Cloud Search Service

FAQ

Issue 01

Date 2026-11-26



HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

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

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

Trademarks and Permissions


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

Notice

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

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

Huawei Cloud Computing Technologies Co., Ltd.

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

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

Contents

1 General Consulting....1
1.1 How Does CSS Ensure Data and Workload Security?....1
1.2 What Storage Options Are Available for a CSS Cluster?....1
1.3 What Files Are Stored on the Disks of a CSS Cluster?....2
1.4 What Data Compression Algorithms Does CSS Use?....2
2 Accessing CSS Clusters....4
2.1 How Do I Reset the Administrator Password of a Security-mode Cluster in CSS?....4
2.2 Are Ports 9200 and 9300 Open for Accessing Elasticsearch Clusters?....5
2.3 How Do I Use a NAT Gateway to Enable Public Network Access for an Elasticsearch/OpenSearch Cluster?....5
2.4 How Do I Connect In-house Developed Kibana to an Elasticsearch Cluster in CSS?....7
2.5 How Do I Connect In-house Developed OpenSearch Dashboards to an OpenSearch Cluster in CSS?....8
3 Migrating CSS Clusters....10
3.1 Can CSS Elasticsearch Clusters Be Migrated Across VPCs?....10
3.2 Can CSS Clusters Be Migrated Across Different Regions?....10
3.3 Example Configuration Files for Migrating Elasticsearch/OpenSearch Clusters Using CSS Logstash....11
3.4 Example Configuration File for Ingesting Data from DIS into an Elasticsearch or OpenSearch Cluster....13
3.5 Can I Export Data from Kibana in CSS?....17
3.6 Can I Export Data Using OpenSearch Dashboards in CSS?....18
3.7 What Should I Do If the Result of a Logstash Connectivity Test Is Unavailable?....20
3.8 Why Does Data Volume Differ After Migration?....21
4 Using CSS Cluster Search Engines....22
4.1 Why Are Newly Created Index Shards Allocated to a Single Node in CSS?....22
4.2 How Do I Create a Type Under an Index in an Elasticsearch 7.*x* Cluster of CSS?....23
4.3 How Do I Change the Number of Replicas for Elasticsearch Indexes in CSS?....23
4.4 What Are the Impacts If an Elasticsearch Cluster of CSS Has Too Many Shards?....24
4.5 How Do I Check the Number of Index Shards and Replicas in a CSS Elasticsearch Cluster?....24
4.6 What Does the Value i for node.roles Mean for Nodes in an Elasticsearch Cluster of CSS?....25
4.7 How Do I Change the Maximum Number of Results Returned in Response to a Single Search Request to a Specific Index in a CSS Elasticsearch Cluster?....26
4.8 How Do I Update Index Lifecycle Policies for an Elasticsearch Cluster of CSS?....26
4.9 How Do I Clear Elasticsearch Indexes in CSS?....28

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4.10 How Do I Clear Elasticsearch Cache in CSS?..... 29
4.11 Why Does Disk Usage Increase After Running delete_by_query to Delete Data in an Elasticsearch Cluster?..... 30
4.12 Do CSS Elasticsearch Clusters Support script dotProduct?..... 30
5 Managing CSS Clusters..... 32
5.1 How Do I Check the AZ or AZs of a CSS Cluster?..... 32
5.2 What Is the Relationship Between the Filebeat Version and Cluster Version in CSS?..... 32
5.3 How Do I Obtain the Security Certificate of a CSS Security-Mode Cluster?..... 32
5.4 How Do I Convert the Format of a CER Security Certificate in CSS?..... 33
5.5 Can I Change the Security Group of a CSS Cluster?..... 34
5.6 Can I Expand the Subnet for an Elasticsearch or OpenSearch Cluster?..... 35
5.7 How Do I Set search.max_buckets for a CSS Cluster?..... 35
5.8 How Do I Modify the TLS Algorithm for a CSS Cluster?..... 36
5.9 How Do I Enable Audit Logs for a CSS Cluster?..... 36
5.10 Can I Stop a CSS Cluster?..... 37
5.11 How Do I Query the Index Size on OBS After the Freezing of Indexes for a CSS Cluster?..... 38
5.12 How Do I Plan the Quantity of Index Shards for a Cluster?..... 38
6 CSS Cluster Backup and Restoration..... 40
6.1 How Do I Query the Snapshot Information of a CSS Cluster?..... 40
6.2 How Do I Back Up Large Volumes of Data Using Snapshots?..... 42
6.3 Can a Deleted CSS Cluster Be Restored?..... 46
7 CSS Cluster Monitoring and O&M..... 48
7.1 How Do I Set Slow Query Log Thresholds for an Elasticsearch Cluster of CSS?..... 48
7.2 What Do I Do If the Average Memory Usage of a CSS Cluster Reaches 98%?..... 48
7.3 How Do I Check the Total Disk Usage of a CSS Cluster?..... 49
7.4 Will CSS Cluster Services Be Affected If the Disk Usage of a Single Node Gets Too High?..... 49

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1 General Consulting

1.1 How Does CSS Ensure Data and Workload Security?

CSS uses network isolation in addition to various host and data security measures.

- Network isolation

  The entire network is divided into two planes: service plane and management plane. The two planes are deployed and isolated physically to ensure the security of the service and management networks.

  - Service plane: This is the network plane of the cluster. It provides service channels for users and delivers data definitions, indexing, and search capabilities.
  - Management plane: This is the management console, where you manage CSS.

- Host security

  CSS provides the following security measures:

  - The VPC security group ensures the security of the hosts in a VPC.
  - Network access control lists (ACLs) allow you to control what data can enter or exit your network.
  - The internal security infrastructure (including the network firewall, intrusion detection system, and protection system) monitors all network traffic that enters or exits the VPC through an IPsec VPN.

- Data security

  CSS uses multiple replicas, cross-AZ deployment of clusters, and third-party (OBS) index backup to ensure user data security.

1.2 What Storage Options Are Available for a CSS Cluster?

CSS uses EVS and local disks to store your indexes. During cluster creation, you can specify the EVS disk type and specifications (the EVS disk size).

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- EVS disk types: common I/O, high I/O, ultra-high I/O, and extreme SSD.
- The EVS disk size varies depending on the node specifications you selected during cluster creation.

When creating a cluster, you can attach a certain number of EVS disks to each node (corresponding to one ECS). The size of the EVS disks depends on the selected node specifications. You can calculate the cluster's total storage capacity based on the number of nodes and the size of the EVS disks attached to each node.

1.3 What Files Are Stored on the Disks of a CSS Cluster?

You can store the following logs and files:

- Log files: cluster logs
- Data files: cluster index files
- Other files: cluster configuration files
- OS: 5% of the storage capacity is reserved for the OS by default.

1.4 What Data Compression Algorithms Does CSS Use?

CSS supports two data compression algorithms: LZ4 (by default) and best_compression.

- **LZ4 algorithm**

  LZ4 is the default compression algorithm for Elasticsearch. This algorithm can compress and decompress data quickly, but its compression ratio is low.

  How this algorithm works: During compression, the algorithm uses a minimum match length of 4 bytes and a hash table to find matching content in the previously encoded data that can be referenced. It scans the input data forward one byte at a time. When it finds duplicate data, it encodes the match as a back-reference instead of storing the data again. This algorithm is well-suited where reads far exceed writes in terms of data volumes.

- **best_compression algorithm**

  This algorithm can be used when a large amount of data is written and the index storage cost is high, such as logs and time series analysis. This algorithm can greatly reduce the index storage cost.

Run the following command to switch the default compression algorithm (LZ4) to best_compression:

```
PUT index-1
{
   "settings": {
      "index": {
         "codec": "best_compression"
      }
   }
}
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

LZ4 is faster in both compression and decompression, while best_compression achieves higher compression ratios but has a slower decompression speed.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2 Accessing CSS Clusters

2.1 How Do I Reset the Administrator Password of a Security-mode Cluster in CSS?

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

1. Log in to the **CSS management console**.
2. In the navigation pane, choose **Clusters** > **Elasticsearch** or **Clusters** > **OpenSearch**.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Overview** tab.
5. In the **Network Information** area, click **Reset** under **Reset Password**. Set and confirm the new administrator password.
   - 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 allowed: **~!@#$%^&*()-_=+\|[{}];:,<.>/?**
   - Do not use the administrator name, or the administrator name spelled backwards, as the password.
   - You are advised to change the password periodically.

Figure 2-1 Reset Password



 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

2.2 Are Ports 9200 and 9300 Open for Accessing Elasticsearch Clusters?

Yes. Port 9200 is used by external systems to access CSS Elasticsearch clusters, and port 9300 is used for communication between nodes.

The methods for accessing port 9300 are as follows:

- If your client is in the same VPC and subnet with the CSS cluster, you can access it directly.
- If your client is in the same VPC but different subnet from the CSS cluster, apply for a route separately.
- If your client and the CSS cluster are in different VPCs and subnets, you must first create a VPC peering connection to connect the two VPCs. Then, configure a route to connect the two subnets.

2.3 How Do I Use a NAT Gateway to Enable Public Network Access for an Elasticsearch/OpenSearch Cluster?

Perform the following operations:

1. **Obtaining Cluster Information**
2. **Configuring a NAT Gateway**
3. **Modifying Security Group Rules for the Cluster**
4. **Accessing a Cluster over the Public Network**

CAUTION

Do not configure a NAT gateway for CSS clusters with security mode disabled. Doing so exposes your data to the public internet.

Obtaining Cluster Information

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

Step 2 In the navigation pane, choose **Clusters** > **Elasticsearch** or **Clusters** > **OpenSearch**.

Step 3 In the cluster list, click the name of the target cluster. The cluster information page is displayed.

Step 4 Click the **Overview** tab.

Step 5 In the **Basic Information** area, obtain the cluster's **Region**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Step 6 In the **Network Information** area, obtain the cluster's **VPC**, **Current Subnet**, and **Private IPv4 Address**.

----End

Configuring a NAT Gateway

Step 1 Create a public NAT gateway to enable public network access for the current cluster.

For details, see **Buying a Public NAT Gateway**. **Table 2-1** describes the key parameters. Set other parameters based on service requirements.

Table 2-1 Configuring a public NAT gateway

Parameter	Description
Region	Use the region of the Elasticsearch/OpenSearch cluster.
VPC	Use the VPC of the Elasticsearch/OpenSearch cluster.
Subnet	Use the subnet of the Elasticsearch/OpenSearch cluster.

Step 2 After a public NAT gateway is created, add DNAT rules to allow the cluster in your VPC to provide services accessible from the Internet.

For details, see **Adding a DNAT Rule**. **Table 2-2** describes the key parameters. Set other parameters based on service requirements.

Table 2-2 Adding a DNAT rule

Parameter	Description
Public IP Address Type	Select **EIP**. Remember the configured IP address, which will be needed for accessing the cluster from the public network.
Public Port	A custom port can be configured. Remember the configured port, which will be needed for accessing the cluster from the public network.
Private IP Address	Enter the cluster's private IPv4 address obtained in **Obtaining Cluster Information**.
Private Port	Enter 9200.

CAUTION

If the cluster has multiple private IPv4 addresses, add multiple DNAT rules.

----End

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Modifying Security Group Rules for the Cluster

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

Step 2 In the navigation pane, choose **Clusters** > **Elasticsearch** or **Clusters** > **OpenSearch**.

Step 3 In the cluster list, click the name of the target cluster. The cluster information page is displayed.

Step 4 Click the **Overview** tab.

Step 5 In the **Network Information** area, find **Security Group**, and click the security group name to go to its details page.

Step 6 Click the **Inbound Rules** tab.

Step 7 Click **Add Rule** to add an inbound rule to allow port 9200.

Step 8 Click **OK**.

----End

Accessing a Cluster over the Public Network

Enter **https://{IP}:{port}** or **http://{IP}:{port}** in the browser address box to access the Elasticsearch or OpenSearch cluster.

- ***IP*** and ***port*** are the EIP and port you set when you added DNAT rules.
- If you have enabled **Security Mode** for the cluster, enter **https://{IP}:{port}** and then enter the username and password for the cluster.
- If you have not enabled **Security Mode** for the cluster, enter **http://{IP}:{port}**.

2.4 How Do I Connect In-house Developed Kibana to an Elasticsearch Cluster in CSS?

Constraints

Only Kibana images of the OSS version can be connected to Elasticsearch clusters in CSS.

Procedure

1. Create an ECS.
   - The ECS must be within the same VPC as the CSS cluster.
   - Port 5601 must be allowed by the security group associated with the ECS.
   - An EIP must be allocated to the ECS.

   For details, see **Purchasing an ECS**.
2. Obtain the address for accessing the Elasticsearch cluster of CSS.
   a. Log in to the **CSS management console**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

b. In the navigation pane on the left, choose **Clusters > Elasticsearch**.

c. In the cluster list, obtain the target cluster's internal network address from the **Internal Network Address** column. Typical address format: *<host>*:*<port>* or *<host>*:*<port>*,*<host>*:*<port>*. Example: **10.62.179.32:9200,10.62.179.33:9200**.

   If the cluster has client nodes, only the IP addresses and ports of all the client nodes are displayed. Otherwise, the IP addresses and ports of all data nodes and cold data nodes are displayed.

3. Install Kibana on the ECS and modify the **kibana.yml** configuration file.

   – The following is an example of the configuration file for a security-mode cluster:

```
elasticsearch.username: "***"    //Username of the security cluster
elasticsearch.password: "***"    //Password of the security cluster
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   //IP address or DNS name of the Kibana server. localhost is recommended.
elasticsearch.hosts: http://10.0.0.xxx:9200   //Address for accessing the Elasticsearch cluster
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
```

   CAUTION

   To access a security-mode cluster, the opendistro_security_kibana plug-in must be installed. For details, see **security-kibana-plugin**. The plug-in version must be the same as that of the cluster. To check the plug-in version, run the **GET _cat/plugins** command.

   – The following is an example of the configuration file for a non-security mode cluster:

```
server.port: 5601
logging.dest: /home/Ruby/log/kibana.log
pid.file: /home/Ruby/run/kibana.pid
server.host: 192.168.xxx.xxx   //IP address or DNS name of the Kibana server. localhost is recommended.
elasticsearch.hosts: http://10.0.0.xxx:9200   //Address for accessing the Elasticsearch cluster
```

4. Use a browser on your local PC to connect to the EIP associated with the ECS. The URL is **http://EIP:5601**. Log in to Kibana to access the Elasticsearch cluster.

2.5 How Do I Connect In-house Developed OpenSearch Dashboards to an OpenSearch Cluster in CSS?

Constraints

Only OpenSearch Dashboards images of the OSS version can be connected to OpenSearch clusters in CSS.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Procedure

1. Create an ECS.
   - The ECS must be within the same VPC as the CSS cluster.
   - Port 5601 must be allowed by the security group associated with the ECS.
   - An EIP must be allocated to the ECS.

   For details, see **Purchasing an ECS**.
2. Obtain the address for accessing the OpenSearch cluster of CSS.
   a. Log in to the **CSS management console**.
   b. In the navigation pane on the left, choose **Clusters > OpenSearch**.
   c. In the cluster list, obtain the target cluster's internal network address from the **Internal Network Address** column. Typical address format: *<host>*:*<port>* or *<host>*:*<port>*,*<host>*:*<port>*. Example: **10.62.179.32:9200,10.62.179.33:9200**.

      If the cluster has client nodes, only the IP addresses and ports of all the client nodes are displayed. Otherwise, the IP addresses and ports of all data nodes and cold data nodes are displayed.
3. Install OpenSearch Dashboards on the ECS and modify the **opensearch_dashboards.yml** configuration file.
   - The following is an example of the configuration file for a security-mode cluster:

     ```
     opensearch.username: "***" //Username of the security cluster
     opensearch.password: "***" //Password of the security cluster
     opensearch.ssl.verificationMode: none
     server.ssl.enabled: false
     server.rewriteBasePath: false
     server.port: 5601
     logging.dest: /home/Ruby/log/opensearch-dashboards.log
     pid.file: /home/Ruby/run/opensearch-dashboards.pid
     server.host: 192.168.xxx.xxx  //IP address or DNS name of the OpenSearch Dashboards server. localhost is recommended.
     opensearch.hosts: http://10.0.0.xxx:9200   //Address for accessing the OpenSearch cluster
     opensearch.requestHeadersWhitelist: ["securitytenant","Authorization"]
     opensearch_security.multitenancy.enabled: true
     opensearch_security.multitenancy.tenants.enable_global: true
     opensearch_security.multitenancy.tenants.enable_private: true
     opensearch_security.multitenancy.tenants.preferred: ["Private", "Global"]
     opensearch_security.multitenancy.enable_filter: false
     ```

     NOTE

     To access a security-mode cluster, the opensearch_security_dashboards plug-in must be installed. For details, see **security-dashboards-plugin**. The plug-in version must be the same as that of the cluster. To check the plug-in version, run the **GET _cat/plugins** command.
   - The following is an example of the configuration file for a non-security mode cluster:

     ```
     server.port: 5601
     logging.dest: /home/Ruby/log/opensearch-dashboards.log
     pid.file: /home/Ruby/run/opensearch-dashboards.pid
     server.host: 192.168.xxx.xxx  //IP address or DNS name of the OpenSearch Dashboards server. localhost is recommended.
     opensearch.hosts: http://10.0.0.xxx:9200   //Address for accessing the OpenSearch cluster
     ```
4. Use a browser on your local PC to connect to the EIP associated with the ECS. The URL is **http://EIP:5601**. Log in to OpenSearch Dashboards to access the OpenSearch cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3 Migrating CSS Clusters

3.1 Can CSS Elasticsearch Clusters Be Migrated Across VPCs?

Use either of the following methods to migrate Elasticsearch clusters across different VPCs:

Method 1

Use the backup and restoration function to migrate cluster data. For details, see **Index Backup and Restoration**.

Method 2

1. Connect the VPC network and establish a VPC peering connection. For details, see **VPC Peering Connection Overview**.
2. After the network is connected, use Logstash to migrate data.

3.2 Can CSS Clusters Be Migrated Across Different Regions?

In CSS, perform the following to migrate Elasticsearch clusters across different regions:

- If the OBS bucket is in the same region as your CSS cluster, migrate the cluster by following the instructions in **Index Backup and Restoration**.
- If the OBS bucket is not in the same region as your CSS cluster, **configure cross-region replication** to back up the cluster to the bucket, and migrate the cluster by following the instructions in **Index Backup and Restoration**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

- Before cross-region replication, ensure the snapshot folder of the destination cluster is empty. Otherwise, the snapshot information cannot be updated to the snapshot list of the destination cluster.
- Before every migration, ensure the folder is empty.

3.3 Example Configuration Files for Migrating Elasticsearch/OpenSearch Clusters Using CSS Logstash

In our example, both the source and destination are Elasticsearch clusters created in CSS, and the migration is performed using the CSS-hosted Logstash service on Elasticsearch or OpenSearch clusters.

Checking the Cluster Type

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, choose **Clusters** > **Elasticsearch**.
3. In the cluster list, find the source or destination Elasticsearch cluster, and click the cluster name to go to the cluster information page.
4. Click the **Overview** tab. In the **Network Information** area, check whether **Security Mode** and **HTTPS Access** are enabled. **Figure 3-1** shows an Elasticsearch cluster with the security mode and HTTPS both enabled.

Figure 3-1 Checking a cluster's security settings



Table 3-1 Example Logstash configuration files for different types of clusters

Scenario	Example Logstash Configuration File
Migrating data between non-security mode clusters	**Example Logstash Configuration File for Non-Security Mode Clusters**
Migrating data between security-mode clusters that use HTTP	**Example Logstash Configuration File for Security-Mode Clusters That Use HTTP**
Migrating data between security-mode clusters that use HTTPS	**Example Logstash Configuration File for Security-Mode Clusters That Use HTTPS**

Depending on the security-mode settings of the source and destination clusters, modify the input and output modules provided in our examples to obtain the configuration file you need.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Example Logstash Configuration File for Non-Security Mode Clusters

The following is an example Logstash configuration file when security mode is disabled for both the source and destination Elasticsearch or OpenSearch clusters.

```
input {
    elasticsearch {
        # Source cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Name of the source index to be migrated
        index => "xxx,xxx"
        docinfo => true
    }
}

filter {
    # Removes some metadata fields automatically added by Logstash
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}

output {
    elasticsearch {
        # Destination cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Target index for writing events. The following configuration retains the source index name,
document type, and document IDs.
        index => "%{[@metadata][_index]}"
        document_type => "%{[@metadata][_type]}"
        document_id => "%{[@metadata][_id]}"
    }
}
```

Example Logstash Configuration File for Security-Mode Clusters That Use HTTP

The following is an example Logstash configuration file for when security mode is enabled for both the source and destination Elasticsearch or OpenSearch clusters but HTTPS is disabled for them.

```
input {
    elasticsearch {
        # Source cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Name of the source index to be migrated
        index => "xxx,xxx"
        docinfo => true
        user => "xxx"           # Username for accessing the cluster
        password => "xxx"       # Password corresponding to the username
    }
}

filter {
    # Removes some metadata fields automatically added by Logstash
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}

output {
    elasticsearch {
        # Destination cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Target index for writing events. The following configuration retains the source index name,
document type, and document IDs.
        index => "%{[@metadata][_index]}"
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
        document_type => "%{[@metadata][_type]}"
        document_id => "%{[@metadata][_id]}"
        user => "xxx"           # Username for accessing the cluster
        password => "xxx"       # Password corresponding to the username
    }
}
```

Example Logstash Configuration File for Security-Mode Clusters That Use HTTPS

The following is an example Logstash configuration file for when security mode and HTTPS are enabled for both the source and destination Elasticsearch or OpenSearch clusters.

```
input {
    elasticsearch {
        # Source cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Name of the source index to be migrated
        index => "xxx,xxx"
        docinfo => true
        user => "xxx"           # Username for accessing the cluster
        password => "xxx"       # Password corresponding to the username
        ssl => true             # Enable SSL connections.
        ca_file => "/opt/logstash/extend/certs"     # Path of the CA certificate used to verify the source cluster
    }
}

filter {
    # Removes some metadata fields automatically added by Logstash
    mutate {
        remove_field => ["@timestamp", "@version"]
    }
}

output {
    elasticsearch {
        # Destination cluster node addresses. No need to include the protocol.
        hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
        # Target index for writing events. The following configuration retains the source index name,
document type, and document IDs.
        index => "%{[@metadata][_index]}"
        document_type => "%{[@metadata][_type]}"
        document_id => "%{[@metadata][_id]}"
        user => "xxx"           # Username for accessing the cluster
        password => "xxx"       # Password corresponding to the username
        ssl => true             # Enable SSL connections.
        cacert => "/opt/logstash/extend/certs"       # Path of the CA certificate used to verify the destination
cluster
        ssl_certificate_verification => false        # Whether to enable SSL certificate verification for the
destination cluster.
    }
}
```

3.4 Example Configuration File for Ingesting Data from DIS into an Elasticsearch or OpenSearch Cluster

In our example, the source is Huawei Cloud's Data Ingestion Service (DIS) and the destination is an Elasticsearch or OpenSearch cluster created in CSS, and the migration is performed using the CSS-hosted Logstash service.

```
# logstash-input-dis
input {
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
  dis {
    # DIS stream name
    streams => ["YOUR_DIS_STREAM_NAME"]
    # Endpoint URL for the region where your DIS service is located
    endpoint => "https://dis.example.com"
    # User's AK/SK
    ak => "YOUR_ACCESS_KEY_ID"
    sk => "YOUR_SECRET_KEY_ID"
    # Region where your DIS service is located
    region => "YOUR_Region"
    # Project ID
    project_id => "YOUR_PROJECT_ID"
    # DIS application name
    group_id => "YOUR_APP_ID"
    # Client ID
    client_id => "YOUR_CLIENT_ID"
    # Start position for data consumption in the stream
    auto_offset_reset => "earliest"
  }
}

filter {
  # Removes some metadata fields automatically added by Logstash
  mutate {
    remove_field => ["@timestamp", "@version"]
  }
}

# logstash-output-elasticsearch
output {
  elasticsearch {
    # Destination cluster node addresses. No need to include the protocol.
    hosts => ["xxx.xxx.xxx.xxx:9200", "xxx.xxx.xxx.xxx:9200"]
    # Name of the index that events are written into.
    index => "xxx"
    # Mandatory fields for a security-mode cluster. (Delete them for a cluster with the security mode
disabled.)
    # user => "xxx"          # Username for accessing the cluster.
    # password => "xxx"      # Password corresponding to the username.
    # If SSL is enabled for the destination cluster, additionally configure the following information:
    # ssl => true
    # cacert => "/opt/logstash/extend/certs"      # Path of the CA certificate used to verify the destination
cluster.
    # ssl_certificate_verification => false       # Whether to enable SSL certificate verification for the
destination cluster.
  }
}
```

Table 3-2 Configuration items

Configuration Item	Mandatory	Description
logstash-input-dis		
streams	Yes	DIS stream name. The value must match the stream name specified on the DIS console during stream creation.
endpoint	Yes	Endpoint for the region where your DIS service is located

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configuration Item	Mandatory	Description
ak	Yes	The user's Access Key (AK). For details, see **Checking Authentication Information**.
sk	Yes	The user's Secret Key (SK). For details, see **Checking Authentication Information**.
region	Yes	Region where your DIS service is located
project_id	Yes	Project ID. 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.
client_id	No	Client ID, which identifies a consumer in a consumer group. If multiple pipelines or Logstash instances are present, configure unique values for different consumers. For example, the value of instance 1 is **client1**, and the value of instance 2 is **client2**.
auto_offset_reset	No	Start position for data consumption in the stream. The value can be: • **earliest**: Starts consuming from the oldest messages. • **latest**: Starts consuming from the most recent messages.
logstash-output-elasticsearch		
hosts	Yes	Destination cluster node addresses. You can configure multiple IP addresses. Value format: ["<Node IP address 1>:<Port number>", "<Node IP address 2>:<Port number>"]

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configuration Item	Mandatory	Description
index	Yes	Name of the index to which events are written. • Single index: Enter the index name, for example, **my_index**. • Multiple indexes: Use dynamic naming (based on event fields) or configure multiple conditional output blocks to route events to different indexes.
user	No	Username for accessing the destination cluster. Mandatory for a security-mode cluster.
password	No	Password for accessing the destination cluster. Mandatory for a security-mode cluster.
ssl	No	Whether SSL is enabled for the destination cluster. The value can be: • **true**: Uses HTTPS to transmit data. • **false**: Uses HTTP to transmit data.
cacert	No	Path of the CA certificate used to verify the destination cluster. Value format: <certificate path><certificate name>, for example, /opt/logstash/extend/certs. • If the destination is a CSS Elasticsearch or OpenSearch cluster, the **certificate name** and **certificate path** of the default CA certificate will be obtained. For details, see **Viewing Default Certificates**. • If the destination is a self-managed or third-party Elasticsearch or OpenSearch cluster, upload the destination cluster's security certificate to Logstash and obtain the **certificate name** and **certificate path**. For details, see **Uploading a Custom Certificate**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Configuration Item	Mandatory	Description
ssl_certificate_verification	No	Whether SSL certificate verification is enabled for the destination cluster. The value can be: • **true** (default): Uses an SSL certificate to verify the destination cluster. • **false**: Ignores SSL certificate verification.

For more information, see **Configuring the DIS Logstash Plugin** as well as **Elasticsearch output plugin** from the official Logstash documentation.

3.5 Can I Export Data from Kibana in CSS?

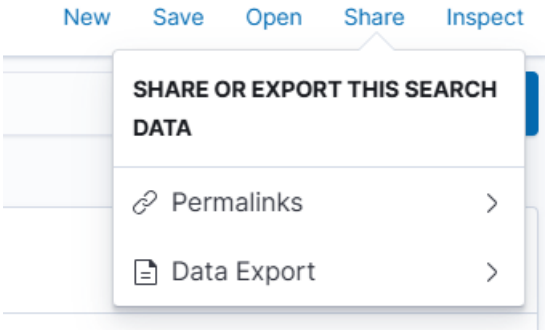
- With Elasticsearch 7.10.2 (with an image version no earlier than 7.10.2_24.3.0_x.x.x), Kibana supports one-click data export in CSV.

  CAUTION

  - A maximum of 10 MB of data can be exported. If the data you want to export exceeds 10 MB, only the first 10 MB is exported.
  - Any special characters such as **=+-@** in exported CSV files may be identified as part of some formulas, leading to data export failures.

  On the **Discover** page of Kibana, click **Save** in the upper right corner to save the data. Then, choose **Share** > **Data Export** and select **Export CSV**.

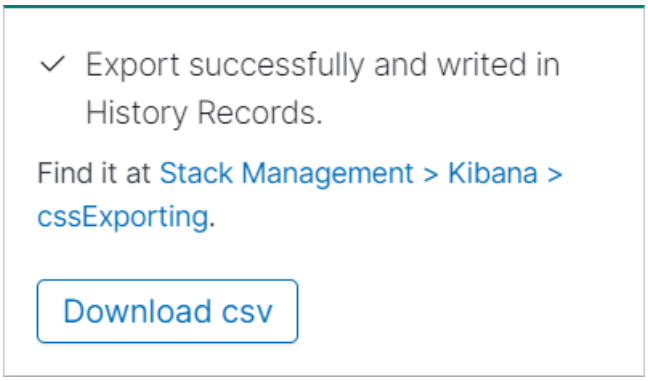
  Figure 3-2 Exporting data

  Wait for a few minutes. Then click **Download csv** in the lower right corner to download the data to the local PC.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 3-3 Downloading data

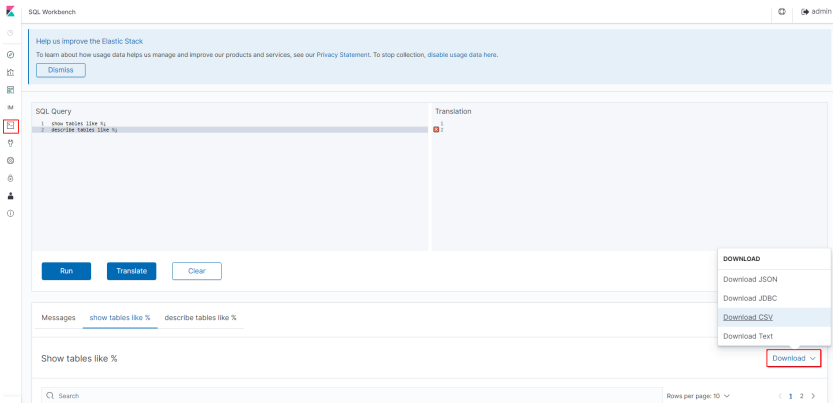


- With Elasticsearch 7.6.2, 7.9.3, and 7.10.2 (the image version is earlier than 24.3.0), the SQL Workbench plug-in is required for exporting data on Kibana.

  In **SQL Workbench** of Kibana, you can run Elasticsearch SQL statements to query data or click **Download** to export data. You can export 1 to 200 records. By default, 200 records are exported.

  For details about the Elasticsearch SQL statements used, see **Elasticsearch SQL**.

Figure 3-4 SQL Workbench



3.6 Can I Export Data Using OpenSearch Dashboards in CSS?

Only OpenSearch 2.19.0 and 3.4.0 clusters allow you to export charts and tables through OpenSearch Dashboards.

CAUTION

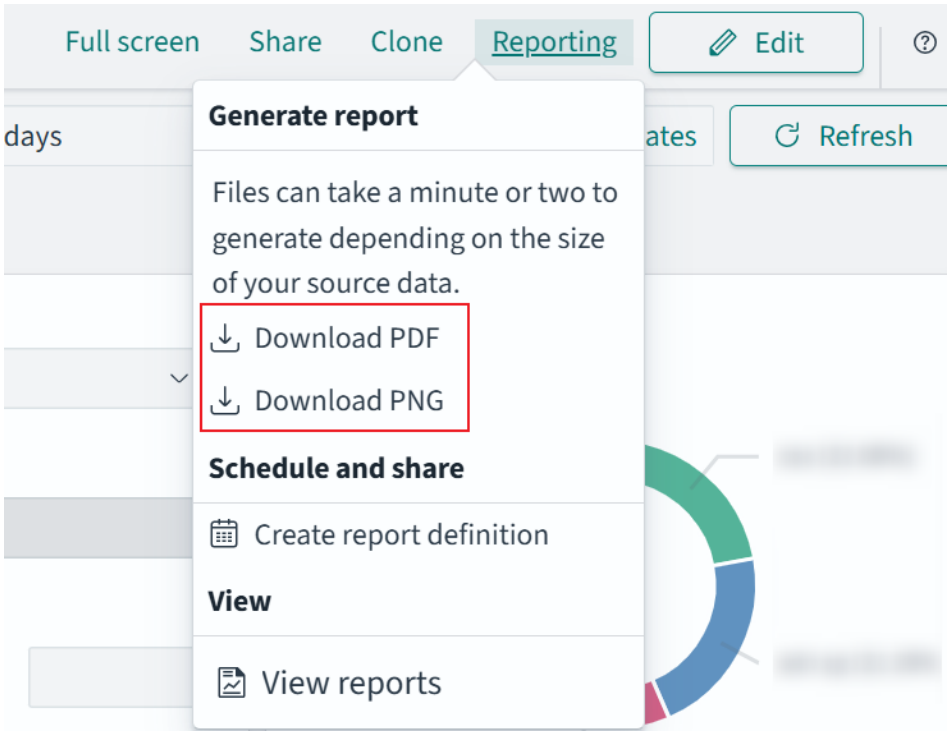
- A maximum of 10 MB of data can be exported. If the data you want to export exceeds 10 MB, only the first 10 MB is exported.
- Any special characters such as **=+-@** in exported CSV files may be identified as part of some formulas, leading to data export failures.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Exporting a dashboard**
  a. Log in to the OpenSearch Dashboards.
  b. Choose **OpenSearch Dashboards > Dashboards**. The **Dashboards** page is displayed.
  c. Find the dashboard you want to export and click its title. The details page is displayed.
  d. Click **Reporting** in the upper right corner, and select a download format. Once the report is generated, select a save location on your local PC to download it.

     Save the dashboard before exporting it, or an error will be returned.

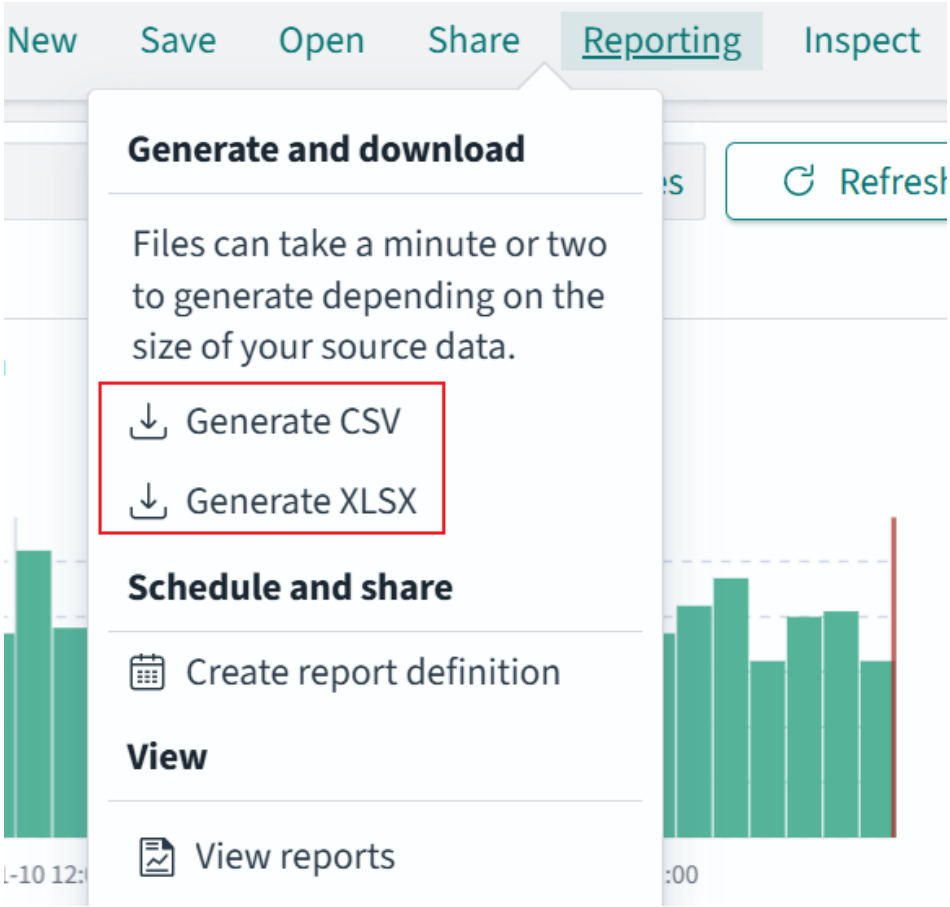
     Figure 3-5 Selecting an export file format



- **Exporting data**
  a. Log in to the OpenSearch Dashboards.
  b. Choose **OpenSearch Dashboards > Discover**. The **Discover** page is displayed.
  c. Click **Save** in the upper right corner to save search data, then click **Reporting**. Select a download format, and once the report is generated, select a save location on your local PC to download it.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 3-6 Selecting an export file format



3.7 What Should I Do If the Result of a Logstash Connectivity Test Is Unavailable?

When using a CSS Logstash cluster to migrate data, if the connectivity test between Logstash and the data source or destination returns Unavailable, the most probable cause is incorrect network configuration. This section describes how to check network configurations in different environments to resolve connectivity issues.

- Within the same VPC

  If Logstash and the data source or destination are in the same VPC, the following conditions must be met:

  - The security group used by the Logstash cluster allows outbound traffic for all protocols.
  - The security group used by the Logstash cluster allows inbound traffic for the ICMP protocol as well as the TCP ports used by the peer addresses.

  For details, see **Adding a Security Group Rule**.

- Within different VPCs

  If Logstash and the data source or destination are in different VPCs, the following conditions must be met:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

  - The security group used by the Logstash cluster allows outbound traffic for all protocols.
  - The security group used by the Logstash cluster allows inbound traffic for the ICMP protocol as well as the TCP ports used by the peer addresses.
  - A route has been configured between the Logstash cluster and the peer VPC.

  For details, see **Adding a Security Group Rule** and **Configuring Routes for a Logstash Cluster**.

- Within different LANs (they must communicate across the public network):

  If Logstash and the data source or destination are in different LANs, the following condition must be met: A public IP address has been configured for the Logstash cluster to ensure it can be accessed from the Internet.

  For details, see **Configuring Public Network Access**.

3.8 Why Does Data Volume Differ After Migration?

After migrating data using tools like Logstash, it's normal for the data volume in the source and destination clusters to differ. We recommend comparing document counts first — differences in volumes do not affect data integrity.

This difference may be attributed to the following factors:

- Elasticsearch storage mechanism

  CSS Elasticsearch uses a shard- and segment-based storage architecture with dynamic management. Each index is divided into multiple shards, and each shard contains multiple segments. During migration, write operations cause the destination cluster to regenerate its segment and shard structures, which can lead to changed data volumes.

- Data bloat during data rewriting

  When migrating in data rewriting mode, the destination cluster rebuilds its storage structure based on the current index configuration and load. Newly generated segments may occupy more storage than those in the original cluster due to differences in parameters such as compression policies and encoding methods. This is especially noticeable in scenarios involving both hot and cold data.

- Index configuration differences

  Index configurations, including the number of replicas, sharding policies, and compression settings, directly affect the final data volumes. For example, if the destination cluster uses a different compression algorithm than the source, data volumes may change in non-linear ways.

To confirm data integrity after migration, compare the document counts rather than data volumes. Document count is a reliable indicator of data consistency.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4 Using CSS Cluster Search Engines

4.1 Why Are Newly Created Index Shards Allocated to a Single Node in CSS?

Possible Causes

The possible causes are as follows:

- Shards were unevenly distributed in previous index allocations, and the predominate parameter in the latest indexed shard allocation was **balance.shard**. To balance the shard distribution across nodes, the new shards were allocated to the node with only a small number of shards.
- After a new node was added to a cluster and before the automatic cluster rebalancing completes, the predominate parameter was **balance.shard**. The shards of a new index are allocated to the new node, where there are no shards yet.

The following two parameters are used to balance the shard allocation in a cluster:

cluster.routing.allocation.balance.index (default value: **0.45f**)

cluster.routing.allocation.balance.shard (default value: **0.55f**)

NOTE

- **balance.index**: A larger value indicates that all the shards of an index are more evenly distributed across nodes. For example, if an index has six shards and there are three data nodes, two shards will be distributed on each node.
- **balance.shard**: A larger value indicates that all the shards of all the indexes are more evenly distributed across nodes. For example, if index **a** has two shards, index **b** has four, and there are three data nodes, two shards will be distributed on each node.
- You can specify both **balance.index** and **balance.shard** to balance the shard allocation.

Solution

To prevent all the shards of an index from being allocated to a single node, use either of the following methods:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. To create an index during cluster scale-out, configure the following parameter:

   ```
   PUT INDEX_NAME/_settings
   {
     "index.routing.allocation.total_shards_per_node": 2
   }
   ```

   That is, allow no more than two shards of an index to be allocated on each node. Determine the maximum number of shards allocated to each node based on the number of data nodes in your cluster and the number of index shards (both primary and secondary).

2. If a node contains relatively few existing shards, newly created index shards may be preferentially allocated to that node. In this case, you can use the **move** command in **POST _cluster/reroute** to relocate shards to other nodes. The rebalance module then automatically relocates other shards as needed to achieve a more balanced shard distribution in the cluster. Adjust the values of **balance.index** and **balance.shard** as appropriate for your workload.

4.2 How Do I Create a Type Under an Index in an Elasticsearch 7.*x* Cluster of CSS?

In Elasticsearch 7.*x* and later versions, types cannot be created for indexes.

If you need to use types, add **include_type_name=true** to the command. Only a single type is supported.

```
PUT index?include_type_name=true
{
  "mappings": {
    "my_type": {
      "properties": {
        "@timestamp": {
          "type": "date"
        }
      }
    }
  }
}
```

After an index is created, run the following command to write data in it:

```
PUT index/my_type/1
{
  "@timestamp":"2019-02-20"
}
```

4.3 How Do I Change the Number of Replicas for Elasticsearch Indexes in CSS?

When creating an index for an Elasticsearch cluster, you can specify the number of shards, that is, the number of primary shards. Once an index is created, the number of primary shards cannot be changed, but the number of replicas can be changed. **Number of replica shards = Number of primary shards x Number of replicas.**

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to Kibana and go to the command execution page. Elasticsearch clusters support multiple access methods. This topic uses Kibana as an example to describe the operation procedures.

   a. Log in to the **CSS management console**.

   b. In the navigation pane on the left, choose **Clusters > Elasticsearch**.

   c. In the cluster list, find the target cluster, and click **Kibana** in the **Operation** column to log in to the Kibana console.

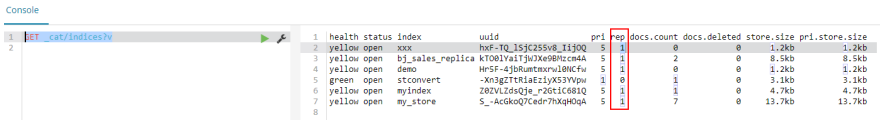
   d. In the left navigation pane, choose **Dev Tools**.

      The left part of the console is the command input box, and the triangle icon in its upper-right corner is the execution button. The right part shows the execution result.

2. On the Kibana console, run the following command to check the number of replicas for each Elasticsearch index:

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

Figure 4-1 Checking the number of replicas



3. Run the following command to configure the number of index replicas:

```
PUT /indexname/_settings
{
  "number_of_replicas" :1       //Number of replicas
}
```

indexname indicates the name of the index to be modified, and **number_of_replicas** indicates the number of replicas to be set.

4.4 What Are the Impacts If an Elasticsearch Cluster of CSS Has Too Many Shards?

1. Shard creation becomes progressively slower as the number of shards in a cluster increases.
2. If automatic index creation is enabled, slow index creation may cause a large number of write requests to be stacked in the memory or result in a cluster breakdown.
3. If there are too many shards and you cannot properly monitor workloads, the number of records in a single shard may exceed the threshold, and write requests may be denied.

4.5 How Do I Check the Number of Index Shards and Replicas in a CSS Elasticsearch Cluster?

1. Log in to Kibana and go to the command execution page. Elasticsearch clusters support multiple access methods. This topic uses Kibana as an example to describe the operation procedures.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

a. Log in to the **CSS management console**.
b. In the navigation pane on the left, choose **Clusters > Elasticsearch**.
c. In the cluster list, find the target cluster, and click **Kibana** in the **Operation** column to log in to the Kibana console.
d. In the left navigation pane, choose **Dev Tools**.

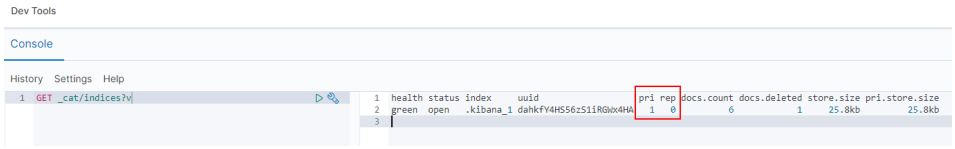
   The left part of the console is the command input box, and the triangle icon in its upper-right corner is the execution button. The right part shows the execution result.

2. Run the following command on Kibana to query the number of shards and replicas of each index:

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

In **Figure 4-2**, the **pri** column indicates the number of index shards, and the **rep** column indicates the number of replicas. Once an index is created, its **pri** value cannot be changed, but its **rep** value can be changed.

Figure 4-2 Querying index information



4.6 What Does the Value i for node.roles Mean for Nodes in an Elasticsearch Cluster of CSS?

Description

If the value of **node.roles** of a client node is **i**, then is this client node an ingest node?

- Are there coordinating-only nodes in your cluster? Are the client requests distributed to coordinating nodes?
- Are ingest nodes in idle state when there are no ingest requests?

Solution

If the value of **node.roles** of a client node is **i**, the ingest node mode is enabled.

- In CSS, Elasticsearch's coordinating-only nodes are called client nodes. If no client nodes are configured in a cluster, all nodes can handle coordinating tasks in addition to their assigned roles, sharing the workload of handling client requests.
- Ingest nodes are used for data transformation (preprocessing) and provide processing capabilities similar to Logstash pipelines. If there are no ingest workloads, client nodes do not sit idle either.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

4.7 How Do I Change the Maximum Number of Results Returned in Response to a Single Search Request to a Specific Index in a CSS Elasticsearch Cluster?

Solution

- Method 1

  Open Kibana and run the following commands on the **Dev Tools** page:

  ```
  PUT _all/_settings?preserve_existing=true
  {
  "index.max_result_window" : "10000000"
  }
  ```

- Method 2

  Run the following command on a server (a non-security mode cluster is used as an example here):

  ```
  curl -k -XPUT 'http://localhost:9200/*/_settings?preserve_existing=true' -d \
  '{
    "index.max_result_window": "10000000"
  }'
  ```

  localhost indicates the address of the Elasticsearch cluster.

CAUTION

This setting affects the consumption of memory and CPU resources. Exercise caution when setting this parameter.

4.8 How Do I Update Index Lifecycle Policies for an Elasticsearch Cluster of CSS?

The lifecycle of Elasticsearch clusters is implemented using the Index State Management (ISM) of Open Distro. For details about how to configure policies related to the ISM template, see the **Open Distro documentation**.

1. When a policy is created, the system writes a record to the **.opendistro-ism-config** index. In the record, **_id** is the policy name, and the content is the policy definition.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Figure 4-3 Writing a data record



```
{
  "_index" : ".opendistro-ism-config",
  "_type" : "_doc",
  "_id" : "policy1",
  "_score" : 1.0,
  "_source" : {
    "policy" : {
      "policy_id" : "policy1",
      "description" : "A simple default policy that changes the replica count between hot and cold states.",
      "last_updated_time" : 1641432150329,
      "schema_version" : 1,
      "error_notification" : null,
      "default_state" : "hot",
      "states" : [
        {
          "name" : "hot",
          "actions" : [ ],
          "transitions" : [
            {
              "state_name" : "delete",
              "conditions" : {
                "min_index_age" : "2d"
              }
            }
          ]
        },
        {
          "name" : "delete",
          "actions" : [
            {
              "delete" : { }
            }
          ],
          "transitions" : [ ]
        }
      ]
    }
  }
]
```


2. After a policy is bound to an index, the system writes another record to the **.opendistro-ism-config** index. The following figure shows the initial status of a record.

Figure 4-4 Initial data status



```
{
  "_index" : ".opendistro-ism-config",
  "_type" : "_doc",
  "_id" : "FABkSF5GSTCmR0QkW41HVw",
  "_score" : 1.0,
  "_source" : {
    "managed_index" : {
      "name" : "data1",
      "enabled" : true,
      "index" : "data1",
      "index_uuid" : "FABkSF5GSTCmR0QkW41HVw",
      "schedule" : {
        "interval" : {
          "start_time" : 1641432652693,
          "period" : 1,
          "unit" : "Minutes"
        }
      },
      "last_updated_time" : 1641432652694,
      "enabled_time" : 1641432652694,
      "policy_id" : "policy1",
      "policy_seq_no" : null,
      "policy_primary_term" : null,
      "policy" : null,
      "change_policy" : null
    }
  }
}
]
```


3. Run the **explain** command. Only a policy ID will be returned.

```
GET _opendistro/_ism/explain/data2
{
  "data2" : {
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
  "index.opendistro.index_state_management.policy_id" : "policy1"
 }
}
```

Open Distro will execute an initialization process to fill the policy content in the record. The following figure shows the initialized data.

Figure 4-5 Initialized data



After the initialization, **min_index_age** in the policy will be copied.

NOTE

The initialized index uses a copy of this policy. Any policy update will not be applied to the index.

4. After the policy is modified, call the **change_policy** API to update the policy.

```
POST _opendistro/_ism/change_policy/data1
{
  "policy_id": "policy1"
}
```

4.9 How Do I Clear Elasticsearch Indexes in CSS?

CAUTION

Before deleting index data, carefully evaluate any potential impact on services.

- Have indexes automatically cleared on a regular basis.

  You can create a scheduled task to execute an index deletion request periodically. CSS supports Open Distro Index State Management. For details

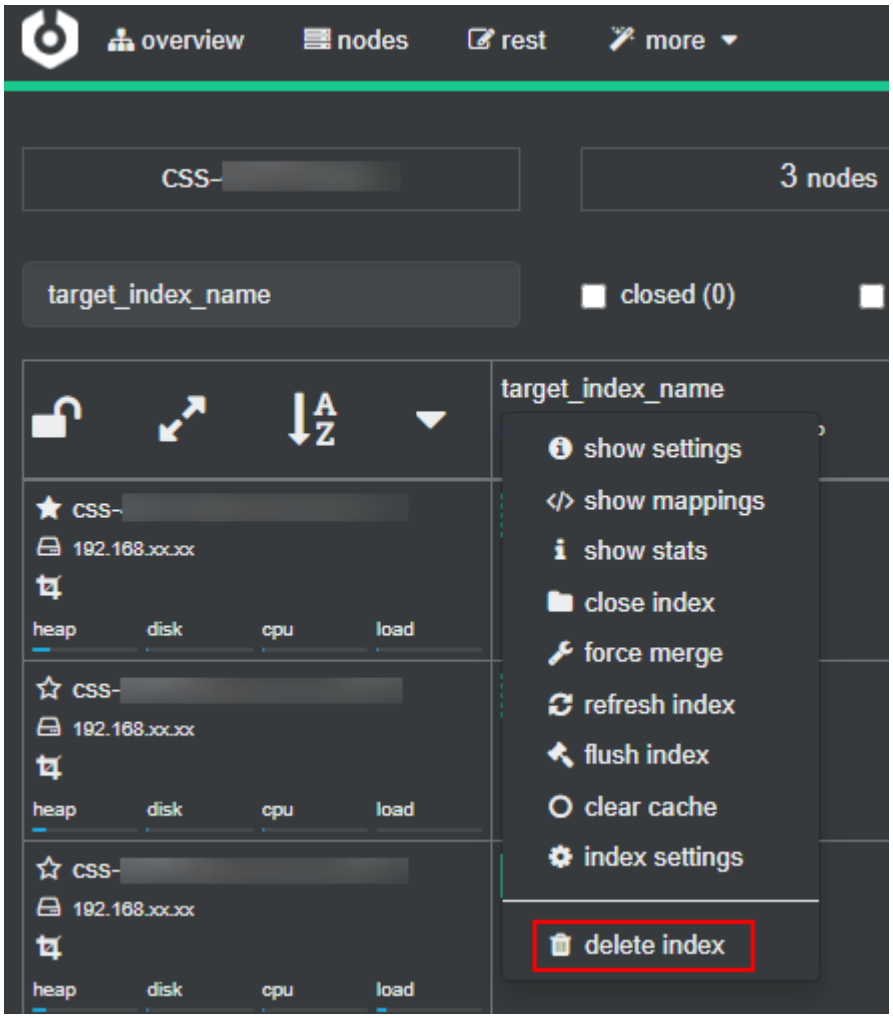
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

about how to clear obsolete indexes periodically, see **Decoupling Index Storage and Compute Through Index Lifecycle Management**.

For details about Open Distro Index State Management, see https://opendistro.github.io/for-elasticsearch-docs/docs/im/ism/.

- Manually clear indexes.
  - Log in to Kibana, go to the **Dev Tools** page, and run the following command to delete a specified index:
    ```
    DELETE /{index_name}
    ```
  - Log in to Cerebro, search for the target index name, click the index name, click **delete index**, and click **Confirm** in the displayed dialog box.

    Figure 4-6 Deleting an index from Cerebro

4.10 How Do I Clear Elasticsearch Cache in CSS?

- **Clear the fielddata**

  During aggregation and sorting, data is converted to the fielddata structure, which consumes extensive memory.

  a. Run the following command on Kibana to query the fielddata cache status:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
GET /_cat/nodes?v&h=name,fielddataMemory
```

b. If the memory usage of **fielddata** is too high, you can run the following command to clear the **fielddata cache** of a specified index or all indexes:

```
POST /test/_cache/clear?fielddata=true
```

In the preceding command, *test* indicates the name of the index whose fielddata occupies a large amount of memory.

```
POST /_cache/clear?fielddata=true
```

- **Clear segments**

  The FST structure of each segment is loaded to the memory and will not be cleared. If the number of index segments gets too large, the memory usage will be high. You are advised to periodically clear the segments.

  a. Run the following command on Kibana to check the number of segments and their memory usage on each node:

```
GET /_cat/nodes?v&h=segments.count,segments.memory&s=segments.memory:desc
```

  b. If segments consume too much memory, you can free up memory by deleting or closing unnecessary indexes, or periodically merging indexes that are no longer being updated.

- **Clear the cache**

  Run the following command on Kibana to clear the cache:

```
POST /_cache/clear
```

4.11 Why Does Disk Usage Increase After Running delete_by_query to Delete Data in an Elasticsearch Cluster?

Running the **delete_by_query** command does not immediately delete data. Instead, it adds a deletion mark to the target data. While marked data is filtered out from search results, it still occupies disk space.

As a result, running the delete_by_query command does not immediately free up disk space. The data marked for deletion is physically removed only during a subsequent segment merge. That is when the space it occupies is finally reclaimed.

Before the merge takes place, the marked data continues to occupy disk space, while the deletion operation generates new segments. Consequently, disk usage may actually increase rather than decrease.

4.12 Do CSS Elasticsearch Clusters Support script dotProduct?

For vector search, CSS Elasticsearch clusters currently do not support Elasticsearch-native vector search scripts (script dotProduct). Therefore, such scripts cannot be executed.

You are advised to use the vector search service provided by CSS. Based on a proprietary vector search engine, CSS's vector search service is deeply integrated with the Elasticsearch plugin architecture, delivering high performance, high precision, low cost, and multi-modality support. It offers an efficient, cost-effective

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

solution to meet diversified, high-dimensional vector search needs. For details, see **Vector Search**.

CAUTION

Only Elasticsearch 7.6.2 and 7.10.2 clusters in CSS support vector search.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

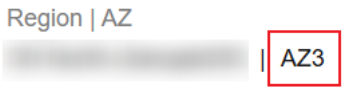
5 Managing CSS Clusters

5.1 How Do I Check the AZ or AZs of a CSS Cluster?

You can check the AZ or AZs (in the case of a multi-AZ deployment) of a cluster on its **Overview** tab.

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Overview** tab. In the **Basic Information** area, find **Region | AZ**, and check the cluster's AZ or AZs.

Figure 5-1 Checking the cluster's AZs



5.2 What Is the Relationship Between the Filebeat Version and Cluster Version in CSS?

- Clusters with the security mode disabled: no restrictions.
- Clusters with the security mode enabled: The Filebeat OSS version must match the cluster version. For the correct Filebeat OSS version to download, see **https://www.elastic.co/downloads/past-releases#filebeat-oss**.

5.3 How Do I Obtain the Security Certificate of a CSS Security-Mode Cluster?

The security certificate (**CloudSearchService.cer**) can be downloaded only for security-mode clusters that have enabled HTTPS access.

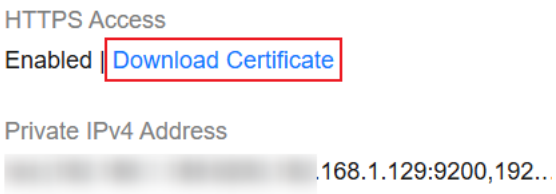
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

NOTE

When using certificate-based authentication in a public network environment, disable hostname verification. Otherwise, certificate verification will fail.

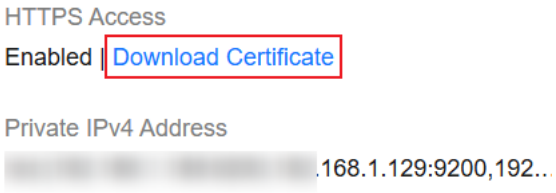
- Obtain the security certificate of an Elasticsearch cluster.
  a. Log in to the **CSS management console**.
  b. In the navigation pane on the left, choose **Clusters > Elasticsearch**.
  c. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
  d. Click the **Overview** tab. In the **Network Information** area, click **Download Certificate** below **HTTPS Access**.

  Figure 5-2 Downloading a security certificate

- Obtain the security certificate of an OpenSearch cluster.
  a. Log in to the **CSS management console**.
  b. In the navigation pane on the left, choose **Clusters > OpenSearch**.
  c. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
  d. Click the **Overview** tab. In the **Network Information** area, click **Download Certificate** below **HTTPS Access**.

  Figure 5-3 Downloading a security certificate

5.4 How Do I Convert the Format of a CER Security Certificate in CSS?

The security certificate (**CloudSearchService.cer**) can be downloaded only for security clusters that have enabled HTTPS access. Most software supports certificates in the **.pem** or **.jks** format. You need to convert the format of the CSS security certificate.

- Run the following command to convert the security certificate from **.cer** to **.pem**:
  ```
  openssl x509 -inform pem -in CloudSearchService.cer -out newname.pem
  ```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- Run the following command to convert the security certificate from **.cer** to **.jks**:
  ```
  keytool -import -alias newname -keystore ./truststore.jks -file ./CloudSearchService.cer
  ```

In the preceding commands, *newname* indicates the user-defined certificate name.

After the command is executed, set the certificate password and confirm the password as prompted. Securely store the password. It will be used for accessing the cluster.

5.5 Can I Change the Security Group of a CSS Cluster?

You can change the security group of an existing CSS cluster.

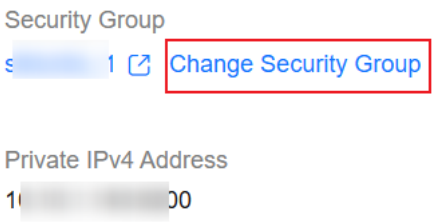
Constraints

- It is advisable to perform this operation during off-peak hours.
- The security group cannot be changed for clusters created before February 2023. To do that, create a new cluster by selecting the desired security group, then perform **Migrating Data Between Elasticsearch Clusters Using Backup and Restoration** to migrate its data to that new cluster.
- Make sure the inbound rules of the new security group allow all the ports required for service access. For Elasticsearch and OpenSearch clusters, port 9200 must be allowed. For Logstash clusters, port 9600 must be allowed.

Changing the Security Group

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Overview** tab.
5. In the **Network Information** area, click **Change Security Group** under **Security Group**.

   Figure 5-4 Changing the security group

6. In the **Change Security Group** dialog box, select a new security group and click **OK**.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

5.6 Can I Expand the Subnet for an Elasticsearch or OpenSearch Cluster?

If the current subnet of an existing Elasticsearch or OpenSearch cluster cannot meet your requirements, you can switch to another subnet.

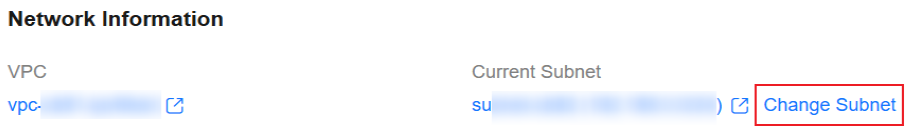
Constraints

- After switching the subnet, the new subnet is used for adding nodes, including scaling, adding dedicated master/client nodes, and enabling VPC Endpoint. IP addresses will be allocated to newly added nodes from the new subnet.
- A cluster can be associated with a maximum of two subnets.
- Subnet switching does not affect the network settings of existing nodes.

Procedure

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Overview** tab.
5. In the **Network Information** area, click **Change Subnet** under **Current Subnet**. In the displayed dialog box, select a new subnet. If there are no subnets that meet your requirements, click **Create Subnet** to go to the networking service console to create a new subnet.

   Figure 5-5 Switching to another subnet

6. Click **OK** after you change the subnet.

5.7 How Do I Set search.max_buckets for a CSS Cluster?

The default maximum number of buckets allowed in an aggregation operation for a CSS Elasticsearch and OpenSearch cluster is 10,000 for versions earlier than Elasticsearch 7.9.3, and 65,535 for Elasticsearch 7.9.3 and later as well as OpenSearch clusters.

To return more buckets than the default maximum allowed in an aggregation operation, you can increase the value of **search.max_buckets**.

CAUTION

Increasing the value of **search.max_buckets** will increase the cluster's load and memory usage. Exercise caution.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Run the following command on the Dev Tools page of Kibana (for Elasticsearch clusters) or OpenSearch Dashboards (for OpenSearch clusters):

```
PUT _cluster/settings
{
   "persistent": {
      "search.max_buckets": 70000
   }
}
```

5.8 How Do I Modify the TLS Algorithm for a CSS Cluster?

The TLS algorithm can be modified for Elasticsearch 7.6.2 and later as well as OpenSearch clusters.

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Choose **Cluster Settings** > **Parameter Settings**.
5. Click the **OpenSearch** or **Elasticsearch** tab, depending on your cluster type.
6. Click **Edit**, expand **Custom**, and click **Add**.
   - For an Elasticsearch cluster, add the following parameter:

     Parameter: opendistro_security.ssl.http.enabled_ciphers

     Value: ['TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256', 'TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384']
   - For an OpenSearch cluster, add the following parameter:

     Parameter: plugins.security.ssl.http.enabled_ciphers

     Value: ['TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256', 'TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384']

   If **Value** contains multiple algorithms/protocols, enclose them using a pair of square brackets ([]). If it contains only a single algorithm/protocol, enclose it using a pair of single quotation marks ('').
7. After the change is complete, click **Submit**. In the displayed dialog box, confirm the settings, 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 change list, the change has been saved.
8. Click **Restart** in the upper right corner to restart the cluster, thus making the change take effect.

5.9 How Do I Enable Audit Logs for a CSS Cluster?

Audit logs can be enabled for security-mode Elasticsearch 7.6.2 clusters as well as security-mode OpenSearch clusters.

Audit logs are disabled for Elasticsearch clusters by default.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Choose **Cluster Settings** > **Parameter Settings**.
5. Click the **OpenSearch** or **Elasticsearch** tab, depending on your cluster type.
6. Click **Edit**, expand **Custom**, and click **Add**.
   - For an Elasticsearch cluster, add the following parameter:

     Parameter: opendistro_security.audit.type

     Value: internal_elasticsearch
   - For an OpenSearch cluster, add the following parameter:

     Parameter: plugins.security.audit.type

     Value: internal_opensearch
7. After the change is complete, click **Submit**. In the displayed dialog box, confirm the settings, 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 change list, the change has been saved.
8. Click **Restart** in the upper right corner to restart the cluster, thus making the change take effect.
9. After cluster restart, check whether audit logs have been enabled.
   a. For an Elasticsearch cluster, click **Kibana** in the **Operation** column to log in to Kibana. For an OpenSearch cluster, click **Dashboards** in the **Operation** column to log in to OpenSearch Dashboards.
   b. Expand the menu in the upper-left corner, and choose **Dev Tools**.
   c. Run the following command. If the result contains indexes whose names contain **.*audit***, audit logs have been enabled.

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

5.10 Can I Stop a CSS Cluster?

To stop a CSS cluster, you delete it. If you want to stop a cluster upon completion of a cluster migration, stop all services in the source cluster, and confirm that all data has been migrated to the destination cluster. Then delete the source cluster. You can stop services in the source cluster in the following ways:

- If the cluster version in use supports the flow control function, you can enable **one-click traffic blocking** to block traffic everywhere except the O&M interface, rejecting all requests.
- If your cluster version in use does not support traffic control, you can disable read and write for all service indexes instead. For example, if all service indexes start with **log**, run the following command on the **Dev Tools** page of Kibana:

```
PUT log*/_settings
{
  "index.blocks.read": true,
  "index.blocks.write": true,
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
  "index.blocks.metadata": true
}
```

5.11 How Do I Query the Index Size on OBS After the Freezing of Indexes for a CSS Cluster?

The size of indexes remains unchanged after freezing. By querying the size of frozen indexes in OBS, you obtain the size of all indexes stored on OBS.

Run the following command to obtain information about all indexes that are being frozen or have already been frozen:

```
GET _cat/freeze_indices?stage=$
```

The output is as follows (as an example only):

```
green open data2 0bNtxWDtRbOSkS4JYaUgMQ 3 0  5 0  7.9kb  7.9kb
green open data3 oYMLvw31QnyasqUNuyP6RA 3 0 51 0 23.5kb 23.5kb
```

The last column of the returned result contains the index size information.

Related Questions

- **Billing for index storage on OBS**

  Storing index data in OBS incurs additional charges. For details, see "Standard cold storage prices" in **Cloud Search Service Pricing Details**.

- **Why are frozen indexes stored in OBS still searchable via commands?**

  Elasticsearch and OpenSearch clusters use local storage by default, and Lucene index files are stored on local disks. Lucene interacts with the underlying storage via the Directory API. Files can be read through the following API:

  ```
  public abstract IndexInput openInput(String name, IOContext context) throws IOException;
  ```

  The storage-compute decoupling feature enables interaction with OBS through the Directory API to read files stored in OBS. This is why information about frozen indexes stored in OBS can be queried using commands.

5.12 How Do I Plan the Quantity of Index Shards for a Cluster?

Before importing data to a cluster, carefully consider your service needs and plan the cluster's data structure and distribution in advance. This includes properly designing indexes and deciding on the appropriate number of index shards. To ensure optimal performance and scalability for a cluster, consider following these best practices:

- **The size of a single shard**: Keep the size of each shard between 10 GB and 50 GB. This helps strike a balance between storage efficiency and query performance.
- **Memory-to-shards ratio**: Limit the number of shards per 1 GB of memory to 20 to 30. This ensures that each shard has sufficient memory resources to respond to indexing and query requests.
- **Number of shards per node**: To prevent node overload, keep the number of shards on each node under 1000. This helps to improve node stability.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- **Relationship between the number of index shards and the number of nodes**: For each index, make sure the number of shards is an integral multiple of the total number of data nodes and cold data nodes in the cluster. This helps improve load balancing and optimize query and indexing performance.
- **Total number of shards in a cluster**: To facilitate management and avoid oversized shards, make sure the total number of shards in a cluster is less than 30,000. This helps maintain the stability and responsiveness of the cluster.

Following these suggestions, you can plan and manage index shards for a CSS cluster more effectively, improving the cluster's overall performance and maintainability.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

6 CSS Cluster Backup and Restoration

6.1 How Do I Query the Snapshot Information of a CSS Cluster?

You can query snapshot information only if cluster snapshots are enabled and some snapshots have been created.

You can check them on the CSS management console or query them by running commands on Kibana or OpenSearch Dashboards.

Querying Snapshots on the CSS Management Console

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Cluster Snapshots** tab.
5. In the cluster snapshot task list, click the target snapshot name, and check snapshot information in the displayed dialog box.

Querying Snapshots by Running Commands

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. For an Elasticsearch cluster, click **Kibana** in the **Operation** column to log in to Kibana. For an OpenSearch cluster, click **Dashboards** in the **Operation** column to log in to OpenSearch Dashboards.
4. Expand the menu in the upper-left corner, and choose **Dev Tools**.
5. Run the following command to query the snapshot repository information, that is, the basic cluster snapshot settings.

   ```
   GET _snapshot/_all
   ```

   Example response:

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
{
  "repo_auto": {
    "type": "obs",
    "settings": {
      "bucket": "123xxx",       // OBS bucket name
      "chunk_size": "2g",       // Chunk size in snapshots (unit: GB)
      "endpoint": "obs.xxx.example.com:443",       // OBS VPC endpoint
      "max_restore_bytes_per_sec": "0MB",       // Maximum recovery rate (per second)
      "compress": "true",       // Whether to enable data compression
      "base_path": "css_repository/css-xxx",       // Backup path
      "region": "xxx",       // Region
      "max_snapshot_bytes_per_sec": "40MB"       // Maximum backup rate (per second)
    }
  }
}
```

6. Run the following command to query the snapshot list:

```
GET _snapshot/repo_auto/_all
```

Example output (two snapshots are returned):

```
{
  "snapshots": [
    {
      "snapshot": "snapshot-2dc3",       // Snapshot name
      "uuid": "VW5y2NBJS9iPh7YcGLxxxx",  // Snapshot ID
      "version_id": xxxxxxx,          // Internal cluster version
      "version": "x.x.x",          // Cluster version
      "indices": [              // Backed up indexes
        ".opendistro_security"
      ],
      "data_streams": [ ],          // Data streams
      "include_global_state": true,      // Whether to include the global cluster status
      "state": "SUCCESS",              // Snapshot status
      "start_time": "2025-08-30T01:41:57.068Z",  // Snapshot start time
      "start_time_in_millis": 1756518117068,      // Snapshot start time in milliseconds
      "end_time": "2025-08-30T01:41:57.469Z",      // Snapshot end time
      "end_time_in_millis": 1756518117469,      // Snapshot end time in milliseconds
      "duration_in_millis": 401,              // Snapshot creation duration (from start to completion) in
milliseconds
      "failures": [ ],                 // Failed shards during snapshot creation
      "shards": {                      // Shard statistics
        "total": 1,                    // Total number of shards
        "failed": 0,                    // Number of failed shards
        "successful": 1                 // Number of successful shards
      }
    },
    {
      "snapshot": "snapshot-dd37",
      "uuid": "FD4VcooLS8yjPY3w0-x-xx",
      "version_id": xxxxxxx,
      "version": "x.x.x",
      "indices": [
        ".kibana",
        ".opendistro_security"
      ],
      "data_streams": [ ],
      "include_global_state": true,
      "state": "SUCCESS",
      "start_time": "2025-08-30T01:54:55.750Z",
      "start_time_in_millis": 1756518895750,
      "end_time": "2025-08-30T01:54:55.950Z",
      "end_time_in_millis": 1756518895950,
      "duration_in_millis": 200,
      "failures": [ ],
      "shards": {
        "total": 2,
        "failed": 0,
        "successful": 2
      }
    }
```

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

```
 ]
}
```

To query information about a specified snapshot, run the following command:

```
GET _snapshot/repo_auto/{snapshot_name}
```

Replace *snapshot_name* with the actual snapshot name. Wildcard matching is supported.

7. (Optional) Run the following command to delete a specified snapshot:

```
DELETE _snapshot/repo_auto/{snapshot_name}
```

Exercise caution when choosing snapshots to delete. Once deleted, they cannot be used to restore cluster data.

6.2 How Do I Back Up Large Volumes of Data Using Snapshots?

When you have large volumes of data and backing up the data may take more than one day, you can use the following methods to improve backup efficiency:

- Back up specified indexes in batches: Manually create snapshots on the console and specify a subset of indexes each time to reduce the amount of data backed up in a single operation.
- Customize the backup rate: Modify the snapshot configuration on the console to adjust the maximum backup rate and improve the efficiency of each backup operation.
- Use a custom snapshot repository: Create a custom repository through Kibana to flexibly configure backup rates and chunk sizes for efficient backup.

Method 1: Backing Up Specified Indexes in Batches

When the data volume is extremely large, you can manually create snapshots by specifying a subset of indexes each time to back up indexes in batches.

1. On the **Dev Tools** page of Kibana or OpenSearch Dashboards, run the following command to query the names of all indexes in the cluster:

```
GET /_cat/indices
```

2. Divide indexes into multiple batches based on index size and priority.
3. Log in to the **CSS management console**.
4. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
5. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
6. Choose the **Cluster Snapshots** tab and enable snapshots.
7. Under **Cluster Snapshot Tasks**, click **Create Snapshot Manually**. In the displayed dialog box, configure the snapshot policy.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 6-1 Manually creating a snapshot

Parameter	Description
Snapshot Name	Set the snapshot name. You can use names like **snapshot-batch1** or **snapshot-batch2** to distinguish different batches.
Index	Specify the indexes to be backed up in this snapshot. If you do not specify this parameter, all indexes in the cluster are backed up by default.
Snapshot Description	Add a snapshot description.

8. Click **OK** to start creating snapshots.
9. In the cluster snapshot task list, when **Snapshot Task Status** changes to **Available**, backup for the current batch is complete.

   Click the expand icon in front of the snapshot name to view the backed-up shards and indexes and verify successful backup.
10. Repeat steps **7** to **9** until all indexes are backed up.

Method 2: Customizing the Backup Rate

You can modify the snapshot configuration on the console to adjust the maximum backup rate and improve the backup speed of each node.

CAUTION

Increasing the backup rate increases the CPU and disk I/O load on the cluster. It is recommended that you perform this operation during off-peak hours and reduce this rate before peak hours.

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Choose the **Cluster Snapshots** tab, and click **Modify Settings** to increase the maximum backup rate.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Table 6-2 Modifying the backup rate

Parameter	Description
Maximum Backup Rate (/s)	The parameter sets the maximum backup rate per node. When it is exceeded, flow control is triggered to prevent excessive resource usage and ensure system stability. **Value format**: number (0–9999) + unit (KB, MB, GB, TB, PB, or B) **Default value**: 40 MB The value 0MB means there is no rate limit. An excessively high backup rate may lead to excessive resource usage, which may impact cluster stability. Configure this parameter carefully to maintain optimal performance. The actual backup rate may not reach the configured value, as it depends on factors such as OBS performance and disk I/O.
Maximum Recovery Rate (/s)	The parameter sets the maximum recovery rate per node. When it is exceeded, flow control is triggered to prevent excessive resource usage and ensure system stability. **Value format**: number (0–9999) + unit (KB, MB, GB, TB, PB, or B) **Default value**: 40 MB for Elasticsearch 7.6.2 or earlier; and 0 MB (no rate limit) for OpenSearch and Elasticsearch versions later than 7.6.2. An excessively high recovery rate may lead to excessive resource usage, which may impact cluster stability. Configure this parameter carefully to maintain optimal performance. For OpenSearch clusters and Elasticsearch versions later than 7.6.2, the actual recovery rate is also affected by the value of **indices.recovery.max_bytes_per_sec**. Whichever is smaller between the two parameters will be used. Additionally, the actual recovery rate may not reach the configured value, as it depends on factors such as OBS performance and disk I/O.

5. Click **OK** to save the change.
6. Manually create a snapshot or wait for an automatic snapshot to be triggered. The backup will be performed based on the new rate.

Method 3: Using a Custom Snapshot Repository

In addition to the repo_auto repository provided by CSS, you can create a custom snapshot repository to flexibly configure parameters such as backup rate and chunk size for more efficient backup.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

1. Log in to the **Dev Tools** page of Kibana or OpenSearch Dashboards.
2. Create a custom snapshot repository.

```
PUT _snapshot/my_backup
{
  "type": "obs",
  "settings": {
    "bucket": "<OBS bucket name>",
    "base_path": "<Backup path>",
    "chunk_size": "2g",
    "endpoint": "obs.<Region name>.example.com",
    "region": "<Region name>",
    "compress": "true",
    "access_key": "<AK>",
    "secret_key": "<SK>",
    "max_restore_bytes_per_sec": "100mb",
    "max_snapshot_bytes_per_sec": "100mb"
  }
}
```

Table 6-3 Custom snapshot repository parameters

Parameter	Mandatory	Description
bucket	Yes	OBS bucket name. The bucket must be in the same region as the cluster, and its storage class must be **Standard**.
base_path	Yes	Path for storing snapshots in the OBS bucket. The path cannot start with **/** or **..**
chunk_size	No	Chunk size for backup. The default value is **2g**. Retain the default value for large data volumes.
endpoint	Yes	OBS endpoint in the format **obs.<Region name>.example.com**.
region	Yes	Region where the cluster is located.
compress	No	Whether to enable data compression. The default value is **true**. Data compression reduces OBS storage usage.
access_key	Yes	AK for accessing OBS. For details about how to obtain the AK, see **How Do I Obtain an Access Key (AK/SK)?**
secret_key	Yes	SK for accessing OBS. For details about how to obtain the SK, see **How Do I Obtain an Access Key (AK/SK)?**
max_snapshot_bytes_per_sec	No	Maximum backup rate per node. You can increase this value during off-peak hours (for example, 100 MB) and reduce it during peak hours.
max_restore_bytes_per_sec	No	Maximum restore rate per node. Adjust this value based on cluster performance.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

3. Create a snapshot using a custom repository.

```
PUT _snapshot/my_backup/<snapshot name>
{
  "indices": "*",
  "ignore_unavailable": true,
  "include_global_state": false
}
```

Table 6-4 Parameters for creating a snapshot

Parameter	Mandatory	Description
<snapshot name>	Yes	Snapshot name. Use names like **snapshot-batch1** to distinguish different batches.
indices	No	Specifies the indexes to back up. Separate multiple indexes with commas. Wildcards (*) are supported. The default value is *, meaning to back up all indexes.
ignore_unavailable	No	Whether to ignore indexes that cannot be found. The default value is **false**. • **true**: Ignore indexes that cannot be found and continue backing up other indexes. • **false**: Return a failure when a target index cannot be found.
include_global_state	No	Whether to save the global cluster state. The default value is **false**. You are advised to retain the default value to prevent the global state from affecting restoration.

4. Query the snapshot status.

```
GET _snapshot/my_backup/<snapshot name>/_status
```

If the value of **state** in the returned result is **SUCCESS**, the backup is successful.

6.3 Can a Deleted CSS Cluster Be Restored?

For a deleted Elasticsearch or OpenSearch, if it still has snapshots stored in OBS, it can be restored using these snapshots. Otherwise, it cannot be restored. Therefore, exercise caution when deciding to delete a cluster.

To restore a deleted cluster using one of its snapshots stored in OBS, perform the following steps:

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. Click **Create Cluster** in the upper-right corner to create a cluster that is of the same type as the deleted cluster. Make sure that:

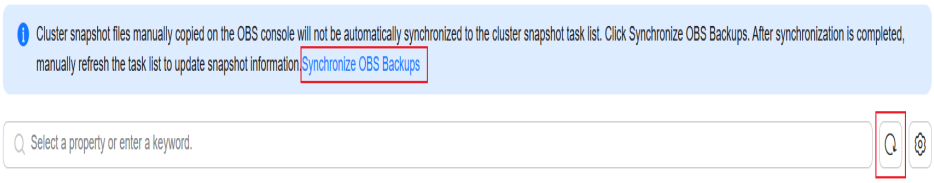
 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- The new cluster is in the same region as the deleted cluster.
- The cluster version is the same as or later than that of the deleted cluster.
- The number of nodes in the new cluster is greater than half of that in the deleted cluster. Ideally, they should have the same number of nodes. This is to prevent data restoration failures.
- Deselect **Automatic Snapshot Creation**.

4. When the status of the new cluster changes to **Available**, enable cluster snapshots.
   a. In the cluster list, click the name of the new cluster. The cluster information page is displayed.
   b. Click the **Cluster Snapshots** tab.
   c. Click **Enable Snapshot**. In the displayed dialog box, configure the necessary settings. Keep **OBS Bucket** and **Backup Path** consistent with those of the deleted cluster.
   d. Click **OK**.
5. Above the cluster snapshot task list, click **Synchronize OBS Backups**. After data synchronization, manually refresh the list to show the updated snapshot information.

   The snapshots of the deleted cluster are displayed in the cluster snapshot task list.

   Figure 6-1 Synchronizing snapshots stored in OBS

6. Restore data from snapshots.
   a. In the cluster snapshot task list, locate the needed snapshot, and click **Restore** in the **Operation** column. In the displayed dialog box, configure index settings and select the current cluster in **Cluster**.
   b. Click **OK**. When the snapshot's **Task Status** changes to Restoration succeeded, data has been restored successfully.

      If there are multiple snapshots, restore all of them in the right order to restore the data of the deleted cluster.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

7 CSS Cluster Monitoring and O&M

7.1 How Do I Set Slow Query Log Thresholds for an Elasticsearch Cluster of CSS?

The slow query log settings of CSS are the same as those of Elasticsearch. You can configure slow query logs via the _settings API. For example, you can run the following command in Kibana to set the index level:

```
PUT /my_index/_settings
{
    "index.search.slowlog.threshold.query.warn": "10s",
    "index.search.slowlog.threshold.fetch.debug": "500ms",
    "index.indexing.slowlog.threshold.index.info": "5s"
}
```

- If a query takes longer than 10 seconds, a WARN log will be generated.
- If retrieval takes longer than 500 milliseconds, a DEBUG log will be generated.
- If an index takes longer than 5 seconds, an INFO log will be generated.

For details, visit the official website: **https://www.elastic.co/guide/cn/elasticsearch/guide/current/logging.html**.

7.2 What Do I Do If the Average Memory Usage of a CSS Cluster Reaches 98%?

Symptom

The cluster monitoring result shows that the average memory usage of a cluster is 98%. Does it affect cluster performance?

Possible Cause

In an Elasticsearch cluster, 50% of the memory is occupied by Elasticsearch and the other 50% is used by Lucene to cache files. It is normal that the average memory usage reaches 98%.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

Solution

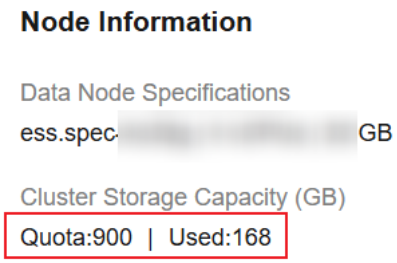
You can monitor the cluster memory usage by checking the maximum JVM heap usage and average JVM heap usage.

7.3 How Do I Check the Total Disk Usage of a CSS Cluster?

1. Log in to the **CSS management console**.
2. In the navigation pane on the left, expand **Clusters**. Select a cluster type based on the target cluster. The cluster list is displayed.
3. In the cluster list, click the name of the target cluster. The cluster information page is displayed.
4. Click the **Overview** tab.
5. In the **Node Information** area, obtain the value of **Cluster Storage Capacity (GB)**.

   Total disk usage of the cluster = Used/Quota

   Figure 7-1 Obtaining a cluster's storage capacity and usage information

7.4 Will CSS Cluster Services Be Affected If the Disk Usage of a Single Node Gets Too High?

Symptom

According to the cluster monitoring information, the disk usage of some individual nodes of an Elasticsearch cluster exceeds 80%. Does it impact cluster performance?

Impact on Services

- When the disk usage of a node exceeds 85%, no new replica shards will be allocated to this node. However, new primary shards can still be allocated to ensure service continuity, though this may impact cluster availability.
- When the disk usage of a node exceeds 90%, a shard migration is automatically triggered to reallocate shards on this node to other data nodes with lower disk usage. During this process, disk space cannot be allocated to new shards. Shard migration and reallocation may increase query delay or temporarily interrupt services, impacting service continuity.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.

- When the disk usage of a node exceeds 95%, the **read_only_allow_delete** attribute is enabled in its indexes. In this case, indexes on this node can only be read or deleted but data cannot be written in.

If per-node disk usage is too high, you can add more nodes or expand the capacity of existing nodes. For details, see **Scaling Out/Up an Elasticsearch Cluster**. After new nodes are added, index shards automatically re-allocate to balance the load. You can launch Cerebro to check the shard allocation process. You can also modify **indices.recovery.max_bytes_per_sec** and **cluster.routing.allocation.cluster_concurrent_rebalance** to speed up this process.

 Copyright © Huawei Cloud Computing Technologies Co., Ltd.