	Minimum available storage space of nodes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
avg_jvm_old_gc_count	Avg. GCs of Old-Generation JVM	Average number of old-generation garbage collections of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_jvm_ol d_gc_time	Avg. GC Duration of Old- Generation JVM	Average old- generation garbage collection duration of nodes in a CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute
avg_jvm_yo ung_gc_cou nt	Avg. GCs of Young- Generation JVM	Average number of young- generation garbage collections of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_jvm_yo ung_gc_tim e	Avg. GC Duration of Young- Generation JVM	Average young- generation garbage collection duration of nodes in a CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute
avg_max_fi le_descript ors	Avg. Maximum Allowed File Descriptors	Average value of the maximum number of allowed file descriptors on each node in a CSS cluster.	≥ 0	CSS cluster	1 minute
avg_mem_f ree_in_byte s	Avg. Available Memory	Average unused memory capacity of nodes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
avg_mem_free_percent	Avg. Available Memory Percentage	Average percentage of unused memory of nodes in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
avg_mem_used_in_bytes	Avg. Used Memory	Average used memory of nodes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
avg_mem_used_percent	Avg. Used Memory Percentage	Average percentage of used memory of nodes in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
max_mem_free_in_bytes	Max. Available Memory	Maximum unused memory of nodes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute
max_mem_free_percent	Max. Available Memory Percentage	Maximum percentage of unused memory of nodes in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
max_mem_used_in_bytes	Max. Used Memory	Maximum used memory of nodes in a CSS cluster. Unit: byte	≥ 0 bytes	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
max_mem_used_percent	Max. Used Memory Percentage	Maximum percentage of used memory of nodes in a CSS cluster. Unit: %	0-100%	CSS cluster	1 minute
sum_jvm_old_gc_count	Total GCs of Old-Generation JVM	Number of old-generation garbage collections of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_jvm_old_gc_time	Total GC Duration of Old-Generation JVM	Total old-generation garbage collection duration of nodes in the CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute
sum_jvm_young_gc_count	Total GCs of Young-Generation JVM	Number of young-generation garbage collections of nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_jvm_young_gc_time	Total GC Duration of Young-Generation JVM	Total young-generation garbage collection duration of nodes in the CSS cluster. Unit: ms	≥ 0 ms	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
sum_current_opened_http_count	Currently Open HTTP Connections	Number of HTTP connections that are open on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
sum_total_opened_http_count	Historical Open HTTP Connections	Number of HTTP connections that were open on nodes in a CSS cluster.	≥ 0	CSS cluster	1 minute
IndexingLatency	Average Index Latency	Average time required for a shard to complete an indexing operation. Unit: ms	≥ 0 ms	CSS cluster	1 minute
IndexingRate	Average Index Rate	Average number of index operations per second in a cluster.	≥ 0	CSS cluster	1 minute
SearchLatency	Average Search Latency	Average time required for a segment to complete the search operation. Unit: ms	≥ 0 ms	CSS cluster	1 minute
SearchRate	Average QPS	Average queries per second (QPS) in a cluster.	≥ 0	CSS cluster	1 minute

Metric ID	Metric	Description	Value Range	Monitored Object	Monitoring Interval (Raw Data)
task_max_running_time	Max. Task Runtime	Duration of the most time-consuming task in the cluster.	≥ 0 ms	CSS cluster	1 minute
number_of_pending_tasks	Number of Pending Tasks in Queue	Number of pending tasks to be processed by the master node in the CSS cluster.	≥ 0	CSS cluster	1 minute

Dimension

Table 7-2 Dimension description

Key	Value
cluster_id	CSS cluster

7.2 Monitoring Metrics of Logstash Clusters

Function

This section describes CSS metrics that can be monitored by Cloud Eye as well as their namespaces and dimensions. You can use the management console or [APIs](#) provided by Cloud Eye to view the monitoring metrics and alarms generated for CSS.

If the pipeline list in the configuration center of the Logstash cluster does not contain the data migration record of the configuration file, the corresponding monitoring record of the Logstash cluster is empty. Go to [Configuration Center](#) to create and start a configuration file for data migration.

Namespaces

SYS.ES

Monitoring Metrics

- [Table 7-3](#) describes the monitoring metrics of CSS clusters.
- Monitored object: CSS clusters This section describes monitoring metrics of Logstash clusters. For details about the monitoring metrics of Elasticsearch clusters, see [Monitoring Metrics of Elasticsearch & OpenSearch Clusters](#).
- Monitoring period (original metric): 1 minute

 **NOTE**

Accumulated value: The value is accumulated from the time when a node is started. After the node is restarted, the value is reset to zero and accumulated again.

Table 7-3 CSS metrics

Metric ID	Metric	Description	Value Range
jvm_heap_usage	JVM Heap Usage	Node JVM heap usage Unit: %	0-100%
jvm_old_gc_count	Total GCs of Old-Generation JVM	Number of old-generation garbage collections of nodes in a CSS cluster	≥ 0
jvm_old_gc_time	Total GC Duration of Old-Generation JVM	Old-generation garbage collection duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
jvm_young_gc_count	Total GCs of Young-Generation JVM	Number of young-generation garbage collections of nodes in a CSS cluster	≥ 0
jvm_young_gc_time	Total GC Duration of Young-Generation JVM	Young-generation garbage collection duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
max_jvm_heap_usage	Max. JVM Heap Usage	Maximum JVM heap usage of nodes in a CSS cluster Unit: %	0-100%

Metric ID	Metric	Description	Value Range
max_jvm_young_gc_time	Max. JVM Young GC Duration	Maximum accumulated JVM Young GC duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
max_jvm_young_gc_count	Max. JVM Young GC Count	Maximum accumulated JVM Young GC count of nodes in a CSS cluster	≥ 0
max_jvm_old_gc_time	Max. JVM Old GC Duration	Maximum accumulated JVM Old GC duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
max_jvm_old_gc_count	Max. JVM Old GC Count	Maximum accumulated JVM Old GC count of nodes in a CSS cluster.	≥ 0
max_cpu_usage	Max. CPU Usage	Maximum node CPU usage in a CSS cluster Unit: %	0-100%
cpu_usage	CPU Usage	CPU usage	0-100%
load_average	Average Load	Average number of queuing tasks per minute on a node	≥ 0
max_load_average	Max. Node Load	Maximum number of average queuing tasks per minute on nodes in a cluster	≥ 0
avg_cpu_usage	Avg. CPU Usage	Average node CPU usage in a CSS cluster. Unit: %	0-100%
avg_load_average	Avg. Node Load	Average number of queuing tasks per minute on nodes in a CSS cluster	≥ 0

Metric ID	Metric	Description	Value Range
avg_jvm_heap_usage	Avg. JVM Heap Usage	Average node JVM heap usage in a CSS cluster Unit: %	0-100%
avg_jvm_old_gc_count	Avg. GCs of Old-Generation JVM	Average number of old-generation garbage collections of nodes in a CSS cluster	≥ 0
avg_jvm_old_gc_time	Avg. GC Duration of Old-Generation JVM	Average old-generation garbage collection duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
avg_jvm_young_gc_count	Avg. GCs of Young-Generation JVM	Average number of young-generation garbage collections of nodes in a CSS cluster	≥ 0
avg_jvm_young_gc_time	Avg. GC Duration of Young-Generation JVM	Average young-generation garbage collection duration of nodes in a CSS cluster Unit: ms	≥ 0 ms
sum_events_in	Total Records Passed Through the Input Plug-in	Total number of records that have passed through the input plugin on all the nodes in a cluster	≥ 0
sum_events_filtered	Total Records Passed Through the Filtered Plug-in	Total number of records that have passed through the filtered plugin on all the nodes in a cluster	≥ 0
sum_events_out	Total Records Passed Through the Out Plug-in	Total number of records that have passed through the out plugin on all the nodes in a cluster	≥ 0

Metric ID	Metric	Description	Value Range
events_in	Records Passed Through the Input Plug-in on Node	Number of data records that have passed through the input plugin on the current node	≥ 0
events_filtered	Records Passed Through the Filtered Plug-in on Node	Number of records that have passed through the filtered plugin on the current node	≥ 0
events_out	Records Passed Through the Out Plug-in on Node	Number of records that have passed through the out plugin on the current node	≥ 0

Dimension

Table 7-4 Dimension description

Key	Value
cluster_id	CSS cluster

7.3 Monitoring Metrics

Function

This topic describes CSS metrics that can be monitored by Cloud Eye as well as their namespaces and dimensions. You can search for the monitoring metrics and alarms generated for CSS by using the Cloud Eye console or [calling APIs](#).

Namespace

SYS.ES

Monitoring Metrics

- [Table 7-5](#) describes the monitoring metrics of CSS clusters.
- Monitored object: Cloud service nodes of CSS clusters
- Monitoring period (original metric): 1 minute

NOTE

Accumulated value: The value is accumulated from the time when a node is started. After the node is restarted, the value is reset to zero and accumulated again.

Table 7-5 CSS metrics

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
jvm_heap_usage	JVM Heap Usage	JVM heap memory usage of a node. Unit: %	0-100 %	CSS cluster - cloud service node	1 minute
cpu_usage	CPU Usage	CPU usage. Unit: %	0-100 %	CSS cluster - cloud service node	1 minute
load_average	Average Load	Average number of queuing tasks per minute on a node	≥ 0	CSS cluster - cloud service node	1 minute
open_file_descriptors	Open File Descriptors	Number of opened file descriptors on a node	≥ 0	CSS cluster - cloud service node	1 minute
max_file_descriptors	Max. Allowed File Descriptors	Maximum number of allowed file descriptors	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_write_queue	Tasks in Write Queue	Number of tasks in queue for the write thread pool	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
thread_pool_search_queue	Tasks in Search Queue	Number of job queues in a search thread pool	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_force_merge_queue	Tasks in ForceMerge Queue	Number of job queues in a force merge thread pool	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_write_rejected	Rejected Tasks in Write Queue	Number of tasks rejected by the write thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_search_rejected	Rejected Tasks in Search Queue	Number of tasks rejected by the search thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_force_merge_rejected	Rejected Tasks in ForceMerge Queue	Number of tasks rejected by the force merge thread pool	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_write_threads	Size of Write Thread Pool	Size of a write thread pool	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_search_threads	Size of Search Thread Pool	Size of a search thread pool	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
thread_pool_force_merge_threads	Size of ForceMerge Thread Pool	Size of a force merge thread pool	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_flush_queue	Tasks in Flush Queue	Number of tasks in queue for the flush thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_flush_rejected	Rejected Tasks in Flush Queue	Number of tasks rejected by the flush thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_flush_threads	Size of Flush Thread Pool	Size of the flush thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_generic_queue	Tasks in Generic Queue	Number of tasks in queue for the generic thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_generic_rejected	Rejected Tasks in Generic Queue	Number of tasks rejected by the generic thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_generic_threads	Size of Generic Thread Pool	Size of the generic thread pool.	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
thread_pool_management_queue	Tasks in Management Queue	Number of tasks in queue for the management thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_management_rejected	Rejected Tasks in Management Queue	Number of tasks rejected by the management thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_management_threads	Size of Management Thread Pool	Size of the management thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_refresh_queue	Tasks in Refresh Queue	Number of tasks in queue for the refresh thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_refresh_rejected	Rejected Tasks in Refresh Queue	Number of tasks rejected by the refresh thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_refresh_threads	Size of Refresh Thread Pool	Size of the refresh thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_searcher_queue	Tasks in OBS Searcher Queue	Number of tasks in queue for the OBS Searcher thread pool.	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
thread_pool_obs_searcher_rejected	Rejected Tasks in OBS Searcher Queue	Number of tasks rejected by the OBS Searcher thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_searcher_threads	Size of OBS Searcher Thread Pool	Size of the OBS Searcher thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_queue	Tasks in OBS Queue	Number of tasks in queue for the OBS thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_rejected	Rejected Tasks in OBS Queue	Number of tasks rejected by the OBS thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_threads	Size of OBS Thread Pool	Size of the OBS thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_upload_queue	Tasks in OBS Upload Queue	Number of tasks in queue for the OBS Upload thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_upload_rejected	Rejected Tasks in OBS Upload Queue	Number of tasks rejected by the OBS Upload thread pool.	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
thread_pool_obs_upload_threads	Size of OBS Upload Thread Pool	Size of the OBS Upload thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_download_queue	Tasks in OBS Download Queue	Number of tasks in queue for the OBS Download thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_download_rejected	Rejected Tasks in OBS Download Queue	Number of tasks rejected by the OBS Upload thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
thread_pool_obs_download_threads	Size of OBS Download Thread Pool	Size of the OBS Download thread pool.	≥ 0	CSS cluster - cloud service node	1 minute
free_fs_size	Available Size of File Systems	Available size of file systems in a CSS cluster Unit: byte	≥ 0 bytes	CSS cluster - cloud service node	1 minute
total_fs_size	Total Size of File Systems	Total size of file systems in a CSS cluster Unit: byte	≥ 0 bytes	CSS cluster - cloud service node	1 minute
jvm_old_gc_count	Total GCs of Old-Generation JVM	Number of old-generation garbage collection times	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
jvm_old_gc_time	Total GC Duration of Old-Generation JVM	Old-generation garbage collection duration. Unit: ms	≥ 0 ms	CSS cluster - cloud service node	1 minute
jvm_young_gc_count	Total GCs of Young-Generation JVM	Number of young-generation garbage collection times	≥ 0	CSS cluster - cloud service node	1 minute
jvm_young_gc_time	GC Duration of Young-Generation JVM	Young-generation garbage collection duration. Unit: ms	≥ 0 ms	CSS cluster - cloud service node	1 minute
mem_free_in_bytes	Available Memory	Unused memory space of a node. Unit: byte	≥ 0 bytes	CSS cluster - cloud service node	1 minute
mem_free_percent	Available Memory Percentage	Percentage of unused memory space on a node.	≥ 0	CSS cluster - cloud service node	1 minute
mem_used_in_bytes	Used Memory	Used memory space of a node. Unit: byte	≥ 0 bytes	CSS cluster - cloud service node	1 minute
current_opened_http_count	Currently Open HTTP Connections	Number of HTTP connections on a node	≥ 0	CSS cluster - cloud service node	1 minute

Metric ID	Metric	Description	Value Range	Monitored Target	Monitoring Interval (Raw Data)
total_opened_http_count	Total Open HTTP Connections	Total number of HTTP connections on a node	≥ 0	CSS cluster - cloud service node	1 minute

Dimension

Table 7-6 Dimension description

Key	Value
cluster_id	CSS cluster

7.4 Configuring Cluster Monitoring

You can use Cloud Eye to monitor the created clusters. After configuring the cluster monitoring, you can log in to the Cloud Eye management console to view cluster metrics.

The procedure for configuring cluster monitoring:

1. **Creating Alarm Rules:** Customize alarm rules for the monitoring metrics. Once a metric exceeds the threshold, the system will notify you by sending emails or HTTP/HTTPS requests.
2. **Configuring Monitoring Metrics:** Configure monitoring metrics for a cluster or a node in the cluster.
3. **Viewing Monitoring Metrics:** View the statistics of the monitoring metrics in specific periods.

Prerequisites

- The cluster is in the **Available** or **Processing** status.
- The cluster has been running properly for more than 10 minutes.

Recommended Monitoring Metrics

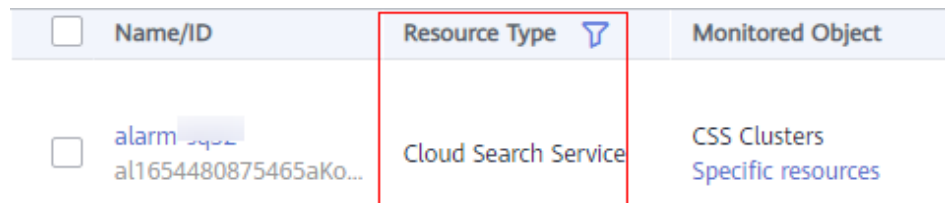
- Cluster CPU and JVM usage. You are advised to configure the following monitoring metrics: average JVM heap usage, maximum JVM heap usage, average CPU usage, and maximum CPU usage.

- Cluster write and query latency and throughput. You are advised to configure the following monitoring metrics: average index latency, average index rate, average search latency, and average QPS.
- Cluster write and query queue and rejected tasks. You are advised to configure the following monitoring metrics: tasks in write queue, tasks in search queue, rejected tasks in write queue, and rejected tasks in search queue.

Creating Alarm Rules

1. Log in to the Cloud Eye console.
2. In the navigation pane on the left, choose **Alarm Management > Alarm Rules**.
3. In the **Resource Type** column, select **Cloud Search Service** as criteria to search for alarm rules that meet the requirements.

Figure 7-1 Viewing alarm rules



If no alarm rules are available, create one by referring to [Creating an Alarm Rule](#). For details about how to set **Resource Type** and **Dimension**, see [Table 7-7](#).

Table 7-7 Alarm rule configuration parameter

Parameter	Description	Remark
Resource Type	Type of the resource that the alarm rule is created for	Select Cloud Search Service .
Dimension	Metric dimension of the selected resource type	CSS supports two dimensions. Select a dimension as required. <ul style="list-style-type: none"> • CSS Clusters: Alarm rules are specified by cluster. • CSS Clusters - CSS Instances: Alarm rules are specified by node in a cluster.

Configuring Monitoring Metrics

1. Create a monitoring panel by referring to [Creating a Dashboard](#). If an available monitoring panel has been created, skip this step.

2. Add CSS monitoring graphs by referring to [Adding a Graph](#).
For details about how to set **Resource Type** and **Dimension**, see [Table 7-8](#).

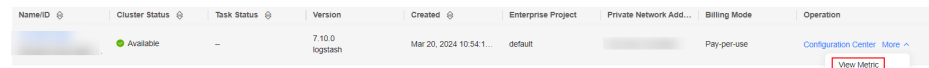
Table 7-8 Graph configuration parameter

Parameter	Description	Remark
Resource Type	Type of the resource to be monitored	Select Cloud Search Service .
Dimension	Metric dimension	<p>CSS supports two dimensions. Select a dimension as required.</p> <ul style="list-style-type: none"> • CSS Clusters: Monitoring is executed by cluster. • CSS Clusters - CSS Instances: Monitoring is executed by node in a cluster.

Viewing Monitoring Metrics

1. Log in to the CSS management console.
2. Choose **Clusters**. Locate the target cluster and choose **More > View Metric** in the **Operation** column.

Figure 7-2 Viewing metrics



3. Select a time range.
4. View the monitoring metrics.

8 Auditing

8.1 Key Operations Recorded by CTS

With CTS, you can record operations associated with CSS for later query, audit, and backtrack operations.

Prerequisites

CTS has been enabled.

Key Operations Recorded by CTS

Table 8-1 Key operations recorded by CTS

Operation	Resource Type	Event Name
Creating a cluster	cluster	createCluster
Deleting a cluster	cluster	deleteCluster
Expanding the cluster capacity	cluster	roleExtendCluster
Restarting a cluster	cluster	rebootCluster
Configuring a custom word dictionary	cluster	loadLexicon
Deleting a custom word dictionary	cluster	deleteLexicon
Performing basic configurations for a cluster snapshot	cluster	updateSnapshotPolicy
Setting the automatic snapshot creation policy	cluster	updateAutoSnapshotPolicy

Operation	Resource Type	Event Name
Upgrading a cluster	cluster	upgradeCluster
Retrying the upgrade	cluster	retryAction
Manually creating a snapshot	snapshot	createSnapshot
Restoring a snapshot	snapshot	restoreSnapshot
Deleting a snapshot	snapshot	deleteSnapshot

8.2 Querying Real-Time Traces

Scenarios

After you enable CTS and the management tracker is created, CTS starts recording operations on cloud resources. After a data tracker is created, the system starts recording operations on data in OBS buckets. CTS stores operation records generated in the last seven days.


This section describes how to query and export operation records of the last seven days on the CTS console.




- [Viewing Real-Time Traces in the Trace List of the New Edition](#)
- [Viewing Real-Time Traces in the Trace List of the Old Edition](#)

Constraints


- Traces of a single account can be viewed on the CTS console. Multi-account traces can be viewed only on the **Trace List** page of each account, or in the OBS bucket or the **CTS/system** log stream configured for the management tracker with the organization function enabled.
- You can only query operation records of the last seven days on the CTS console. To store operation records for more than seven days, you must configure an OBS bucket to transfer records to it. Otherwise, you cannot query the operation records generated seven days ago.
- After performing operations on the cloud, you can query management traces on the CTS console 1 minute later and query data traces on the CTS console 5 minutes later.



Viewing Real-Time Traces in the Trace List of the New Edition

1. Log in to the management console.
2. Click  in the upper left corner and choose **Management & Governance > Cloud Trace Service**. The CTS console is displayed.
3. Choose **Trace List** in the navigation pane on the left.
4. On the **Trace List** page, use advanced search to query traces. You can combine one or more filters.

- **Trace Name:** Enter a trace name.
 - **Trace ID:** Enter a trace ID.
 - **Resource Name:** Enter a resource name. If the cloud resource involved in the trace does not have a resource name or the corresponding API operation does not involve the resource name parameter, leave this field empty.
 - **Resource ID:** Enter a resource ID. Leave this field empty if the resource has no resource ID or if resource creation failed.
 - **Trace Source:** Select a cloud service name from the drop-down list.
 - **Resource Type:** Select a resource type from the drop-down list.
 - **Operator:** Select one or more operators from the drop-down list.
 - **Trace Status:** Select **normal**, **warning**, or **incident**.
 - **normal:** The operation succeeded.
 - **warning:** The operation failed.
 - **incident:** The operation caused a fault that is more serious than the operation failure, for example, causing other faults.
 - Time range: Select **Last 1 hour**, **Last 1 day**, or **Last 1 week**, or specify a custom time range.
5. On the **Trace List** page, you can also export and refresh the trace list, and customize the list display settings.
- Enter any keyword in the search box and press Enter to filter desired traces.
 - Click **Export** to export all traces in the query result as an .xlsx file. The file can contain up to 5000 records.
 - Click  to view the latest information about traces.
 - Click  to customize the information to be displayed in the trace list. If **Auto wrapping** is enabled (), excess text will move down to the next line; otherwise, the text will be truncated. By default, this function is disabled.
6. For details about key fields in the trace structure, see [Trace Structure](#) and [Example Traces](#).
7. (Optional) On the **Trace List** page of the new edition, click **Go to Old Edition** in the upper right corner to switch to the **Trace List** page of the old edition.

Viewing Real-Time Traces in the Trace List of the Old Edition

1. Log in to the management console.
2. Click  in the upper left corner and choose **Management & Governance > Cloud Trace Service**. The CTS console is displayed.
3. Choose **Trace List** in the navigation pane on the left.
4. Each time you log in to the CTS console, the new edition is displayed by default. Click **Go to Old Edition** in the upper right corner to switch to the trace list of the old edition.

5. Set filters to search for your desired traces. The following filters are available:
 - **Trace Type, Trace Source, Resource Type, and Search By:** Select a filter from the drop-down list.
 - If you select **Resource ID** for **Search By**, specify a resource ID.
 - If you select **Trace name** for **Search By**, specify a trace name.
 - If you select **Resource name** for **Search By**, specify a resource name.
 - **Operator:** Select a user.
 - **Trace Status:** Select **All trace statuses, Normal, Warning, or Incident.**
 - Time range: You can query traces generated during any time range in the last seven days.
 - Click **Export** to export all traces in the query result as a CSV file. The file can contain up to 5000 records.
6. Click **Query**.
7. On the **Trace List** page, you can also export and refresh the trace list.
 - Click **Export** to export all traces in the query result as a CSV file. The file can contain up to 5000 records.
 - Click  to view the latest information about traces.
8. Click  on the left of a trace to expand its details.

Trace Name	Resource Type	Trace Source	Resource ID	Resource Name	Trace Status	Operator	Operation Time	Operation
createDockerConfig	dockerlogincmd	SWR	--	dockerlogincmd	normal		Nov 16, 2023 10:54:04 GMT+08:00	View Trace

request

trace_id: [redacted]

code: 200

trace_name: createDockerConfig

resource_type: dockerlogincmd

trace_rating: normal

api_version:

message: createDockerConfig, Method: POST Url=/v2/manage/utlils/secret, Reason:

source_ip: [redacted]

domain_id: [redacted]

trace_type: ApiCall

9. Click **View Trace** in the **Operation** column. The trace details are displayed.

View Trace ×

```

{
  "request": "",
  "trace_id": "[redacted]",
  "code": "200",
  "trace_name": "createDockerConfig",
  "resource_type": "dockerlogincmd",
  "trace_rating": "normal",
  "api_version": "",
  "message": "createDockerConfig, Method: POST Url=/v2/manage/utlils/secret, Reason:",
  "source_ip": "[redacted]",
  "domain_id": "[redacted]",
  "trace_type": "ApiCall",
  "service_type": "SWR",
  "event_type": "system",
  "project_id": "[redacted]",
  "response": "",
  "resource_id": "",
  "tracker_name": "system",
  "time": "Nov 16, 2023 10:54:04 GMT+08:00",
  "resource_name": "dockerlogincmd",
  "user": {
    "domain": {
      "name": "[redacted]",
      "id": "[redacted]"
    }
  }
}

```

10. For details about key fields in the trace structure, see [Trace Structure](#) and [Example Traces](#) in the *CTS User Guide*.

11. (Optional) On the **Trace List** page of the old edition, click **New Edition** in the upper right corner to switch to the **Trace List** page of the new edition.