ALIBABA CLOUD

Alibaba Cloud

Elasticsearch

Instances

Document Version: 20201222

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Cross-zone instance deployment 08
1.1. Precautions 08
1.2. Perform a switchover 10
1.3. Perform a recovery 11
2.Manage instances 12
2.1. Create clusters 12
2.2. Clusters 12
2.3. Manage cluster tags 14
2.4. Restart a cluster or node 16
2.5. View the progress of a cluster task 20
2.6. Refresh a cluster 21
2.7. View basic information of a cluster 21
2.8. Set a maintenance window 25
2.9. View node information 25
2.10. View the configuration of an Elasticsearch cluster 27
2.11. Release a cluster 28
3.Data migration 29
3.1. Migrate nodes in a zone 29
4.Upgrade 31
4.1. Upgrade the version of a cluster 31
4.2. Check for and modify incompatible configurations before ... 33
5.Upgrade or downgrade a cluster 41
5.1. Update the kernel of a cluster 41
5.2. Scale in an Elasticsearch cluster 43
5.3. Upgrade the configuration of a cluster 49
6.Cluster configuration 52

6.1. Overview 52
6.2. Configure synonyms 52
6.2.1. Configuration rules 52
6.2.2. Upload a synonym dictionary file 56
6.2.3. Configure synonyms 58
6.3. Configure a garbage collector 68
6.4. Modify the YML configuration 69
6.5. Configure YML 73
6.5.1. Configure CORS 73
6.5.2. Recreate indexes by calling the Reindex operation 75
6.5.3. Configure the audit log indexing feature 78
6.5.4. Configure queue sizes 80
6.6. Perform scenario-based configuration 81
6.6.1. Use a scenario-based template to modify the configura... 81
6.6.2. Modify the dynamic settings of a cluster 84
6.6.3. Modify the index template of a cluster 85
6.6.4. Modify the index lifecycle configurations of a cluster 87
7.Plug-ins 89
7.1. Overview of plug-ins 89
7.2. Built-in plug-ins 89
7.2.1. Install and remove a built-in plug-in 89
7.2.2. Use the analysis-ik plug-in 91
7.2.3. Use the aliyun-sql plug-in 96
7.2.3.1. Use method 96
7.2.3.2. Query syntax 102
7.2.4. Use the physical replication feature of the apack plug-... 121
7.2.5. Use the analysis-aliws plug-in 124
7.2.6. Use the aliyun-knn plug-in for vector search 132

7.2.7. Use the aliyun-qos plug-in 139
7.2.8. Use the codec-compression plug-in of the beta version 144
7.2.9. Use the faster-bulk plug-in 146
7.2.10. Use the gig plug-in 148
7.3. Upload and install a custom plug-in 151
8.Cluster monitoring and alerting 153
8.1. Enable the alerting feature 153
8.2. Configure the monitoring and alerting feature in Cloud M... 154
8.3. View cluster monitoring data 157
8.4. Monitoring metrics 158
8.5. Configure X-Pack Watcher 163
8.6. Configure monitoring indexes 170
9.Query logs 173
10.Security 176
10.1. Configure a whitelist to access an Elasticsearch cluster ov... 176
10.2. Reset the access password for an Elasticsearch cluster 177
10.3. Enable HTTPS 178
10.4. Connect two Elasticsearch clusters 180
11.Data backup 186
11.1. View the snapshot feature 186
11.2. Create automatic snapshots and restore data from autom... 188
11.3. Query snapshot status 193
11.4. Commands for creating snapshots and restoring data 197
11.5. Configure a shared OSS repository 206
12.Data visualization 210
12.1. Kibana 210
12.1.1. Log on to the Kibana console 210
12.1.2. Configure the language of the Kibana console 211

12.1.3. Configure a whitelist for access to the Kibana console... 212
12.1.4. Install a Kibana plug-in 213
12.1.5. Use the bsearch_querybuilder plug-in to query data 215
12.1.6. Use the bsearch_label plug-in to label data 219
13.FAQ 223
13.1. Incorrect configuration selected for an Alibaba Cloud Elas... 223
13.2. Access to an Alibaba Cloud Elasticsearch cluster from the... 224
13.3. Kibana console password 229
13.4. Installation errors of a custom plug-in 229
13.5. FAQ about Alibaba Cloud Elasticsearch clusters 231
13.6. FAQ about open source Elasticsearch 244

1.Cross-zone instance deployment

1.1. Precautions

Cross-zone Elasticsearch cluster

A cross-zone Elasticsearch cluster offers improved disaster recovery capabilities. This topic describes the precautions for deploying and using a cross-zone Elasticsearch cluster.

When you purchase an Elasticsearch cluster, you can select the number of zones for it. If you select two or three zones, the system deploys the cluster across these zones. During deployment, you do not need to select each zone. The system automatically selects the zones. For more information, see Create an Elasticsearch cluster.

> **Notice** Currently, you can deploy an Elasticsearch cluster across three zones only in the China (Hangzhou), China (Beijing), China (Shanghai), or China (Shenzhen) region.

Nodes

- You must purchase dedicated master nodes.
- The numbers of data nodes, warm nodes, and client nodes must be a multiple of the number of zones. For more information about zones, see Regions and zones.

Index replicas

- If your Elasticsearch cluster is deployed across two zones but one zone fails, the remaining zone needs to continue providing services. Therefore, you must configure at least one replica for each index.

  By default, five primary shards and one replica are enabled for each index. If you do not have specific requirements on read performance, you can use the default setting.

- If your Elasticsearch cluster is deployed across three zones but one or two of them fail, the remaining zones need to continue providing services. Therefore, you must configure at least two replicas for each index.

  By default, five primary shards and one replica are enabled for each index. Therefore, you must modify the index template to adjust the default number of replicas. For more information, see Index Templates. The following example code demonstrates how to modify the index template to set the number of replicas to 2:

```
PUT _template/template_1
{
 "template": "*",
 "settings": {
  "number_of_replicas": 2
 }
}
```

Cluster configuration

The system automatically configures shard allocation awareness policies for cross-zone Elasticsearch clusters. For more information, see Shard allocation awareness.

The following table lists the parameter configurations of an Elasticsearch cluster deployed across the cn-hangzhou-f and cn-hangzhou-g zones.

Parameter	Description	Example value
cluster.routing.allocation.awareness.attributes	**Notice** Do not call an Elasticsearch API operation to change the value of this parameter. Otherwise, exceptions may occur. Specifies the node attributes that are used to configure a shard allocation awareness policy for a cross-zone Elasticsearch cluster. The cluster identifies the zone of a node based on the Enode.attr.zone_id parameter added to the startup parameter of the node. This parameter is fixed to zone_id . **Note** When you use a cross-zone Elasticsearch cluster, the system adds -Enode.attr.zone_id to the startup parameter of each node in the cluster. For example, if a node is deployed in the cn-hangzhou-g zone, the system adds -Enode.attr.zone_id=cn-hangzhou-g to the startup parameter of the node.	"zone_id"
cluster.routing.allocation.awareness.force.zone_id.values	Specifies whether forced awareness is enabled for shard allocation. Assume that the index of an Elasticsearch cluster deployed across the cn-hangzhou-f and cn-hangzhou-g zones contains one primary shard and three replica shards. Based on the shard allocation awareness policy, the system allocates two shards to each of the cn-hangzhou-f and cn-hangzhou-g zones. If the cluster.routing.allocation.awareness.force.zone_id.values parameter is specified, when the cn-hangzhou-f zone fails, forced awareness prevents the system from reallocating the shards of the cn-hangzhou-f zone to the cn-hangzhou-g zone. **Note** By default, this parameter is not specified. You can specify it as required.	["cn-hangzhou-f", "cn-hangzhou-g"]

Switchover and recovery

After a cross-zone Elasticsearch cluster is deployed, you can perform the following operations for it:

- If the nodes of your Elasticsearch cluster in a zone become faulty, you can perform a switchover to remove these nodes. For more information, see Perform a switchover.
- After the faulty nodes recover, you can perform a recovery to add the nodes to the zone again. For more information, see Perform a recovery.

1.2. Perform a switchover

Perform a switchover

If your Elasticsearch cluster is deployed across zones and the nodes in a zone become faulty, you can perform a switchover. The system removes the nodes from this zone and transmits the network data sent from clients only to nodes in the other zones that are in the Enabled state. This topic describes how to perform a switchover.

switchover

Prerequisites

- A cross-zone Elasticsearch cluster is created.

  For more information, see Create an Elasticsearch cluster. When you create a cluster, set **Number of Zones** to **2-AZ** or **3-AZ**.

  > **Notice** Currently, you can deploy an Elasticsearch cluster across three zones only in the China (Hangzhou), China (Beijing), China (Shanghai), or China (Shenzhen) region.

- Replicas are configured for the indexes of the cross-zone Elasticsearch cluster. This ensures normal read and write operations on the cluster after a switchover.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the lower part of the **Basic Information** page that appears, click the **Node Visualization** tab. Then, move the pointer over the zone where you want to perform a switchover and click **Switch Over**.
5. In the **Confirm Operation** dialog box that appears, click **OK**.
   The system then restarts your Elasticsearch cluster to make the switchover take effect. After the switchover succeeded, the state of the zone changes from Enabled to Disabled.

   > **Note** To ensure that your Elasticsearch cluster has sufficient computing resources and the read and write operations on indexes are not affected, the system adds nodes to the zones that are in the Enabled state during a switchover. These nodes may include dedicated master nodes, client nodes, and data nodes.

What's next

If your indexes has replicas before you perform a switchover, the state of your cluster becomes

abnormal (indicated by the color yellow) after the switchover. After you confirm that the switchover succeeded, you can run a command in the Kibana console to set cluster parameters so that the shards in the zone can be allocated to the zones that are in the Enabled state. After the shards are allocated, the state of the cluster becomes normal (indicated by the color green). Example command:

```
PUT /_cluster/settings
{
  "persistent" : {
    "cluster.routing.allocation.awareness.force.zone_id.values" : {"0": null, "1": null, "2": null}
  }
}
```

Note For more information about how to log on to the Kibana console, see Log on to the Kibana console.

1.3. Perform a recovery

Perform a recovery

After the faulty nodes in the zone where a switchover was performed recover, you can perform a recovery. The system adds the nodes that were removed during the switchover to the zone again and transmits the network data sent from clients to nodes in all zones that are in the Enabled state. This topic describes how to perform a recovery.

recovery

Prerequisites

A switchover is performed. For more information, see Perform a switchover.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the lower part of the **Basic Information** page that appears, click the **Node Visualization** tab. Then, move the pointer over the zone where you want to perform a recovery and click **Switch Back**.
5. In the **Confirm Operation** dialog box that appears, click **OK**.
   The system then restarts your Elasticsearch cluster to make the recovery take effect. After the recovery succeeded, the state of the zone changes from Disabled to Enabled.

   Note During the recovery, the system removes the nodes that were added during the switchover. These nodes may include dedicated master nodes, client nodes, and data nodes. In addition, the system migrates the data stored on these nodes to other data nodes.

2.Manage instances

2.1. Create clusters

This topic describes how to create an Alibaba Cloud Elasticsearch cluster.

Prerequisites

- An Alibaba cloud account is created.

  For more information, see Create an Alibaba Cloud account.
- A Virtual Private Cloud (VPC) and a VSwitch are created.

  For more information, see Create an IPv4 VPC network.

Procedure

1. Access the buy page of Alibaba Cloud Elasticsearch.
2. Complete the cluster configurations in the wizard.For more information, see Parameters on the buy page.

   > Note
   > - **Pay-As-You-Go**: We recommend that you purchase pay-as-you-go clusters for testing purposes in the development and testing stages.
   > - **Subscription**: Elasticsearch offers a promotional discount for subscription clusters based on the subscription duration.
3. Click **Next: Confirm Order** to preview the cluster configurations.
4. Read and agree to the service agreement by selecting the check box, and click **Buy Now**.
5. After the page shows that the Elasticsearch cluster is created, click **Alibaba Cloud Elasticsearch console** to access the **Overview** page.
6. In the left-side navigation pane, click **Elasticsearch Clusters**. On the **Clusters** page, the created Alibaba Cloud Elasticsearch cluster is displayed.

2.2. Clusters

The Clusters page of Alibaba Cloud Elasticsearch displays basic information of Elasticsearch clusters. This page also provides a number of features such as cluster creation, alerting configuration, tag binding, cluster status updating, cluster list exporting, cluster list customization, and cluster management.

Log on to the Alibaba Cloud Elasticsearch console. The **Clusters** page shows all Elasticsearch clusters in the current region under your account. You can perform the following tasks on this page.

Operation	Description

Operation	Description
View the cluster list	You can choose columns you want to display in the cluster list. The cluster list includes these columns: **Cluster ID/Name**, **Tag**, **Status**, **Version**, **Cluster Type**, **Data Nodes**, **Cluster Specification**, **Zone**, **Billing Method**,**Network Type**, and **Created At**.
View basic cluster information	Click a cluster ID in the **Cluster ID/Name** column to open the **Basic Information** page. For more information, see View basic information of a cluster.
Create a cluster	Click **Create** to open the buy page and then create a cluster. For more information, see Create clusters.
Enable/disable alerting	By clicking **Alarms**, you can enable the alerting feature for Alibaba Cloud Elasticsearch in the CloudMonitor console. By default, this feature is disabled for Alibaba Cloud Elasticsearch. After alerting is enabled for Elasticsearch, rules are created to detect exceptions such as abnormal cluster status, high disk usage (> 75%), and high JVM heap memory (> 85%). These rules are applied to all Elasticsearch clusters under your Alibaba Cloud account. For more information, see Configure the monitoring and alerting feature in Cloud Monitor.
Bind tags to a cluster	If you have a large number of Alibaba Cloud Elasticsearch clusters, you can bind tags to them for easy management. A tag is composed of a key-value pair. You can use keys and values to further classify clusters. For more information, see Manage cluster tags.
Update cluster status	Click the **Refresh** button in the upper-right corner to update the status of the clusters. After an Elasticsearch cluster is created, it is in the **Initializing** state. You can click **Refresh** to update the status of the cluster. After the **Status** of the cluster changes to **Active**, you can use the cluster.
Export the cluster list	The export feature allows you to customize a cluster list and then export it. Click the  icon in the upper-right corner. In the **Export** dialog box, set **Export Mode** and **Translate Heading**, select the columns you want to export, and click **OK**. • **Export Mode**: the mode is used to export the cluster list. Valid values: All and Selected. Default value: All. • **Translate Heading**: specifies whether to translate the heading. Valid values: Yes and No. Default value: Yes. If you set this parameter to**Yes**, exported headings are in Chinese. If you set this parameter to **No**, exported headings are in English, such as instanceId and description. • **Custom**: customizes the columns for exporting: **Cluster ID**, **Cluster Name**, **Tag**, **Version**, **Cluster Type**, **Data Nodes**, **Specification**, **Zone**, **Billing Method**, **Created At**, **VPC ID**, and **Network Type**. By default, all these columns are exported. You can also select columns as required.

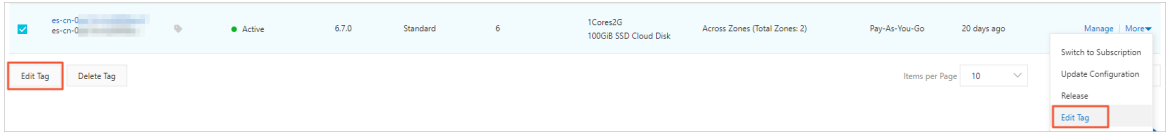
Operation	Description
Customize a cluster list	Click the icon in the upper-right corner. In the **Select Filters** dialog box, select columns as required. **Cluster ID/Name** and **Actions** are dimmed. You cannot select or clear them.
Manage clusters	Click **Manage** in the **Actions** column. On the page that appears, you can perform tasks such as upgrading the cluster, viewing logs, configuring security settings, and configuring plug-ins.
Switch to subscription	This feature only applies to pay-as-you-go clusters. Click **More** in the **Actions** column and select **Switch to Subscription**. On the **Confirm Order** page, find the target cluster and change the billing method of the cluster.
Update the configuration of a cluster	Click **More** in the **Actions** column, select **Update Configuration** to open the **Update** page, and then modify the cluster configuration. For more information, see Upgrade the configuration of a cluster.
Release a cluster	Click **More** in the **Actions** column and select **Release**. In the **Release** message, find the target cluster and click OK. **Warning** After a cluster is released, all the data on the cluster is lost and cannot be recovered. Exercise caution when your release a cluster.

2.3. Manage cluster tags

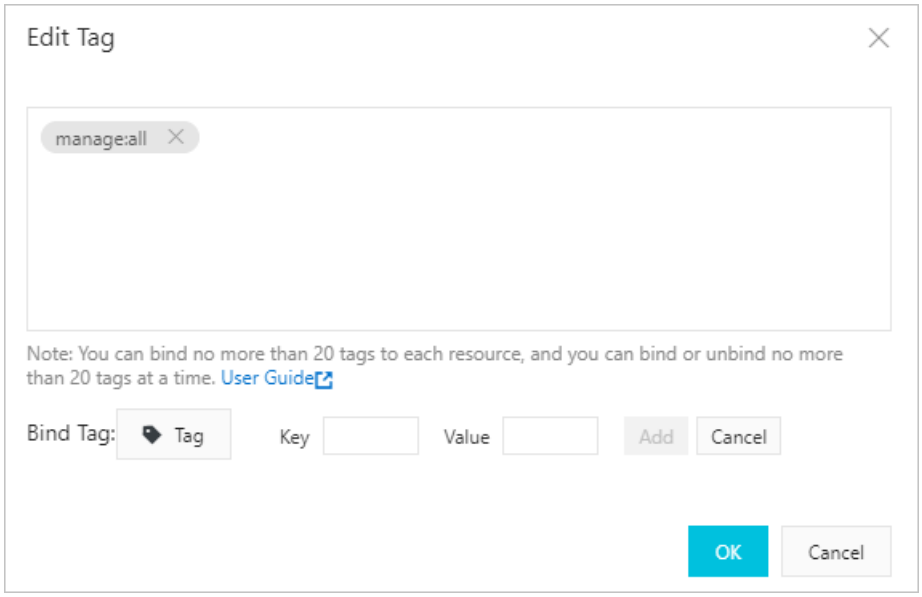
If you have a large number of Alibaba Cloud Elasticsearch clusters, you can bind tags to them for easy management. A tag is composed of a key-value pair. You can use keys and values to further classify clusters.

Create a tag

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, create tags for one or more clusters.



   - To create a tag for a single cluster, click **More** in the **Actions** column that corresponds to the target cluster and select **Edit Tag**.
   - To create tags for multiple clusters, select target clusters and click **Edit Tag**.
4. In the **Edit Tag** dialog box, click **Add Tag**.
5. Set **Key** and **Value** and click **Add**.



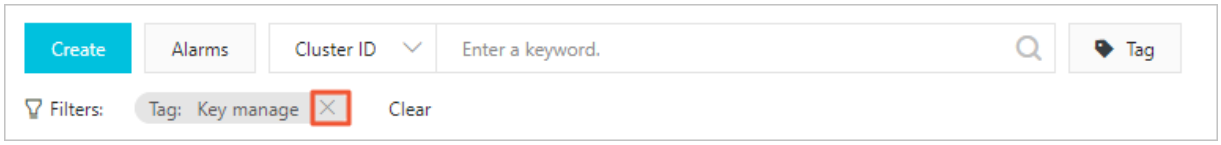
> **Note**
>
> - You can bind no more than 20 tags to each cluster. Tag keys must be unique. Two or more tags with the same key overwrite each other.
> - You can bind tags to up to 50 clusters at a time.
> - Clusters deployed in different regions do not share the same tag.
> - After you unbind a tag, the tag is deleted if it is not bound to other clusters.

6. Click **OK**.

Use tags to filter Elasticsearch clusters

After you bind tags to clusters, you can click the **Tag** button next to the search box on the **Clusters** page and select a key and a value to filter the clusters.

You can click the **X** icon next to a tag to delete the filter condition.



Delete a tag

> **Notice**
>
> - After you unbind a tag, the tag is deleted if it is not bound to other clusters.
> - After you unbind a tag, the system automatically deletes this tag two hours later.
> - You can unbind no more than 20 tags at a time.

In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, use one of the following methods to delete tags:

- Delete a tag from a single cluster

i. Find the target cluster, click **More** in the **Actions** column, and select **Edit Tag**.

ii. In the **Edit Tag** dialog box, click the **X** icon next to a tag.

iii. Click **OK**.

- Delete tags from multiple clusters

i. Select target clusters and click **Delete Tag**.

ii. In the **Delete Tag** dialog box, specify **Keys** you want to unbind from the clusters.

iii. Click **OK**.

2.4. Restart a cluster or node

After you modify the configuration of an Elasticsearch cluster or node or perform other operations on them, you may need to manually restart the cluster for the changes to take effect. This topic describes how to restart an Elasticsearch cluster or node in the console.

restart an Elasticsearch cluster restart an Elasticsearch node

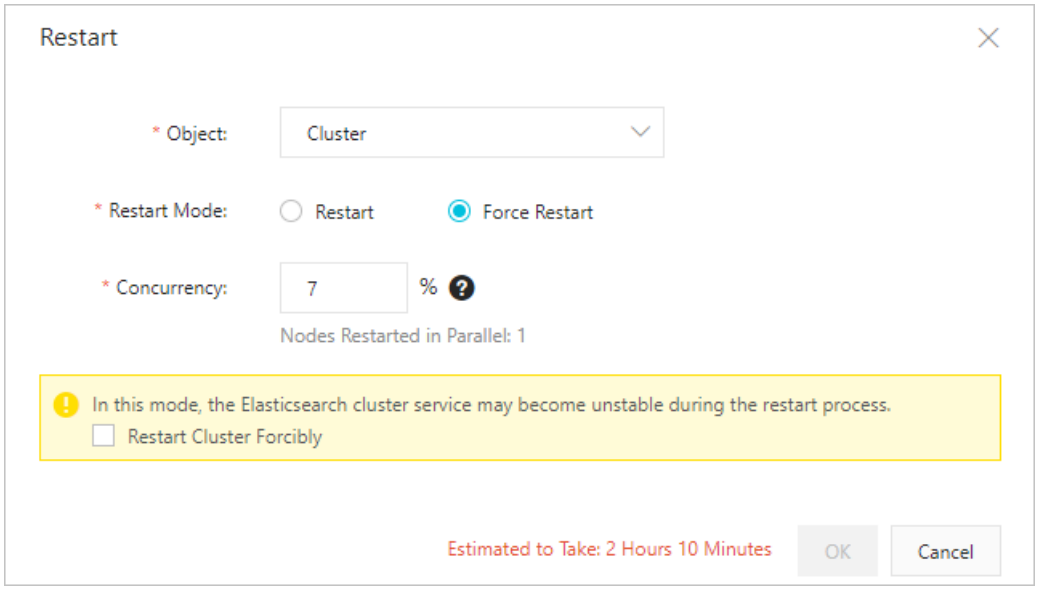
Prerequisites

Before you restart a cluster, make sure that the **status** of the cluster is **Active** (green), the index has at least one replica, and the resource usage is not high.

> **Note** You can view the resource usage on the **Cluster Monitoring** page. For example, the value of **NodeCPUUtilization(%)** is about 80%, the value of **NodeHeapMemoryUtilization** is about 50%, and the value of **NodeLoad_1m** is lower than the number of vCPUs of the current data node. For more information, see Monitoring metrics.

Restart a cluster

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the upper-right corner of the **Basic Information** page, click **Restart**.
5. In the **Restart** dialog box, set parameters to restart a cluster.



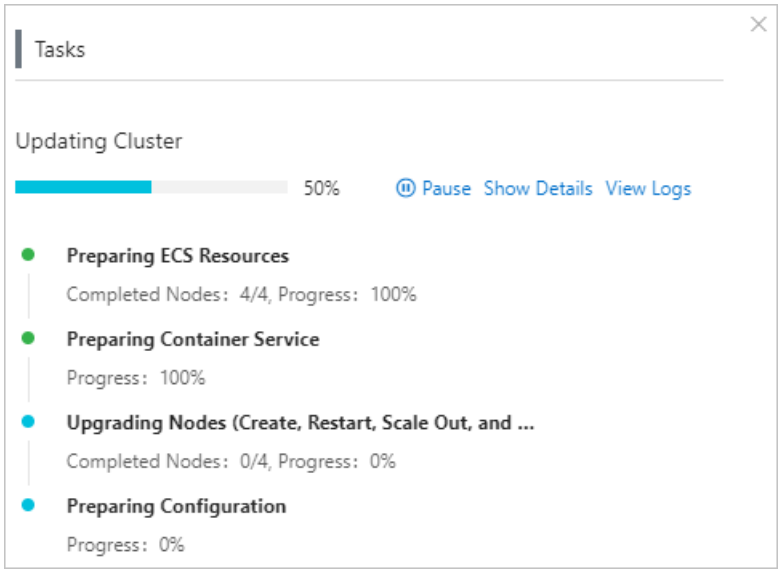
Parameter	Description
Object	**Cluster** and **Node** are supported. ○ **Cluster**: Restarts all nodes in a cluster. ○ **Node**: Restarts a single node. For more information, see Restart a node.

Parameter	Description
Restart mode	Alibaba Cloud Elasticsearch provides two restart modes: **Restart** and **Force Restart**. ○ **Restart**: You can only restart a cluster whose **status** is **Active** (green). Otherwise, you need to forcibly restart the cluster. This restart mode does not affect the service of your cluster but is time-consuming. Notice ■ During the restart, the CPU utilization and memory usage of the nodes in your Elasticsearch cluster will increase sharply. This may affect the stability of your services for a short period of time. ■ The specific amount of time that is used to restart a cluster depends on the volume of data stored on the cluster and the numbers of nodes, indexes, and shards in the cluster. You can view the progress of a restart task in the Tasks dialog box. ○ **Force Restart**: If your Elasticsearch cluster is unhealthy (yellow or red), you can only use the Force Restart mode to restart the cluster. Notice If the disk usage exceeds the value of `cluster.routing.allocation.disk.watermark.low`, the **Status** of your Elasticsearch cluster may be displayed yellow or red. If your Elasticsearch cluster is abnormal, do not perform the following operations on the cluster: node addition, node capacity expansion, disk space expansion, restart, password reset, or other operations that may change the configuration of the cluster. Perform the operations only after the **Status** of the cluster is **Active** (green).
Concurrency	You can set the concurrency of the cluster to improve the restart speed. The higher the concurrency, the faster the force restart. The default concurrency is calculated by dividing one by the total number of nodes in the cluster.
Estimated to Take	This value is calculated by multiplying the average time of the previous restart tasks by the total number of nodes. The actual restart time takes precedence.

6. Click **OK**.

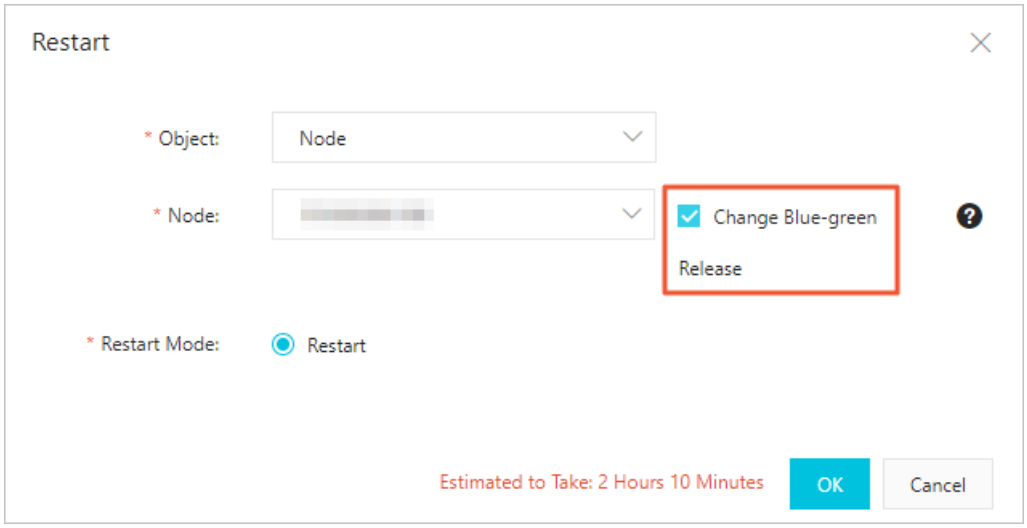
> **Note** If you choose Force Restart, select **Restart Cluster Forcibly** to confirm the restart.

During the restart, the **Status** field is **Initializing** (yellow). You can view details in the **Tasks** dialog box. After the cluster is restarted, the **Status** field is **Active** (green).



Restart a node

Restart a single node. The procedure and precautions for restarting a node are similar to those in Restart a cluster. The differences are as follows:



- In the **Restart** dialog box, set **Object** to **Node**.
- Select the node you want to restart.

> **Notice** If your Elasticsearch cluster is abnormal, you need to forcibly restart nodes in the cluster.

- The system provides the **Change Blue-green Release** feature. If you select it, Elasticsearch adds a node to your Elasticsearch cluster, migrates the data on the original node to the new node, and

removes the original node. If a node experiences a hardware failure, you can use the **Change Blue-green Release** feature to remove the node.

> **Warning**
> - If you use the **Change Blue-green Release** feature, make sure that your Elasticsearch cluster is in the **Active** state (green) and do not set **Restart Mode** to **Force Restart**.
> - If you select **Change Blue-green Release**, the IP address of the node changes after the restart.

2.5. View the progress of a cluster task

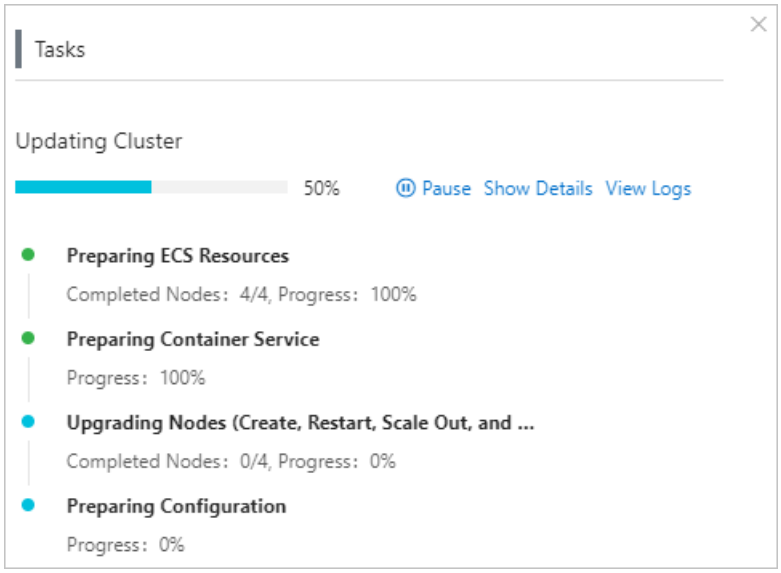
You can click the Tasks icon to view the progress of a running task, such as a cluster creation or restart task.

Prerequisites

The cluster is in the **Initializing** state.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. Click the ☰ icon in the upper-right corner.
5. In the **Tasks** dialog box, view the progress of the cluster updating.
6. Click **Show Details** to view detailed information about the cluster updating task.



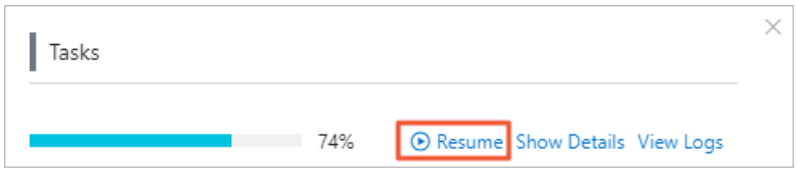
To pause a task, follow these steps:

7. Pause a task.

i. Click **Pause** in the Tasks dialog box.

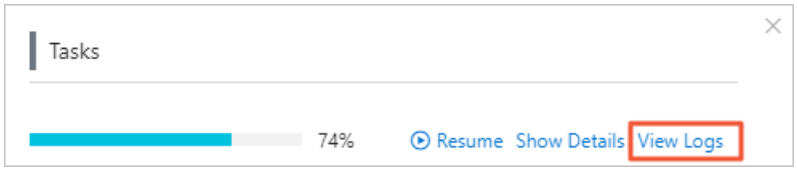
ii. In the **Pause Updates** dialog box, select **I Have Read and Understand the Risks** after you read the agreement, and click **OK**.

8. Resume the task. After you pause a task, you can click **Resume** in the **Tasks** dialog box to resume the task.



> **Notice**
>
> - An Elasticsearch cluster is in the **Paused** state if it has a task paused. If your services running on the cluster are affected, resume the task or run the task again. You can only resume a cluster configuration upgrade task or plug-in management task.
> - After you click **Resume**, Elasticsearch restarts all the nodes in the cluster, which may take a short period of time.

9. View task logs. Click **View Logs** to access the **Logs** page. On this page, you can view the operation logs of the cluster. For more information, see Query logs.



2.6. Refresh a cluster

If the cluster information in the Alibaba Cloud Elasticsearch console is not updated promptly, you can use this feature to manually update the cluster information. For example, if the status of a cluster displays Failed after you create the cluster, you can manually update its status.

Procedure

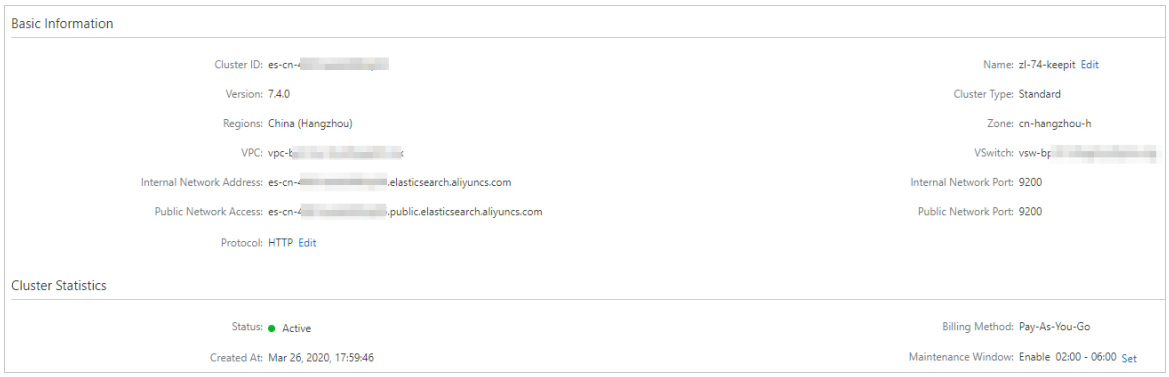
1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the upper-right corner of the **Basic Information** page, click **Refresh**. After the page is refreshed, the information is updated. If the error persists, contact Alibaba Cloud technical support.

2.7. View basic information of a cluster

This topic describes the content displayed on the Basic Information page for an Alibaba Cloud Elasticsearch cluster.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. On the **Basic Information** page, view the basic information and status of a cluster.



Parameter	Description
Cluster ID	The unique ID of the Elasticsearch cluster.
Name	The name of the Elasticsearch cluster. By default, the name of a cluster is the same as its ID. Cluster names are configurable. You can search for clusters by name.
Version	The version of the Elasticsearch cluster. Valid values: **5.5.3**, **5.6.0**, **6.3.2**, **6.7.0**, **6.8.0**, and **7.4.0**. **Notice** If you want to upgrade the cluster version, you can use the version upgrade feature of Alibaba Cloud Elasticsearch. You can only upgrade Alibaba Cloud Elasticsearch clusters from V6.3.2 to V6.7.0. For more information, see Upgrade the version of a cluster.
Cluster Type	The type of the Elasticsearch cluster. Valid values: **Standard** and **Advanced**.
Region	The region where the Elasticsearch cluster resides.
Zone	The zone where the Elasticsearch cluster resides.
VPC	The Virtual Private Cloud (VPC) to which the Elasticsearch cluster belongs.
VSwitch	The VSwitch to which the Elasticsearch cluster belongs.
Tag	The tag bound to the Elasticsearch cluster. You can use tags to classify and manage Elasticsearch clusters. For more information, see Manage cluster tags.

Parameter	Description
Protocol	The protocol used by the Elasticsearch cluster. Default value: **HTTP**. Valid values: **HTTP** and **HTTPS**. .
Internal Network Address	The internal endpoint of the Elasticsearch cluster. In a VPC, you can use an Elastic Compute Service (ECS) instance to connect to the internal endpoint of an Elasticsearch cluster. **Notice** For security and stability purposes, we recommend that you do not connect to an Elasticsearch cluster over the Internet. You can purchase an ECS instance that resides in the same VPC as your Elasticsearch cluster and then use the ECS instance to connect to the internal endpoint of the Elasticsearch cluster.
Internal Network Port	The port used to connect to the Elasticsearch cluster over a private network. Alibaba Cloud Elasticsearch supports the following ports: ○ Port **9200** for HTTP and HTTPS. ○ Port **9300** for TCP. Only Alibaba Cloud Elasticsearch V5.5.3 supports this port. **Note** For Alibaba Cloud Elasticsearch V6.X and later, you cannot use the transport client to access port `9300` .
Public Network Access	The public endpoint of the Elasticsearch cluster. To connect to an Elasticsearch cluster by using its public endpoint, turn on **Public Network Access**. You must enable public network access first on the Network Settings page. For more information, see Configure a whitelist to access an Elasticsearch cluster over the Internet or a VPC.

Parameter	Description
Public Network Port	The port used to connect to the Elasticsearch cluster over the Internet. This parameter is only displayed after you enable **Public Network Access**. Alibaba Cloud Elasticsearch supports the following ports: ○ Port **9200** for HTTP and HTTPS. ○ Port **9300** for TCP. Only Alibaba Cloud Elasticsearch V5.5.3 supports this port. **Note** ■ For Alibaba Cloud Elasticsearch V6.X and later, you cannot use the transport client to access port **9300**. ■ You must configure the public IP address whitelist. By default, all IP addresses are prohibited from accessing Elasticsearch. For more information, see Configure a whitelist to access an Elasticsearch cluster over the Internet or a VPC.
Status	The state of the Elasticsearch cluster. An Elasticsearch cluster has the following states: **Active** (green), **Initializing** (yellow), **Unhealthy** (red), **Paused** (red), and **Expired** (gray).
Billing Method	The billing method of the Elasticsearch cluster. Valid values: **Pay-As-You-Go** and **Subscription**.
Created At	The time when the Elasticsearch cluster was created.
Maintenance Window	The maintenance window for the Elasticsearch cluster. The default value of this parameter is set to **02:00 ~ 06:00**. You can set **Maintenance Window** based on your business needs. For more information, see Set a maintenance window.
Renew	This parameter is only displayed when **Billing Method** is **Subscription**. You can click **Renew** on the right side of **Basic Information** to renew the cluster. The minimum renewal period of a cluster is one month.
Switch to Subscription	This parameter is only displayed when **Billing Method** is **Pay-As-You-Go**. You can click **Switch to Subscription** in the lower-right corner of the **Basic Information** section to change the billing method. With the **Switch to Subscription** feature, you can change the billing method of an Elasticsearch cluster from **Pay-As-You-Go** to **Subscription**. However, no discount is offered when you change the billing method.

2.8. Set a maintenance window

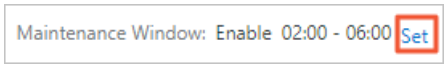
The backend system of Alibaba Cloud Elasticsearch maintains clusters to improve their stability. You can specify a maintenance window within which the backend system maintains your cluster. The default maintenance window is from 02:00 to 06:00. We recommend that you set the maintenance window to the off-peak hours of your business to avoid impacts on your business.

Precautions

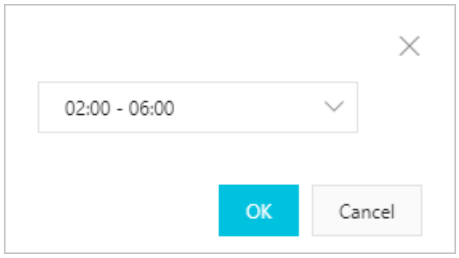
- Before maintenance is performed, the system sends SMS messages and emails to the contacts listed in your Alibaba Cloud account.
- To ensure smooth maintenance, your Alibaba Cloud Elasticsearch cluster enters the **Initializing** state prior to the maintenance window. In this case, you can still access the cluster and perform query operations such as performance monitoring. However, you cannot perform modification operations such as restart and configuration upgrades for the cluster.
- Make sure that you configured automatic reconnection policies for your applications because the cluster may experience transient disconnections within the maintenance window.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. On the **Basic Information** page, click **Set** on the right side of **Maintenance Window** in the **Cluster Statistics** section.



5. In the dialog box that appears, select a maintenance window and click **OK**.



> **Note** The maintenance window is in UTC+8. The duration of the maintenance window is four hours and cannot be customized.

2.9. View node information

Alibaba Cloud Elasticsearch provides the node visualization feature. You can use this feature to view the information of all the nodes in your Elasticsearch cluster on a node diagram.

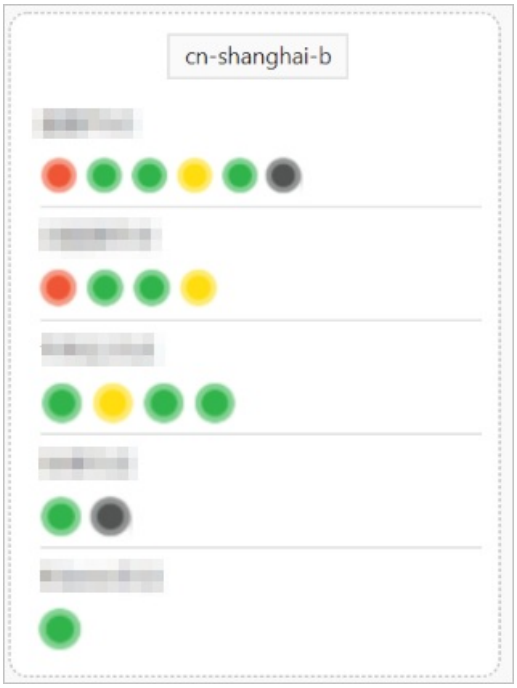
Access the Node Visualization tab

1. Log on to the Alibaba Cloud Elasticsearch console.

2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. On the **Basic Information** page, click the **Node Visualization** tab.

View node status

On the Data Visualization tab, you can view the color of each node in the cluster and check whether the nodes in a cluster are healthy based on their colors.

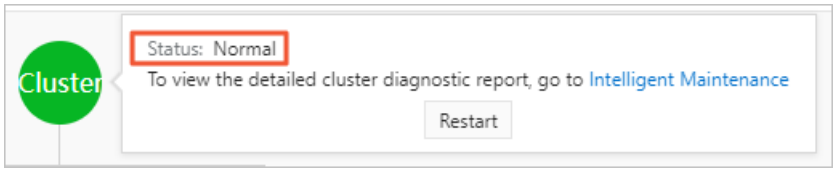


> **Note** The color of a node is determined by the resource usage of the node. The resource usage thresholds are the same as those in CloudMonitor. For more information, see Monitoring metrics

- Red: Warning.
- Yellow: Alert.
- Green: Normal.
- Gray: Unknown. This state indicates that the system has failed to retrieve the node information for a long period of time.

View cluster status

1. Access the Node Visualization tab. For more information, see Access the Node Visualization tab.
2. Move the pointer to **Cluster**.
3. In the message that appears, view the **Status** of the cluster.

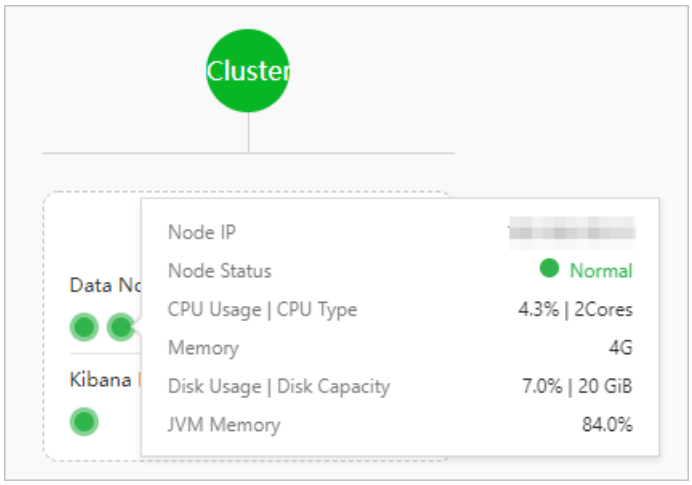


View the cluster diagnosis report

1. Access the Node Visualization tab. For more information, see Access the Node Visualization tab.
2. Move the pointer to **Cluster**.
3. Click **Intelligent Maintenance**.
4. On the **Intelligent Maintenance** page, you can view the cluster diagnosis report. For more information about how to view the cluster diagnosis report, see Overview.

View node information

1. Access the Node Visualization tab. For more information, see Access the Node Visualization tab.
2. Move the pointer to a node.
3. In the dialog box that appears, view the information of the node.



> **Note** If a node is in red or yellow color, the system displays the **The node is disconnected, and we recommend that you use Intelligent Maintenance** message. If a node is in gray color, the system displays the **The node status is unhealthy, and we recommend that you use Intelligent Maintenance** message. You can click **Intelligent Maintenance** in the message and choose **Intelligent Maintenance > Cluster Diagnosis** in the left-side navigation pane of the page that appears to diagnose the cluster. For more information, see Overview.

Restart a node

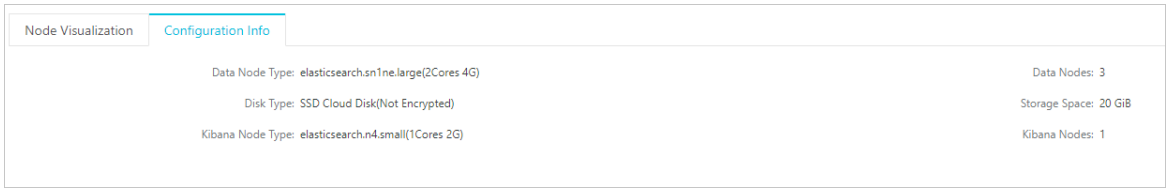
1. Access the Node Visualization tab. For more information, see Access the Node Visualization tab.
2. Move the pointer to a node.
3. Click **Restart**.
4. In the **Restart** dialog box, set parameters. For more information, see Restart a cluster or node.

2.10. View the configuration of an Elasticsearch cluster

You can view the configuration of an Alibaba Cloud Elasticsearch cluster on the Configuration Info tab. The cluster configuration includes the numbers and specifications of data nodes and Kibana nodes.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. On the **Basic Information** page, click the **Configuration Info** tab.
5. On this tab, you can view the configuration of the cluster.



For more information about parameter descriptions, see Parameters on the buy page.

2.11. Release a cluster

Release an Elasticsearch cluster

You can only release pay-as-you-go clusters or subscription clusters that have expired. If a subscription cluster has not expired, you must claim a refund before you can release the cluster. This topic describes how to release a pay-as-you-go cluster.

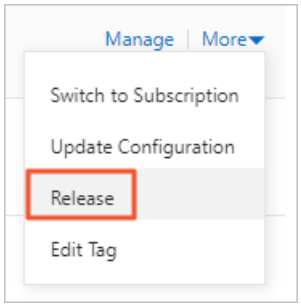
release an Elasticsearch cluster

Prerequisites

After a cluster is released, its data cannot be restored. We recommend that you back up data before you release a cluster. For more information, see Commands for creating snapshots and restoring data.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. On the **Clusters** page, find the target cluster, click **More** in the **Actions** column, and select **Release**.



4. In the message that appears, click **OK**.

3.Data migration

3.1. Migrate nodes in a zone

Migrate Elasticsearch nodes

If the zone where your Elasticsearch cluster resides has insufficient Elastic Compute Service (ECS) instances for a configuration upgrade, you can migrate the nodes in the zone to another zone before the upgrade.

migrate Elasticsearch nodes

Prerequisites

- A zone with sufficient resources exists under the current account.

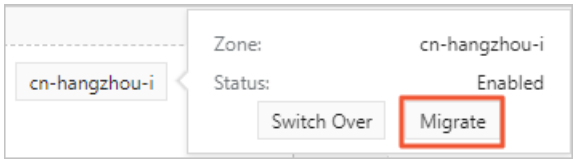
  We recommend that you select new zones from bottom to top in alphabetical order because these zones may have sufficient resources. For example, if both **cn-hangzhou-e** and **cn-hangzhou-h** are available, select **cn-hangzhou-h**. After you migrate nodes to another zone, you must manually update the configuration of your Elasticsearch cluster. For more information, see Upgrade the configuration of a cluster.

- The cluster is in the Active state.

  You can run the `GET _cat/health?v` command to check the status of your Elasticsearch cluster.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. On the **Basic Information** page, click the **Node Visualization** tab. Then, move your pointer to the zone and click **Migrate**.

5. In the **Node Migration** dialog box, set **Target Zone** and **VSwitch**.

Parameter	Description
Target Zone	The target zone may also have insufficient ECS instances. We recommend that you select new zones from bottom to top in alphabetical order. For example, if both **cn-hangzhou-e** and **cn-hangzhou-h** are available, select **cn-hangzhou-h**.
VSwitch	If your Elasticsearch cluster is deployed in one zone, you must specify a new **VSwitch**. If your Elasticsearch cluster is deployed across zones or is an Alibaba Finance Cloud cluster, you do not need to specify a new **VSwitch**.

> Notice
>
> - After migration, the IP addresses of the nodes in the cluster change. If you specified the IP addresses in the cluster configuration, update them after the migration.
> - Node migration triggers a cluster restart, but the cluster can still provide services during the restart. However, this may cause service instability. We recommend that you perform this operation during off-peak hours.

6. Read and agree the terms of data migration, and click **OK**. The system then restarts the cluster.
   After the cluster is restarted, its nodes are migrated to the target zone.
   After the migration is complete, the zone specified by **Target Zone** is used.

4.Upgrade

4.1. Upgrade the version of a cluster

This topic describes how to upgrade the version of an Elasticsearch cluster. Currently, you can upgrade an Elasticsearch cluster only from V6.3.2 to V6.7.0 with one click.

Elasticsearch cluster version upgrade

Prerequisites

A version upgrade check is performed.

For more information about relevant check items, see Check items before a version upgrade.

Precautions

- When you upgrade the version of an Elasticsearch cluster, you can continue to read data from or write data to the cluster, but you cannot cancel the upgrade or make other changes. We recommend that you perform a version upgrade during off-peak hours.
- To upgrade the version of a cluster, add nodes of the target version to the cluster, migrate data stored on nodes of the source version to the new nodes, and remove the nodes of the source version. This causes the IP addresses of nodes to change.

Check items before a version upgrade

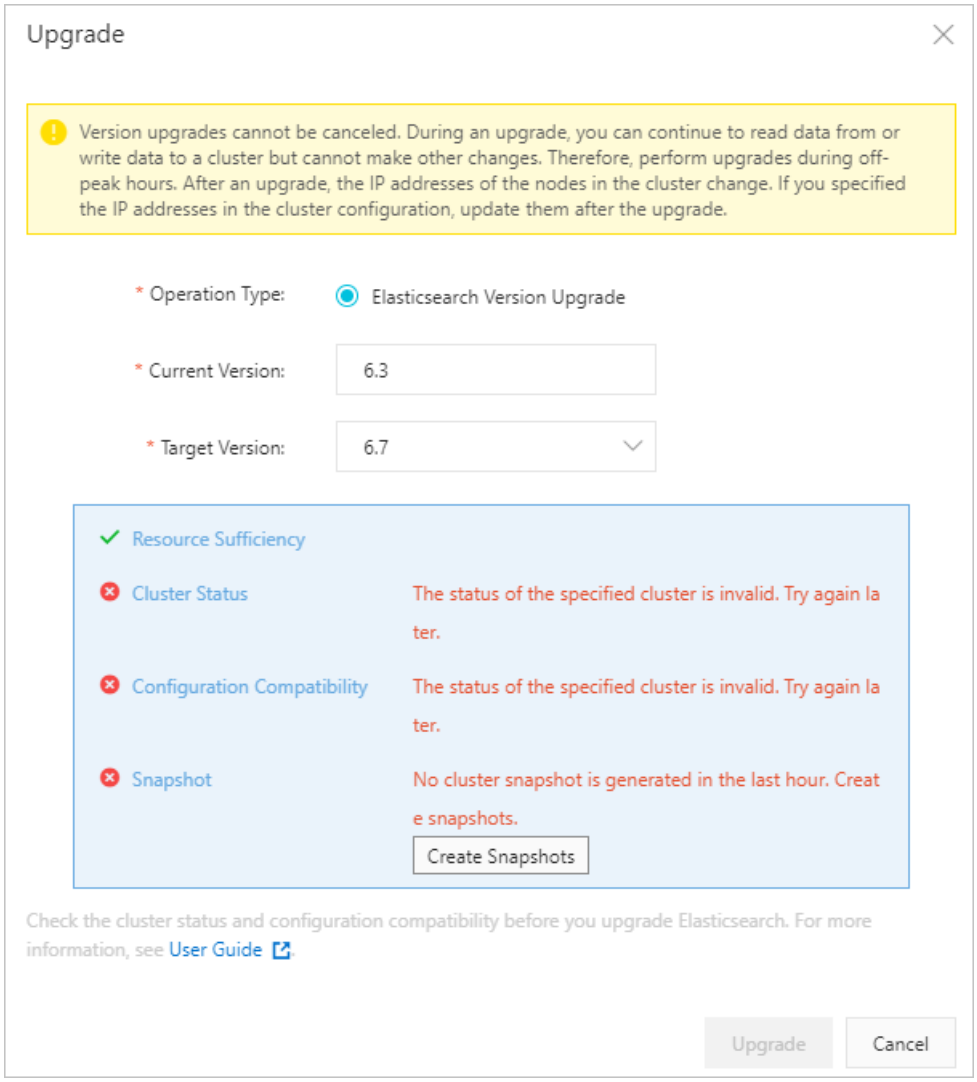
Before you perform a version upgrade, check items listed in the following table. You can only upgrade the versions of Elasticsearch clusters that are in normal states.

Check item	Normal state
Cluster state	The cluster is in the Active state (indicated by the color green).
JVM heap memory usage	The JVM heap memory usage of the cluster is less than 75%.
Disk usage	The disk usage of nodes is less than the value of `cluster.routing.allocation.disk.watermark.low`.
Replica	All indexes are configured with replicas.
Snapshot	The cluster created snapshots within the last hour.
Custom plug-in	The cluster does not have custom plug-ins installed.
ECS instance in the zone where the cluster resides	The zone where the cluster resides contains sufficient ECS instances. **Note** To upgrade the version of a cluster, add nodes of the target version to the cluster, migrate data stored on nodes of the source version to the new nodes, and remove the nodes of the source version. Therefore, before the upgrade, ensure that the zone where the cluster resides has sufficient ECS instances.

Check item	Normal state
YML file configuration	The cluster of the target version is compatible with the YML file configuration of the source version.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the upper-right corner of the **Basic Information** page, click **Update and Upgrade**.
5. In the **Upgrade** dialog box, select the target version.
6. Click **Precheck**. The system then checks the configuration compatibility, status, snapshots, and basic resources of the cluster.



After the check is complete, handle exceptions as prompted. For example, if the cluster has not created snapshots within the last hour, you can click **Create Snapshots** to trigger the snapshot operation.

7. After the check is successful, click **Upgrade**. During the upgrade, you can view the upgrade progress in the Tasks dialog box.
   After the upgrade, you can view the cluster version on the **Basic Information** page.

   > **Notice** After the upgrade, the IP addresses of nodes change. If you specified the IP addresses in the cluster configuration, update them after the upgrade.

4.2. Check for and modify incompatible configurations before you perform an upgrade from V5.6 to V6.3

A later version may be incompatible with some configurations in an earlier version. If you do not modify these configurations before a version upgrade, your services may be affected after the upgrade. Therefore, before the upgrade, you must check for incompatible configurations and modify the configurations as required. This topic describes the manual checks that you must perform and the automatic check items before an upgrade from V5.6.16 to V6.3.2. It also provides the methods that can be used to modify the incompatible configurations.

Context

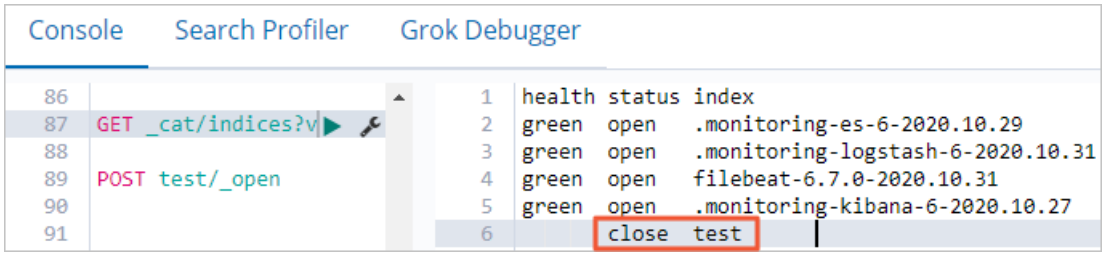
When you check for incompatible configurations, take note of the following points:

- Before a version upgrade, a precheck is performed. You must identify incompatible configurations based on the precheck results and modify the configurations as required. For more information, see Upgrade the version of a cluster.
- The commands provided in this topic can be executed in the Kibana console. For more information about how to log on to the Kibana console, see Log on to the Kibana console.

Perform manual checks

1. Split each multi-type index on the Elasticsearch V5.X cluster into multiple single-type indexes.Elasticsearch clusters of V6.X or later do not support multi-type indexes. If the V5.X cluster contains multi-type indexes, you can write data to the indexes but cannot create multi-type indexes after the cluster version is upgraded to V6.X. If you create multi-type indexes, errors are reported. Therefore, before the upgrade, we recommend that you split each multi-type index into single-type indexes.
2. Check whether the cluster contains indexes that are in the close state.

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

   - Yes: Run the following command to open the indexes that are in the close state:

```
POST test/_open
```

- No: Go to the next step.

3. Check whether the cross-cluster search feature is enabled for the cluster.

```
GET _cluster/settings
```

- Yes: Disable the feature. You can enable it after the upgrade.

```
PUT _cluster/settings
{
 "persistent": {
  "search.remote.*": null
 },
 "transient": {
  "search.remote.*": null
 }
}
```

> **Notice** The search.remote parameter is used to configure the cross-cluster search feature in V5.X, whereas the cluster.remote parameter is used in V6.X.

- No: No actions are required.

Automatic check items

Before you upgrade the version of your cluster, you must click **Precheck** to check the cluster. For more information, see Upgrade the version of a cluster. Then, the system performs a compatibility check based on the following table.

Compatibility check items

The following table lists the parameters that are deprecated in V6.0 and later. For more information, see Breaking changes in 6.0.

> **Notice** If the index template contains the related configurations listed in the following table after the version upgrade, the template cannot be used to create indexes.

No.	Configuration level	Category	Parameter
1	Cluster	Snapshot settings	cluster.routing.allocation.snapshot.relocation_enabled
2		Storage throttling settings	indices.store.throttle.type and indices.store.throttle.max_bytes_per_sec
3		Similarity settings	index.similarity.base

No.	Configuration level	Category	Parameter
4	Index	Shadow replica settings	index.shared_filesystem and index.shadow_replicas
5		Index storage settings	index.store.type
6		Storage throttling settings	index.store.throttle.type and index.store.throttle.max_bytes_per_sec
7		include_in_all setting in the mappings configuration of an index	include_in_all **Note** The indexes that are created before an upgrade from V5.X to V6.X and have this parameter configured can still be used after the upgrade. The indexes that are created after the upgrade do not support this parameter.
8		Version settings for index creation	index.version.created **Note** This parameter specifies that indexes cannot be upgraded across major versions. For example, you cannot directly upgrade the indexes created in V5.X to V7.X. Before an upgrade from V5.X to V7.X, you must call the reindex operation to migrate data in the indexes to a V7.X cluster. Then, delete the indexes from the V5.X cluster and upgrade the version of the V5.X cluster.
9		Similarity settings	index.similarity.base
10		Shadow replica settings	index.shared_filesystem and index.shadow_replicas
11		Index storage settings	index.store.type
12		Storage throttling settings	index.store.throttle.type and index.store.throttle.max_bytes_per_sec

No.	Configuration level	Category	Parameter
13	Index template	include_in_all setting in the mappings configuration of the index template	include_in_all
14		_all setting in the mappings configuration of the index template	_all
15		Type settings in the mappings configuration of the index template	None **Note** Check whether the mappings configuration in the index template contains multiple type settings.

Note The preceding check items are at the CRITICAL level. If a CRITICAL check item is reported, the cluster fails the compatibility check, and its version cannot be upgraded. A later version is incompatible with the configurations indicated by CRITICAL check items. If a WARNING check item is reported, the cluster fails the compatibility check, but its version can still be upgraded. The configurations indicated by WARNING check items are ignored after the upgrade.

Modify incompatible configurations

This section provides the methods that are used to modify incompatible configurations.

- Cluster-level incompatible configurations

  For cluster-level incompatible configurations, you can disable the configurations.

Configuration category	Command to disable the configuration

Configuration category	Command to disable the configuration

Snapshot settings	PUT _cluster/settings
	{
	"persistent": {
	"cluster.routing.allocation.snapshot.relocation_enabled": null
	},
	"transient": {
	"cluster.routing.allocation.snapshot.relocation_enabled": null
	}
	}

Storage throttling settings	PUT _cluster/settings
	{
	"persistent": {
	"indices.store.throttle.type": null,
	"indices.store.throttle.max_bytes_per_sec": null
	},
	"transient": {
	"indices.store.throttle.type": null,
	"indices.store.throttle.max_bytes_per_sec": null
	}
	}

- Index-level incompatible configurations

  For index-level incompatible configurations, you can disable the configurations.

Configuration category	Command to disable the configuration	Additional information

<table>
<tr><th>Configuration category</th><th>Command to disable the configuration</th><th>Additional information</th></tr>
<tr><td>Similarity settings</td><td><pre>PUT test_index/_settings
{
  "index.similarity.base.*": null
}</pre></td><td rowspan="3">These configurations can be modified only after indexes are closed. You cannot read data from or write data to closed indexes. After the modifications, you can open the indexes. The following example shows how to open and close the test_index index:
◦ Close the index
<pre>POST test_index/_close</pre>◦ Open the index
<pre>POST test_index/_open</pre></td></tr>
<tr><td>Shadow replica settings</td><td><pre>PUT test_index/_settings
{
  "index.shared_filesystem": null,
  "index.shadow_replicas": null
}</pre></td></tr>
<tr><td>Index storage settings</td><td><pre>PUT test_index/_settings
{
  "index.store.type": null
}</pre></td></tr>
<tr><td>Storage throttling settings</td><td><pre>PUT test_index/_settings
{
 "settings": {
  "index.store.throttle.type": null,
  "index.store.throttle.max_bytes_
per_sec": null
 }
}</pre></td><td>None.</td></tr>
</table>

Note Indexes that have the include_in_all parameter configured can still be used in the later version. You do not need to modify this parameter.

- Index template-level incompatible configurations

  The following example shows how to modify configurations in the test_template index template:

  i. Run the `GET _template/test_template` command to query the test_template index template.

  The query result shows that the index template contains the following incompatible configurations: index storage settings, _all, and include_in_all.

```
{
 "test_template": {
  "order": 0,
  "template": "test_*",
  "settings": {
   "index": {
    "store": {
     "throttle": {
      "max_bytes_per_sec": "100m"
     }
    }
   }
  },
  "mappings": {
   "test_type": {
    "_all": {
     "enabled": true
    },
    "properties": {
     "test_field": {
      "type": "text",
      "include_in_all": true
     }
    }
   }
  },
  "aliases": {}
 }
}
```

ii. Delete the incompatible configurations and run the `PUT _template/test_template` command to update the index template.

```
PUT _template/test_template
{
  "order": 0,
  "template": "test_*",
  "settings": {
  },
  "mappings": {
   "test_type": {
    "properties": {
     "test_field": {
      "type": "text"
     }
    }
   }
  },
  "aliases": {}
}
```

5.Upgrade or downgrade a cluster

5.1. Update the kernel of a cluster

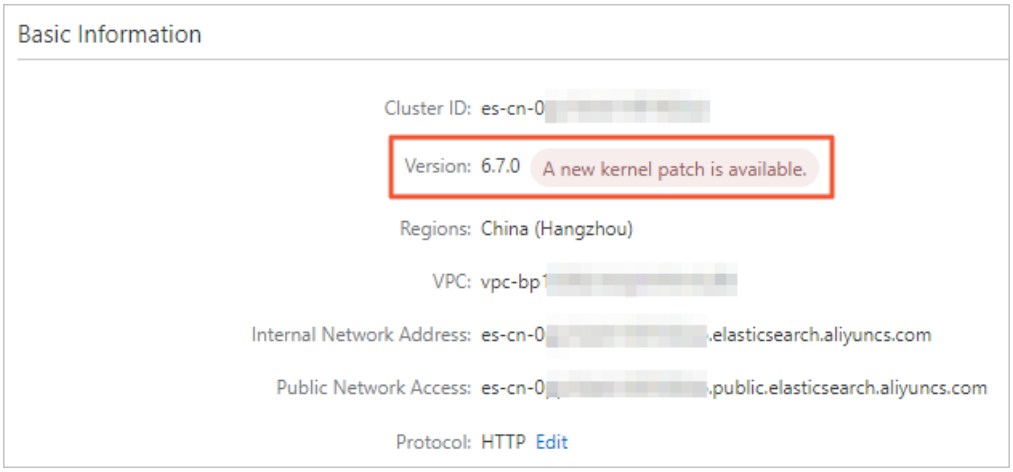
This topic describes how to update the kernel of an Alibaba Cloud Elasticsearch cluster to support the new features developed based on the open-source Elasticsearch kernel. Only the kernels of Alibaba Cloud Elasticsearch V6.7 clusters can be updated.

kernel update

Prerequisites

- A new kernel version is available.

  You can go to the Basic Information page of a cluster to check whether a new kernel version is available for the cluster.



- A check is performed for the kernel update.

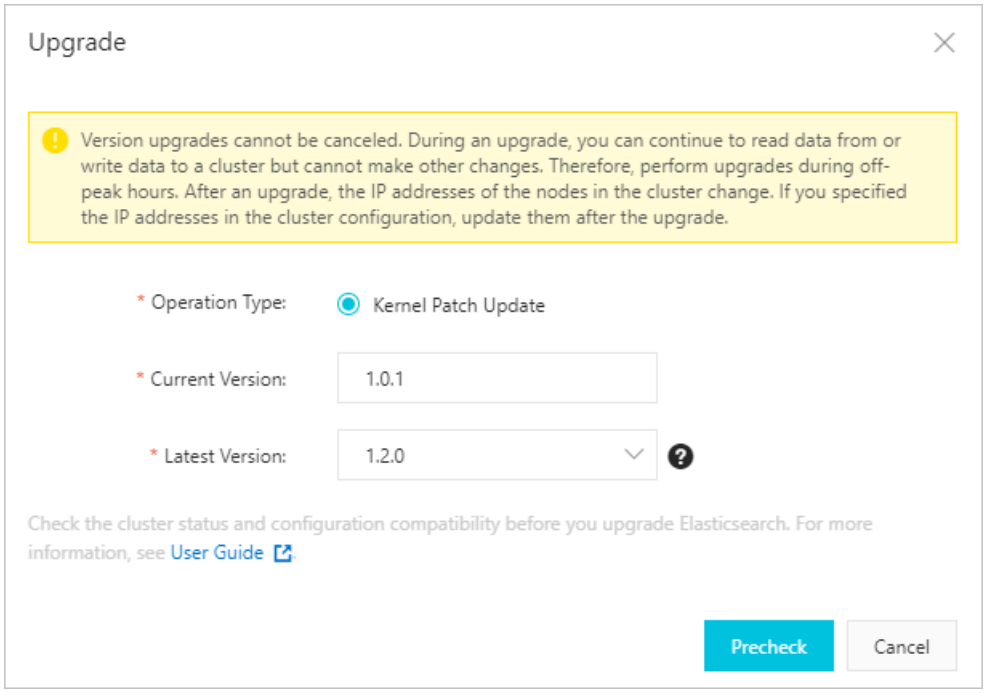
  For more information about relevant check items, see Upgrade the version of a cluster.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the upper-right corner of the **Basic Information** page, click **Update and Upgrade**.

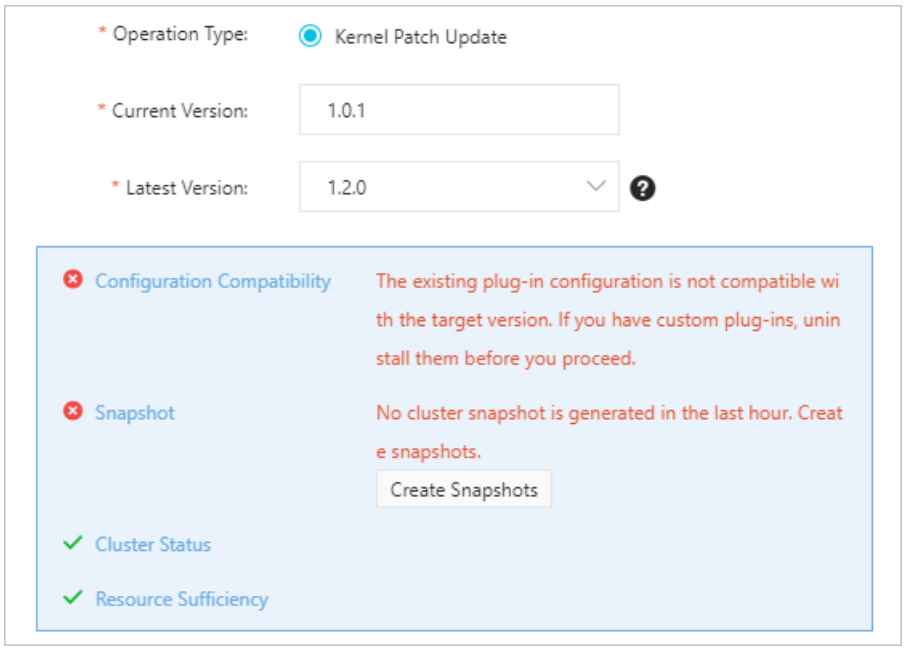
   > **Notice** Updating a kernel does not change the version of the cluster. The **Update and Upgrade**AliES release notes.

5. In the **Upgrade** dialog box, select the target version.



> **Note** When you update the kernel for the first time, the value of **Current Version** is **None** by default.

6. Click **Precheck**. The system then checks the configuration compatibility, status, snapshots, and basic resources of the cluster.



After the check is complete, handle exceptions as prompted. For example, if the cluster has not created snapshots within the last hour, you can click **Create Snapshots** to trigger the snapshot operation.

7. After the check is successful, click **Upgrade**. During the kernel update, you can view the update progress in the Tasks dialog box.
After the kernel is updated, the **A new kernel patch is available.** message next to **Version** and

the **Update and Upgrade** button are no longer displayed.

5.2. Scale in an Elasticsearch cluster

Alibaba Cloud Elasticsearch allows you to remove data nodes from an Elasticsearch cluster to scale in it.

Precautions

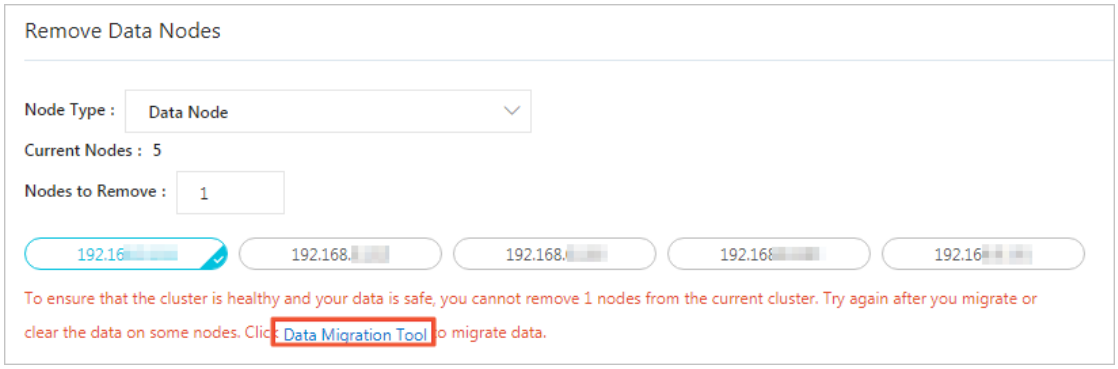
After you remove data nodes from an Elasticsearch cluster, the system restarts the cluster. Therefore, before the removal, make sure that the restart does not affect your business.

Remove data nodes

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the lower-right corner of the **Basic Information** page, choose **Configuration Update > Remove Data Nodes**.
5. In the **Remove Data Nodes** section of the page that appears, set **Node Type**.
6. Select the data nodes that you want to remove.

   > **Notice** The number of reserved nodes must be greater than two and greater than half of the existing nodes.

7. Migrate data.For security purposes, make sure that the data nodes you want to remove store no data. If these data nodes store data, the system prompts you to migrate the data. After the data is migrated, no index data is stored on the data nodes and no index data will be written to them.
   i. Click **Data Migration Tool** in the message that appears.

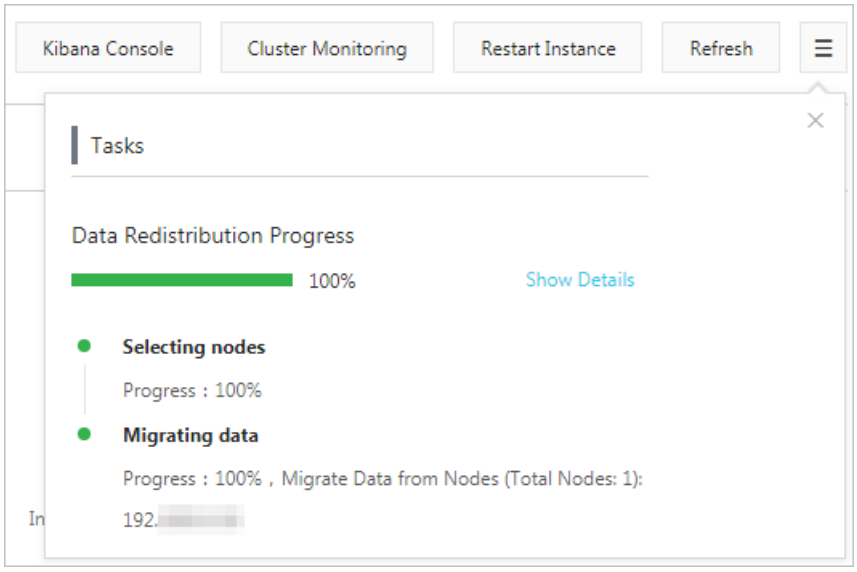
   

   ii. In the **Migrate Data** dialog box, select a data migration method.

Parameter	Description
Smart Migration	The system automatically selects data nodes for data migration.
Custom Migration	You must select the data nodes whose data you want to migrate.

   iii. Read and agree to the terms of data migration, and click **OK**.

iv. Click **OK**.
The system then restarts the cluster. During the restart, you can view the data migration progress in the Tasks dialog box. After the cluster is restarted, the data stored on the data nodes that you want to remove is migrated.



Note During the data migration, you can click **Pause** in the **Tasks** dialog box to stop the migration.

8. In the lower-right corner of the **Basic Information** page, choose **Configuration Update > Remove Data Nodes** again.
9. In the **Remove Data Nodes** section of the page that appears, select the data nodes whose data is migrated and click **OK**.
The system then restarts the cluster. During the restart, you can view the data migration progress in the Tasks dialog box. After the cluster is restarted, the data nodes are removed from the cluster.

Roll back data migration

Data migration is time-consuming. Cluster status changes or data modifications may result in a data migration failure. You can view detailed information in the Tasks dialog box. To roll back data migration, perform the following steps:

1. Log on to the Kibana console of the current cluster.For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**. On the **Console** tab of the page that appears, run the following command to obtain the IP addresses of the data nodes whose data is migrated:

```
GET _cluster/settings
```

If the command is successfully executed, the following result is returned:

```
{
 "transient": {
  "cluster": {
   "routing": {
    "allocation": {
     "exclude": {
      "_ip": "192.168.xx.xx,192.168.xx.xx,192.168.xx.xx"
     }
    }
   }
  }
 }
}
```

3. Roll back data.
   - Roll back the data on some data nodes. Use the exclude parameter to exclude the data nodes whose data you do not want to roll back.

     ```
     PUT _cluster/settings
     {
      "transient": {
       "cluster": {
        "routing": {
         "allocation": {
          "exclude": {
           "_ip": "192.168.xx.xx,192.168.xx.xx"
          }
         }
        }
       }
      }
     }
     ```

   - Roll back the data on all data nodes.

```
PUT _cluster/settings
{
 "transient": {
  "cluster": {
   "routing": {
    "allocation": {
     "exclude": {
      "_ip": null
     }
    }
   }
  }
 }
}
```

4. Run the following command to check whether the data is rolled back:

```
GET _cluster/settings
```

If the IP addresses of the data nodes whose data is rolled back are not contained in the command output, the rollback is successful. You can also check the rollback progress based on whether shards are reallocated to the data nodes.

> **Note** To check the status of a data migration or rollback task, run the `GET _cat/shards?v` command.

FAQ

- What do I do if the "This operation may cause a shard distribution error or insufficient storage, CPU, or memory resources." message appears?

  Cause

  - Insufficient resources

    After data nodes are removed, the cluster does not have sufficient disk capacity, memory, or CPU resources to store system data or handle workloads.

  - Shard allocation errors

    Based on Lucene principles, Elasticsearch does not migrate two or more replica shards of the same index on a data node to the same data node. In this case, after data nodes are removed, the number of replica shards in a cluster may be greater than or equal to the number of data nodes. This results in shard allocation errors.

  Solution

- Insufficient resources

  `Run the GET _cat/indices?v` command to check whether the resource usage, such as the disk usage, is greater than the threshold. You must make sure that the cluster has sufficient resources to store data or process requests. If these requirements are not met, upgrade the configuration of the cluster. For more information, see Upgrade the configuration of a cluster.

- Shard allocation errors

  Run the `GET _cat/indices?v` command to check whether the number of replica shards in the cluster is less than the number of data nodes after data nodes are removed. If this requirement is not met, change the number of replica shards. For more information, see Index Templates. The following sample code demonstrates how to modify the index template to set the number of replica shards to 2:

  ```
  PUT _template/template_1
  {
   "template": "*",
   "settings": {
    "number_of_replicas": 2
   }
  }
  ```

- What do I do if the "The cluster is running tasks or in an error status. Try again later." message appears?

  Solution: Run the `GET _cluster/health` command to check the status of the cluster or go to the pages under **Intelligent Maintenance** to view the cause.

- What do I do if the "The nodes in the cluster contain data. You must migrate the data first." message appears?

  Solution: Migrate data. For more information, see Remove data nodes.

- What do I do if the "The number of nodes that you reserve must be more than two and more than half of the existing nodes." message appears?

  Cause: To ensure the reliability and stability of the cluster, the number of data nodes you reserve during data node removal or data migration must be greater than two and greater than half of the existing data nodes.

  Solution: If the preceding requirements are not met, re-select the data nodes or upgrade the configuration of the cluster. For more information about how to upgrade the configuration of a cluster, see Upgrade the configuration of a cluster.

- What do I do if the "The current Elasticsearch cluster configuration does not support this operation. Check the Elasticsearch cluster configuration first." message appears?

  Solution: Run the `GET _cluster/settings` command to view the cluster configuration. Then, check whether the cluster configuration contains the settings that do not allow data allocation.

- What do I do if data nodes fail to be removed or data fails to be migrated due to the `auto_expand_replicas` index setting?

Cause: Some users may use the access control feature provided by the X-Pack plug-in. In earlier Elasticsearch versions, this feature applies the `"index.auto_expand_replicas" : "0-all"` setting to the .security index by default. This causes errors when you migrate data or remove data nodes.

Solution:

i. Query index settings.

```
GET .security/_settings
```

The following result is returned:

```
{
 ".security-6" : {
  "settings" : {
   "index" : {
    "number_of_shards" : "1",
    "auto_expand_replicas" : "0-all",
    "provided_name" : ".security-6",
    "format" : "6",
    "creation_date" : "1555142250367",
    "priority" : "1000",
    "number_of_replicas" : "9",
    "uuid" : "9t2hotc7S5OpPuKEIJ****",
    "version" : {
     "created" : "6070099"
    }
   }
  }
 }
}
```

ii. Use one of the following methods to modify the auto_expand_replicas index setting:

- Method 1:

```
PUT .security/_settings
{
 "index" : {
  "auto_expand_replicas" : "0-1"
 }
}
```

- Method 2:

```
PUT .security/_settings
{
 "index" : {
  "auto_expand_replicas" : "false",
  "number_of_replicas" : "1"
 }
}
```

Notice Set `number_of_replicas` based on your business requirements. Make sure that the number of replica shards enabled for each index is greater than or equal to one but no more than the number of available data nodes.

5.3. Upgrade the configuration of a cluster

Upgrade the configuration of an Elasticsearch cluster

As your business develops, you may have higher requirements for the configuration of your Elasticsearch cluster. If the current configuration of your Elasticsearch cluster cannot meet your business needs, you can upgrade its configuration. This topic describes how to upgrade the configuration of an Elasticsearch cluster and the related precautions.

Prerequisites

The specifications and storage capacity of your cluster are evaluated. For more information, see Evaluate specifications and storage capacity.

Precautions

- Specification upgrade
  - For each upgrade, you can upgrade the configuration for only one type of node. The node types include data nodes, warm nodes, client nodes, dedicated master nodes, and Kibana nodes.
  - You cannot change the disk types of nodes when you upgrade the configuration of your Elasticsearch cluster. You can only increase the storage space of nodes.
  - You cannot reduce disk space or downgrade nodes to downgrade the configuration of your Elasticsearch cluster.

    Note You can remove data nodes from your cluster to implement the downgrade. For more information about how to remove data nodes and the related limits, see Scale in an Elasticsearch cluster.

- Impact on services
  - If your Elasticsearch cluster is abnormal (indicated by the color yellow or red), you must select **Forced Update**. This may affect services.
  - In most cases, the system restarts your Elasticsearch cluster after a configuration upgrade to make the changes take effect. However, if your cluster contains dedicated master nodes and you only change the **number of nodes**, the system does not restart the cluster.

- Version upgrade

  You can only upgrade your Elasticsearch cluster from V6.3.2 to V6.7.0. A version upgrade cannot be performed during a configuration upgrade. For more information, see Upgrade the version of a cluster.

  > **Notice** If you perform a version upgrade during a configuration upgrade, the system displays the "UpgradeVersionMustFromConsole" error message.

- Changes in billing

  After you submit a configuration upgrade order, your Elasticsearch cluster is charged based on the new configuration.

  > **Note** During a configuration upgrade, you can check the price of your order on the **Update** page in real time.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the lower-right corner of the **Basic Information** page, choose **Configuration Update > Upgrade**.
5. On the Update page, change the configuration of the cluster based on the following instructions:

   > **Note** The **Current Config** section on the Update page shows the current configuration of the cluster. You can use this as a reference during the upgrade.

   Follow the instructions on the Update page to upgrade the configuration of your cluster based on your business requirements. For more information about the parameters on the Update page, see Parameters on the buy page. The following table describes only some of the parameters.

   > **Notice** If the zone where your cluster resides has insufficient resources for a configuration upgrade, you can migrate the nodes in the zone to another before the upgrade. For more information, see Migrate nodes in a zone.

Parameter	Description
Node Storage	If **Category** is set to **Cloud Disk**, you can increase the value of **Node Storage** for **data nodes**. The maximum storage space supported by a single node depends on the **disk type** of the node. You can check specific limits in the Elasticsearch console.

Parameter	Description
Forced Update	If your Elasticsearch cluster is abnormal (indicated by the color red or yellow) and your services are severely affected, you must immediately upgrade the cluster configuration. In this case, we recommend that you select **Forced Update**. The system will perform a forced update regardless of the cluster status. The update requires only a short period. **Notice** ◦ After a **forced update**, the system restarts your cluster. During the restart, the services running on the cluster may be unstable. ◦ If you do not select **Forced Update**, the system uses the **default mode** to restart your cluster to make the changes take effect. For more information, see Restart a cluster or node. ◦
Dedicated Master Node	You can purchase or upgrade dedicated master nodes. **Notice** If the specifications of your dedicated master nodes are **1 vCPU and 2 GiB of memory**, you can upgrade these nodes on the **Update** page. After the upgrade, the cluster is charged based on the new specifications. If your dedicated master nodes are free of charge, you are charged for these nodes after you upgrade them.
Client Node	You can purchase or upgrade client nodes.
Warm Node	You can purchase or upgrade warm nodes.
Kibana Node	You can upgrade your Kibana node. **Notice** When you purchase a cluster, Alibaba Cloud provides a Kibana node for you free of charge. This Kibana node offers 1 vCPU and 2 GiB of memory. You can upgrade the Kibana node on the Update page.

6. Read and agree to the terms of cluster configuration upgrades. Then, click **Buy Now** and complete the payment as prompted.
   After you complete the payment, the system restarts the cluster to make the changes take effect.

6.Cluster configuration

6.1. Overview

Configure an Elasticsearch cluster

The cluster configuration feature of Alibaba Cloud Elasticsearch allows you to customize the configurations of synonyms, garbage collectors, and YML files.

configure synonyms configure garbage collectors configure YML files
You can perform the following operations by using the cluster configuration feature of Elasticsearch:

- Configure synonyms

  Upload a synonym dictionary file that is tailored to your business requirements. After the file is uploaded, the system automatically updates the synonym dictionary of your Elasticsearch cluster based on the file. The updated synonym dictionary can be used to accelerate your queries.

- Configure a garbage collector

  Switch between the CMS and G1 garbage collectors.

- Configure YML files

  Modify the YML settings of your Elasticsearch cluster. For example, modify the settings to allow automatic index creation, specified index deletion, audit log indexing configuration, watcher enabling, and other configurations.

6.2. Configure synonyms

6.2.1. Configuration rules

Synonym configuration rules

Alibaba Cloud Elasticsearch allows you to use a filter to configure synonyms. The filter supports two synonym formats: Solr and WordNet.

Elasticsearch synonym configuration Solr WordNet

Configuration example

```
PUT /test_index
{
  "settings": {
    "index" : {
      "analysis" : {
        "analyzer" : {
          "synonym" : {
            "tokenizer" : "whitespace",
            "filter" : ["synonym"]
            }
          },
          "filter" : {
            "synonym" : {
              "type" : "synonym",
              "synonyms_path" : "analysis/synonym.txt",
              "tokenizer" : "whitespace"
             }
            }
          }
        }
    }
}
```

In this example, a `synonym` filter that contains synonym dictionary file path `analysis/synonym.txt` is configured in `filter`. This path indicates the location of config. For more information about parameters, see Synonym Token Filter in the open-source Elasticsearch documentation.

Solr synonyms

Configuration example:

```
# Blank lines and lines starting with pound are comments.
# Explicit mappings match any token sequence on the LHS of "=>"
# and replace with all alternatives on the RHS.  These types of mappings
# ignore the expand parameter in the schema.
# Examples:
i-pod, i pod => ipod,
sea biscuit, sea biscit => seabiscuit
# Equivalent synonyms may be separated with commas and give
# no explicit mapping.  In this case the mapping behavior will
# be taken from the expand parameter in the schema.  This allows
# the same synonym file to be used in different synonym handling strategies.
# Examples:
ipod, i-pod, i pod
foozball , foosball
universe , cosmos
lol, laughing out loud
# If expand==true, "ipod, i-pod, i pod" is equivalent
# to the explicit mapping:
ipod, i-pod, i pod => ipod, i-pod, i pod
# If expand==false, "ipod, i-pod, i pod" is equivalent
# to the explicit mapping:
ipod, i-pod, i pod => ipod
# Multiple synonym mapping entries are merged.
foo => foo bar
foo => baz
# is equivalent to
foo => foo bar, baz
```

You can also define synonyms in the filter, but you must use `synonyms` rather than `synonyms_path`. The sample code is as follows:

```
PUT /test_index
{
  "settings": {
    "index" : {
      "analysis" : {
        "filter" : {
          "synonym" : {
            "type" : "synonym",
            "synonyms" : [
              "i-pod, i pod => ipod",
              "begin, start"
            ]
          }
        }
      }
    }
  }
}
```

Note We recommend that you use `synonyms_path` to define large synonym sets in the file.

WordNet synonyms

Configuration example:

```
PUT /test_index
{
  "settings": {
    "index" : {
      "analysis" : {
        "filter" : {
          "synonym" : {
            "type" : "synonym",
            "format" : "wordnet",
            "synonyms" : [
              "s(100000001,1,'abstain',v,1,0).",
              "s(100000001,2,'refrain',v,1,0).",
              "s(100000001,3,'desist',v,1,0)."
            ]
          }
        }
      }
    }
  }
}
```

This example uses `synonyms` to define WordNet synonyms. You can also use `synonyms_path` to define WordNet synonyms.

6.2.2. Upload a synonym dictionary file

Configure a synonym dictionary file

This topic describes how to upload a custom synonym dictionary file to your Alibaba Cloud Elasticsearch cluster. After you upload a custom file and the synonym dictionary of your Elasticsearch cluster is updated, you can search new indexes by using the updated synonym dictionary.

Elasticsearch synonym

Context

Note the following points when you upload a synonym dictionary file:

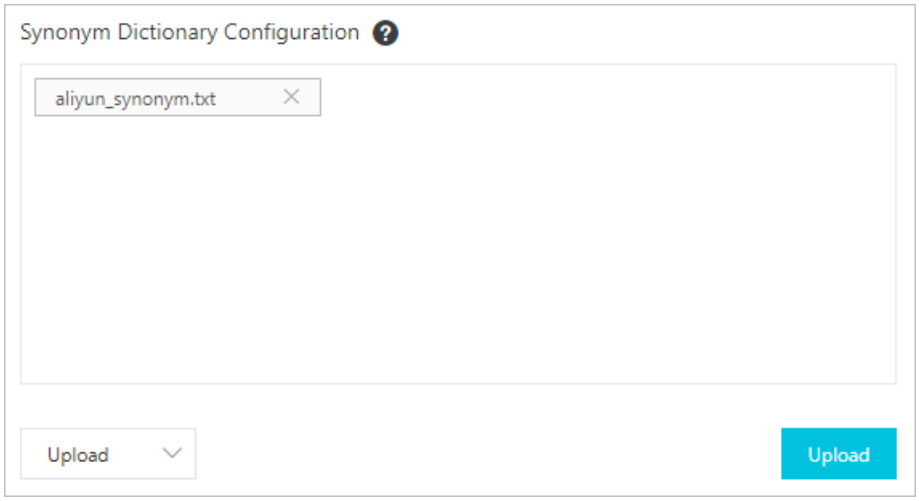
- After you upload a synonym dictionary file to your Elasticsearch cluster, the cluster is restarted. During the restart process, the system updates the synonym dictionary file to all nodes in the cluster. The time that is required for the updated dictionary to take effect is based on the number of nodes.
- Assume that the `index-aliyun` index is created based on the aliyun.txt synonym file. If you have uploaded a new synonym dictionary file to overwrite the existing dictionary file, the existing index cannot automatically load the new dictionary file. You must recreate the indexes after you update the synonym dictionary. Otherwise, the updated synonym dictionary only takes effect on new indexes.

- A synonym dictionary file must be a `TXT` file encoded by using `UTF-8`. Each line can contain only one synonym expression.

```
ipod, i-pod, i pod => ipod, i-pod, i pod
foo => foo bar
```

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Cluster Configuration**.
5. On the **Cluster Configuration** page, click **Synonym Dictionary Configuration** on the right side of **Word Splitting**.
6. In the **Synonym Dictionary Configuration** pane, select the mode that you want to upload a synonym dictionary file. Then, upload the file that is generated based on the rules in Configuration rules.



   - **Upload**: If you select this mode, click **Upload** and select the local synonym dictionary file that you want to upload.
   - **Add OSS File**: If you select this mode, specify Bucket Name and File Name and click **Add**.

     Make sure that the specified bucket is in the same region as the Elasticsearch cluster and the specified file is a TXT file.
7. Click **Save**.

What's next

After the status of the Elasticsearch cluster becomes Active, log on to the Kibana console to create indexes, verify synonyms, and upload test data to perform a search test. When you create an index, you must configure `setting` and `mapping` and configure `"synonyms_path": "analysis/your_dict_name.txt"` in `setting`. For more information, see Using Synonyms in the open-source Elasticsearch documentation and Configure synonyms.

6.2.3. Configure synonyms

Configure Elasticsearch synonyms

This topic describes how to configure synonyms. After you complete the synonym configuration, you can upload the updated synonym dictionary file, apply the file to Alibaba Cloud Elasticsearch clusters, and use the new dictionary for searches.

configure Elasticsearch synonyms synonym

Prerequisites

A synonym dictionary file is uploaded. If you have not uploaded a file, perform the operations described in Upload a synonym dictionary file first.

Configuration example 1

The following example uses a filter to configure synonyms. It also uses the `aliyun_synonyms.txt` file as the test file to configure `begin, start` .

1. Log on to the Kibana console of your Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab, run the following command to create an index:

```
PUT /aliyun-index-test
{
 "settings": {
  "index":{
   "analysis": {
     "analyzer": {
      "by_smart": {
       "type": "custom",
       "tokenizer": "ik_smart",
       "filter": ["by_tfr","by_sfr"],
       "char_filter": ["by_cfr"]
      },
      "by_max_word": {
       "type": "custom",
       "tokenizer": "ik_max_word",
       "filter": ["by_tfr","by_sfr"],
       "char_filter": ["by_cfr"]
      }
     },
     "filter": {
       "by_tfr": {
        "type": "stop",
        "stopwords": [" "]
        },
      "by_sfr": {
        "type": "synonym",
        "synonyms_path": "analysis/aliyun_synonyms.txt"
        }
     },
     "char_filter": {
      "by_cfr": {
       "type": "mapping",
       "mappings": ["| => |"]
      }
     }
   }
  }
 }
}
```

4. Configure the `title` synonym field.
   - Run the following command if your Elasticsearch version is earlier than V7.0:

```
PUT /aliyun-index-test/_mapping/doc
{
"properties": {
 "title": {
  "type": "text",
  "analyzer": "by_max_word",
  "search_analyzer": "by_smart"
 }
}
}
```

   - Run the following command if your Elasticsearch version is V7.0 or later:

```
PUT /aliyun-index-test/_mapping/
{
"properties": {
 "title": {
  "type": "text",
  "analyzer": "by_max_word",
  "search_analyzer": "by_smart"
 }
}
}
```

   > **Note** In open-source Elasticsearch 7.0 and later, the type parameter is deprecated and its function is replaced by `_doc`. You do not need to specify the type when you configure the index mapping. If you specify the type, an error is returned.

5. Run the following command to verify synonyms:

```
GET /aliyun-index-test/_analyze
{
"analyzer": "by_smart",
"text":"begin"
}
```

   If the command is successfully executed, the following result is returned:

```
{
"tokens": [
 {
  "token": "begin",
  "start_offset": 0,
  "end_offset": 5,
  "type": "ENGLISH",
  "position": 0
 },
 {
  "token": "start",
  "start_offset": 0,
  "end_offset": 5,
  "type": "SYNONYM",
  "position": 0
 }
]
}
```

6. Add data for further testing.
   - Run the following command if your Elasticsearch version is earlier than V7.0:

```
PUT /aliyun-index-test/doc/1
{
"title": "Shall I begin?"
}
```

```
PUT /aliyun-index-test/doc/2
{
"title": "I start work at nine."
}
```

   - Run the following command if your Elasticsearch version is V7.0 or later:

```
PUT /aliyun-index-test/_doc/1
{
"title": "Shall I begin?"
}
```

```
PUT /aliyun-index-test/_doc/2
{
"title": "I start work at nine."
}
```

7. Run the following command to perform a search test and verify synonyms:

```
GET /aliyun-index-test/_search
{
 "query" : { "match" : { "title" : "begin" }},
 "highlight" : {
   "pre_tags" : ["<red>", "<bule>"],
   "post_tags" : ["</red>", "</bule>"],
   "fields" : {
     "title" : {}
   }
 }
}
```

If the command is successfully executed, the following result is returned:

```
{
"took": 11,
"timed_out": false,
"_shards": {
 "total": 5,
 "successful": 5,
 "failed": 0
},
"hits": {
 "total": 2,
 "max_score": 0.41048482,
 "hits": [
  {
   "_index": "aliyun-index-test",
   "_type": "doc",
   "_id": "2",
   "_score": 0.41048482,
   "_source": {
    "title": "I start work at nine."
   },
   "highlight": {
```

```
    "title": [
      "I <red>start</red> work at nine."
    ]
   }
  },
  {
   "_index": "aliyun-index-test",
   "_type": "doc",
   "_id": "1",
   "_score": 0.39556286,
   "_source": {
    "title": "Shall I begin?"
   },
   "highlight": {
    "title": [
     "Shall I <red>begin</red>?"
    ]
   }
  }
 ]
}
}
```

Configuration example 2

The following example references synonyms and uses the IK dictionary for word splitting.

1. Log on to the Kibana console of your Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab, run the following command to create an index:

```
PUT /my_index
{
 "settings": {
   "analysis": {
     "analyzer": {
       "my_synonyms": {
         "filter": [
           "lowercase",
           "my_synonym_filter"
         ],
         "tokenizer": "ik_smart"
       }
     },
     "filter": {
       "my_synonym_filter": {
         "synonyms": [
           "begin,start"
         ],
         "type": "synonym"
       }
     }
   }
 }
}
```

You can run this command to execute the following tasks:

i. Configure the `my_synonym_filter` synonym filter and a synonym dictionary.

ii. Configure the `my_synonyms` analyzer and use the `ik_smart` IK analyzer to split words.

iii. The `ik_smart` IK analyzer splits words and then converts all letters into lowercase.

4. Configure the `title` synonym field.
   - Run the following command if your Elasticsearch version is earlier than V7.0:

```
PUT /my_index/_mapping/doc
{
"properties": {
 "title": {
  "type": "text",
  "analyzer": "my_synonyms"
 }
}
}
```

- Run the following command if your Elasticsearch version is V7.0 or later:

```
PUT /my_index/_mapping/
{
"properties": {
 "title": {
  "type": "text",
  "analyzer": "my_synonyms"
 }
}
}
```

> **Note** In open-source Elasticsearch 7.0 and later, the type parameter is deprecated and its function is replaced by `_doc`. You do not need to specify the type when you configure the index mapping. If you specify the type, an error is returned.

5. Run the following command to verify synonyms:

```
GET /my_index/_analyze
{
 "analyzer":"my_synonyms",
 "text":"Shall I begin?"
}
```

If the command is successfully executed, the following result is returned:

```
{
"tokens": [
{
 "token": "shall",
 "start_offset": 0,
 "end_offset": 5,
 "type": "ENGLISH",
 "position": 0
},
{
 "token": "i",
 "start_offset": 6,
 "end_offset": 7,
 "type": "ENGLISH",
 "position": 1
},
{
 "token": "begin",
 "start_offset": 8,
 "end_offset": 13,
 "type": "ENGLISH",
 "position": 2
},
{
 "token": "start",
 "start_offset": 8,
 "end_offset": 13,
 "type": "SYNONYM",
 "position": 2
}
]
}
```

6. Add data for further testing.
   - Run the following command if your Elasticsearch version is earlier than V7.0:

```
PUT /my_index/doc/1
{
"title": "Shall I begin?"
}
```

```
PUT /my_index/doc/2
{
"title": "I start work at nine."
}
```

- Run the following command if your Elasticsearch version is V7.0 or later:

```
PUT /my_index/_doc/1
{
"title": "Shall I begin?"
}
```

```
PUT /my_index/_doc/2
{
"title": "I start work at nine."
}
```

7. Run the following command to perform a search test and verify synonyms:

```
GET /my_index/_search
{
"query" : { "match" : { "title" : "begin" }},
"highlight" : {
 "pre_tags" : ["<red>", "<bule>"],
 "post_tags" : ["</red>", "</bule>"],
 "fields" : {
    "title" : {}
 }
}
}
```

If the command is successfully executed, the following result is returned:

```
{
"took": 11,
"timed_out": false,
"_shards": {
 "total": 5,
 "successful": 5,
 "failed": 0
},
"hits": {
 "total": 2,
 "max_score": 0.41913947
```

```
 "max_score" : 0.41913947,
"hits": [
 {
  "_index": "my_index",
  "_type": "doc",
  "_id": "2",
  "_score": 0.41913947,
  "_source": {
   "title": "I start work at nine."
  },
  "highlight": {
   "title": [
    "I <red>start</red> work at nine."
   ]
  }
 },
 {
  "_index": "my_index",
  "_type": "doc",
  "_id": "1",
  "_score": 0.39556286,
  "_source": {
   "title": "Shall I begin?"
  },
  "highlight": {
   "title": [
    "Shall I <red>begin</red>?"
   ]
  }
 }
]
}
}
```


6.3. Configure a garbage collector

Elasticsearch garbage collector

Alibaba Cloud Elasticsearch V6.7.0 and later allows you to configure and switch between garbage collectors if the memory size per data node in your Elasticsearch cluster is greater than or equal to 32 GiB. Supported garbage collectors include **CMS** and **G1**.

configure Elasticsearch garbage collectors CMS garbage collector G1 garbage collector

Prerequisites

Prerequisites

The version of your Elasticsearch cluster is V6.7.0 or later. The memory size per data node in your cluster is greater than or equal to 32 GiB. Otherwise, you must upgrade the cluster specifications. For more information, see Upgrade the configuration of a cluster.

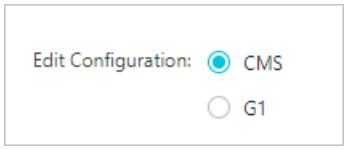
> **Note** Elasticsearch clusters that do not meet these requirements can use only the **CMS garbage collector** and cannot switch to the **G1 garbage collector**.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Cluster Configuration**.
5. On the **Cluster Configuration** page, click **Edit Configuration** on the right side of **Garbage Collectors**.

   > **Warning** After you change the garbage collector type, the system will restart the cluster. Make sure that the restart does not affect your services.

6. In the **Edit Configuration** pane, select **G1** and click **OK** to switch to the G1 garbage collector.

   After you confirm the operation, the cluster is restarted. After the cluster is restarted, the garbage collector is switched to G1.

6.4. Modify the YML configuration

Modify the YML configuration

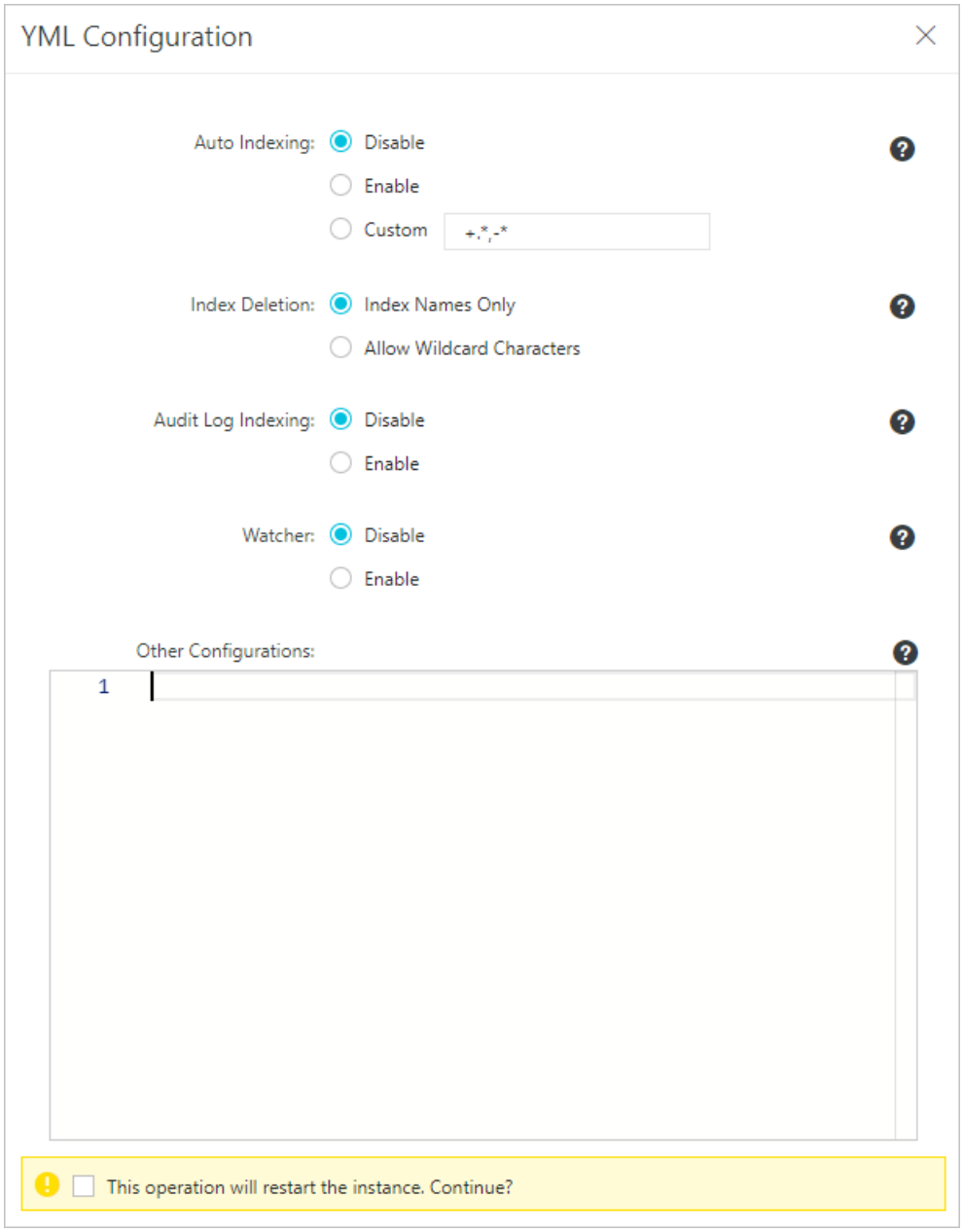
This topic describes how to modify the YML configuration of your Alibaba Cloud Elasticsearch cluster. For example, you can enable Auto Indexing, Index Deletion, Audit Log Indexing, and Watcher, and specify Other Configurations.

Elasticsearch YML configuration Auto Indexing Index Deletion Audit Log Indexing Watcher

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Cluster Configuration**.
5. On the **Cluster Configuration** page, click **Modify Configuration** on the right side of **YML Configuration**.

6. In the **YML Configuration** pane, specify the required parameters.



Parameter	Description
Auto Indexing	This parameter specifies whether to automatically create an index if a new file is uploaded to an Elasticsearch cluster but no index has been created. We recommend that you disable Auto Indexing because indexes created by this feature may not meet your business requirements. The configuration item in the YML file for this parameter is `action.auto_create_index`. The default value is `false`.

Parameter	Description
Index Deletion	This parameter specifies whether to specify the index name when you delete an index. If you select **Allow Wildcards**, you can use wildcards to delete multiple indexes at a time. You cannot restore the indexes that have been deleted. Exercise caution when configuring this item. The configuration item in the YML file for this parameter is `action.destructive_requires_name`. The default value is `false`.
Audit Log Indexing	If you enable Audit Log Indexing, index logs are generated when you create, delete, modify, or search an index in your Elasticsearch cluster. These logs consume disk space and affect cluster performance. We recommend that you disable Audit Log Indexing. Exercise caution when configuring this item. **Note** This parameter is unavailable for Elasticsearch V7.4.0 clusters. The configuration item in the YML file for this parameter is `xpack.security.audit.enabled`. The default value is `false`.
Watcher	If you enable Watcher, you can use the X-Pack Watcher feature. Make sure that you clear the `.watcher-history*` index at regular intervals to save disk space. The configuration item in the YML file for this parameter is `xpack.watcher.enabled`. The default value is `false`.

Parameter	Description

Other Configurations	The following content lists some supported configuration items. These items are available for Elasticsearch V5.X, V6.X, and V7.X unless otherwise specified.
	◦ Configure CORS
	▪ `http.cors.enabled`
	▪ `http.cors.allow-origin`
	▪ `http.cors.max-age`
	▪ `http.cors.allow-methods`
	▪ `http.cors.allow-headers`
	▪ `http.cors.allow-credentials`
	◦ Recreate indexes by calling the Reindex operation
	`reindex.remote.whitelist`
	◦ Configure the audit log indexing feature
	▪ `xpack.security.audit.enabled`
	▪ `xpack.security.audit.index.bulk_size`
	▪ `xpack.security.audit.index.flush_interval`
	▪ `xpack.security.audit.index.rollover`
	▪ `xpack.security.audit.index.events.include`
	▪ `xpack.security.audit.index.events.exclude`
	▪ `xpack.security.audit.index.events.emit_request_body`
	◦ Configure queue sizes
	▪ `thread_pool.bulk.queue_size` (available for Elasticsearch V5.X)
	▪ `thread_pool.write.queue_size` (available for Elasticsearch V6.X and V7.X)
	▪ `thread_pool.search.queue_size`
	◦ Custom SQL plug-in configuration
	`xpack.sql.enabled`
	Elasticsearch uses the X-Pack built-in SQL plug-in by default. To upload a custom SQL plug-in, set `xpack.sql.enabled` to `false`.

> **Warning** After you modify **YML Configuration** of your Elasticsearch cluster, the system performs a rolling restart on the cluster for the modifications to take effect. If replicas are configured for the indexes in your cluster, the rolling restart does not affect your services. If replicas are not configured for the indexes in your cluster, the rolling restart may affect your services. Therefore, make sure that you want to proceed with the modifications. We recommend that you modify the YML configuration during off-peak hours.

7. In the lower part of the pane, select the **This operation will restart the cluster. Continue?** check box and click **OK**. The Elasticsearch cluster restarts. You can view the restart progress in the Tasks dialog box. After the cluster is restarted, the YML configuration is updated.

6.5. Configure YML

6.5.1. Configure CORS

Configure cross-origin access for Elasticsearch

This topic describes how to configure cross-origin resource sharing (CORS) for Alibaba Cloud Elasticsearch. CORS can be configured to allow browsers on other origins to access your clusters.

Elasticsearch CORS configuration Elasticsearch cross-origin access configuration

> **Notice**
> - The configuration items in the following table are custom configurations provided by Elasticsearch to support HTTP.
> - The configuration items in the following table support only static configuration. For the configurations to take effect, you must add the configurations to the *elasticsearch.yml* file.
> - The configuration items in the following table depend on the network settings of an Elasticsearch cluster.

Configuration item	Description

Configuration item	Description
http.cors.enabled	The CORS configuration item. This item is used to specify whether to allow browsers on other origins to access Elasticsearch. Valid values: true and false. • If you set the value to true, CORS is enabled, and then Elasticsearch can process OPTIONS CORS requests. If the origin in a request is declared in http.cors.allow-origin, Elasticsearch returns a response that has the Access-Control-Allow-Origin header included. • The default value is false. If you set the value to false, CORS is disabled. In this case, Elasticsearch ignores the origin in the request header and returns a response that does not have the Access-Control-Allow-Origin header included. If a client cannot send a pre-flight request that has origin information included in the request header or does not validate the Access-Control-Allow-Origin header in the response that is returned from the server, the cross-origin security is compromised. If CORS is disabled for Elasticsearch, a client can only send an OPTIONS request to check whether the Access-Control-Allow-Origin header exists.
http.cors.allow-origin	The origin configuration item. This item specifies the origins from which requests are allowed. By default, no origin is allowed. If you add a forward slash (/) to the start and end of the value, this item is treated as a regular expression. This allows you to use regular expressions to support HTTP and HTTPS requests. For example, /https?:\/\/localhost(:[0-9]+)?/ indicates that Elasticsearch responds to all requests that match the regular expression. Note The asterisk (*) is a valid character but considered as a security risk because it indicates that an Elasticsearch cluster is open to all origins. We recommend that you do not use asterisks.
http.cors.max-age	Browsers can send OPTIONS requests to query the CORS configuration. This item specifies the cache time of the retrieved CORS configuration. The default value is 1728000 seconds (20 days).
http.cors.allow-methods	The item that is used to configure the request method. Valid values: OPTIONS, HEAD, GET, POST, PUT, and DELETE.
http.cors.allow-headers	The item that is used to configure the request header. Valid values: X-Requested-With, Content-Type, and Content-Length.

Configuration item	Description
http.cors.allow-credentials	The credential configuration item. This item specifies whether Elasticsearch is allowed to return the Access-Control-Allow-Credentials header. The default value is false, which indicates that Elasticsearch is not allowed to return the header. You can set the value to true. This allows Elasticsearch to return the header.

6.5.2. Recreate indexes by calling the Reindex operation

Recreate indexes by calling the Reindex operation

This topic describes how to recreate indexes by calling the Reindex operation. After you recreate indexes in the current Alibaba Cloud Elasticsearch cluster, you can migrate data of indexes in an Elasticsearch cluster of earlier versions to a cluster of the newly released Elasticsearch version.

Elasticsearch reindex recreate indexes

Configuration example

To recreate indexes in the current Elasticsearch cluster by calling the Reindex operation, you must configure the reindex.remote.whitelist item in the *elasticsearch.yml* file of the current cluster. This item is used to add the access address of a remote Elasticsearch cluster (an original cluster) to the remote access whitelist of the current cluster.

An access address in the whitelist can be a combination of host and port. Separate the configurations of multiple hosts with commas (,), for example: otherhost:9200,another:9200,127.0.10.**:9200,localhost:**. Only host and port are used to configure security policies because the whitelist ignores the protocol information.

> **Notice** To configure the whitelist, use <Elasticsearch cluster domain>:9200 if a remote Elasticsearch cluster is deployed in a single zone. Otherwise, use the combinations of the IP addresses and ports of all data nodes in the remote Elasticsearch cluster if the cluster is deployed across zones.

After you configure the whitelist, you can call the Reindex operation to recreate indexes. The sample code is as follows:

```
POST _reindex
{
 "source": {
  "remote": {
   "host": "http://otherhost:9200",
   "username": "user",
   "password": "pass"
  },
  "index": "source",
  "query": {
   "match": {
    "test": "data"
   }
  }
 },
 "dest": {
  "index": "dest"
 }
}
```

- `host` is the address of the remote cluster. The address must include the protocol, domain, and port, for example. `https://otherhost:9200`.

  > **Notice**
  >
  > - If a remote Elasticsearch cluster is deployed in a single zone, set `host` of the cluster to a value in the format of `<Elasticsearch cluster domain>:<9200>` and perform the operations described in Connect two Elasticsearch clusters.
  > - If a remote Elasticsearch cluster is deployed across zones, set `host` of the cluster to a value in the format of `<IP address of any data node in the Elasticsearch cluster>:<9200>` and perform the operations described in Connect two Elasticsearch clusters.

- The `username` and `password` parameters are optional. If the requested Elasticsearch service uses `basic authentication`, provide the required information in the request. `Basic authentication` must be implemented over HTTPS. Otherwise, the password will be sent in plaintext. For more information about other parameters, see Reindex API.

> **Note**
>
> - If the access address of a remote Elasticsearch cluster is added to the whitelist of the current cluster, the current cluster directly sends requests to the remote cluster without the need to verify or modify the request parameters.
> - Recreating indexes from a remote Elasticsearch cluster does not support **manual slicing** or **automatic slicing**. For more information, see Manual slicing or Automatic slicing.

Set the batch size

Indexing from a remote Elasticsearch cluster uses on-heap buffer. Default maximum batch size: 100 MB. If the index in the remote cluster contains large documents, you must adjust the batch size to a small value.

In the following example, `size` is set to 10.

```
POST _reindex
{
 "source": {
  "remote": {
   "host": "http://otherhost:9200"
  },
  "index": "source",
  "size": 10,
  "query": {
   "match": {
    "test": "data"
   }
  }
 },
 "dest": {
  "index": "dest"
 }
}
```

Set timeout periods

Use `socket_timeout` to set the `socket` read timeout period. Default value: `30s`. Use `connect_timeout` to set the connection timeout period. Default value: `1s`.

In the following example, the `socket` read timeout period is set to one minute and the connection timeout period is set to 10 seconds.

```
POST _reindex
{
 "source": {
  "remote": {
   "host": "http://otherhost:9200",
   "socket_timeout": "1m",
   "connect_timeout": "10s"
  },
  "index": "source",
  "query": {
   "match": {
    "test": "data"
   }
  }
 },
 "dest": {
  "index": "dest"
 }
}
```

6.5.3. Configure the audit log indexing feature

Configure audit log indexing

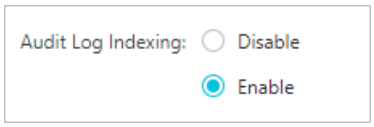
This topic describes how to enable the audit log indexing feature for your Alibaba Cloud Elasticsearch cluster, view the auditing event log file, and configure the related parameters.

configure auditing log indexing

Enable audit log indexing

> **Note** The audit log indexing feature is unavailable for Elasticsearch V7.4.0 clusters.

By default, Elasticsearch does not allow you to view the auditing event log file that contains request information. If you want to view the log file, you must log on to the Elasticsearch console and enable the audit log indexing feature for your Elasticsearch cluster. After the audit log indexing feature is enabled, auditing events are saved to your cluster and added to the index that has prefix `.security_audit_log-*` .



Configure audit log indexing

After the audit log indexing feature is enabled, you can customize the configuration of this feature. The sample code is as follows:

```
xpack.security.audit.index.bulk_size: 5000
xpack.security.audit.index.events.emit_request_body: false
xpack.security.audit.index.events.exclude: run_as_denied,anonymous_access_denied,realm_authentication_failed,access_denied,connection_denied
xpack.security.audit.index.events.include: authentication_failed,access_granted,tampered_request,connection_granted,run_as_granted
xpack.security.audit.index.flush_interval: 180s
xpack.security.audit.index.rollover: hourly
xpack.security.audit.index.settings.index.number_of_replicas: 1
xpack.security.audit.index.settings.index.number_of_shards: 10
```

Configuration item	Default value	Description
xpack.security.audit.index.bulk_size	1000	Specifies the number of auditing events when you write them into a single auditing log index in batches.
xpack.security.audit.index.flush_interval	1s	Specifies the frequency of flushing buffered auditing events to the index.
xpack.security.audit.index.rollover	daily	Specifies the frequency of rolling over to a new index. Valid values: hourly , daily , weekly , and monthly .
xpack.security.audit.index.events.include	access_denied, access_granted, anonymous_access_denied, authentication_failed, connection_denied, tampered_request, run_as_denied, run_as_granted	Specifies the types of auditing events to be included in indexing. For more information about the auditing event types, see Audit Event Types.
xpack.security.audit.index.events.exclude	null , which indicates that no auditing event is processed	Specifies the types of auditing events to be excluded from indexing.
xpack.security.audit.index.events.emit_request_body	false	Specifies whether to include the body of REST requests upon specific events, such as authentication_failed .

Configuration item	Default value	Description

Notice

- If an auditing event contains the `request body`, sensitive data in plaintext may be compromised.
- After the audit log indexing feature is enabled, auditing events are saved to your cluster and added to the index that has prefix `.security_audit_log-*`. This index consumes the storage space of your cluster. Elasticsearch does not automatically clear expired indexes. You must manually clear expired auditing log indexes.

You can also use `xpack.security.audit.index.settings` to configure the indexes in which the auditing events are stored. The following example shows you how to set both the numbers of shards and replicas to `1` for auditing log indexes.

```
xpack.security.audit.index.settings:
 index:
  number_of_shards: 1
  number_of_replicas: 1
```

Note If you want to configure custom values for auditing log indexes, add the preceding settings to the YML configuration after you set `xpack.security.audit.enabled` to `true` to enable audit log indexing. After the configuration takes effect, auditing log indexes are created in your Elasticsearch cluster. If you do not customize the configuration, your Elasticsearch cluster uses default settings `number_of_shards: 5` and `number_of_replicas: 1` to create the indexes.

For more information, see Auditing Security Settings.

6.5.4. Configure queue sizes

You can customize the queue sizes to adjust the sizes of the document write queue and document search queue.

This topic describes how to configure the `thread_pool.bulk.queue_size`, `thread_pool.write.queue_size`, and `thread_pool.search.queue_size` settings to specify the sizes of the document write queue and document search queue.

The following examples show you how to set the sizes of both the document write queue and document search queue to `500`. You can adjust the values as required.

thread_pool.bulk.queue_size: 500 (available for Elasticsearch V5.X)

thread_pool.write.queue_size: 500 (available for Elasticsearch V6.X and V7.X)

thread_pool.search.queue_size: 500

6.6. Perform scenario-based configuration

6.6.1. Use a scenario-based template to modify the configurations of a cluster

Perform scenario-based configuration

Alibaba Cloud Elasticsearch provides a scenario-based configuration feature. This feature allows you to use scenario-based templates that are provided by the system to modify the configurations of your Alibaba Cloud Elasticsearch cluster. These templates help you achieve optimal cluster and index configurations and avoid cluster exceptions and performance issues that are caused by incorrect configurations. Before you use these templates, you must specify a scenario. General, data analysis, database acceleration, search, and logging scenarios are supported. This topic describes how to use a scenario-based template to modify the configurations of a cluster.

Context

Before you use a scenario-based template to modify the configurations of a cluster, take note of the following items:

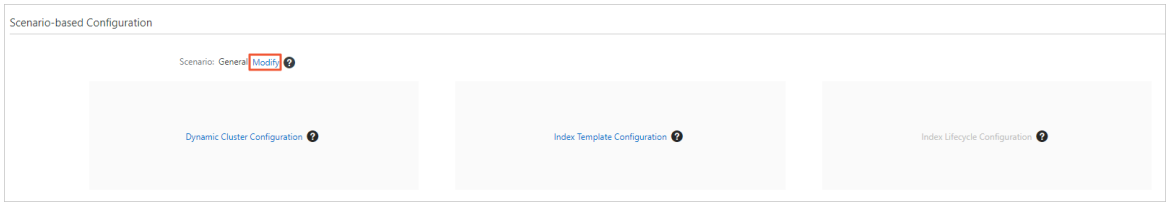
- Different versions and types of clusters support different templates. The templates available in the console take precedence.
- Elasticsearch clusters of the Standard Edition support general, data analysis, database acceleration, and search scenarios. Whereas, Elasticsearch clusters of the Advanced Edition support only logging scenarios.
- The default index template is named aliyun_default_index_template. This template has a low order value and does not affect your custom index templates.
- The policy defined in the default index lifecycle template is named aliyun_default_ilm_policy. This policy is already applied to aliyun_default_index_template.
- When you purchase a cluster, you can select a scenario on the buy page. The default scenario for an Elasticsearch cluster of the Standard Edition is General and that for an Elasticsearch cluster of the Advanced Edition is Logging. After a cluster is purchased, the system automatically applies the configurations in related templates to the cluster.

> **Notice** If your cluster is purchased before the scenario-based configuration feature is launched, the feature is disabled for the cluster. You can manually enable the feature as needed. After the feature is enabled, you must manually apply the configurations in the templates to your cluster.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.

3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the page that appears, click **Cluster Configuration**.
5. In the **Scenario-based Configuration** section, click **Modify** next to **Scenario**.



6. In the **Select Scenario** dialog box, select a scenario from the **Scenario** drop-down list and click **OK**.

> **Notice** Modifications in the Scenario-based Configuration section take effect immediately without the need to restart your cluster.

7. Use scenario-based templates to modify the configurations of your cluster.

   - Dynamic Cluster Configuration: allows you to modify the dynamic settings of your cluster. Dynamic Cluster Configuration functions the same as the `PUT /_cluster/settings` command. For more information, see Cluster Update Settings.
   - Index Template Configuration: allows you to modify the index template that is automatically used when an index is created. Index Template Configuration functions the same as the `PUT _template/aliyun_default_index_template` command. Modifying the index template has no impact on existing indexes. For more information, see Index Templates.
   - Index Lifecycle Configuration: allows you to modify the index lifecycle configurations of an Elasticsearch cluster of V6.7 or later that has warm nodes. Index Lifecycle Configuration functions the same as the `PUT _ilm/policy/aliyun_default_ilm_policy` command. For more information, see Setting up a new policy.

   The following substeps demonstrate how to modify the configurations in an index template.

      i. Click **Index Template Configuration**.

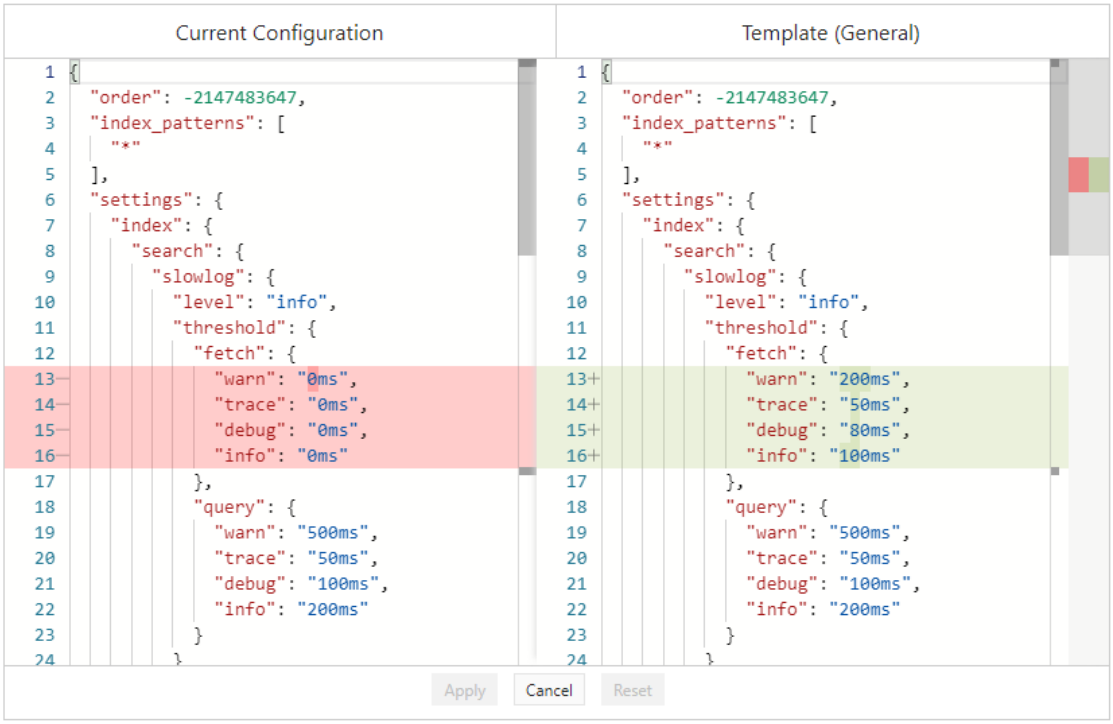
ii. In the **Index Template Configuration** pane, click **Apply**.

The following features are provided:

- **Apply**: applies the configurations in the Template section to the Current Configuration section. Then, you can modify the configurations.
- **Compare**: compares the configurations in the Current Configuration and Template sections. If no modifications are made, this feature is unavailable. During the comparison, you cannot modify configurations.
- **Reset**: restores configurations in the Current Configuration section to the original configurations.

iii. In the **Current Configuration** section, modify modifications as needed.

iv. Click **Compare**.



v. Click **Cancel**.

8. Click **Submit**. Then, the modifications take effect immediately for your cluster.

6.6.2. Modify the dynamic settings of a cluster

Modify the dynamic settings of an Elasticsearch cluster

After you enable the scenario-based configuration feature, you can use the related template to modify the dynamic settings of a cluster. This topic describes how to modify the dynamic settings of a cluster.

Elasticsearch cluster configuration cluster.routing.allocation.balance.index cluster.routing.allocation.balance.shard search.max_buckets
For more information about how to modify the dynamic settings of a cluster, see Use a scenario-based template to modify the configurations of a cluster. The following table describes the related parameters.

Parameter	Description
cluster.routing.allocation.balance.index	The weight factor for the number of shards per index that are allocated to a specific node. Default value: 0.55f. A large value indicates a high tendency to balance the number of shards per index among all nodes in a cluster. For example, you increase the value to 0.8f for search scenarios. This achieves more even distribution of shards for each index among nodes and improves query performance.
cluster.routing.allocation.balance.shard	The weight factor for the total number of shards that are allocated to a node. Default value: 0.45f. A large value indicates a high tendency to balance the number of shards among all nodes in a cluster.

Parameter	Description
search.max_buckets	The maximum number of buckets that are allowed in a single response. This parameter is available in Elasticsearch 6.2 and later. The default value of this parameter is -1, which indicates that the maximum number is not limited. However, if a response contains more than 10,000 buckets, a deprecation alert is logged. In Elasticsearch 7.x, the default value of this parameter is 10000.

6.6.3. Modify the index template of a cluster

After you enable the scenario-based configuration feature, you can dynamically modify the index template of a cluster. This topic describes how to modify the index template of a cluster.

For more information about how to modify the index template of a cluster, see Use a scenario-based template to modify the configurations of a cluster. The following table describes the related parameters.

> **Notice** The default index template is named aliyun_default_index_template. The default order value in the template is Integer.MIN_VALUE plus 1. This value is less than the order values of your custom index templates. We recommend that you do not change this value. This default index template provides configurations that are suitable for your selected scenario but does not affect your custom index templates.

Parameter	Description
index_patterns	The index pattern used by the index template to match indexes. Wildcards are supported. Default value: `*`. **Notice** Alibaba Cloud Elasticsearch allows you to change this default value to adjust the impact scope of the default index template. However, we recommend that you do not change this value.
index.search.slowlog.level	The level of a slow fetch or query log for a search request.
index.search.slowlog.threshold.fetch.warn	The time threshold used to define a slow fetch log at the warn level.
index.search.slowlog.threshold.fetch.info	The time threshold used to define a slow fetch log at the info level.
index.search.slowlog.threshold.fetch.debug	The time threshold used to define a slow fetch log at the debug level.
index.search.slowlog.threshold.fetch.trace	The time threshold used to define a slow fetch log at the trace level.

Parameter	Description
index.search.slowlog.threshold.query.warn	The time threshold used to define a slow query log at the warn level.
index.search.slowlog.threshold.query.trace	The time threshold used to define a slow query log at the trace level.
index.search.slowlog.threshold.query.info	The time threshold used to define a slow query log at the info level.
index.search.slowlog.threshold.query.debug	The time threshold used to define a slow query log at the debug level.
index.refresh_interval	The interval at which a refresh operation is performed. Default value: 1s. For scenarios that do not have high requirements for real-time performance, you can increase the value of this parameter to reduce refresh overheads and improve cluster performance.
index.unassigned.node_left.delayed_timeout	The delayed time for reallocating replica shards after a node is removed from a cluster. Default value: 1m. You can increase the value of this parameter to accelerate cluster recovery.
index.indexing.slowlog.threshold.index.warn	The time threshold used to define a slow indexing log at the warn level.
index.indexing.slowlog.threshold.index.info	The time threshold used to define a slow indexing log at the info level.
index.indexing.slowlog.threshold.index.debug	The time threshold used to define a slow indexing log at the debug level.
index.indexing.slowlog.threshold.index.trace	The time threshold used to define a slow indexing log at the trace level.
index.indexing.slowlog.level	The level of a slow indexing log.
index.indexing.slowlog.source	The number of characters in the source that the system records in a slow log.
index.number_of_shards	The number of primary shards for an index. In versions earlier than Elasticsearch 7.x, the default value of this parameter is 5. In Elasticsearch 7.x and later, the default value of this parameter is 1. Setting this parameter to 1 effectively limits the number of primary shards on a cluster and prevents excessive workloads caused by numerous primary shards.

Parameter	Description
index.translog.durability	Specifies whether a translog is synchronized to a disk and then committed after every indexing, deletion, update, or bulk request. Valid values: • request: The translog is synchronized to a disk and then committed after every request. This ensures that data in the translog is not lost if a node becomes abnormal. • async: The translog is synchronized to a disk and then committed on a regular basis. This improves write performance but deteriorates data reliability.
index.merge.policy.segments_per_tier	The allowed number of segments per tier. A small value results in more merging operations but lower indexing performance. Default value: 10. We recommend that the value of this parameter is greater than or equal to that of index.merge.policy.max_merge_at_once. Otherwise, numerous merging operations occur, which lowers cluster performance.
index.merge.policy.max_merged_segment	The maximum size of a merged segment during indexing. The value of this parameter is an approximate value. Default value: 5GB. The size of a merged segment is calculated by using the following formula: Size of a merged segment = Total size of the segments that form the merged segment - Total size of the documents that are deleted from these segments
index.lifecycle.name	The index lifecycle policy.
mappings._default_._all.enabled	If you set this parameter to false, the _all field is disabled. In Elasticsearch 5.x, the default value of this parameter is true. We recommend that you set this parameter to false. In Elasticsearch 6.x, the default value of this parameter is false. In Elasticsearch 7.x, this parameter is deprecated.

6.6.4. Modify the index lifecycle configurations of a cluster

After you enable the scenario-based configuration feature, you can use the related template to modify the index lifecycle configurations of a cluster. This topic describes how to modify the index lifecycle configurations of a cluster.

For more information about how to modify the index lifecycle configurations of a cluster, see Use a scenario-based template to modify the configurations of a cluster. The following table describes the related parameters.

Notice

You can use the related template to modify the index lifecycle configurations of an Alibaba Cloud Elasticsearch cluster of V6.7 or later that has warm nodes.

The policy defined in the default index lifecycle template is named aliyun_default_ilm_policy. This policy is already applied to aliyun_default_index_template.

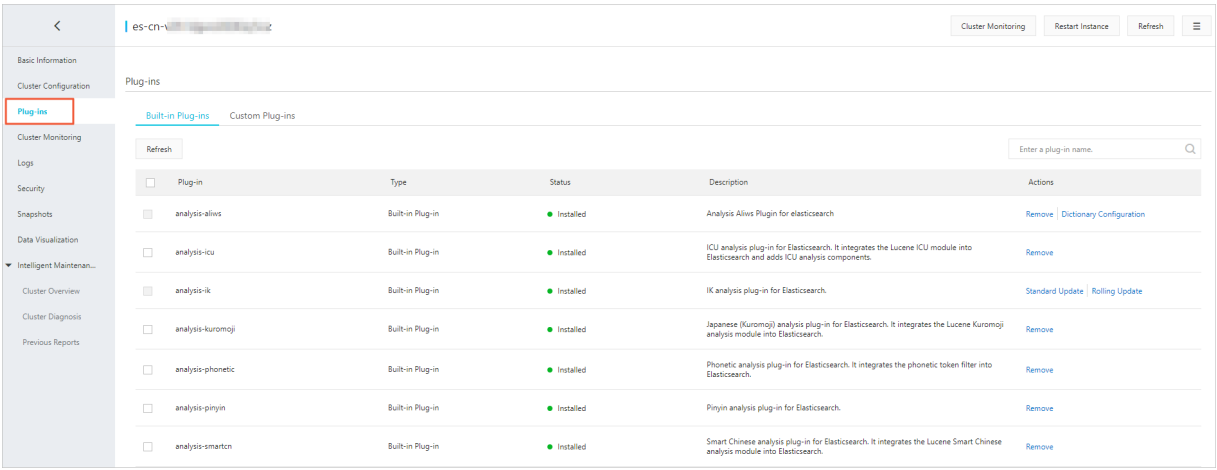
Parameter	Description
phases.hot.min_age	The time required for an index to enter the hot phase.
phases.hot.actions.set_priority.priority	The priority of an index in the hot phase.
phases.warm.min_age	The time required for an index to enter the warm phase.
phases.warm.actions.allocate.number_of_replicas	The number of replica shards for an index in the warm phase.
phases.warm.actions.allocate.require.box_type	The shard allocation policy in the warm phase. For example, the system allocates shards to warm nodes.
phases.warm.actions.set_priority.priority	The priority of an index in the warm phase.
phases.cold.min_age	The time required for an index to enter the cold phase.
phases.cold.actions.set_priority.priority	The priority of an index in the cold phase.

7.Plug-ins

7.1. Overview of plug-ins

Alibaba Cloud Elasticsearch supports all the plug-ins of open-source Elasticsearch and also provides some self-developed plug-ins. This topic provides an overview of these plug-ins.

Alibaba Cloud Elasticsearch provides two types of plug-ins: built-in and custom plug-ins.



- Built-in plug-ins

  In most cases, you can install or remove a built-in plug-in of Alibaba Cloud Elasticsearch as required. For more information, see Install and remove a built-in plug-in.

  However, you cannot remove the **analysis-ik** and **elasticsearch-repository-oss** plug-ins. You can use the standard update or rolling update method to update custom dictionaries with the **analysis-ik** plug-in. For more information, see Use the analysis-ik plug-in.

- Custom plug-ins

  You can upload, install, and remove custom plug-ins to meet your business requirements. For more information, see Upload and install a custom plug-in.

7.2. Built-in plug-ins

7.2.1. Install and remove a built-in plug-in

Install a built-in plug-in

After you purchase an Alibaba Cloud Elasticsearch cluster, the Elasticsearch console displays the built-in plug-ins on the Built-in Plug-ins tab. You can install or remove these plug-ins as required. This topic describes how to install and remove a built-in plug-in for Alibaba Cloud Elasticsearch.

Alibaba Cloud Elasticsearch built-in plug-in

Context

analysis-ik and **elasticsearch-repository-oss** are both built-in plug-ins of Alibaba Cloud Elasticsearch, but they cannot be removed.

- **analysis-ik**: an IK analyzer plug-in. In addition to its open-source functionalities, this plug-in supports the dynamic loading of dictionaries stored on Object Storage Service (OSS). You can use the

standard update or rolling update method to update dictionaries.

- **elasticsearch-repository-oss**: In addition to its open-source functionalities, this plug-in provides OSS storage while you create and restore index snapshots.

Precautions

The installation or removal of a built-in plug-in triggers the cluster to restart. Alibaba Cloud Elasticsearch removes only the plug-in that you select. You must confirm the operation before you can proceed.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Plug-ins**.
5. On the **Built-in Plug-ins** tab, find the target plug-in, and click **Install** or **Remove** in the **Actions** column.
6. Read the message that appears and click **OK**. The system then restarts the cluster. While the cluster restarts, you can view its task progress in the Tasks dialog box.

Additional information

The following table describes the built-in plug-ins that Alibaba Cloud Elasticsearch supports.

Plug-in	Default status	Description	Supported operation
analysis-aliws	Not Installed	The aliws analysis plug-in for Elasticsearch.	Install, Remove, and Dictionary Configuration
analysis-icu	Installed	The ICU analysis plug-in for Elasticsearch. This plug-in integrates the Lucene ICU module into Elasticsearch and adds ICU analysis components.	Install and Remove
analysis-ik	Installed	The IK analysis plug-in for Elasticsearch. This plug-in cannot be removed.	Standard Update and Rolling Update
analysis-kuromoji	Installed	The Japanese (Kuromoji) analysis plug-in for Elasticsearch. This plug-in integrates the Lucene Kuromoji analysis module into Elasticsearch.	Install and Remove
analysis-phonetic	Installed	The phonetic analysis plug-in for Elasticsearch. This plug-in integrates the phonetic token filter into Elasticsearch.	Install and Remove
analysis-pinyin	Installed	The Pinyin analysis plug-in for Elasticsearch.	Install and Remove

Plug-in	Default status	Description	Supported operation
analysis-smartcn	Installed	The smart Chinese analysis plug-in for Elasticsearch. This plug-in integrates the Lucene smart Chinese analysis module into Elasticsearch.	Install and Remove
analysis-stconvert	Not Installed	The analysis plug-in that allows you to convert text between simple and traditional Chinese characters.	Install and Remove
elasticsearch-repository-oss	Installed	The plug-in that allows you to use Alibaba Cloud OSS to store Elasticsearch snapshots. This plug-in cannot be removed.	None
ingest-attachment	Installed	The ingest processor for Elasticsearch. This plug-in uses Apache Tika to extract content.	Install and Remove
mapper-murmur3	Installed	The plug-in that allows you to both compute the hash values of fields when you create an index and store those values in the index.	Install and Remove
mapper-size	Installed	The plug-in that allows you to record the size of documents before they are compressed when you create an index.	Install and Remove
repository-hdfs	Installed	The plug-in that provides support for Hadoop Distributed File System (HDFS) repositories.	Install and Remove

7.2.2. Use the analysis-ik plug-in

analysis-ik is an IK analysis plug-in of Alibaba Cloud Elasticsearch. This plug-in cannot be removed. In addition to open source features, the plug-in can dynamically load the dictionaries that are stored in Object Storage Service (OSS). The plug-in also allows you to use the standard or rolling update method to update dictionaries. This topic describes how to use the plug-in.

Context

The analysis-ik plug-in supports two update methods for IK dictionaries: standard update and rolling update. For more information, see Perform a standard update for IK dictionaries and Perform a rolling update for IK dictionaries. The following table provides more details about the two methods.

Update method	Application mode	Loading mode	Description

Update method	Application mode	Loading mode	Description
Standard update	This method updates the dictionaries on all nodes in an Elasticsearch cluster. It requires a restart of the cluster for the update to take effect.	The system sends an uploaded dictionary file to all nodes in an Elasticsearch cluster, modifies the *IKAnalyzer.cfg.xml* file, and then restarts the nodes to load the file.	You can use the standard update method to update the built-in IK main dictionary and stopword list of the analysis-ik plug-in. In the Standard Update pane, you can view the built-in main dictionary `SYSTEM_MAIN.dic` and the built-in stopword list `SYSTEM_STOPWORD.dic`.
Rolling update	The first time you upload a dictionary file, the dictionaries on all nodes in an Elasticsearch cluster are updated. The cluster needs to be restarted for the update to take effect. If the dictionary file that you upload has the same name as the existing dictionary file, the cluster does not need to be restarted. The dictionaries are directly loaded while the cluster is running.	If the content of a dictionary file changes, you can use this method to update the dictionaries on all nodes in an Elasticsearch cluster. After you upload the latest dictionary file, the nodes automatically load the file. If the dictionary file list changes when you perform a rolling update, all nodes in the cluster need to reload dictionary configurations. For example, when you upload a new dictionary file or delete an existing dictionary file, the changes are synchronized to the *IKAnalyzer.cfg.xml* file.	When you upload a dictionary file for the first time, the system modifies the *IKAnalyzer.cfg.xml* file. After the dictionaries are updated, the cluster must be restarted for the update to take effect.

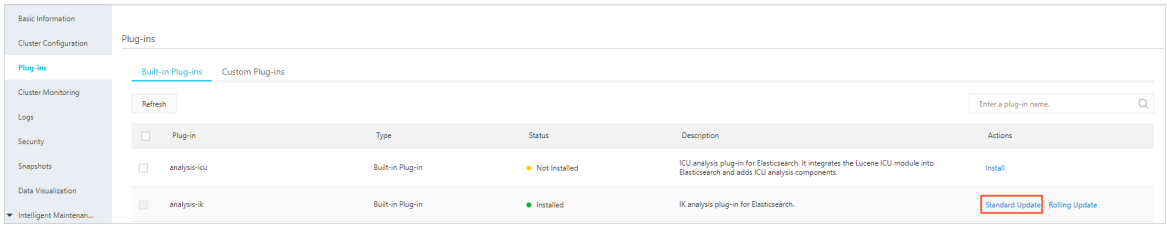
> **Notice**
>
> New dictionaries apply only to data that is inserted after a standard or rolling update. If you want to apply the new dictionaries to both the existing data and new data, you must reindex the existing data.
>
> If you choose the standard update method, you can modify the built-in main dictionary or stopword list. However, you cannot delete the built-in main dictionary or stopword list. The following modification methods are available:
>
> - If you want to update the built-in main dictionary, upload a dictionary file named **SYSTEM_MAIN.dic**. The new dictionary file automatically overwrites the existing file. For more information, see IK Analysis for Elasticsearch.
> - If you want to update the built-in stopword list, upload a file named **SYSTEM_STOPWORD.dic**. The new file automatically overwrites the existing file. For more information, see IK Analysis for Elasticsearch and Configure a stopword list.

Prerequisites

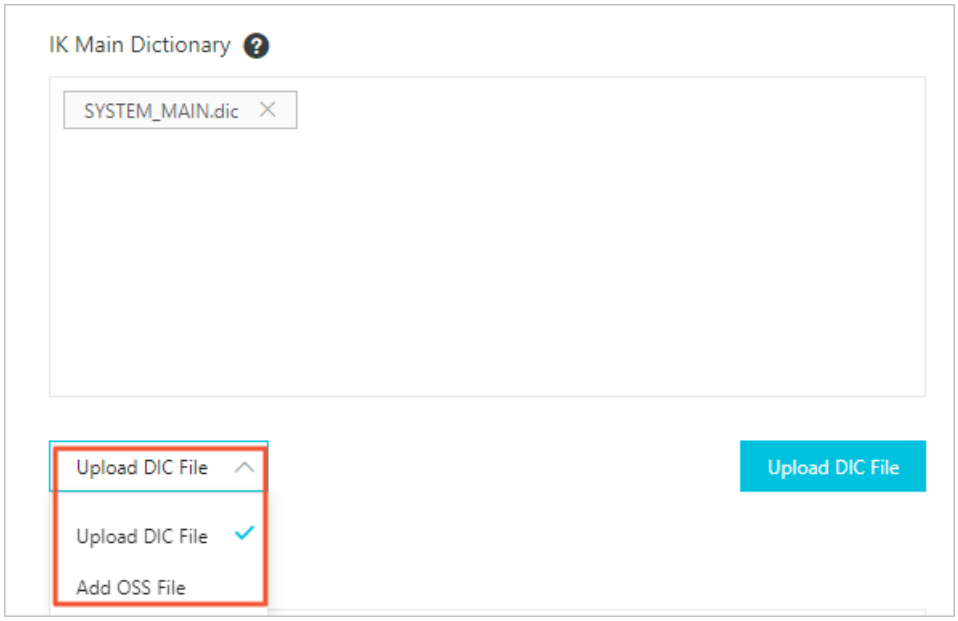
Your Elasticsearch cluster is in a normal state. You can check the cluster status on the Basic Information page.

Perform a standard update for IK dictionaries

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Plug-ins**.
5. On the **Built-in Plug-ins** tab, find the analysis-ik plug-in and click **Standard Update** in the **Actions** column.



6. In the **Standard Update** pane, click **Configure** in the lower-right corner.
7. Select a method to upload a dictionary file from the drop-down list that is below the **IK Main Dictionary** section. Then, upload the dictionary file based on the following instructions.



You can select the **Upload DIC File** or **Add OSS File** method.

- **Upload DIC File**: If you select this method, click **Upload DIC File** and select the local file that you want to upload.
- **Add OSS File**: If you select this method, specify Bucket Name and File Name, and click **Add**.

  Make sure that the bucket resides in the same region as your Elasticsearch cluster and the file to upload is a DIC file. If the content of the dictionary that is stored in OSS changes, you must manually upload the dictionary file again.

> **Warning** The following operation restarts your Elasticsearch cluster. Before you perform this operation, make sure that the restart does not affect your business.

8. Scroll down to the lower part of the pane, select **This operation will restart the cluster. Continue?**, and click **Save**.If you choose the standard update method, the system restarts your cluster no matter whether you upload a new dictionary file, remove a dictionary file, or update dictionary content.
9. After the cluster is restarted, log on to the Kibana console of the cluster and run the following command to check whether the new dictionary file takes effect.

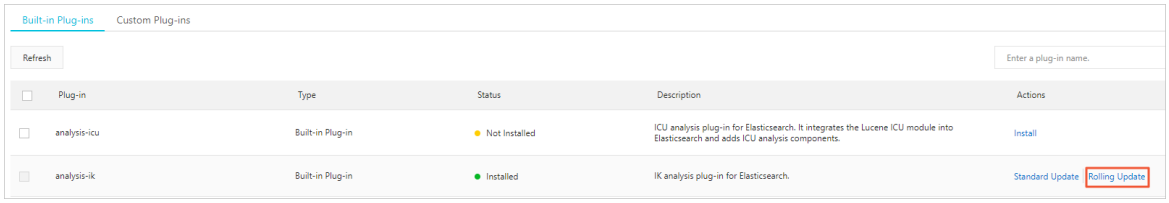
> **Note** For more information about how to log on to the Kibana console, see Log on to the Kibana console.

```
GET _analyze
{
"analyzer": "ik_smart",
"text": ["Tokens in the new dictionary file"]
}
```

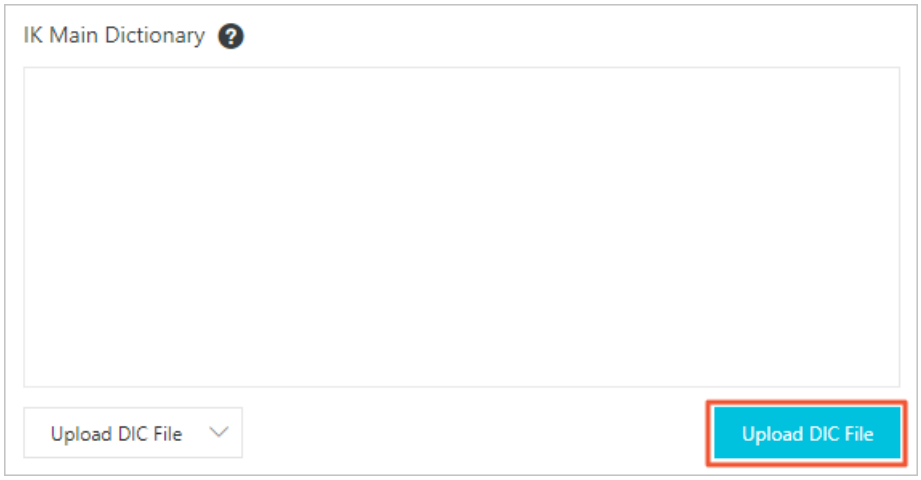
Perform a rolling update for IK dictionaries

1. On the **Built-in Plug-ins** tab, find the analysis-ik plug-in and click **Rolling Update** in the **Actions**

column.



2. In the **Rolling Update** pane, click **Configure** in the lower-right corner.
3. Select a method to upload a dictionary file from the drop-down list that is below the **IK Main Dictionary** section. Then, upload the dictionary file based on the following instructions.



> **Note** You cannot use the rolling update method to modify the built-in main dictionary. If you want to modify the built-in main dictionary, use the standard update method.

You can select the **Upload DIC File** or **Add OSS File** method.

- **Upload DIC File**: If you select this method, click **Upload DIC File** and select the local file that you want to upload.
- **Add OSS File**: If you select this method, specify Bucket Name and File Name, and click **Add**.

  Make sure that the bucket resides in the same region as your Elasticsearch cluster and the file to upload is a DIC file. The `dic_0.dic` file is used in the following operations. If the content of the dictionary that is stored in OSS changes, you must manually upload the dictionary file again.

> **Warning** The following operation restarts your Elasticsearch cluster. Before you perform this operation, make sure that the restart does not affect your business.

4. Scroll down to the lower part of the pane, select **This operation will restart the instance. Continue?**, and click **Save**. If this is the first time that you upload a dictionary file, the system automatically restarts the cluster.

   After you click Save, the system performs a rolling update for the cluster. After the rolling update is completed, the new dictionary takes effect.

   If you want to add or remove tokens from dictionaries, perform the following steps to modify the `dic_0.dic` file:

5. In the Rolling Update pane, delete the existing `dic_0.dic` file and upload a new dictionary file. The new dictionary file must have the same name.This operation changes the content of the existing dictionary file in the cluster and uploads a new file that has the same name. The system does not need to restart the cluster for the update to take effect.
6. Click **Save**.The analysis-ik plug-in on the nodes of the Elasticsearch cluster automatically loads the dictionary file. The time that is required by each node to load the dictionary file varies. It requires about two minutes for all nodes to load the dictionary file. You can log on to the Kibana console of the Elasticsearch cluster and run the following command multiple times to verify the new dictionary file.

   > **Note** For more information about how to log on to the Kibana console, see Log on to the Kibana console.

   ```
   GET _analyze
   {
   "analyzer": "ik_smart",
   "text": ["Tokens in the new dictionary file"]
   }
   ```

Configure a stopword list

Alibaba Cloud Elasticsearch provides a built-in stopword list. The list contains the following predefined tokens: a, an, and, are, as, at, be, but, by, for, if, in, into, is, it, no, not, of, on, or, such, that, the, their, then, there, these, they, this, to, was, will, with.

You can perform the following steps to remove tokens from the stopword list:

1. Download the default IK configuration file from the official website of open source Elasticsearch.
2. Decompress the downloaded package and open the *stopword.dic* dictionary file in the config folder.
3. Remove the tokens that you no longer require and save the dictionary file.
4. Change the name of the *stopword.dic* dictionary file to *SYSTEM_STOPWORD.dic*.
5. Upload the *SYSTEM_STOPWORD.dic* file to your Elasticsearch cluster. The file automatically overwrites the existing stopword list.
6. After the cluster is restarted, the new stopword list takes effect.

7.2.3. Use the aliyun-sql plug-in

7.2.3.1. Use method

The aliyun-sql plug-in is developed based on Apache Calcite and deployed on a server. It is used to parse SQL queries. After you install this plug-in on your Alibaba Cloud Elasticsearch cluster, you can execute SQL statements to query data in the cluster the same way as in common databases. This greatly reduces the training and usage costs of Elasticsearch.

Prerequisites

You have completed the following operations:

- An Alibaba Cloud Elasticsearch cluster is created. The cluster version is 6.7.0 or later.

  For more information, see Create an Elasticsearch cluster.

  > **Notice** The aliyun-sql plug-in is available only for Alibaba Cloud Elasticsearch clusters of V6.7.0 or later.

- The aliyun-sql plug-in is installed.

  By default, the aliyun-sql plug-in is installed on the Elasticsearch cluster. You can check whether the plug-in is installed on the plug-in configuration page. If the plug-in is not installed, follow the instructions provided in Install and remove a built-in plug-in to install the plug-in.

Context

The aliyun-sql plug-in offers more features than open source SQL plug-ins. The following table compares the aliyun-sql plug-in with open source SQL plug-ins.

SQL plug-in	SQL parser	Paged query	Join	Nested	Common function	Case Function	Extended UDF	Optimization of execution plans
x-pack-sql (6.x)	Antlr	Supported.	Not supported.	Supported. The syntax is a.b.	Supports abundant functions.	Not supported.	Not supported.	Provides a large number of rules for the optimization of execution plans.
opendistro-for-elasticsearch	Druid	Not supported. The maximum number of data entries that can be queried is determined by the `max_result_window` parameter.	Supported.	Supported. The syntax is `nested(message.info)`.	Supports a few functions.	Not supported.	Not supported.	Provides a few rules for the optimization of execution plans.

SQL plug-in	SQL parser	Paged query	Join	Nested	Common function	Case Function	Extended UDF	Optimization of execution plans
aliyun-sql	Javacc	Supported.	Supported. The plug-in provides the truncation feature. This feature allows you to dynamically configure the number of data entries that you can query from a single table. For more information, see Syntax overview.	Supported. The syntax is a.b.	Supports abundant functions. For more information, see Other functions and expressions.	Supported.	Supported. For more information, see UDFs.	Provides a large number of rules for the optimization of execution plans and uses Calcite to optimize execution plans.

Note The CASE statement in the preceding table refers to the CASE WHEN THEN ELSE syntax.

Precautions

- Before you use the plug-in, make sure that the `aliyun.sql.enabled` parameter is set to true for your Elasticsearch cluster. You can set the parameter in the Kibana console. For more information, see Log on to the Kibana console.
- You can manually remove the plug-in. Before you remove the plug-in, run the following command in the Kibana console to disable it. To disable the plug-in, set `aliyun.sql.enabled` to null. The following example shows the configuration:

```
PUT _cluster/settings
{
 "persistent": {
  "aliyun.sql.enabled": null
 }
}
```

Plug-in removal restarts your cluster. If you do not disable the plug-in before you remove it, your cluster remains stuck in the restart. In this case, you must run the following command to clear the archiving configurations and resume the restart:

```
PUT _cluster/settings
{
 "persistent": {
  "archived.aliyun.sql.enabled": null
 }
}
```

Syntax overview

The aliyun-sql plug-in uses the syntax of MySQL 5.0 and supports a wide range of functions and expressions. For more information, see Other functions and expressions.

- Basic queries

```
SELECT [DISTINCT] (* | expression) [[AS] alias] [, ...]
FROM table_name
[WHERE condition]
[GROUP BY expression [, ...]
 [HAVING condition]]
[ORDER BY expression [ ASC | DESC ] [, ...]]
[LIMIT [offset, ] size]
```

- JOIN queries

```
SELECT
 expression
FROM table_name
JOIN table_name
 ON expression
[WHERE condition]
```

> **Notice**
>
> - When you perform a JOIN query, Alibaba Cloud Elasticsearch limits the maximum number of data entries that you can query from a single table. The default number is 10,000. You can specify the maximum number of queries by setting `max.join.size`.
> - The JOIN query you performed is an INNER JOIN query. Actually, the aliyun-sql plug-in performs a merge join for the query. When you perform a JOIN query, make sure that the field values in the tables you want to join change with the document IDs of Elasticsearch. JOIN queries can be performed only for fields of a numeric data type.

Procedure

1. Log on to the Kibana console of your Elasticsearch cluster.For more information, see Log on to the Kibana console.
2. Enable the aliyun-sql plug-in.

```
PUT _cluster/settings
{
 "transient": {
  "aliyun.sql.enabled": true
 }
}
```

3. Initiate a write request.

> **Note** The aliyun-sql plug-in supports only query requests. Therefore, the following code uses a bulk request to write data.

- Data of student information

```
PUT stuinfo/_doc/_bulk? refresh
{"index":{"_id":"1"}}
{"id":572553,"name":"xiaoming","age":"22","addr":"addr1"}
{"index":{"_id":"2"}}
{"id":572554,"name":"xiaowang","age":"23","addr":"addr2"}
{"index":{"_id":"3"}}
{"id":572555,"name":"xiaoliu","age":"21","addr":"addr3"}
```

- Data of student rankings

```
PUT sturank/_doc/_bulk? refresh
{"index":{"_id":"1"}}
{"id":572553,"score":"90","sorder":"5"}
{"index":{"_id":"2"}}
{"id":572554,"score":"92","sorder":"3"}
{"index":{"_id":"3"}}
{"id":572555,"score":"86","sorder":"10"}
```

4. Execute an SQL statement.Perform a JOIN query to query the name and ranking of a student.

```
POST /_alisql
{
 "query":"select stuinfo.name,sturank.sorder from stuinfo join sturank on stuinfo.id=sturank.id"
}
```

If the statement is successfully executed, the aliyun-sql plug-in returns table information. The `columns` field contains column names and data types. The `rows` field contains row data.

```
{
 "columns" : [
  {
   "name" : "name",
   "type" : "text"
  },
  {
   "name" : "sorder",
   "type" : "text"
  }
 ],
 "rows" : [
  [
   "xiaoming",
   "5"
  ],
  [
   "xiaowang",
   "3"
  ],
  [
   "xiaoliu",
   "10"
  ]
 ]
}
```

7.2.3.2. Query syntax

This topic describes the query syntax of the aliyun-sql plug-in. This plug-in supports basic queries, queries with cursors, JSON-formatted queries, translate queries, special queries, user-defined functions (UDFs), other functions, and expressions.

> **Note** After you understand the query syntax of the aliyun-sql plug-in, you can test and use the plug-in in the Kibana console. For more information, see Use method.

Basic queries

- Common query

```
POST /_alisql?pretty
{
 "query": "select * from monitor where host='100.80.xx.xx' limit 5"
}
```

- Query in which the number of data entries to return is specified

```
POST /_alisql?pretty
{
  "query": "select * from monitor",
  "fetch_size": 3
}
```

- Query in which parameters are specified

```
POST /_alisql?pretty
{
 "query": "select * from monitor where host= ? ",
 "params": [{"type":"STRING","value":"100.80.xx.xx"}],
 "fetch_size": 1
}
```

Type	Parameter	Required?	Example	Description
URL parameter	`pretty`	No	None.	Formats query results.
Request body parameter	`query`	Yes	`select * from monitor where host='100.80.xx.xx' limit 5`	Specifies the SQL statement that you want to execute.
	`fetch_size`	No	`3`	Specifies the number of data entries to return. The default value is 1000. The maximum value is 10000. If you set this parameter to a value greater than 10000, the system still regards the value as 10000.

Type	Parameter	Required?	Example	Description
	params	No	[{"type":"STRING","value":"100.80.xx.xx"}]	This parameter implements the features of the PreparedStatement interface.

- Query results

  When you execute an SQL statement to query large amounts of data for the first time, the number of data entries to return is determined by the `fetch_size` parameter. The query results also include cursors.

```
{
 "columns": [
  {
   "name": "times",
   "type": "integer"
  },
  {
   "name": "value2",
   "type": "float"
  },
  {
   "name": "host",
   "type": "keyword"
  },
  {
   "name": "region",
   "type": "keyword"
  },
  {
   "name": "measurement",
   "type": "keyword"
  },
  {
   "name": "timestamp",
   "type": "date"
  }
 ],
 "rows": [
  [
   572575,
   4649800.0,
   "100.80.xx.xx",
   "china-dd",
   "cpu",
   "2018-08-09T08:18:42.000Z"
  ]
 ],
 "cursor": "u5HzAgJzY0BEWEYxWlhKNVFXNWtS****"
}
```

Parameter	Description
columns	The names and data types of the fields that you queried.
rows	The query results.
cursor	The cursor that is used for the next query.

Notice A maximum number of 1,000 data entries are returned by default. If the number of data entries that you want to query is greater than 1,000, you can continually use cursors to query additional data entries until no cursors or data entries are returned.

Queries with cursors

- Query request

```
POST /_alisql?pretty
{
  "cursor": "u5HzAgJzY0BEWEYxWlhKNVFXNWtS****"
}
```

Type	Parameter	Required?	Description
URL parameter	`pretty`	No	Formats query results.
Request body parameter	`cursor`	Yes	The cursor that is used to query specific data.

- Query results

```
{
 "rows": [
  [
   572547,
   3.327459E7,
   "100.80.xx.xx",
   "china-dd",
   "cpu",
   "2018-08-09T08:19:12.000Z"
  ]
 ],
 "cursor": "u5HzAgJzY0BEWEYxWlhKNVFXNWtS****"
}
```

The query results do not include the columns field. This reduces network transmission latency. Other fields in the query results are similar to those in the query results of a basic query.

JSON-formatted queries

- Query request (JOIN statements are not supported.)

```
POST /_alisql?format=org
{
 "query": "select * from monitor where host= ? ",
 "params": [{"type":"STRING","value":"100.80.xx.xx"}],
 "fetch_size": 1
}
```

`format=org` indicates that query results are in the JSON format. Other parameters are the same as those in a basic query.

- Query results

```
{
 "_scroll_id": "DXF1ZXJ5QW5kRmV0Y2gBAAAAAAAAAAsWYXNEdlVJZzJTSXFfOGluOVB4Q3Z****",
 "took": 18,
 "timed_out": false,
 "_shards": {
  "total": 1,
  "successful": 1,
  "skipped": 0,
  "failed": 0
 },
 "hits": {
  "total": 2,
  "max_score": 1.0,
  "hits": [
   {
    "_index": "monitor",
    "_type": "_doc",
    "_id": "2",
    "_score": 1.0,
    "_source": {
     "times": 572575,
     "value2": 4649800,
     "host": "100.80.xx.xx",
     "region": "china-dd",
     "measurement": "cpu",
     "timestamp": "2018-08-09T16:18:42+0800"
    }
   }
  ]
 }
}
```


The query results are in the same format as those of domain-specific language (DSL) statements. The `_scroll_id` parameter in the query results is used for paged queries.

Translate queries

You can use translate queries to convert requested SQL statements to DSL statements that Elasticsearch supports.

- Query request (JOIN statements are not supported.)

```
POST _alisql/translate
{
 "query": "select * from monitor where host= '100.80.xx.xx' "
}
```

- Query results

```
{
 "size": 1000,
 "query": {
  "constant_score": {
   "filter": {
    "term": {
     "host": {
      "value": "100.80.xx.xx",
      "boost": 1.0
     }
    }
   },
   "boost": 1.0
  }
 },
 "_source": {
  "includes": [
   "times",
   "value2",
   "host",
   "region",
   "measurement",
   "timestamp"
  ],
  "excludes": []
 }
}
```

Special queries

The aliyun-sql plug-in allows you to query data based on fields of the nested and text types.

1. Create a table schema.

```
PUT user_info/
{
  "mappings":{
    "_doc":{
      "properties":{
        "addr":{
          "type":"text"
        },
        "age":{
          "type":"integer"
        },
        "id":{
          "type":"integer"
        },
        "name":{
          "type":"nested",
          "properties":{
            "first_name":{
              "type":"keyword"
            },
            "second_name":{
              "type":"keyword"
            }
          }
        }
      }
    }
  }
}
```

2. Insert large amounts of data at a time.

```
PUT user_info/_doc/_bulk?refresh
{"index":{"_id":"1"}}
{"addr":"467 Hutchinson Court","age":80,"id":"1","name":[{"first_name":"lesi","second_name" : "Adams"},{"first_name":"chaochaosi","second_name" : "Aams"}]}
{"index":{"_id":"2"}}
{"addr":"671 Bristol Street","age":21,"id":"2","name":{"first_name":"Hattie","second_name" : "Bond"}}
{"index":{"_id":"3"}}
{"addr":"554 Bristol Street","age":23,"id":"3","name":{"first_name":"Hattie","second_name" : "Bond"}}
```

3. Query user information based on the `second_name` field of the nested type.

```
POST _alisql
{
 "query": "select * from user_info where name.second_name='Adams'"
}
```

The following results are returned:

```
{
 "columns" : [
  {
   "name" : "id",
   "type" : "integer"
  },
  {
   "name" : "addr",
   "type" : "text"
  },
  {
   "name" : "name.first_name",
   "type" : "keyword"
  },
  {
   "name" : "age",
   "type" : "integer"
  },
  {
   "name" : "name.second_name",
   "type" : "keyword"
  }
 ],
 "rows" : [
  [
   1,
   "467 Hutchinson Court",
   "lesi",
   80,
   "Adams"
  ]
 ]
}
```

4. Query user information based on the `addr` field of the text type.

```
POST _alisql
{
 "query": "select * from user_info where addr='Bristol'"
}
```

The following results are returned:

```
{
 "columns" : [
  {
   "name" : "id",
   "type" : "integer"
  },
  {
   "name" : "addr",
   "type" : "text"
  },
  {
   "name" : "name.first_name",
   "type" : "keyword"
  },
  {
   "name" : "age",
   "type" : "integer"
  },
  {
   "name" : "name.second_name",
   "type" : "keyword"
  }
 ],
 "rows" : [
  [
   2,
   "671 Bristol Street",
   "Hattie",
   21,
   "Bond"
  ],
  [
   3,
   "554 Bristol Street",
   "Hattie",
   23,
   "Bond"
  ]
 ]
}
```

UDFs

UDFs can only be added during the initialization of the aliyun-sql plug-in. The following operations demonstrate how to add the `date_format` UDF:

1. Define the `DateFormat` class.

```
/**
 * DateFormat.
 */
public class DateFormat extends UDF {
  public String eval(DateTime time, String toFormat) {
    if (time == null || toFormat == null) {
      return null;
    }
    Date date = time.toDate();
    SimpleDateFormat format =  new SimpleDateFormat(toFormat);
    return format.format(date);
  }
}
```

2. Add the `DateFormat` class to the initialization method of the plug-in.

```
udfTable.add(KeplerSqlUserDefinedScalarFunction
        .create("date_format"
            , DateFormat.class
            , (JavaTypeFactoryImpl) typeFactory));
```

3. Use the date_format UDF to query data.

```
select date_format(date_f,'yyyy') from date_test
```

Other functions and expressions

Type	Name	Example	Description
	ABS	`SELECT ABS(num_field) FROM table`	Returns the absolute value of a number.
	ACOS	`SELECT ACOS(num_field) FROM table`	Returns the arccosine of a number.
	ASIN	`SELECT ASIN(num_field) FROM table`	Returns the arcsine of a number.

Type	Name	Example	Description
Numeric function	ATAN	`SELECT ATAN(num_field) FROM table`	Returns the arctangent of a number.
	ATAN2	`SELECT ATAN2(num_field1,num_field2) FROM table`	Returns the arctangent of two numbers.
	CEIL	`SELECT CEIL(num_field) FROM table`	Returns the smallest integer that is greater than or equal to a number.
	CBRT	`SELECT CBRT(num_field) FROM table`	Returns the double-precision cube root of a number.
	COS	`SELECT COS(num_field) FROM table`	Returns the cosine of a number.
	COT	`SELECT COT(num_field) FROM table`	Returns the cotangent of a number.
	DEGREES	`SELECT DEGREES(num_field) FROM table`	Converts radians into degrees.
	EXP or EXPM1	`SELECT EXP(num_field) FROM table`	Returns the value of e raised to the power of a number.
	FLOOR	`SELECT FLOOR(num_field) FROM table`	Returns the largest integer that is less than or equal to a number.
	SIN	`SELECT SIN(num_field) FROM table`	Returns the sine of a number.
	SINH	`SELECT SINH(num_field) FROM table`	Returns the hyperbolic sine of a number.
	SQRT	`SELECT SQRT(num_field) FROM table`	Returns the positive square root of a number.
	TAN	`SELECT TAN(num_field) FROM table`	Returns the tangent of a number.

Type	Name	Example	Description
	ROUND	`SELECT ROUND(num_field,2) FROM table`	Rounds a number to a specific decimal place.
	RADIANS	`SELECT RADIANS (num_field) FROM table`	Converts an angle in degrees to its equivalent in radians.
	RAND	`SELECT RAND() FROM table`	Returns a double-precision number that includes a plus sign. The number must be greater than or equal to 0.0 and less than 1.0.
	LN	`SELECT LN (num_field) FROM table`	Returns the natural logarithm of a number.
	LOG10	`SELECT LOG10 (num_field) FROM table`	Returns the base 10 logarithm of a number.
	PI	`SELECT PI() FROM table`	Returns the value of PI.
	POWER	`SELECT POWER (num_field,2) FROM table`	Returns the result of a number raised to a power.
	TRUNCATE	`SELECT TRUNCATE (num_field,2) FROM table`	Truncates a number to a specific decimal place.
Arithmetic operation	+	`SELECT (v1 + v2) as v FROM table`	Returns the sum of two numbers.
	-	`SELECT(v1 - v2) as v FROM table`	Returns the difference of two numbers.
	*	`SELECT(v1 * v2) as v FROM table`	Returns the product of two numbers.
	/	`SELECT(v1 / v2) as v FROM table`	Returns the quotient of a number divided by another.
	%	`SELECT(v1 % v2) as v FROM table`	Returns the remainder of a number divided by another.

Type	Name	Example	Description
Logical operation	AND	`SELECT * FROM table WHERE condition AND condition`	Returns data for a query in which the AND operation is performed on two conditions.
	OR	`SELECT * FROM table WHERE condition OR condition`	Returns data for a query in which the OR operation is performed on two conditions.
	NOT	`SELECT * FROM table WHERE NOT condition`	Returns data for a query in which a condition is excluded.
	IS NULL	`SELECT * FROM table WHERE field IS NULL`	Returns data for a query in which the value of a specific field is null.
	IS NOT NULL	`SELECT * FROM table WHERE field IS NOT NULL`	Returns data for a query in which the value of a specific field is not null.
String function	ASCII	`SELECT ASCII(str_field) FROM table`	Returns the ASCII value of a character.
	LCASE or LOWER	`SELECT LCASE(str_field) FROM table`	Converts a string to lowercase.
	UCASE or UPPER	`SELECT UCASE(str_field) FROM table`	Converts a string to uppercase.
	CHAR_LENGTH or CHARACTER_LENGTH	`SELECT CHAR_LENGTH(str_field) FROM table`	Returns the length of a string, in bytes.
	TRIM	`SELECT TRIM(str_field) FROM table`	Trims a string by removing leading and trailing spaces from it.
	SPACE	`SELECT SPACE(num_field) FROM table`	Returns a string that includes the specific number of spaces.
	LEFT	`SELECT LEFT(str_field, 3) FROM table`	Returns the specific number of leftmost characters from a string.
	RIGHT	`SELECT RIGHT(str_field, 3) FROM table`	Returns the specific number of rightmost characters from a string.

Type	Name	Example	Description
	REPEAT	`SELECT REPEAT(str_field, 3) FROM table`	Repeats a string the specific number of times and returns the result string.
	REPLACE	`SELECT REPLACE(str_field, "SQL", "HTML") FROM table`	Replaces a substring with a new substring within a string.
	POSITION	`SELECT POSITION("test" IN str_field) FROM table`	Returns the position where a substring appears within a string for the first time.
	REVERSE	`SELECT REVERSE(str_test) from table`	Reverses a string and returns the result string.
	LPAD	`SELECT LPAD(str_field, 20, "ABC") FROM table`	Prepends specified characters to a string based on a specific length.
	CONCAT	`SELECT CONCAT(str_field,'test') FROM table`	Concatenates two or more strings and returns the result string.
	SUBSTRING	`SELECT SUBSTRING(str_field, 5, 3) FROM table`	Returns a substring that is extracted from a string based on the specified character position.
	CURRENT_DATE	`SELECT CURRENT_DATE() FROM table`	Returns the current date.
	CURRENT_TIME	`SELECT CURRENT_TIME() FROM table`	Returns the current time.
	CURRENT_TIMESTAMP	`SELECT CURRENT_TIMESTAMP() FROM table`	Returns the current date and time.
	DAYNAME	`SELECT DAYNAME(date_field) FROM table`	Returns the day of the week for a date.
	DAYOFMONTH	`SELECT DAYOFMONTH(date_field) FROM table`	Returns the index of the day of the month for a date.

Type	Name	Example	Description
Date function	DAYOFYEAR	`SELECT DAYOFYEAR(date_field) FROM table`	Returns the index of the day of the year for a date.
	DAYOFWEEK	`SELECT DAYOFWEEK(date_field) FROM table`	Returns the index of the day of the week for a date.
	HOUR	`SELECT HOUR(date_field) FROM table`	Returns the hour part of a date.
	MINUTE	`SELECT MINUTE(date_field) FROM table`	Returns the minute part of a time or datetime.
	SECOND	`SELECT SECOND(date_field) FROM table`	Returns the seconds part of a time or datetime.
	YEAR	`SELECT YEAR(date_field) FROM table`	Returns the year part of a date.
	MONTH	`SELECT MONTH(date_field) FROM table`	Returns the month part of a date.
	WEEK	`SELECT WEEK(date_field) FROM table`	Returns the index of the week in which a date falls. Valid values for the aliyun-sql plug-in: 1 to 54. Valid values for MySQL: 0 to 53.
	MONTHNAME	`SELECT MONTHNAME(date_field) FROM table`	Returns the name of the month for a date.
	LAST_DAY	`SELECT LAST_DAY(date_field) FROM table`	Returns the last day of the month for a date.
	QUARTER	`SELECT QUARTER(date_field) FROM table`	Returns the quarter of the year for a date.
	EXTRACT	`SELECT EXTRACT(MONTH FROM date_field) FROM table`	Returns one or more separate parts of a date or time. For example, this function can return the year, month, day, hour, or minute part of a date or time.

Type	Name	Example	Description
	DATE_FORMAT	`SELECT DATE_FORMAT(date_field,'yyyy') from date_test`	Formats a date or time.
Aggregation function	MIN	`SELECT MIN(num_field) FROM table`	Returns the minimum value among a set of values.
	MAX	`SELECT MAX(num_field) FROM table`	Returns the maximum value among a set of values.
	AVG	`SELECT AVG(num_field) FROM table`	Returns the average of a set of values.
	SUM	`SELECT SUM(num_field) FROM table`	Returns the sum of a set of values.
	COUNT	`SELECT COUNT(num_field) FROM table`	Returns the number of records that meet the specified conditions.
Advanced function	CASE	`SELECT * FROM table ORDER BY(CASE WHEN exp1 THEN exp2 ELSE exp3 END)`	The syntax is CASE WHEN THEN ELSE END. If the condition specified in the WHEN clause is met, the value specified in the THEN clause is returned. If the condition is not met, the value specified in the ELSE clause is returned. The syntax of the CASE statement is similar to that of the IF THEN ELSE statement.

7.2.4. Use the physical replication feature of the apack plug-in

Use the physical replication feature of the Elasticsearch apack plug-in

The apack plug-in is developed by the Alibaba Cloud Elasticsearch team. This plug-in provides the physical replication and vector retrieval features. This topic describes only the physical replication feature. This feature greatly reduces CPU overheads and improves write performance in scenarios such as logging and time series analytics. In these scenarios, replica shards are configured for indexes, large amounts of data are written, and data visibility is latency-insensitive.

Prerequisites

- An Alibaba Cloud Elasticsearch cluster of V6.7.0 is created. The kernel version of the cluster is 1.2.0 or later. For more information, see Create an Elasticsearch cluster.

- The apack plug-in is installed.

  Only Alibaba Cloud Elasticsearch clusters of V6.7.0 support this plug-in. If the kernel version of your cluster is earlier than 1.2.0, you must update the kernel before you use the plug-in. For more information about how to update the kernel of a cluster, see Update the kernel of a cluster. If the kernel version of your cluster is 1.2.0 or later, the plug-in is already installed on your cluster and cannot be removed. You can go to the Plug-ins page to check whether the plug-in is installed.

  > **Note** After the apack plug-in is installed, you can use both the physical replication and vector retrieval features. For more information about how to use the vector retrieval feature, see Use the aliyun-knn plug-in for vector search.

Context

Basic principle of the physical replication feature:

If the feature is disabled, the system writes index data to a primary shard when the node where the primary shard resides receives a write request. Then, the system synchronizes the request to the nodes where the replica shards of the primary shard reside and writes the index data to the replica shards. This process is the same as that in open source Elasticsearch. In this process, index data is written to not only the primary shard and its replica shards but also their translogs.

After the feature is enabled, index data is written only to the primary shard, its translog, and the translogs of its replica shards. This ensures data reliability and consistency. Each time the primary shard is refreshed, the system copies incremental index data to the replica shards of the primary shard over the network. This feature delays data visibility but significantly improves the write performance of a cluster.

Performance testing of the physical replication feature:

- Test environment
  - Node configuration: 5 data nodes (each with 8 vCPUs and 32 GiB of memory) and one 2-TiB standard SSD
  - Dataset: 74-GiB nyc_taixs of Rally provided by open source Elasticsearch
  - Index configuration: five primary shards and one replica shard for each primary shard (default configuration)
- Test result

Service	Write speed (document/s)
Open source Elasticsearch 6.7.0	127,305
Alibaba Cloud Elasticsearch V6.7.0 (with the physical replication feature enabled)	184,592

- Test conclusion

  Alibaba Cloud Elasticsearch with the physical replication feature enabled delivers a write performance 45% better than open source Elasticsearch.

> **Note** You can run the commands provided in this topic in the Kibana console. For more information, see Log on to the Kibana console.

Precautions

- The physical replication feature of the apack plug-in works on indexes. By default, this feature is disabled for indexes created before the plug-in is installed and is enabled for indexes created after the plug-in is installed. If your indexes are created before the plug-in is installed, you must enable the feature before you can use it.
- You can disable the physical replication feature for an index. However, before you disable this feature, disable the index.
- Before you enable the physical replication feature for an index, disable the index and set the number of replica shards for the index to 0.

Enable the physical replication feature for a new index

When you create an index, use the settings configuration to enable the physical replication feature for the index.

```
PUT index-1
{
"settings": {
 "index.replication.type" : "segment"
 }
}
```

Disable the physical replication feature for an index

1. Disable the index.

   ```
   POST index-1/_close
   ```

2. Disable the physical replication feature.

   ```
   PUT index-1/_settings
   {
   "index.replication.type" : null
   }
   ```

3. Enable the index.

   ```
   POST index-1/_open
   ```

Enable the physical replication feature for an existing index

1. Set the number of replica shards for the index to 0.

   ```
   PUT index-1/_settings
   {
    "index.number_of_replicas": 0
   }
   ```

2. Disable the index.

```
POST index-1/_close
```

3. Enable the physical replication feature.

```
PUT index-1/_settings
{
"index.replication.type" : "segment"
}
```

4. Enable the index.

```
POST  index-1/_open
```

5. Set the number of replica shards to 1.

```
PUT index-1/_settings
{
 "index.number_of_replicas": 1
}
```

7.2.5. Use the analysis-aliws plug-in

analysis-aliws is a built-in plug-in of Alibaba Cloud Elasticsearch. This plug-in integrates an analyzer and a tokenizer into Elasticsearch to implement document analysis and retrieval. The plug-in allows you to upload a tailored dictionary file to it. After the upload, the system automatically performs a rolling update for your cluster. Alibaba Cloud Elasticsearch clusters of V5.X do not support this plug-in.

Prerequisites

The analysis-aliws plug-in is installed. It is not installed by default.

If it is not installed, install it. Make sure that your Elasticsearch cluster offers at least 4 GiB of memory. If your cluster is in a production environment, it must offer at least 8 GiB of memory. For more information about how to install the plug-in, see Install and remove a built-in plug-in.

> **Notice** If the memory capacity of your cluster does not meet the preceding requirements, upgrade the configuration of your cluster. For more information, see Upgrade the configuration of a cluster.

Context

After the analysis-aliws plug-in is installed, the following analyzer and tokenizer are integrated into your Elasticsearch cluster:

- Analyzer: aliws, which does not return function words, function phrases, or symbols
- Tokenizer: aliws_tokenizer

You can use the analyzer and tokenizer to search for documents. You can also upload a tailored dictionary file to the plug-in. For more information, see Search for a document and Configure a dictionary.

Search for a document

1. Log on to the Kibana console of your Elasticsearch cluster.For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab of the page that appears, run the following command to create an index:

```
PUT /index
{
  "mappings": {
    "fulltext": {
      "properties": {
        "content": {
          "type": "text",
          "analyzer": "aliws"
        }
      }
    }
  }
}
```

The preceding command creates an index named `index`. The type of the index is `fulltext`. The index contains the `content` property. The type of the property is `text`. The command also adds the `aliws` analyzer.

If the command is successfully executed, the following result is returned:

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

4. Run the following command to add a document:

```
POST /index/fulltext/1
{
 "content": "I like go to school."
}
```

The preceding command adds a document named `1` and sets the value of the `content` field in the document to `I like go to school.`

If the command is successfully executed, the following result is returned:

```
{
 "_index": "index",
 "_type": "fulltext",
 "_id": "1",
 "_version": 1,
 "result": "created",
 "_shards": {
  "total": 2,
  "successful": 2,
  "failed": 0
 },
 "_seq_no": 0,
 "_primary_term": 1
}
```

5. Run the following command to search for the document:

```
GET /index/fulltext/_search
{
 "query": {
  "match": {
   "content": "school"
  }
 }
}
```

The preceding command uses the `aliws` analyzer to analyze all `fulltext`-type documents, and returns the document that has `school` contained in the `content` field.

If the command is successfully executed, the following result is returned:

```
{
 "took": 5,
 "timed_out": false,
 "_shards": {
  "total": 5,
  "successful": 5,
  "skipped": 0,
  "failed": 0
 },
 "hits": {
  "total": 1,
  "max_score": 0.2876821,
  "hits": [
   {
    "_index": "index",
    "_type": "fulltext",
    "_id": "2",
    "_score": 0.2876821,
    "_source": {
     "content": "I like go to school."
    }
   }
  ]
 }
}
```

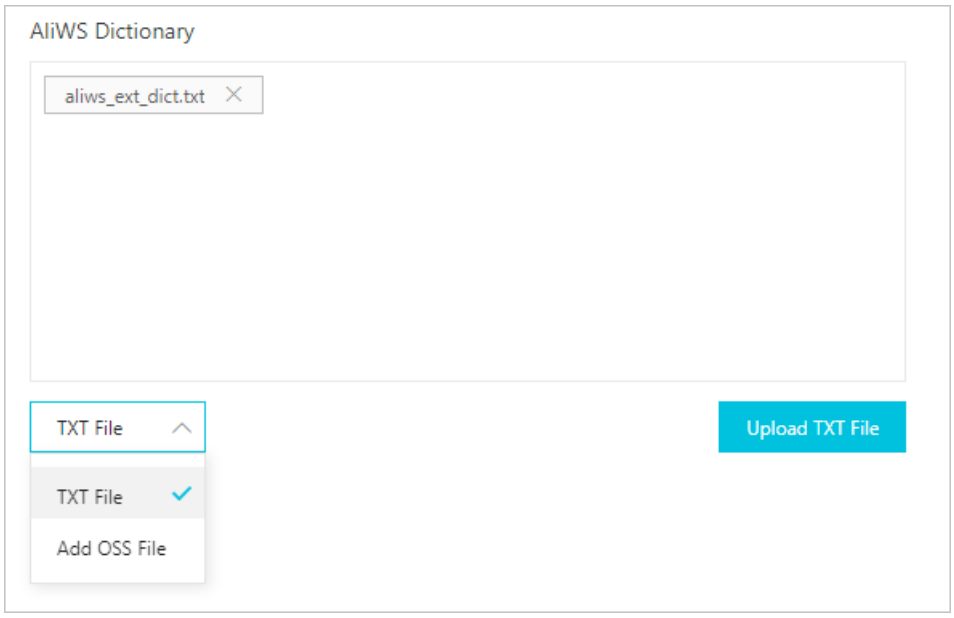

Note If unexpected results are returned, find the cause by following the instructions provided in Test the analyzer and Test the tokenizer.

Configure a dictionary

The analysis-aliws plug-in allows you to upload a tailored dictionary file to it. After the upload, all nodes in your Elasticsearch cluster automatically load the file. In this case, the system does not need to restart the cluster.

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Plug-ins**.
5. On the **Built-in Plug-ins** tab, find the **analysis-aliws** plug-in and click **Dictionary Configuration** in the **Actions** column.

6. In the **Plug-ins** pane, click **Configure** in the lower-right corner.
7. Select the method to upload the dictionary file. Then, upload the dictionary file based on the following instructions.



The dictionary file must meet the following requirements:

- Name: aliws_ext_dict.txt.
- Encoding format: UTF-8.
- Content: Each row contains one word and ends with `\n` (line break in UNIX or Linux). No whitespace characters are used before or after this word. If the dictionary file that you want to upload is generated in Windows, you must use the dos2unix tool to convert the file before the upload.

You can select the **TXT File** or **Add OSS File** method.

- **TXT File**: If you select this method, click **Upload TXT File** and select the local file that you want to upload.
- **Add OSS File**: If you select this method, specify Bucket Name and File Name, and click **Add**.

  Make sure that the bucket you specify resides in the same region as your Elasticsearch cluster. If the content of the dictionary that is stored in OSS changes, you must manually upload the dictionary file again.

8. Click **Save**.The system does not restart your cluster but performs a rolling update to make the uploaded dictionary file take effect. The update requires about 10 minutes.

Test the analyzer

Run the following command to test the aliws analyzer:

```
GET _analyze
{
 "text": "I like go to school.",
 "analyzer": "aliws"
}
```

If the command is successfully executed, the following result is returned:

```
{
 "tokens" : [
  {
   "token" : "i",
   "start_offset" : 0,
   "end_offset" : 1,
   "type" : "word",
   "position" : 0
  },
  {
   "token" : "like",
   "start_offset" : 2,
   "end_offset" : 6,
   "type" : "word",
   "position" : 2
  },
  {
   "token" : "go",
   "start_offset" : 7,
   "end_offset" : 9,
   "type" : "word",
   "position" : 4
  },
  {
   "token" : "school",
   "start_offset" : 13,
   "end_offset" : 19,
   "type" : "word",
   "position" : 8
  }
 ]
}
```

Test the tokenizer

Run the following command to test the aliws_tokenizer tokenizer:

```
GET _analyze
{
 "text": "I like go to school.",
 "tokenizer": "aliws_tokenizer"
}
```

If the command is successfully executed, the following result is returned:

```
{
 "tokens" : [
  {
   "token" : "I",
   "start_offset" : 0,
   "end_offset" : 1,
   "type" : "word",
   "position" : 0
  },
  {
   "token" : " ",
   "start_offset" : 1,
   "end_offset" : 2,
   "type" : "word",
   "position" : 1
  },
  {
   "token" : "like",
   "start_offset" : 2,
   "end_offset" : 6,
   "type" : "word",
   "position" : 2
  },
  {
   "token" : " ",
   "start_offset" : 6,
   "end_offset" : 7,
   "type" : "word",
   "position" : 3
  },
```

```
{
 "token" : "go",
 "start_offset" : 7,
 "end_offset" : 9,
 "type" : "word",
 "position" : 4
},
{
 "token" : " ",
 "start_offset" : 9,
 "end_offset" : 10,
 "type" : "word",
 "position" : 5
},
{
 "token" : "to",
 "start_offset" : 10,
 "end_offset" : 12,
 "type" : "word",
 "position" : 6
},
{
 "token" : " ",
 "start_offset" : 12,
 "end_offset" : 13,
 "type" : "word",
 "position" : 7
},
{
 "token" : "school",
 "start_offset" : 13,
 "end_offset" : 19,
 "type" : "word",
 "position" : 8
},
{
 "token" : ".",
 "start_offset" : 19,
 "end_offset" : 20,
 "type" : "word",
 "position" : 9
```

```
  }
 ]
}
```

7.2.6. Use the aliyun-knn plug-in for vector search

The aliyun-knn plug-in is a vector search engine designed by the Alibaba Cloud Elasticsearch team. It uses the vector databases of Proxima, a vector search engine designed by Alibaba DAMO Academy. This plug-in allows you to use vector spaces in different search scenarios, such as searching images, performing video fingerprinting, conducting both facial and speech recognition, and recommending commodities based on your preferences.

Prerequisites

You have completed the following operations:

- Install the aliyun-knn plug-in. Whether this plug-in is installed by default is determined by the Elasticsearch cluster version and kernel version.
  - If the cluster version is V6.7.0 and the kernel version is V1.2.0 or later, the plug-in is integrated into the apack plug-in. The apack plug-in is installed by default. If you want to install or remove the aliyun-knn plug-in, you must perform operations on the apack plug-in. For more information, see Use the physical replication feature of the apack plug-in.
  - If the cluster version is later than V6.7.0, or the cluster version is V6.7.0 and the kernel version is earlier than V1.2.0, you must manually install the aliyun-knn plug-in. For more information, see Install and remove a built-in plug-in.

  > **Notice**
  > - Only Elasticsearch clusters of V6.7.0 or later support the aliyun-knn plug-in.
  > - Before you install the aliyun-knn plug-in, make sure that each data node offers at least 2 vCPUs and 8 GiB of memory. These specifications are only for tests. For a production environment, the minimum specifications are 4 vCPUs and 16 GiB of memory. If your Elasticsearch cluster does not meet these requirements, upgrade the data nodes. For more information, see Upgrade the configuration of a cluster.

- Perform index planning and cluster sizing.

Context

The vector search engine of Alibaba Cloud Elasticsearch is used in numerous production scenarios inside Alibaba Group, such as Pailitao, Image Search, Youku video fingerprinting, Qutoutiao video fingerprinting, Taobao commodity recommendation, customized searches, and Crossmedia searches.

Alibaba Cloud Elasticsearch provides the aliyun-knn plug-in for you to use its vector search engine. This plug-in is compatible with all open source Elasticsearch versions. Therefore, you do not need to learn how to use the engine. In addition to real-time incremental synchronization and near-real-time (NRT) searches, this engine supports other features of open source Elasticsearch in distributed searches. The features include multi-replica, restoration, and snapshots.

The vector search engine supports the Hierarchical Navigable Small World (HNSW) and Linear Search algorithms. These algorithms are suitable for processing small amounts of data from in-memory storage. The following table compares the performance of the two algorithms.

Comparison between the performance of HNSW and Linear Search

The performance of the two algorithms is measured on an Alibaba Cloud Elasticsearch V6.7.0 cluster. Test environment:

- Node configuration: two data nodes (each with 16 vCPUs and 64 GiB of memory) and one 100-GiB standard SSD
- Datasets: SIFT 128-dimensional float type vectors
- Total data records: 20 million
- Index settings: default settings

Performance metric	HNSW	Linear Search
Top-10 recall ratio	98.6%	100%
Top-50 recall ratio	97.9%	100%
Top-100 recall ratio	97.4%	100%
Latency (p99)	0.093s	0.934s
Latency (p90)	0.018s	0.305s

Note p is short for percentage. For example, latency (p99) indicates how many seconds it requires to respond to 99% of queries.

Index planning

Algorithm	Use scenario	In-memory storage?	Remarks

Algorithm	Use scenario	In-memory storage?	Remarks
HNSW	• Each node stores only small volumes of data. • A low response latency is required. • A high recall ratio is required.	Yes	• HNSW is based on the greedy search algorithm and obeys the triangle inequality. The triangle inequality states that the total sum of costs from A to B and from B to C must be greater than the costs from A to C. Inner product space does not obey the triangle inequality. Therefore, you must convert it to Euclidean space or spherical space before you apply the HNSW algorithm. • After you write data to Elasticsearch, we recommend that you regularly call the force merge API operation during off-peak hours to merge segments in shards. This can reduce the response latency.
Linear Search	• Brute-force search. • A recall ratio of 100% is required. • The latency increases with the volume of data processed. • Effect comparison.	Yes	None.

Cluster sizing

Item	Description
Data node specifications (required)	The minimum data node specifications for a production environment are 4 vCPUs and 16 GiB of memory. The specifications of 2 vCPUs and 8 GiB of memory can be used only for tests.
Maximum volume of data per node	The maximum volume of data stored on each data node equals 50% of the total memory space of the data node.

Item	Description
Write throttling	Vector indexing is a CPU-intensive job. We recommend that you do not maintain a high write throughput. A peak write throughput lower than 5,000 TPS is recommended for a data node with 16 vCPUs and 64 GiB of memory. TPS is short for transactions per second. When Elasticsearch processes queries, it loads all index files to node memory. If nodes are out of memory, Elasticsearch reallocates shards. Therefore, we recommend that you do not write large amounts of data to Elasticsearch when it is processing queries.

Procedure

1. Log on to the Kibana console of your Elasticsearch cluster.For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab, run the following command to create an index:

> **Notice** The following sample code is applicable only to Alibaba Cloud Elasticsearch V6.7.0. For sample code that is applicable to Alibaba Cloud Elasticsearch V7.4.0, see open source Elasticsearch documentation.

```
PUT test
{
 "settings": {
  "index.codec": "proxima",
  "index.vector.algorithm": "hnsw"
 },
 "mappings": {
  "_doc": {
   "properties": {
    "feature": {
     "type": "proxima_vector",
     "dim": 2
    },
    "id": {
     "type": "keyword"
    }
   }
  }
 }
}
```

Parameter	Description
index.vector.algorithm	The algorithm. Valid values: hnsw and linear.
type	The field type. Set the value to proxima_vector to specify a vector-type field.
dim	The number of vector dimensions. Valid values: 1 to 2048.

The preceding sample code creates an index named test. The type of the index is _doc. The index contains two fields: feature and id. You can rename the index and fields as required.

4. Run the following command to add a document:

```
POST test/_doc
{
 "feature": [1.0, 2.0],
 "id": 1
}
```

Notice The value of the feature field must be a float array. The length of the array must be the same as that specified in the dim parameter in mapping.

5. Run the following command to search for the document:

```
GET test/_search
{
 "query": {
  "hnsw": {
   "feature": {
    "vector": [1.5, 2.5],
    "size": 10
   }
  }
 }
}
```

Parameter	Description
hnsw	The value must be the same as that of the algorithm parameter specified when you create the index.
vector	A float array. The length of the array must be the same as that specified in the dim parameter in mapping.

Parameter	Description
size	The number of the top-ranked documents to return.

Parameters

Algorithm parameters

Parameter	Description	Default value
index.vector.algorithm	The algorithm that you use to create an index. Valid values: hnsw and linear.	hnsw

Write parameters for HNSW

Parameter	Description	Default value
index.vector.hnsw.builder.max_scan_num	The maximum number of the nearest neighbors that you want to scan when a graph is created under the worst case.	100000
index.vector.hnsw.builder.neighbor_cnt	The maximum number of the nearest neighbors that each node can have at layer 0. We recommend that you set this parameter to 100. The quality of a graph increases with the value of this parameter. However, inactive indexes consume more storage resources.	100
index.vector.hnsw.builder.upper_neighbor_cnt	The maximum number of the nearest neighbors that each node can have on a layer other than layer 0. We recommend that you set this parameter to 50% of neighbor_cnt.	50
index.vector.hnsw.builder.efconstruction	The number of the nearest neighbors that you want to scan when a graph is created. The quality of a graph increases with the value of this parameter. However, a longer time period is required to create indexes. We recommend that you set this parameter to 400.	400
index.vector.hnsw.builder.max_level	The total number of layers, which includes layer 0. For example, you have 10 million documents and the scaling_factor parameter is set to 30. Use 30 as the base number and then round up the logarithm of 10,000,000 to the nearest integer. The result is 5.	6

Parameter	Description	Default value
index.vector.hnsw.builder.scaling_factor	A scaling factor. The volume of data on a layer equals the volume of data on its upper layer multiplied by the scaling factor. Valid values: 10 to 100. The number of layers decreases with the value of scaling_factor. We recommend that you set this parameter to 50.	50

Search parameters for HNSW

Parameter	Description	Default value
ef	The number of the nearest neighbors that are scanned during an online search. A large value increases the recall ratio but slows down the search. Valid values: 100 to 1000.	100

Sample request:

```
GET test/_search
{
 "query": {
  "hnsw": {
   "feature": {
    "vector": [1.5, 2.5],
    "size": 10,
    "ef": 100
   }
  }
 }
}
```

Circuit breaker parameters

Parameter	Description	Default value
indices.breaker.vector.native.indexing.limit	If the off-heap memory usage exceeds the value specified by this parameter, write operations are suspended. After Elasticsearch creates indexes and releases the memory, it resumes the write operations. If the circuit breaker is triggered, the system memory consumption is high. We recommend that you throttle the write throughput. If you are a beginner, we recommend that you use the default value.	70%

Parameter	Description	Default value
indices.breaker.vector.native.total.limit	The maximum proportion of off-heap memory used to create vector indexes. If the actual off-heap memory usage exceeds the value specified by this parameter, Elasticsearch may reallocate the shards. If you are a beginner, we recommend that you use the default value.	80%

FAQ

- Q: How do I evaluate the recall ratio of documents?

  A: You can create two indexes. One uses the HNSW algorithm and the other uses the Linear Search algorithm. Keep other index settings consistent for the two indexes. Add the same vector data to the indexes from your client and refresh the indexes. Compare the document IDs returned by the HNSW index and the Linear Search index after the same query vector is used. Then, find out the same document IDs that are returned by both indexes.

  > **Note** Divide the number of document IDs returned by both indexes by the total number of returned document IDs to calculate the recall ratio of the documents.

- Q: How do I resolve a circuitBreakingException error when I write data to Elasticsearch?

  A: This error indicates that the off-heap memory usage exceeds the proportion specified by the indices.breaker.vector.native.indexing.limit parameter and that the write operation is suspended. The default proportion is 70%. In most cases, after Elasticsearch creates indexes and releases memory, the write operation is automatically resumed. We recommend that you add a retry mechanism to the data write script on your client.

- Q: Why is the CPU still working after the write operation is suspended?

  A: Elasticsearch creates vector indexes during both the refresh and flush processes. The vector index creation task may be still running even if the write operation is suspended. Computing resources are released after the final refresh is complete.

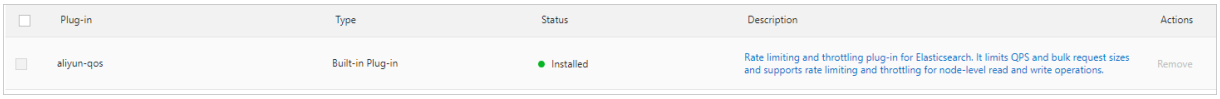
7.2.7. Use the aliyun-qos plug-in

aliyun-qos is a throttling plug-in developed by Alibaba Cloud Elasticsearch to improve cluster stability. It implements node-level read/write throttling and reduces the priority of a specified index if required. You can use the aliyun-qos plug-in to reduce the priorities of services based on the rules predefined by the plug-in. This applies, if you cannot implement throttling on your upstream services, especially on read requests.

aliyun-qos plug-in Elasticsearch cluster throttling

Precautions

aliyun-qos is a built-in plug-in. By default, the throttling feature is disabled, and this plug-in cannot be removed. aliyun-qos is designed to improve cluster stability. It does not perform a precise measurement of the read and write traffic.



> **Note** If your Elasticsearch cluster is created before the plug-in is released, you must install the plug-in on the plug-in configuration page. For more information, see Install and remove a built-in plug-in. The plug-in cannot be uninstalled after it is installed.

Evaluate thresholds

To ensure the execution efficiency of read and write requests, the aliyun-qos plug-in performs throttling only on a single node, and does not perform a precise measurement of the read and write traffic on all nodes in the cluster. This may cause an inconsistency between the calculated threshold and the actual threshold. Before you can use the aliyun-qos plug-in, evaluate throttling thresholds as follows:

- Query requests

  Throttling threshold for query requests = Number of limited query requests in a cluster/Number of client nodes or data nodes in the cluster

  For example, if a cluster has five client nodes and its query requests are limited to 1,000, the throttling threshold for query requests is 200.

  > **Notice** The aliyun-qos plug-in does not synchronize query traffic between nodes, and 200 is only an approximate value. In actual situations, the query traffic is not evenly distributed between nodes, and the value needs to be adjusted.

- Write requests

  The throttling threshold for write requests is calculated by using a similar method to that for query requests, and it needs to be adjusted based on the number of replicas.

  For example, a cluster has two data nodes and one index, the index has one shard and one replica, and 10 MiB of data is written in each time. In this case, each data node is written with 10 MiB of data for each time because the index has a replica. In addition, X-Pack Monitor, Audit, and Watcher tasks also generate write traffic. You must consider these tasks when you set the throttling threshold.

Enable throttling

1. Log on to the Kibana console of your Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab, run the following command to enable the throttling feature of aliyun-qos:

   ```
   PUT _cluster/settings
   {
     "transient" : {
       "apack.qos.ratelimit.enabled":"true"
     }
   }
   ```

   > **Note** By default, the throttling feature of aliyun-qos is disabled.

   After you enable the throttling feature, proceed with the following operations.

Set QPS

You can define the `index_patterns` parameter and set the queries per second (QPS) for a specific index or indexes specified by using a wildcard.

- Set the QPS for a specific index

```
PUT _qos/_ratelimit/<limitName>
{
  "search.index_patterns" : "twitter",
  "search.max_times_sec" : 1000
}
```

- Set the QPS for indexes with a specified prefix

```
PUT _qos/_ratelimit/<limitName>
{
  "search.index_patterns" : "nginx-log-*",
  "search.max_times_sec" : 1000
}
```

- Set the QPS for all indexes

```
PUT _qos/_ratelimit/<limitName>
{
  "search.index_patterns" : "*",
  "search.max_times_sec" : 2000
}
```

Note You can define multiple rules to trigger throttling. If a request hits one of the rules, throttling is triggered.

When you query data on your client or in the Kibana console, the system displays the following error message if the QPS exceeds the value specified by `search.max_times_sec`. You must reduce the QPS.

```
{
 "error": {
  "root_cause": [
   {
    "type": "rate_limited_exception",
    "reason": "request indices:data/read/search rejected, limited by [l1:t*:1.0]"
   }
  ],
  "type": "rate_limited_exception",
  "reason": "request indices:data/read/search rejected, limited by [l1:t*:1.0]"
 },
 "status": 429
}
```

Set bulk.max_bytes_sec

You can set bulk.max_bytes_sec to limit the maximum number of bytes to write and receive per second by client nodes for all bulk requests. For more information, visit Bulk API.

```
PUT _qos/_ratelimit/<limitName>
{
  "bulk.max_bytes_sec" : 1000000
}
```

Note You can define multiple rules to trigger throttling. If a request hits one of the rules, throttling is triggered.

When you write data on your client or in the Kibana console, the system displays the following error message if the number of bytes to write per second exceeds `bulk.max_bytes_sec`. You must reduce the number.

```
{
 "error": {
  "root_cause": [
   {
    "type": "rate_limited_exception",
    "reason": "request indices:data/write/bulk rejected, limited by [b2:ByteSizePreSeconds:992.0]"
   }
  ],
  "type": "rate_limited_exception",
  "reason": "request indices:data/write/bulk rejected, limited by [b2:ByteSizePreSeconds:992.0]"
 },
 "status": 413
}
```

Set bulk.max_bytes_pre

You can set bulk.max_bytes_pre to limit the maximum number of bytes to write and receive by client nodes for a single bulk request. For more information, visit Bulk API.

```
PUT _qos/_ratelimit/<limitName>
{
  "bulk.max_bytes_pre" : 1000
}
```

Note You can define multiple rules to trigger throttling. If a request hits one of the rules, throttling is triggered.

Obtain throttling rules

- Obtain all throttling rules

  ```
  GET _qos/_ratelimit
  ```

- Obtain a specified throttling rule

  ```
  GET _qos/_ratelimit/<limitName>
  ```

- Obtain specified throttling rules

  ```
  GET _qos/_ratelimit/<limitName1,limitName2>
  ```

Delete a throttling rule

```
DELETE _qos/_ratelimit/<limitName>
```

Disable throttling

```
PUT _cluster/settings
{
  "transient" : {
    "apack.qos.ratelimit.enabled":"false"
  }
}
```

```
PUT _cluster/settings
{
  "transient" : {
    "apack.qos.ratelimit.enabled":null
  }
}
```

7.2.8. Use the codec-compression plug-in of the beta version

Elasticsearch index compression

codec-compression is an index compression plug-in developed by Alibaba Cloud Elasticsearch. It supports brotli and zstd compression algorithms and provides a higher compression ratio for indexes. This significantly reduces index storage costs.

Elasticsearch index compression codec-compression brotli zstd

Prerequisites

You have completed the following operations:

- An Alibaba Cloud Elasticsearch V6.7.0 cluster is created. For more information, see Create an Elasticsearch cluster.

  > **Notice** The codec-compression plug-in is available only in Alibaba Cloud Elasticsearch V6.7.0.

- The codec-compression plug-in is installed. By default, this plug-in is installed for new clusters.

  If the cluster is created before the plug-in is released, you must manually install the plug-in. For more information, see Install and remove a built-in plug-in.

Context

The codec-compression plug-in supports brotli and zstd compression algorithms. It is suitable for scenarios where a large volume of data needs to be written or the storage costs for indexes are high, such as in logging and time series data analysis. A performance test is as follows:

- Test environment

- Node configuration: 3 data nodes (each with 16 vCPUs and 64 GiB of memory) + 2-TiB standard SSD
- Datasets: 74-GiB nyc_taixs of Rally provided by open-source Elasticsearch
- Index settings: default settings (force merge after data writing)

- Test result

Compression algorithm	Index size (GiB)	TPS (document/s)
LZ4 (default compression algorithm of Elasticsearch)	35.5	202,682
best_compression (DEFLATE)	26.4	181,686
brotli	24.4	182,593
zstd	24.6	181,393

- Test conclusion

  When codec-compression uses brotli and zstd, it achieves a 45% higher compression ratio and experiences a write performance loss of 10% compared with when it uses LZ4. However, it achieves an 8% higher compression ratio and maintains the same write performance compared with when it uses best_compression (DEFLATE).

Procedure

1. Log on to the Kibana console. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab, run the following commands to specify different compression algorithms for an index:
   - brotli compression algorithm

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

   - zstd compression algorithm

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

7.2.9. Use the faster-bulk plug-in

The faster-bulk plug-in is developed by the Alibaba Cloud Elasticsearch team. This plug-in aggregates bulk write requests in batches based on the specified maximum request size and aggregation interval. This prevents small bulk requests from blocking the write queue, improves write throughput, and reduces write request rejections. This topic introduces the use scenarios of the plug-in and describes how to use it.

Scenarios

The faster-bulk plug-in is ideal for scenarios with high write throughput and numerous index shards. The plug-in can improve the write throughput by more than 20% in these scenarios.

> **Notice** The faster-bulk plug-in aggregates bulk write requests in batches before it writes data to shards. Therefore, we recommend that you do not use this plug-in in latency-sensitive scenarios.

- Test environment
  - Node configuration: 3 data nodes and 2 independent client nodes. Each node offers 16 vCPUs and 64 GiB of memory.
  - Dataset: nyc_taixs provided by Rally in open source Elasticsearch. The size of a single document is 650 bytes.
  - Parameter setting: The apack.fasterbulk.combine.interval parameter is set to 200ms.
  - Translog status: Tests are performed in each of the synchronous and asynchronous states. If the index.translog.durability parameter is set to request, translogs are in the synchronous state. If the index.translog.durability parameter is set to async, translogs are in the asynchronous state.
- Test results

Translog status	Write performance of an open source Elasticsearch cluster without faster-bulk (document/s)	Write performance of an Alibaba Cloud Elasticsearch cluster with faster-bulk (document/s)	Improved By
Synchronous	182,314	226,242	23%
Asynchronous	218,732	241,060	10%

- Test conclusion

  The write performance is improved in both the synchronous state (default state) and asynchronous state after the faster-bulk plug-in is used. In the synchronous state, the write performance is improved by 23%.

Prerequisites

- An Alibaba Cloud Elasticsearch V6.7.0 cluster is created.

  For more information, see Create an Elasticsearch cluster.

  > **Note** Only Alibaba Cloud Elasticsearch V6.7.0 clusters of the Standard or Advanced Edition support the faster-bulk plug-in.

- The faster-bulk plug-in is installed.

  For more information, see Install and remove a built-in plug-in. After the plug-in is installed, the bulk request aggregation feature is disabled by default. Before you use this plug-in, enable this feature.

Enable the bulk request aggregation feature

1. Log on to the Kibana console of your Alibaba Cloud Elasticsearch cluster.For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**.
3. On the **Console** tab of the page that appears, run the following command to enable the bulk request aggregation feature:

```
PUT _cluster/settings
{
  "transient" : {
    "apack.fasterbulk.combine.enabled":"true"
  }
}
```

> **Note** You can also use the cURL tool or a third-party visualizer to run the preceding command.

Configure the maximum request size and aggregation interval

Run the following command to configure the maximum request size and aggregation interval. If the total size of bulk requests or the aggregation interval on a single data node reaches the configured threshold, the system writes data to shards.

```
PUT _cluster/settings
{
  "transient" : {
    "apack.fasterbulk.combine.flush_threshold_size":"1mb",
    "apack.fasterbulk.combine.interval":"50"
  }
}
```

- apack.fasterbulk.combine.flush_threshold_size: the maximum size of bulk requests. Default value: 1mb.
- apack.fasterbulk.combine.interval: the maximum interval at which bulk requests are aggregated. Default value: 50. Unit: ms.

> **Note** To process highly concurrent bulk requests and prevent the requests from blocking the write queue, you can increase the maximum request size or aggregation interval based on your business requirements.

Disable the bulk request aggregation feature

Run the following command to disable the bulk request aggregation feature:

```
PUT _cluster/settings
{
  "transient" : {
    "apack.fasterbulk.combine.enabled":"false"
  }
}
```

7.2.10. Use the gig plug-in

gig is a plug-in developed by Alibaba Cloud Elasticsearch to implement throttling for client nodes in an Elasticsearch cluster. This plug-in integrates the core throttling capabilities possessed by the Taobao team to handle searches. The gig plug-in can perform a switchover within seconds if query jitters caused by accidental node exceptions occur. This minimizes the probability that query jitters occur and ensures the stability of queries. In addition, this plug-in detects traffic to handle query latency surges caused by enabled warm nodes and achieve query warm-up for online business. This topic describes how to use the gig plug-in.

Background information

This section describes how the gig plug-in works.

- The gig plug-in runs on client nodes. For applications that require high query QPS, you can increase the number of replica shards for each primary shard to scale out the cluster. This helps achieve a linear increase in query throughput. The gig plug-in can help client nodes select the most appropriate replica shards to provide query services.

- The plug-in determines the service capabilities of nodes based on query latency and coordinates the nodes that provide services by using the proportion integral differential (PID) algorithm. This ensures rapid and accurate coordination. If exceptions such as surging query latency or rising error rates occur on nodes, the gig plug-in can collect and analyze the metrics of the nodes in real time by using the PID algorithm. Then, the plug-in rapidly isolates anomalous nodes and performs a switchover within seconds.
- When new nodes join the cluster, the plug-in samples online query traffic in real time, replicates some query traffic to the new nodes, and discards query results. The traffic that is replicated is detection traffic. This avoids direct transmission of traffic to nodes that cannot provide services and reduces query latency. If the detection results and metrics show that the latency of the new nodes is in a normal range, the plug-in transmits online query traffic to these nodes. Then, these nodes can provide online services.

Precautions

- The gig plug-in is available for Alibaba Cloud Elasticsearch V6.7.0 clusters that have a kernel version of 1.3.0. Before you use this plug-in, make sure that the kernel version of your Elasticsearch cluster is 1.3.0. Otherwise, upgrade the kernel. You can upgrade only the kernels of Standard Edition clusters whose kernel versions are V0.3.0, V1.0.2, or V1.2.0.
- The gig plug-in is integrated into kernel V1.3.0. After you upgrade the kernel of your cluster, the plug-in is automatically installed. After the plug-in is installed, the throttling feature of the plug-in is disabled by default. If you want to use the feature, you must enable it.
- Before you use the gig plug-in, make sure that sufficient resources are reserved for the data nodes in the cluster. If exceptions occur on one of the data nodes, the query traffic is transmitted to other data nodes. This increases the load of these nodes. Therefore, you must reserve sufficient resources for data nodes to ensure business stability.
- All commands in this topic can be run in the Kibana console. For more information about how to log on to the Kibana console, see Log on to the Kibana console.

Procedure

1. Enable the throttling feature for the gig plug-in.

```
PUT test/_settings
{
 "index.flow_control.enabled": true
}
```

> **Note** If you want to disable the feature, set index.flow_control.enabled to null or false.

2. Configure thresholds for query latency in the gig plug-in. If one of the thresholds are met, the plug-in performs throttling.

```
PUT test/_settings
{
  "index.flow_control.search":{
      "latency_upper_limit_extra": "10s",
      "latency_upper_limit_extra_percent": "1.0",
      "probe_percent": "0.2",
      "full_degrade_error_percent": "0.5",
      "full_degrade_latency": "10s"
  }
}
```

Parameter	Default value	Description
latency_upper_limit_extra	10s	The threshold for the absolute value of the difference between the actual query latency and average query latency. This parameter is represented by using the following formula: |Actual query latency - Average query latency|. The default value is 10s. This indicates that if the average query latency of three data nodes in the cluster is 2s, when the query latency of one of the three data nodes reaches 13s, the gig plug-in performs throttling.
latency_upper_limit_extra_percent	1.0	The threshold for the proportion of the absolute value of the difference between the actual query latency and average query latency to the average query latency. This parameter is represented by using the following formula: (|Actual query latency - Average query latency|)/Average query latency. The default value is 1.0. This indicates that if the average query latency of three data nodes in the cluster is 2s, when the actual query latency of one of the three data nodes reaches 4s, the gig plug-in performs throttling.
probe_percent	0.2	The threshold for the proportion of detection traffic to the actual query traffic. The default value is 0.2. This indicates that if the proportion of detection traffic to the actual query traffic is greater than 0.2, the gig plug-in performs throttling.
full_degrade_error_percent	0.5	The threshold for the proportion of query exceptions. The default value is 0.5. This indicates that if the error rate of query responses of a data node in the cluster reaches 50%, the gig plug-in performs throttling.

Parameter	Default value	Description
full_degrade_latency	10s	The threshold for the query latency. The default value is 10s. This indicates that if the query latency is greater than 10s, the gig plug-in performs throttling.

Notice You can adjust the values of these parameters based on your business requirements.

7.3. Upload and install a custom plug-in

Alibaba Cloud Elasticsearch allows you to upload and install custom plug-ins.

Elasticsearch custom plug-in upload and install a custom plug-in

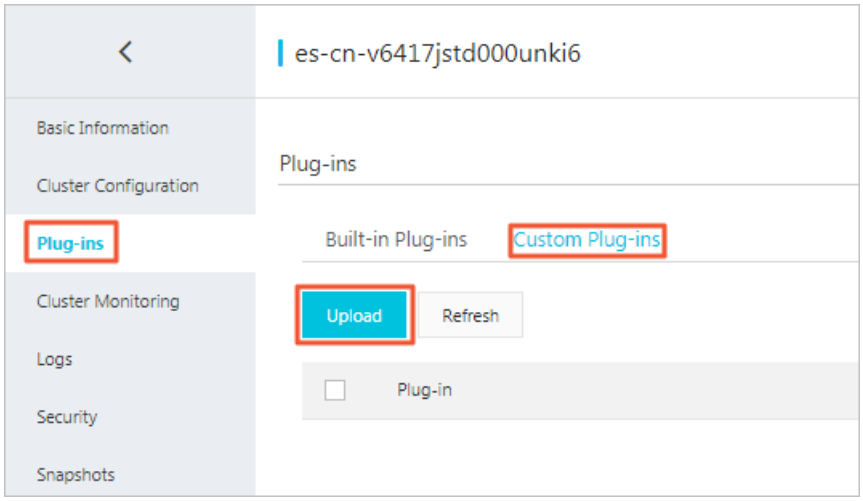
Prerequisites

If you want to upload a custom SQL plug-in, make sure that the `xpack.sql.enabled` parameter in the YML configuration files of your Elasticsearch cluster is set to `false`.

For more information, see Modify the YML configuration.

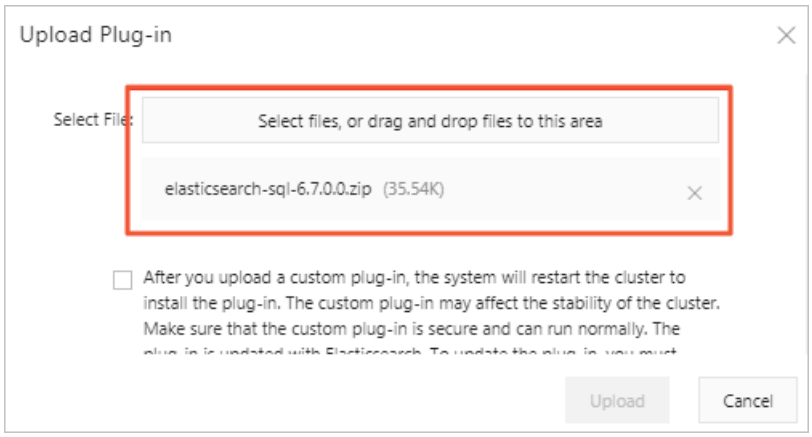
Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Plug-ins**.
5. On the **Plug-ins** page, click the **Custom Plug-ins** tab. Click **Upload**.



> **Warning** Uploading a custom plug-in triggers a restart of the Elasticsearch cluster, and the plug-in may affect the stability of the cluster. Make sure that the custom plug-in is secure and available.

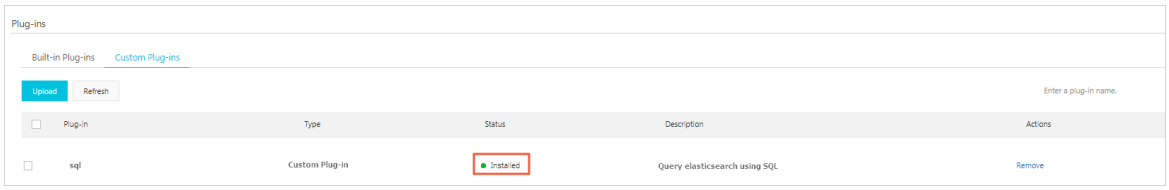
6. In the **Upload Plug-in** dialog box, click **Select files, or drag and drop files to this area**. Then, select the custom plug-in that you want to upload.



You can also drag and drop a custom plug-in file to this area and upload the plug-in. As shown in the preceding figure, the plug-in file **Elasticsearch-sql-6.7.0.0** is added.

> **Note** You can repeat this step to add more custom plug-ins.

7. Carefully read the agreement, select the check box, and click **Upload**.
Your Alibaba Cloud Elasticsearch cluster is restarted when the plug-in is uploaded. After the cluster is restarted, you can check the plug-in on the **Custom Plug-ins** tab. The **Status** of the plug-in that you upload displays **Installed**. This indicates that the plug-in is successfully uploaded and installed.



> **Note** The plug-in is not automatically updated with Elasticsearch. If you want to update the plug-in, you must manually upload a new version of the plug-in.

What's next

If you no longer need the plug-in, find it on the Custom Plug-in tab and click **Remove** in the Actions column. For more information, see Install and remove a built-in plug-in.

8.Cluster monitoring and alerting

8.1. Enable the alerting feature

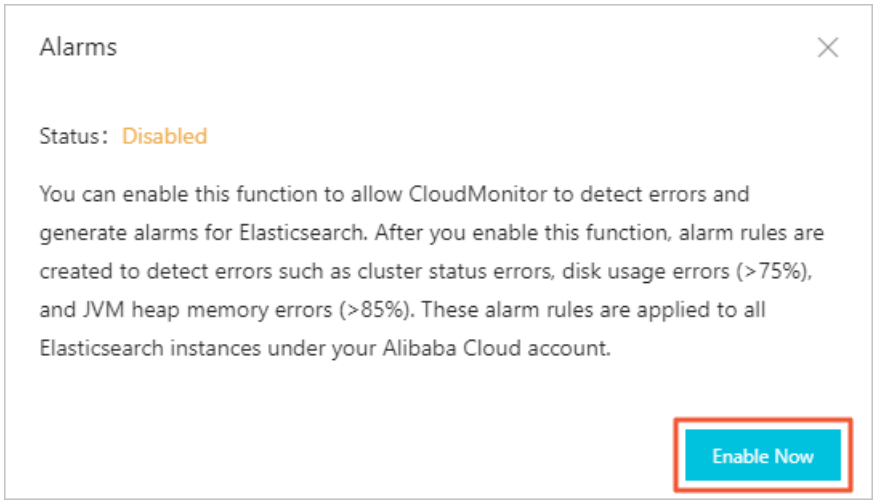
Elasticsearch alerting

This topic describes how to enable the alerting feature for your Alibaba Cloud Elasticsearch cluster. This feature is provided by CloudMonitor. After this feature is enabled, alert rules are created to detect errors, such as abnormal cluster status, high disk usage (greater than 75%), and high JVM heap memory usage (greater than 85%). These rules apply to all Elasticsearch clusters under your Alibaba Cloud account.

Elasticsearch alerting

Procedure

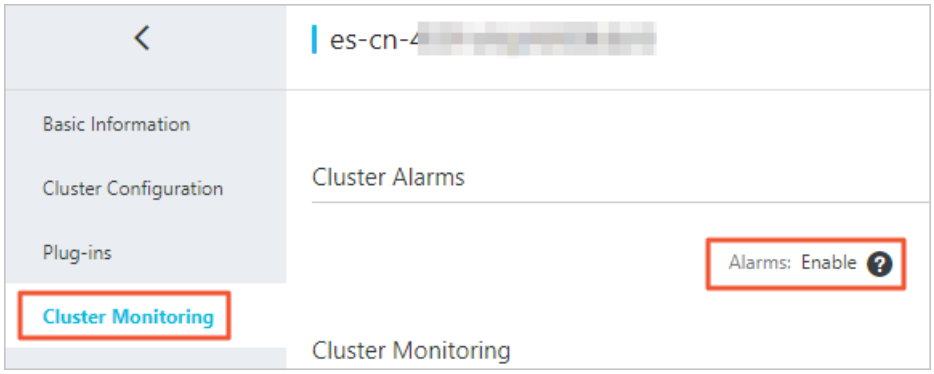
1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click Elasticsearch Clusters. On the **Clusters** page, click **Alarms**.
4. In the **Alarms** message, click **Enable Now**. The Elasticsearch alerting feature is disabled by default.



5. In the CloudMonitor console, turn on the **Initiative Alarm** switch for Elasticsearch.

6. Go to the Alibaba Cloud Elasticsearch console to check whether **Alarms** is enabled.
   i. On the **Clusters** page, find the target cluster and click its ID in the **Cluster ID/Name** column.
   ii. In the left-side navigation pane of the cluster details page, click **Cluster Monitoring**.

iii. In the **Cluster Alarms** section, check the status of **Alarms**. If the status of **Alarms** is **Enable**, you have enabled the Elasticsearch alerting feature.



8.2. Configure the monitoring and alerting feature in Cloud Monitor

Elasticsearch monitoring and alerting

Alibaba Cloud Elasticsearch allows you to monitor clusters and customize alert thresholds. If an alert is detected, the system notifies you of the alert by sending a text message. To ensure the stability of your Elasticsearch cluster, we recommend that you configure monitoring and alerting. This way, the system can monitor items in real time, such as cluster status and disk usage. You can check text messages in time and take measures in advance.

Context

Elasticsearch allows you to configure monitoring and alerting for the metrics that are described in the following table.

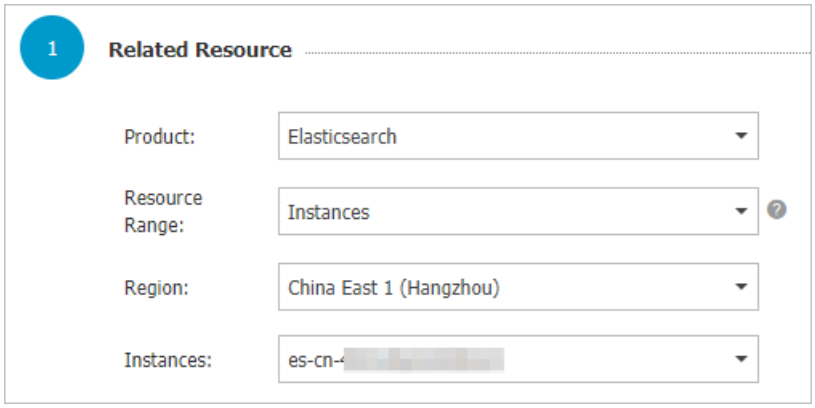
Metric	Description
ClusterStatus	Required. This metric checks the cluster status. Green indicates that an Elasticsearch cluster is in a normal state. Yellow or red indicate that an Elasticsearch cluster is in an abnormal state.
NodeDiskUtilization(%)	Required. Set the threshold to a value that is less than 75%. The upper limit is 80%.
NodeHeapMemoryUtilization(%)	Required. Set the threshold to a value that is less than 85%. The upper limit is 90%.
NodeCPUUtilization(%)	Optional. Set the threshold to a value that is less than or equal to 95%.
NodeLoad_1m	Optional. Set the threshold to a value that is 80% of the number of CPU cores per node.
ClusterQueryQPS(Count/Second)	Optional. Set the threshold based on the actual test result.
ClusterIndexQPS(Count/Second)	Optional. Set the threshold based on the actual test result.

Note The monitoring and alerting feature is enabled for your Elasticsearch cluster by default. You can view historical monitoring data on the Cluster Monitoring page of your cluster. Only monitoring information that is generated over the last month is displayed.

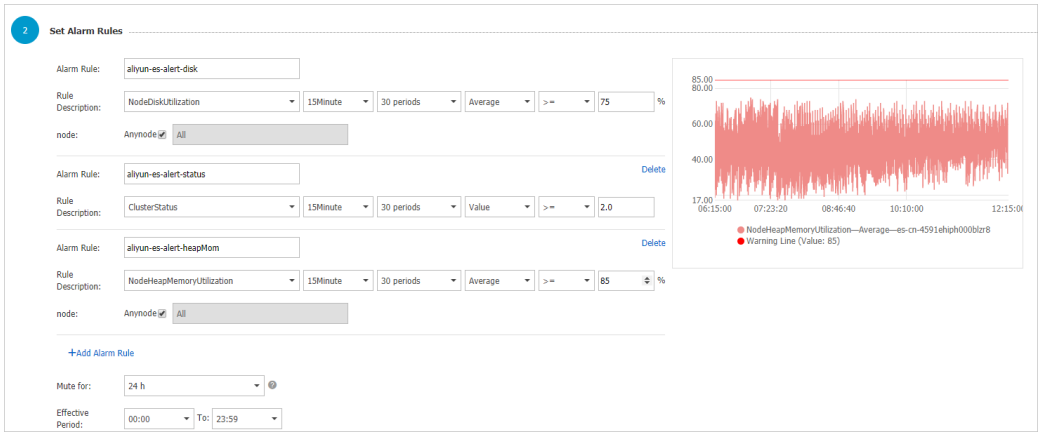
Procedure

1. Go to the Elasticsearch page of the Cloud Monitor console.You can use one of the following methods:
   - From the Elasticsearch console
     a. Log on to the Alibaba Cloud Elasticsearch console.
     b. In the top navigation bar, select the region where your cluster resides.
     c. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
     d. On the **Basic Information** page, click **Cluster Monitoring** in the upper-right corner.
   - From the Cloud Monitor console
     a. Log on to the Cloud Monitor console.
     b. Select the region where your cluster resides.
     c. On the **Instances** tab, find your cluster and click its ID in the **Instance ID** column.
2. In the upper-right corner, click **Alert Rules**.
3. On the page that appears, click **Create Alert Rule**.
4. On the **Create Alert Rule** page, specify **alert rules**.The following example demonstrates how to specify alert rules for **NodeDiskUtilization(%)**, **ClusterStatus**, and **NodeHeapMemoryUtilization(%)**.

   Related Resource step



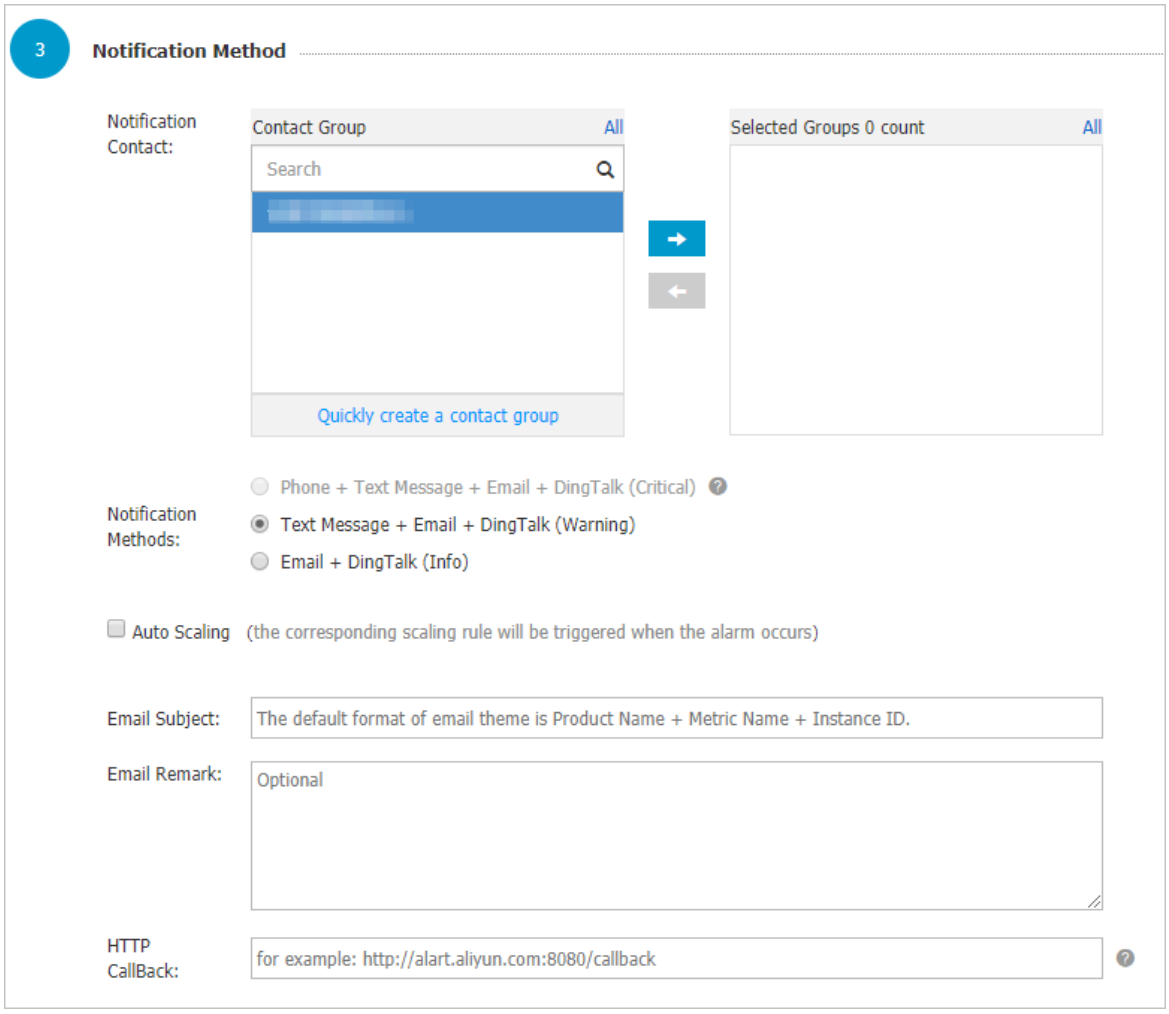
   Set Alert Rules step



- The values for cluster states **Green**, **Yellow**, and **Red** are **0.0**, **1.0**, and **2.0**. Reference these values and set a suitable threshold for the **ClusterStatus** metric.
- The **Mute for** parameter specifies the interval at which an alert notification is re-sent when a threshold is reached.

> **Note** For more information about other parameters, see Create a threshold-triggered alert rule.

5. In the **Notification Method** step, select **Default Contact Group** from the Contact Group section and click the rightwards arrow to add it to the Selected Groups section.If you do not have an alert group, click **Quickly create a contact group** to create a group.



> **Note** In the **HTTP CallBack** field, enter a URL that can be accessed from the Internet. Cloud Monitor delivers a POST request to this URL to push the alert notification. Only HTTP is supported.

6. Click **Confirm**.
   Then, the system starts to monitor your Elasticsearch cluster and displays the monitoring data on the Cluster Monitoring page of the cluster.

8.3. View cluster monitoring data

View cluster monitoring data

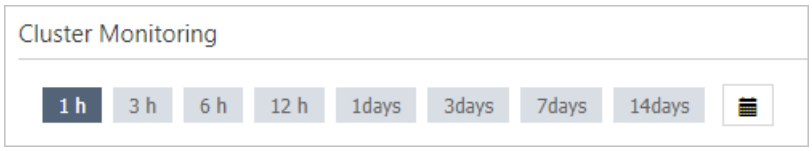
This topic describes how to view cluster monitoring data to query the running status of an Elasticsearch cluster in real time.

Elasticsearch cluster monitoring

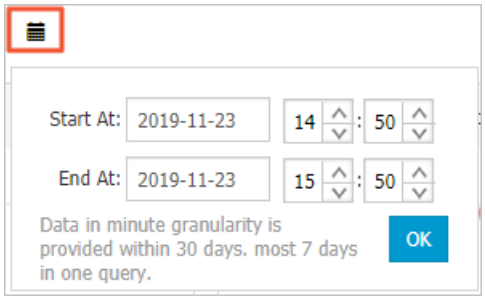
Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.

4. In the left-side navigation pane of the cluster details page, click **Cluster Monitoring**.
5. In the **Cluster Monitoring** section, select a time period to view the detailed monitoring data that is generated within this period.



6. Click the **Custom** icon, set **Start At** and **End At**, and then click **OK** to view the detailed monitoring data that is generated within the customized time period.



> **Notice** You can query monitoring data that is accurate to the minute within the last 30 days.

What's next

Address potential issues in a timely manner based on the results to ensure the stable running of the cluster. For more information about monitoring metrics, see Monitoring metrics.

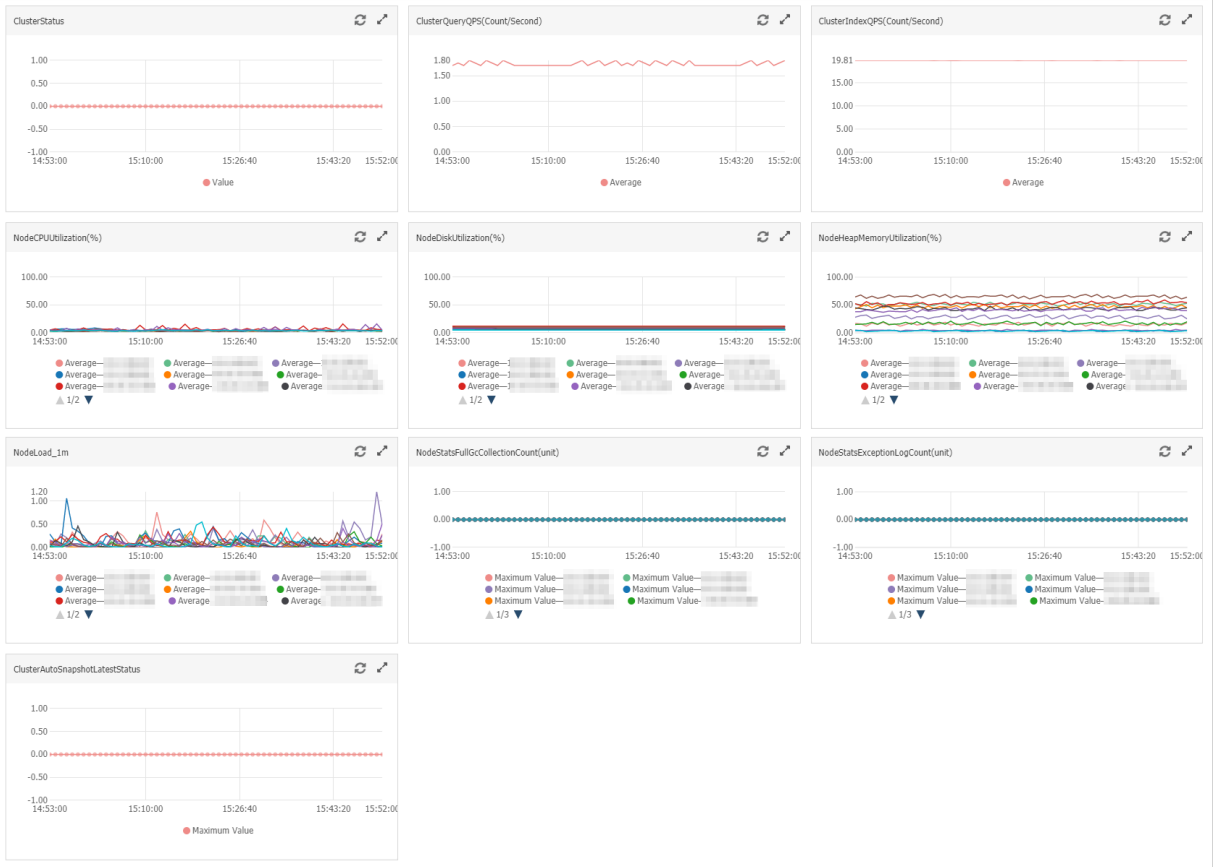
8.4. Monitoring metrics

Elasticsearch monitoring metrics

This topic describes the monitoring metrics for Alibaba Cloud Elasticsearch. These metrics are used to monitor the running status of the clusters, and include ClusterStatus, ClusterQueryQPS(Count/Second), NodeCPUUtilization(%), and NodeDiskUtilization(%). You can learn about the running status of the clusters in real time and address potential issues to ensure the stable running of the clusters.

Elasticsearch monitoring metric Elasticsearch cluster status

Overview



The following monitoring metrics are provided: ClusterStatus, ClusterQueryQPS(Count/Second), ClusterIndexQPS(Count/Second), NodeCPUUtilization(%), NodeDiskUtilization(%), NodeHeapMemoryUtilization(%), NodeLoad_1m, NodeStatsFullGcCollectionCount(unit), NodeStatsExceptionLogCount(unit), and ClusterAutoSnapshotLatestStatus.

ClusterStatus

The **ClusterStatus** metric indicates the status of an Elasticsearch cluster. Value 0.00 indicates that the cluster is properly running. This metric is required. For more information, see Configure the monitoring and alerting feature in Cloud Monitor.

If the status is not displayed in green on the **Basic Information** page, the value of the metric is not 0.00, which indicates that the cluster is abnormal. The following list provides a few common reasons for this issue:

- The CPU usage or heap memory usage of the nodes in the cluster is too high or reaches 100%.
- The disk usage of the nodes in the cluster is too high (higher than 85%) or reaches 100%.
- The minute-average node workload (NodeLoad_1m) is too high.
- The status of the indexes in the cluster is abnormal (not green).

The following table describes the ClusterStatus values.

Value	Color	Status	Description
2.00	Red	Not all of the shards are available.	One or more indexes have unassigned shards.

Value	Color	Status	Description
1.00	Yellow	All shards are available, but not all of the replicas are available.	One or more indexes have unassigned replicas.
0.00	Green	All shards and replicas are available.	All indexes in the cluster are healthy. No unassigned shards or replicas exist.

Note The colors in the table indicate those of **Status** on the Basic Information page of a cluster.

ClusterQueryQPS(Count/Second)

Notice If the **query QPS** spikes, the CPU usage, heap memory usage, or minute-average node workload may reach high levels. This may affect your services that run in the cluster.

The **ClusterQueryQPS(Count/Second)** metric indicates the number of **queries processed by the cluster per second**.

The number of queries processed per second varies with the number of shards of the index that is queried. For example, if an index has five shards, the cluster can process requests to this index at a rate of five queries per second.

ClusterIndexQPS(Count/Second)

Notice If the **write QPS** spikes, the CPU usage, heap memory usage, or minute-average node workload may reach high levels. This may affect your services that run in the cluster.

The **ClusterIndexQPS(Count/Second)** metric is calculated based on the number of write requests that a cluster receives per second and the number of documents that these requests write.

If a cluster receives only one write request in one second and the request only writes one document, the **write QPS** is 1. The value of the metric increases with the number of write requests received per second.

If the cluster receives a _bulk request that writes multiple documents in one second, the **write QPS** equals the number of documents to be written. The value of the metric increases with the number of _bulk requests received per second.

NodeCPUUtilization(%)

The **NodeCPUUtilization(%)** metric indicates the CPU usage of each data node in an Elasticsearch cluster. When the CPU usage is high or close to 100%, services that run in the cluster are affected.

If the CPU usage spikes or fluctuates significantly, an error occurs. The following list provides a few common reasons for this issue:

- The **query QPS** or **write QPS** spikes or fluctuates significantly.
- The cluster receives a few slow queries or write requests.

In this case, you may not find spikes or fluctuations in the query QPS and write QPS. You can log on to the Elasticsearch console, go to the **Logs** page of the cluster, and then click the **Search Slow Log** tab to analyze the log data.

- The cluster has a large number of indexes or shards.

  Elasticsearch monitors indexes in the cluster and records index changes in the log. If the cluster has too many indexes or shards, the CPU usage, heap memory usage, or minute-average node workload may reach high levels.

- Merge operations are performed on the cluster.

  Merge operations consume CPU resources. However, the number of segments on the corresponding node decreases significantly. You can check the number of segments on the Overview page of the node in the Kibana console.

- Garbage collection operations are performed on the cluster.

  Garbage collection operations, such as full garbage collection, can be used to release memory resources. However, these operations consume CPU resources. As a result, the CPU usage may spike.

- Scheduled tasks, such as backup tasks or customized tasks, are performed on the cluster.

NodeDiskUtilization(%)

The **NodeDiskUtilization(%)** metric indicates the disk usage of each data node in an Elasticsearch cluster. The disk usage must be less than 85%. We recommend that you configure an alert rule for this metric. Otherwise, the following situations may occur, which can affect your services that run in the cluster:

- By default, if the disk usage of a data node exceeds 85%, new shards cannot be allocated to the data node.
- By default, if the disk usage of a data node exceeds 90%, Elasticsearch attempts to move the shards on this node to data nodes with low disk usage.
- By default, if the disk usage of a data node exceeds 95%, Elasticsearch adds the `read_only_allow_delete` attribute to all indexes in the cluster. As a result, the indexes cannot be written. These indexes can only be read or deleted.

> **Notice** Do not set the threshold for this metric to a value greater than 80%. We recommend that you set the threshold to a value less than 75%. When alerts are triggered, you can resize disks, add nodes, or clear index data to ensure that your services are not affected.

NodeHeapMemoryUtilization(%)

The **NodeHeapMemoryUtilization(%)** metric indicates the heap memory usage of each data node in an Elasticsearch cluster. If the heap memory usage is high or a large object is stored in the memory, your services that run in the cluster are affected. This also triggers a garbage collection.

If the heap memory usage spikes or fluctuates significantly, an error occurs. The following list provides a few common reasons for this issue:

- The **query QPS** or **write QPS** spikes or fluctuates significantly.
- The cluster receives a few slow queries or write requests.

In this case, you may not find spikes or fluctuations in the query QPS and write QPS. You can log on to the Elasticsearch console, go to the **Logs** page of the cluster, and then click the **Search Slow Log** tab to analyze the log data.

- The cluster receives a large number of slow queries or write requests.

  In this case, you may find spikes or fluctuations in the query QPS and write QPS. You can log on to the Elasticsearch console, go to the **Logs** page of the cluster, and then click the **Indexing Slow Log** tab to analyze the log data.

- The cluster has a large number of indexes or shards.

  Elasticsearch monitors indexes in the cluster and records index changes in the log. If the cluster has too many indexes or shards, the CPU usage, heap memory usage, or minute-average node workload may reach high levels.

- Merge operations are performed on the cluster.

  Merge operations consume CPU resources. However, the number of segments on the corresponding node decreases significantly. You can check the number of segments on the **Overview** page of the node in the Kibana console.

- Garbage collection operations are performed on the cluster.

  Garbage collection operations, such as full garbage collection, can be used to release memory resources. However, these operations consume CPU resources. As a result, the heap memory usage decreases significantly.

- Scheduled tasks, such as backup tasks or customized tasks, are performed on the cluster.

NodeLoad_1m

The **NodeLoad_1m** metric indicates the workload of a data node within one minute. You can reference this metric to determine whether a node is busy. You must set this metric to a value lower than the number of CPU cores on the node.

If the value exceeds the number of CPU cores on the node, an error occurs. The following list provides a few common reasons for this issue:

- The CPU usage or heap memory usage is high or reaches 100%.
- The **query QPS** or **write QPS** spikes or fluctuates significantly.
- Slow queries are received.

  A few or a large number of slow queries are received. You can log on to the Elasticsearch console, go to the **Logs** page of the cluster, and click the required tab to analyze the log data.

The following example uses a single-core node. The values for this metric are described in the following list:

- NodeLoad_1m < 1: No pending processes exist.
- NodeLoad_1m = 1: The system does not have idle resources to run more processes.
- NodeLoad_1m > 1: Processes are queued for resources.

NodeStatsFullGcCollectionCount(unit)

Warning If the full garbage collection is frequently triggered, your services that run in the cluster are affected.

The **NodeStatsFullGcCollectionCount(unit)** metric indicates the number of times that the full garbage collection is triggered within one minute.

If the value is not 0, an error occurs. The following list provides a few common reasons for this issue:

- The heap memory usage is high.
- Large objects are stored in the memory.

NodeStatsExceptionLogCount(unit)

The **NodeStatsExceptionLogCount(unit)** metric indicates the number of warning-level entries generated in an Elasticsearch cluster log within one minute.

If the value is not 0, an error occurs. The following list provides a few common reasons for this issue:

- The cluster receives abnormal queries.
- The cluster receives abnormal write requests.
- Errors occur when the cluster runs tasks.
- Garbage collection operations are performed on the cluster.

> **Note**
>
> - Log on to the Elasticsearch console, go to the **Logs** page of the cluster, and click the **Cluster Log** tab. On the **Cluster Log** tab, find exceptions that occurred in the specific time and analyze the causes.
> - The **NodeStatsExceptionLogCount(unit)** metric also counts the garbage collection times recorded in **Cluster Log**.

ClusterAutoSnapshotLatestStatus

The **ClusterAutoSnapshotLatestStatus** metric indicates the status of the **Auto Snapshot** feature of an Elasticsearch cluster. If the value is -1 or 0, auto snapshot is running normally.

If the value is 2, an error occurs. The following list provides a few common reasons for this issue:

- The disk usage of the nodes is high or close to 100%.
- The cluster is abnormal.

The values of this metric are described as follows:

- 0: Snapshots are created.
- -1: No snapshot is created.
- 1: The system is creating a snapshot.
- 2: The system failed to create snapshots.

8.5. Configure X-Pack Watcher

Configure X-Pack Watcher

This topic describes how to configure X-Pack Watcher for Elasticsearch. X-Pack Watcher allows you to trigger specific actions when specified conditions are met. For example, you can create a watch for Elasticsearch to search the logs index for errors and send alerts through emails or DingTalk messages. X-Pack Watcher is a monitoring and alerting service based on Elasticsearch.

X-Pack Watcher Elasticsearch monitoring and alerting service

Prerequisites

- An Elasticsearch cluster that is deployed in a single zone is created.

  For more information, see Create an Elasticsearch cluster.

  > **Note** X-Pack Watcher is available only for an Elasticsearch cluster that is deployed in a single zone.

- X-Pack Watcher is enabled for an Elasticsearch cluster. It is disabled by default.

  For more information, see Modify the YML configuration.

- An Elastic Compute Service (ECS) instance is created.

  The ECS instance must be accessible over the Internet and located in the same region and Virtual Private Cloud (VPC) as the Elasticsearch cluster. For more information, see Step 1: Create an ECS instance.

  > **Note** The X-Pack Watcher feature of Elasticsearch cannot directly access the Internet. You must use the internal endpoint of an Elasticsearch cluster in a VPC to access the Internet. Therefore, you must create an ECS instance that can access both the Internet and the Elasticsearch cluster and use it as a proxy to perform actions.

Context

X-Pack Watcher allows you to create watches. A watch consists of a trigger, input, condition, and actions.

- Trigger

  Determines when the watch is executed. You must configure a trigger for each watch. X-Pack Watcher allows you to create various types of triggers. For more information, see Schedule Trigger.

- Input

  Loads data into the payload of a watch. Inputs are used as filters to match the specified type of index data. For more information, see Inputs.

- Condition

  Controls whether the actions of a watch are performed.

- Actions

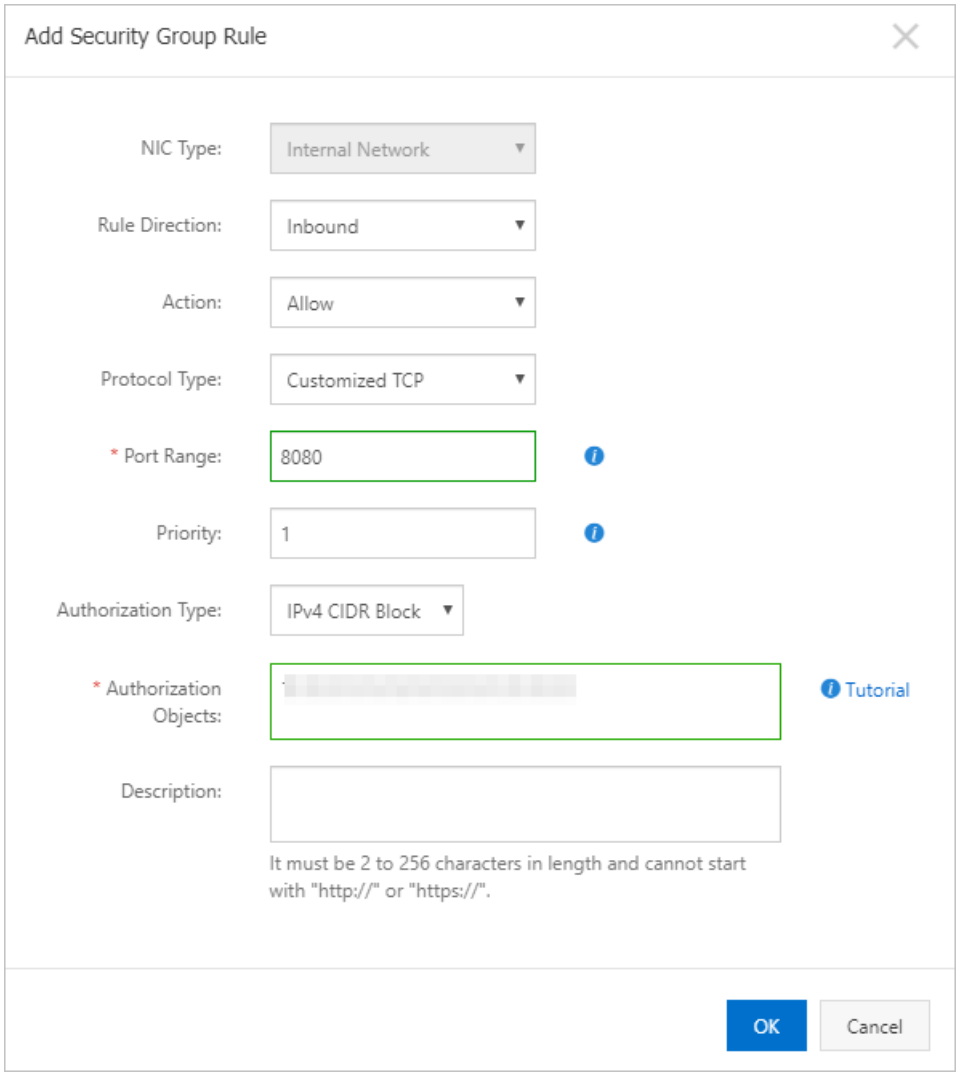
  Determines the actions to be performed when the specified condition is met. This topic uses the webhook action as an example.

Procedure

1. Configure a security group rule for the ECS instance.
   i. Log on to the Alibaba Cloud Elastic Compute Service console. In the left-side navigation pane, click Instances & Images. Then, click **Instances**.
   ii. On the **Instances** page, find the target instance, click **More**, and then choose **Network and Security Group > Configure Security Group** in the **Actions** column.
   iii. On the Security Groups tab, find the target security group and click **Add Rules** in the **Actions** column.

iv. On the **Security Group Rules** page, click **Add Security Group Rule**.

v. Specify the required parameters.



Parameter	Description
Rule Direction	Select **Inbound**.
Action	Select **Allow**.
Protocol Type	Select **Custom TCP**.
Priority	Retain the default value.
Port Range	Set the port to your frequently used port. This parameter is required for NGINX configurations. In this example, port 8080 is used.
Authorization Type	Select **IPv4 CIDR Block**.

Parameter	Description
Authorization Object	Add the IP addresses of all nodes in your Elasticsearch cluster. **Note** To query the IP addresses of the nodes, log on to the Kibana console of your Elasticsearch cluster based on Log on to the Kibana console, click **Monitoring** in the left-side navigation pane, and then click **Nodes**.

vi. Click **OK**.

2. Configure an NGINX proxy.
   i. Install NGINX on the ECS instance.
   ii. Configure the nginx.conf file. Replace the `server` configuration in the nginx.conf file with the following content:

```
server
 {
  listen 8080;# Listening port
  server_name localhost;# Domain name
  index index.html index.htm index.php;
  root /usr/local/webserver/nginx/html;# Website directory
   location ~ . *\.(php|php5)? $
  {
   #fastcgi_pass unix:/tmp/php-cgi.sock;
   fastcgi_pass 127.0.0.1:9000;
   fastcgi_index index.php;
   include fastcgi.conf;
  }
  location ~ . *\.(gif|jpg|jpeg|png|bmp|swf|ico)$
  {
   expires 30d;
   # access_log off;
  }
  location / {
   proxy_pass <Webhook address of the DingTalk Chatbot>;
  }
  location ~ . *\.(js|css)? $
  {
   expires 15d;
   # access_log off;
  }
  access_log off;
 }
```

`<Webhook address of the DingTalk Chatbot>`: Replace it with the webhook address of the DingTalk Chatbot that is used to receive alert notifications.

> **Note** To query the webhook address of the DingTalk Chatbot, create an alert group in DingTalk. Then, in the DingTalk group, click the More icon in the upper-right corner, click ChatBot, and then select Custom to add a ChatBot that is connected by using webhooks. You can then view the webhook address of the DingTalk Chatbot.

iii. Reload the NGINX configuration file and restart NGINX.

```
/usr/local/webserver/nginx/sbin/nginx -s reload       # Reload the NGINX configuration file.
/usr/local/webserver/nginx/sbin/nginx -s reopen       # Restart NGINX.
```

3. Create a watch.

i. Log on to the Kibana console of your Elasticsearch cluster.

> **Note** For more information, see Log on to the Kibana console.

ii. In the left-side navigation pane, click **Dev Tools**.

iii. On the **Console** tab, run the following command to create a watch: The following example shows how to create a watch named `log_error_watch` to search the `logs` index for `errors` every `10 seconds`. If more than `0` errors are found, an alert is triggered.

```
PUT _xpack/watcher/watch/log_error_watch
{
 "trigger": {
  "schedule": {
   "interval": "10s"
  }
 },
 "input": {
  "search": {
   "request": {
    "indices": ["logs"],
    "body": {
     "query": {
      "match": {
       "message": "error"
      }
     }
    }
   }
  }
 },
 "condition": {
  "compare": {
   "ctx.payload.hits.total": {
    "gt": 0
   }
  }
 },
 "actions" : {
 "test_issue" : {
  "webhook" : {
   "method" : "POST",
   "url" : "http://<Private IP address of your ECS instance>:8080",
   "body" : "{\"msgtype\": \"text\", \"text\": { \"content\": \"An error has been found. Handle the issue immediately.\"}}"
  }
 }
}
}
```

> **Notice**
>
> - `url` specified in `actions` must contain the private IP address of your ECS instance that is deployed in the same region and VPC as your Elasticsearch cluster. You must also create a security group rule for the ECS instance by following the preceding procedure. Otherwise, the instance cannot connect to X-Pack Watcher.
> - If error `No handler found for uri [/_xpack/watcher/watch/log_error_watch_2] and method [PUT]` is returned when you run the preceding command, X-Pack Watcher is disabled for your Elasticsearch cluster. In this case, enable X-Pack Watcher and then run the command. For more information, see Modify the YML configuration.

If you no longer need this watch, run the following command to delete the watch:

```
DELETE _xpack/watcher/watch/log_error_watch
```

8.6. Configure monitoring indexes

Monitoring logs

This topic describes how to configure monitoring indexes for your Alibaba Cloud Elasticsearch cluster. After the configuration, you can view monitoring log data. This configuration prevents your storage space from being exhausted by the log data.

Elasticsearch monitoring log Elasticsearch monitoring index Monitoring log

Prerequisites

An Elasticsearch cluster is created. If no cluster is created, first perform the operations described in Create an Elasticsearch cluster. This topic describes how to create an Elasticsearch V6.7 cluster of the Standard Edition.

Context

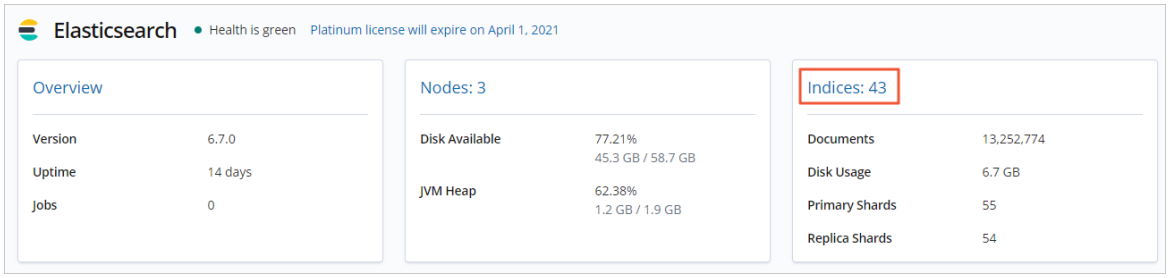
By default, the X-Pack monitoring component collects the monitoring data every 10 seconds and saves the data to the indexes that have prefix `.monitoring-*` in your Elasticsearch cluster.

The `.monitoring-es-6-*` and `.monitoring-kibana-6-*` indexes are used to store the monitoring data. The Elasticsearch cluster rolls over to a new index each day. The name of a `.monitoring-es-6-` index ends with the date when the monitoring data is saved.

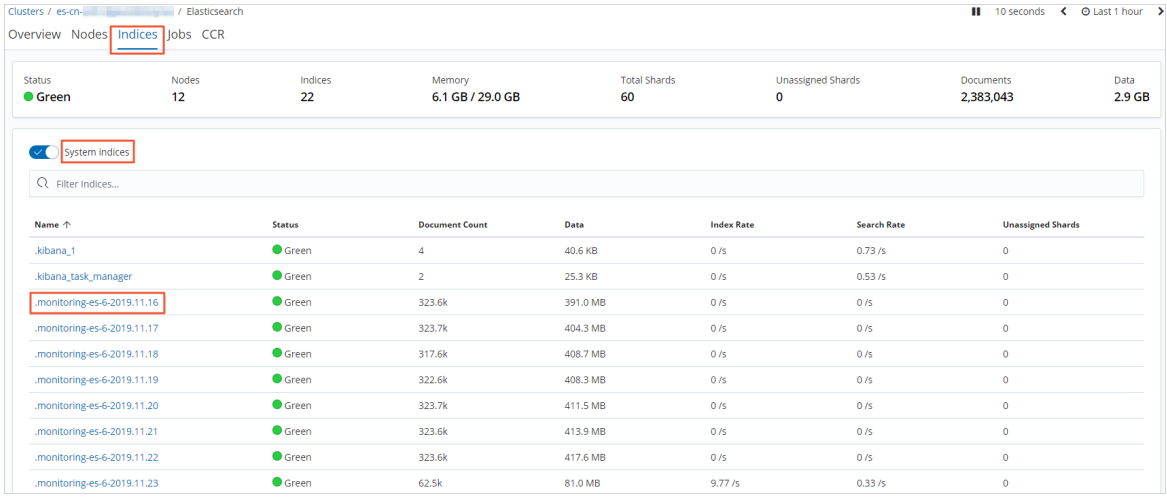
A `.monitoring-es-6-*` index stores information about the cluster status, cluster statistics, node statistics, and index statistics, which consumes a large amount of disk space.

Procedure

1. Log on to the Kibana console of your Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Monitoring**.
3. In the **Elasticsearch** section, click Indices.



4. On the **Indices** tab, turn on the **System indices** switch to query the storage space that is consumed by monitoring indexes.



5. In the left-side navigation pane, click **Dev Tools**.
6. On the **Console** tab, create a monitoring index. By default, the system retains the indexes that are created in the last seven days. Indexes that store the monitoring data, such as `.monitoring-es-6-*`, consume the storage space of your Elasticsearch cluster. The size of each index depends on the numbers of nodes and indexes that include system indexes in the cluster. You can use one or both of the following methods to prevent the indexes that store the monitoring data from consuming a large amount of disk space for your cluster.
   - Configure the index retention period

     ```
     PUT _cluster/settings
     {"persistent": {"xpack.monitoring.history.duration":"1d"}}
     ```

     You can specify the index retention period as required. The minimum retention period is one day.
   - Specify the indexes to be collected

     Call the API operation to specify the inclusion or exclusion list for index collection. This allows you to reduce the size of the `.monitoring-es-6-*` indexes. The following example shows how to create an exclusion list for collecting system indexes.

     ```
     PUT _cluster/settings
     {"persistent": {"xpack.monitoring.collection.indices": "*,-.*"}}
     ```

> **Note** The excluded indexes are not collected or displayed on the **Monitoring** page in the Kibana console. However, these indexes are listed in the index list retrieved by calling the `GET _cat/indices` operation. The status of open or close for the indexes is also displayed.

9.Query logs

This topic describes how to query logs of your Alibaba Cloud Elasticsearch cluster on the Cluster Log, Search Slow Log, Indexing Slow Log, and GC Log tabs of the Logs page. You can enter a keyword and set a time range to search for specific log entries.

query Elasticsearch log cluster log slow log GC log

Context

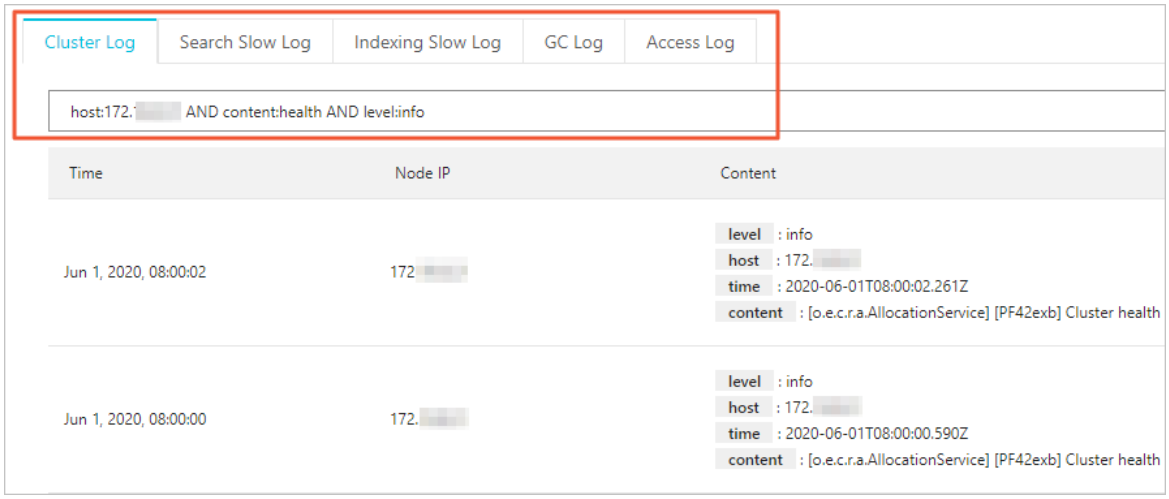
You can query log entries of up to seven consecutive days. By default, the log entries are displayed by time in descending order. The Lucene query syntax is supported. For more information, see Query string syntax.

> **Note** Each query returns up to 10,000 log entries. If the returned log entries do not contain the expected log data, you can specify a time range when you query the log data.

Procedure

This example searches the Elasticsearch cluster log. Log entries are returned if they meet all of these conditions: the `content` field contains the `health` keyword, the `level` field is `info`, and the `host` field is `172.16.xx.xx`.

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Logs**.
5. On the **Logs** page, click the **Cluster Log** tab.
6. Enter the query string in the search bar.



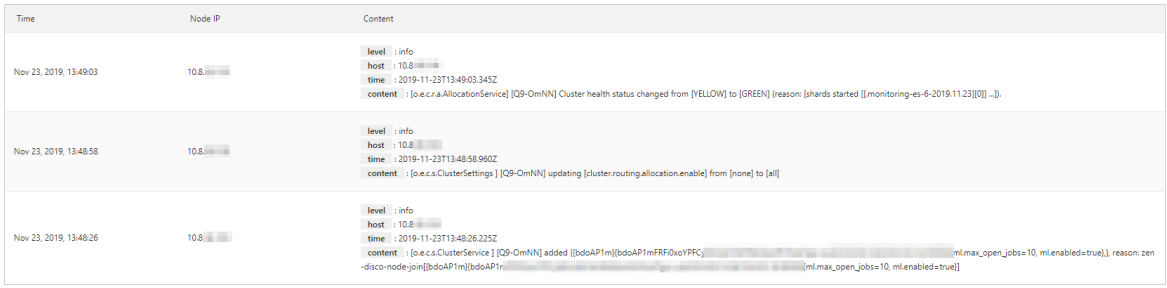
In this example, the query string is `host:172.16.xx.xx AND content:health AND level:info`.

> **Notice** `AND` in the query string must be in uppercase.

7. Specify a time range and click **Search**.

> **Notice**
> - If you do not select an end time, the current system time is specified.
> - If you do not select a start time, the default start time is the time that is one hour earlier than the end time.

After you click Search, Elasticsearch returns the log entries that match your query string and displays them on the **Logs** page. The returned log data contains the following information: **Time**, **Node IP**, and **Content**.



- **Time**: the time when the log entry was generated.
- **Node IP**: the IP address of the node in your Elasticsearch cluster.
- **Content**: includes `level`, `host`, `time`, and `content`.

Field	Description
`level`	The level of the log entry. Log levels include trace, debug, info, warn, and error. A GC log does not contain the `level` field.
`host`	The IP address of the node in your Elasticsearch cluster. To query the IP address of the node, log on to the Kibana console, click **Monitoring** in the left-side navigation pane, and then click **Nodes** in the **Elasticsearch** section. **Note** For more information about how to log on to the Kibana console, see Log on to the Kibana console.
`time`	The time when the log entry was generated.
`content`	The content of the log entry.

Configure slow logs

Elasticsearch logs only read and write operations that take between 5 to 10 seconds to complete as slow logs. This mechanism does not help identify problems, such as unbalanced loads, read and write exceptions, and slow data processing. After you create an Elasticsearch cluster, log on to the Kibana console and run the following command to reduce the timestamp precision of the log entry for capturing more logs:

> **Note** For more information about how to log on to the Kibana console, see Log on to the Kibana console.

```
PUT _settings
{
    "index.indexing.slowlog.threshold.index.debug" : "10ms",
    "index.indexing.slowlog.threshold.index.info" : "50ms",
    "index.indexing.slowlog.threshold.index.warn" : "100ms",
    "index.search.slowlog.threshold.fetch.debug" : "100ms",
    "index.search.slowlog.threshold.fetch.info" : "200ms",
    "index.search.slowlog.threshold.fetch.warn" : "500ms",
    "index.search.slowlog.threshold.query.debug" : "100ms",
    "index.search.slowlog.threshold.query.info" : "200ms",
    "index.search.slowlog.threshold.query.warn" : "1s"
}
```

After the configuration is complete, if the time to run a read or write task is exceeded, you can query related logs on the **Logs** page of your cluster.

10.Security

10.1. Configure a whitelist to access an Elasticsearch cluster over the Internet or a VPC

Configure the whitelists

This topic describes how to add the IP address of your host to the whitelist when you access your Alibaba Cloud Elasticsearch cluster over the Internet or a VPC.

configure whitelists configure a public IP address whitelist configure a VPC whitelist

Prerequisites

An Elasticsearch cluster is created. For more information, see Create an Elasticsearch cluster.

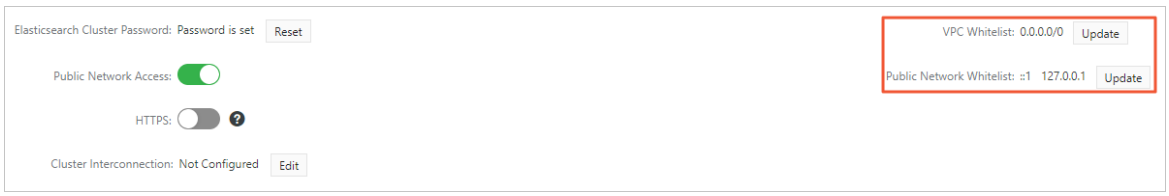
Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Security**.
5. In the **Network Settings** section of the page that appears, turn on the **Public Network Access** switch, which is turned off by default.

   > **Note** If Public Network Access is enabled or you only need to configure the VPC whitelist, skip this step.

   After this feature is enabled, the Public Network Access switch is in **green**. By default, the switch is in gray, which indicates that Public Network Access is disabled. To access your Elasticsearch cluster over the Internet, you must enable Public Network Access.

6. Click **Update** and enter the IP address that you want to add to the whitelist.



   You can enter both IP addresses and CIDR blocks in the public IP address and VPC whitelists. For example, enter `192.168.0.1` or `192.168.0.0/24`. Separate multiple IP addresses and CIDR blocks with commas (,). You can enter `127.0.0.1` to deny requests from all IPv4 addresses or enter `0.0.0.0/0` to allow requests from all IPv4 addresses. A whitelist can contain up to 300 IP addresses or CIDR blocks. The following table describes the differences between the two whitelists.

Category	Description
Public IP address whitelist	◦ If your Elasticsearch cluster is deployed in the China (Hangzhou) region, you can add IPv6 addresses to the whitelist. For example, enter `2401:b180:1000:24::5` or `2401:b180:1000::/48`. Enter `::1` to deny requests from all IPv6 addresses or `::/0` to allow requests from all IPv6 addresses. ◦ By default, requests from all public IP addresses are denied.
VPC whitelist	By default, requests from all IPv4 addresses within the VPC in which the Elasticsearch cluster resides are allowed.

7. Click **OK**.

10.2. Reset the access password for an Elasticsearch cluster

Reset the password

This topic describes how to reset the password of the elastic user that is used to access your Elasticsearch cluster. After you reset the password, if you use the elastic user to log on to your Elasticsearch cluster and Kibana console, you must use the new password of the elastic user.

reset the password to access an Elasticsearch cluster reset the password to access the Kibana console

Prerequisites

An Elasticsearch cluster is created. For more information, see Create an Elasticsearch cluster.

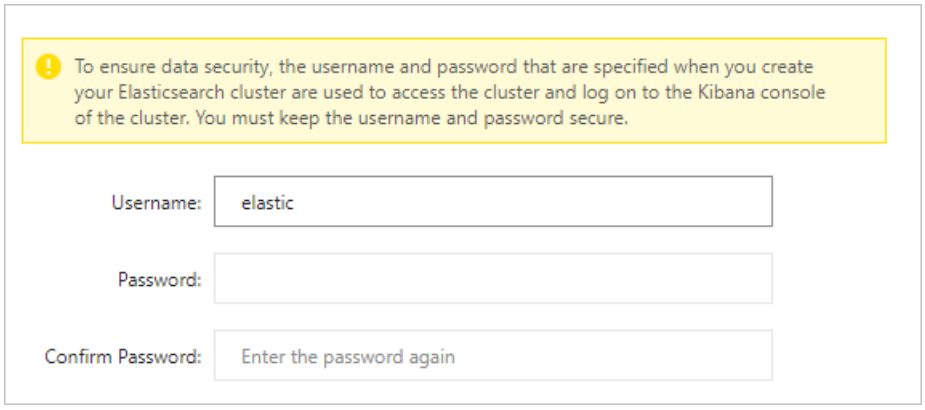
Context

Note the following points when you reset the password:

- This operation only resets the password of the elastic user. We recommend that you use a custom account that is granted the required permissions instead of the elastic account to log on to your Elasticsearch cluster.
- After you confirm the operations, the system does not restart the Elasticsearch cluster for the new password to take effect.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Security**.
5. In the **Network Settings** section of the page that appears, click **Reset** next to **Elasticsearch Cluster Password**.
6. In the **Reset** pane, enter the new password for **elastic** and confirm the password.



7. Click **OK**.
   After you reset the password, the new password takes effect in about five minutes.

10.3. Enable HTTPS

Enable HTTPS

Hypertext Transfer Protocol Secure (HTTPS) is a security-enhanced version of HTTP. HTTPS works with Secure Socket Layer (SSL) to ensure the security of data transmission. HTTPS uses HTTP for communications. SSL is used to encrypt the data. To ensure data security, we recommend that you enable HTTPS.

HTTPS for Elasticsearch Elasticsearch security

Prerequisites

- An Alibaba Cloud Elasticsearch cluster is created.

  For more information, see Create an Elasticsearch cluster.
- A client node is available.

  You can purchase a client node during the Elasticsearch cluster creation or upgrade. For more information, see Upgrade the configuration of a cluster.
- The code of the client that is used to access your Elasticsearch cluster is modified. Otherwise, you cannot use client programs to access your Elasticsearch cluster.

  Use the REST client of the open-source Elasticsearch as an example. After you enable HTTPS, you must include the `https` parameter in `HttpHost`, for example, `new HttpHost("es-cn-xxxxx.elasticsearch.aliyuncs.com", 9200, "https"));`. The sample code is as follows:

- The code before HTTPS is enabled

```
final CredentialsProvider credentialsProvider = new BasicCredentialsProvider();
    credentialsProvider.setCredentials(AuthScope.ANY,
      new UsernamePasswordCredentials("elastic", "Your password"));
RestClientBuilder restClientBuilder = RestClient.builder(
      new HttpHost("es-cn-xxxxx.elasticsearch.aliyuncs.com", 9200));
    RestClient restClient = restClientBuilder.setHttpClientConfigCallback(
      new RestClientBuilder.HttpClientConfigCallback() {
        @Override
        public HttpAsyncClientBuilder customizeHttpClient(HttpAsyncClientBuilder httpClientBuilder) {
          return httpClientBuilder.setDefaultCredentialsProvider(credentialsProvider);
        }
      }).build();
```

- The code after HTTPS is enabled

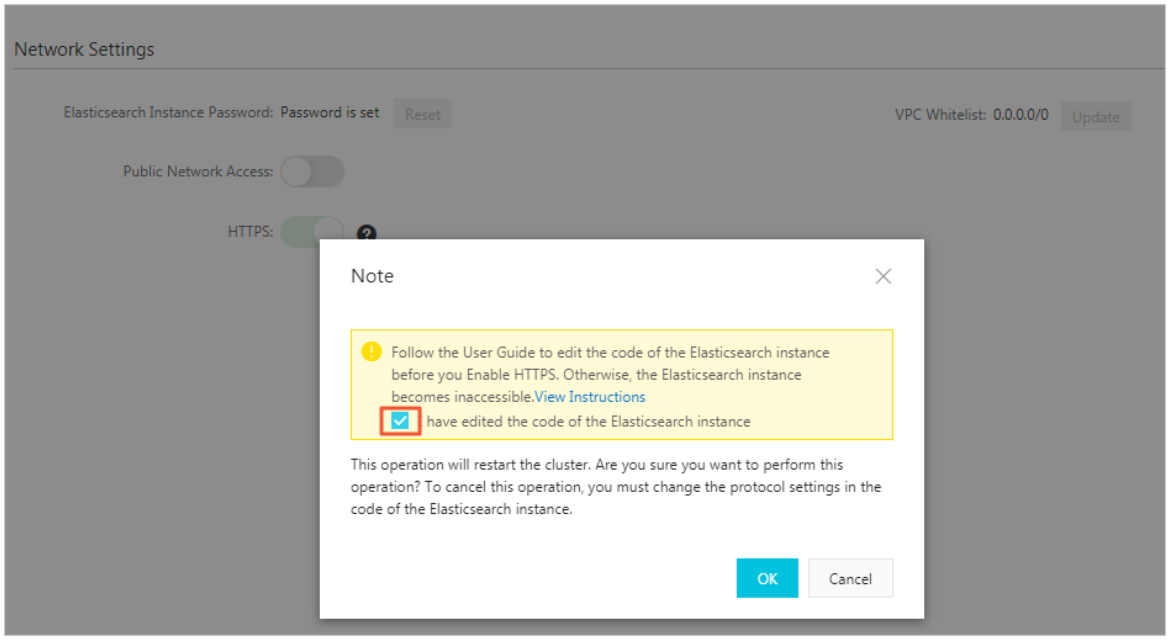
```
final CredentialsProvider credentialsProvider = new BasicCredentialsProvider();
    credentialsProvider.setCredentials(AuthScope.ANY,
      new UsernamePasswordCredentials("elastic", "Your password"));
RestClientBuilder restClientBuilder = RestClient.builder(
      new HttpHost("es-cn-xxxxx.elasticsearch.aliyuncs.com", 9200, "https"));
    RestClient restClient = restClientBuilder.setHttpClientConfigCallback(
      new RestClientBuilder.HttpClientConfigCallback() {
        @Override
        public HttpAsyncClientBuilder customizeHttpClient(HttpAsyncClientBuilder httpClientBuilder) {
          return httpClientBuilder.setDefaultCredentialsProvider(credentialsProvider);
        }
      }).build();
```

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Security**.
5. In the **Network Settings** section of the page that appears, turn on the **HTTPS** switch.

   Warning During the process of enabling or disabling HTTPS, the services that run in the cluster are interrupted and the Elasticsearch cluster is restarted. Make sure that the operation does not affect your services.

6. In the **Note** message, select the **I have created an HTTPS client** check box and click **OK**.



Note If you have not purchased client nodes, the system prompts you to purchase client nodes when you try to turn on the **HTTPS** switch. You must follow the instructions to purchase client nodes.

After you confirm the operation, the Elasticsearch cluster restarts. You can check the restart progress in the Tasks dialog box. After the Elasticsearch cluster is restarted, you can then access the cluster over HTTPS.

10.4. Connect two Elasticsearch clusters

This topic describes how to connect two Alibaba Cloud Elasticsearch clusters. After the clusters are connected, you can search for data across these clusters. Elasticsearch clusters are isolated from each other by default, which ensures data security.

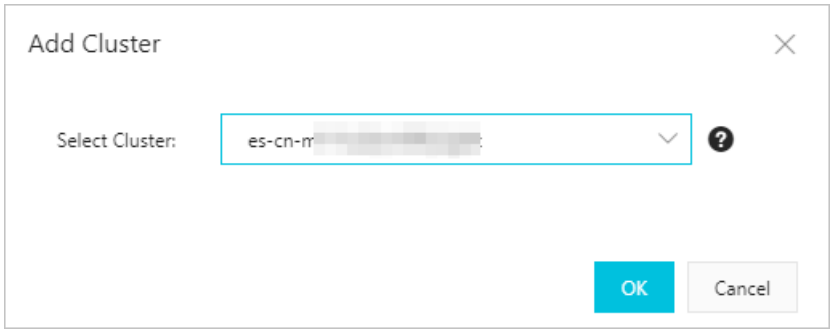
Prerequisites

- The two Elasticsearch clusters are of the same Elasticsearch version.
- The two Elasticsearch clusters belong to the same Alibaba Cloud account.
- The two Elasticsearch clusters are deployed in the same Virtual Private Cloud (VPC).
- The two Elasticsearch clusters use the same deployment method.

Configure the connection between two Elasticsearch clusters

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the cluster details page, click **Security**.
5. On the page that appears, click **Edit** on the right side of **Cluster Interconnection**.
6. In the **Edit Configuration** pane, click **Add Cluster**.

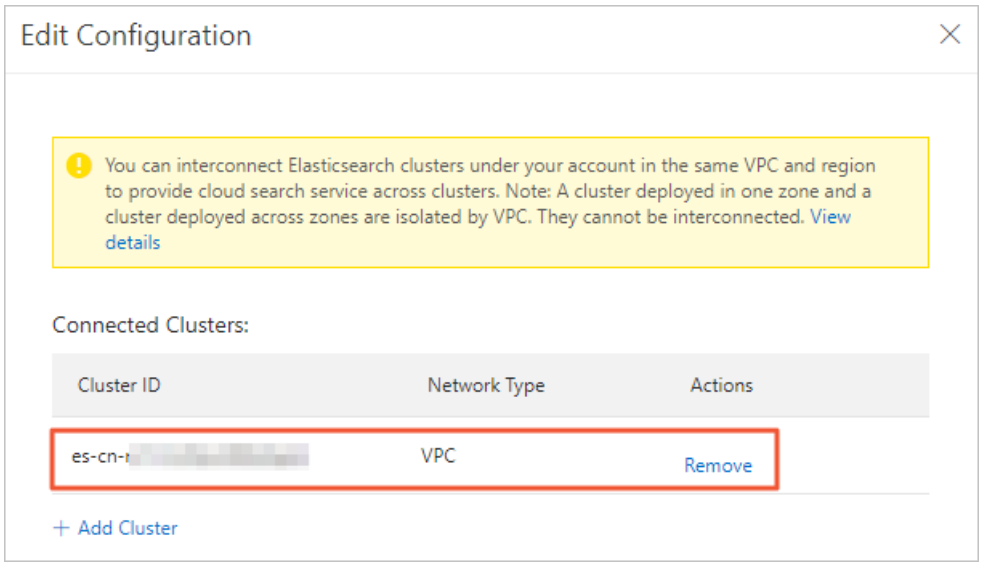
7. In the **Add Cluster** dialog box, select the ID of the remote Elasticsearch cluster that you want to connect.



If one or more Elasticsearch clusters meet the prerequisites, you can select these clusters in the **Add Cluster** dialog box.

> **Notice**
> - RAM users of an Alibaba Cloud account can query all Elasticsearch clusters that belong to the Alibaba Cloud account only when RAM users are granted the ListInstance permission. For more information, see Resource types.
> - After you connect the current Elasticsearch cluster to a remote cluster, you can view the ID of the current cluster on the **Cluster Interconnection** page of the remote cluster. This indicates that the communication between the two clusters is bidirectional.

8. Click **OK**.
After you add a cluster, you can find the added cluster in the **Connected Clusters** section of the **Edit Configuration** pane.



> **Note** If you no longer require the added cluster, click **Remove** to remove the cluster.

After you complete the configuration, you must perform the operations described in Configure the cross-cluster search feature to search for data of the remote Elasticsearch cluster in the current Elasticsearch cluster.

Configure the cross-cluster search feature

1. Log on to the Kibana console of the remote Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. Create an index, add a document to the index, and insert data into the document in the remote Elasticsearch cluster.
   - Create an index

```
PUT /twitter
{
  "settings" : {
    "index" : {
      "number_of_shards" : 3,
      "number_of_replicas" : 2
    }
  }
}
```

   - Create a document and insert data

```
POST twitter/doc/
{
  "user" : "kimchy",
  "post_date" : "2009-11-15T14:12:12",
  "message" : "trying out Elasticsearch"
}
```

   > **Note** The index and document are used to test the cross-cluster search feature.

3. Log on to the Kibana console of the current Elasticsearch cluster. For more information, see Log on to the Kibana console.
4. Use one of the following methods to configure the cross-cluster search feature in the current Elasticsearch cluster. The following methods use an Elasticsearch V6.7 cluster as an example. The methods that are used to configure the cross-cluster search in other Elasticsearch versions are similar. For more information, see Cross-cluster search in open-source Elasticsearch 7.X, Cross-cluster search in open-source Elasticsearch 6.3, and Cross-cluster search in open-source Elasticsearch 5.5.
   - Method 1: Use the internal endpoint of the remote Elasticsearch cluster

```
PUT _cluster/settings
{
 "persistent": {
  "cluster": {
   "remote": {
    "cluster_one": {
     "seeds": [
      "es-cn-o4xxxxxxxxxxx4f1.elasticsearch.aliyuncs.com:9300"
     ]
    }
   }
  }
 }
}
```

- Method 2: Use the IP addresses of the nodes in the remote Elasticsearch cluster

```
PUT _cluster/settings
{
 "persistent": {
  "cluster": {
   "remote": {
    "cluster_one": {
     "seeds": [
      "10.8.xx.xx:9300",
      "10.8.xx.xx:9300",
      "10.8.xx.xx:9300"
     ]
    }
   }
  }
 }
}
```

> **Notice**
>
> - If the Elasticsearch cluster is deployed in a single zone, you can use both Method 1 and Method 2. If the Elasticsearch cluster is deployed across zones, you can use only Method 2. Multiple remote Elasticsearch clusters can be connected to the current Elasticsearch cluster.
> - You can query the index data of the remote Elasticsearch cluster in the current Elasticsearch cluster. This only applies if you have configured the domain name or node IP addresses of the remote Elasticsearch cluster in the current Elasticsearch cluster for cross-cluster search. You cannot run similar commands in the remote Elasticsearch cluster to search for data of the current Elasticsearch cluster. To search for data of the current Elasticsearch cluster in the remote Elasticsearch cluster, you must add the domain name or node IP addresses of the current cluster in the remote cluster.

5. Run the following command to check whether the cross-cluster search feature is successfully configured:

```
POST /cluster_one:twitter/doc/_search
{
 "query": {
  "match_all": {}
 }
}
```

If the cross-cluster search feature is successfully configured, the following result is returned:

```
{
 "took" : 78,
 "timed_out" : false,
 "_shards" : {
  "total" : 3,
  "successful" : 3,
  "skipped" : 0,
  "failed" : 0
 },
 "_clusters" : {
  "total" : 1,
  "successful" : 1,
  "skipped" : 0
 },
 "hits" : {
  "total" : 1,
  "max_score" : 1.0,
  "hits" : [
   {
    "_index" : "cluster_one:twitter",
    "_type" : "doc",
    "_id" : "qudxxxxxxxxx_7ie6J",
    "_score" : 1.0,
    "_source" : {
     "user" : "kimchy",
     "post_date" : "2009-11-15T14:12:12",
     "message" : "trying out Elasticsearch"
    }
   }
  ]
 }
}
```

11.Data backup

11.1. View the snapshot feature

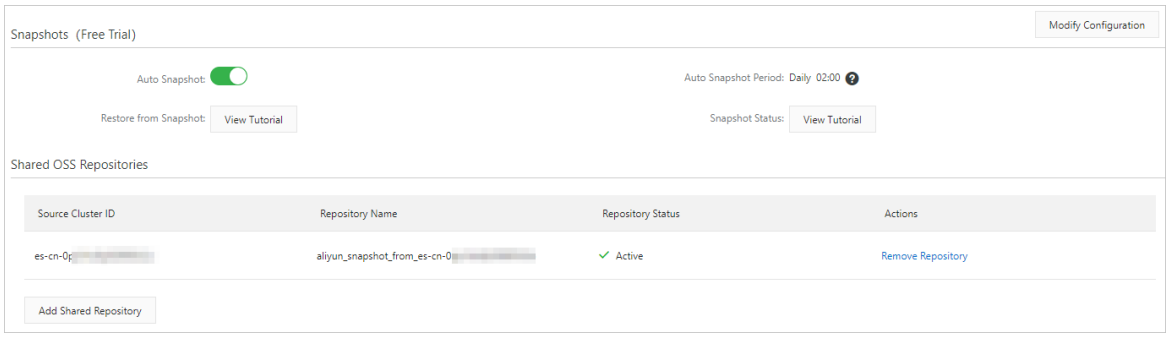
View the snapshot feature of Elasticsearch

This topic describes the snapshot feature of Alibaba Cloud Elasticsearch. This feature allows you to enable the auto snapshot feature, set a snapshot creation interval, and add shared OSS repositories for an Elasticsearch cluster.

Elasticsearch snapshot

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Snapshots**. The **Snapshots** page appears.



The **Snapshots** page contains the **Snapshots (Free Trial)** and **Shared OSS Repositories** sections. The following tables describe the parameters in these sections.

Snapshots (Free Trial) section

Parameter	Description
Auto Snapshot	If the **Auto Snapshot** switch is green, auto snapshot is enabled. By default, the switch is turned off.
Auto Snapshot Period	If auto snapshot is disabled, the system displays the "**You must enable auto snapshot first.**" message. If auto snapshot is enabled, the system time of the region where your Elasticsearch cluster resides is used to create snapshots. **Notice** Do not perform snapshot operations when the system is creating snapshots.

Parameter	Description
Modify Configuration	If auto snapshot is enabled, you can click **Modify Configuration** in the upper-right corner to open the **Auto Snapshot Configuration** pane and then set **Frequency**. Valid values of **Frequency**: ○ **Every 30 Minutes**: The system creates snapshots at 30-minute intervals. ○ **Daily**: The system creates snapshots every day. You can customize the creation time. ○ **Custom**: The system creates snapshots based on the interval you specify. **Notice** The system stores only snapshots that were created in the last five to seven days.
Restore from Snapshot	Click **View Tutorial** to learn how to restore data from a snapshot.
Snapshot Status	Click **View Tutorial** to learn more information about snapshot status.

Shared OSS Repositories section

Parameter	Description
Source Cluster ID	The ID of the Elasticsearch cluster to which the shared OSS repository belongs.
Repository Name	The name of the shared OSS repository.
Repository Status	The status of the shared OSS repository. Valid values: ○ **Active**: indicates that the repository is available. ○ **Inactive**: indicates that the specified cluster or repository does not exist.
Actions	The **Remove Repository** feature is provided.
Add Shared Repository	Click **Add Shared Repository** to add a shared OSS repository. For more information, see Configure a shared OSS repository. **Notice** If you add a shared OSS repository for the first time, the **Add Shared Repository** button does not appear. You need to click **Add Now** to add a shared OSS repository.

11.2. Create automatic snapshots and restore data from automatic snapshots

Create automatic snapshots and restore data from automatic snapshots

Alibaba Cloud Elasticsearch provides the auto snapshot feature. This feature allows you to specify a snapshot creation interval and specific creation time. After you specify them, the system automatically creates snapshots based on the interval and time to ensure data security. These snapshots are called automatic snapshots. You can then quickly restore data from the snapshots to the Elasticsearch cluster in which the snapshots are created. This topic describes how to enable the auto snapshot feature and restore data from automatic snapshots.

Prerequisites

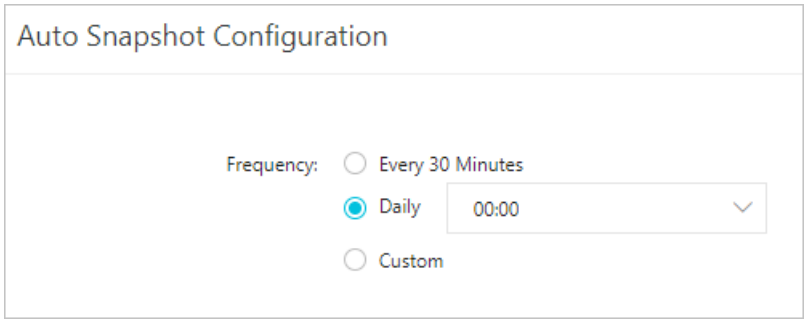
An Alibaba Cloud Elasticsearch cluster is created. For more information, see Create an Elasticsearch cluster.

Enable the auto snapshot feature

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane of the page that appears, click **Snapshots**.
5. In the **Snapshots (Free Trial)** section, turn on Auto Snapshot.
6. Click **Modify Configuration** on the right side of Snapshots (Free Trial).

   > **Notice** To ensure system security, the value of **Frequency** for Elasticsearch clusters of the Advanced Edition is generated by the system and cannot be changed.

7. In the **Auto Snapshot Configuration** pane, set **Frequency**.



Frequency	Description
Every 30 Minutes	The system creates snapshots at 30-minute intervals.
Daily	The system creates snapshots every day. You can customize the specific creation time.

Frequency	Description
Custom	The system creates snapshots based on the interval and time you specify.

> **Notice** The system uses the system time of the region where your Elasticsearch cluster resides to create snapshots.

8. Click **OK**.

Restore data from automatic snapshots

If you enable the auto snapshot feature for an Elasticsearch cluster, the system creates snapshots for the cluster based on the snapshot creation interval and time that you specified. You can call the snapshot operation to restore data to the cluster from the created snapshots.

> **Note**
>
> - The first snapshot is a full copy of the data in an Elasticsearch cluster. Subsequent snapshots store only incremental data. Therefore, it requires longer time to create the first snapshot than a subsequent snapshot.
> - Snapshots store only index data. The following information of your Elasticsearch cluster is not stored in snapshots: monitoring data (such as indexes whose names start with `.monitoring` or `.security_audit`), metadata, translogs, configurations, software packages, built-in and custom plug-ins, and logs.
> - You can restore data from automatic snapshots only to the Elasticsearch cluster in which the snapshots are created.
> - An automatic snapshot repository is created when the first snapshot is created.
> - If you restore indexes whose names start with `.` from snapshots, you may fail to access Kibana. Such indexes are system indexes. We recommend that you do not restore these indexes from snapshots.

1. Log on to the Kibana console of your Elasticsearch cluster.For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Dev Tools**. On the **Console** tab of the page that appears, run a command to perform a specific operation.
   - Query all snapshot repositories

     Run the `GET _snapshot` command to query all snapshot repositories.

     If the command is successfully executed, the following result is returned:

```
{
 "aliyun_auto_snapshot": {
  "type": "oss",
  "settings": {
   "compress": "true",
   "base_path": "xxxx",
   "endpoint": "xxxx"
  }
 }
}
```

Parameter	Description
aliyun_auto_snapshot	The name of the repository.
type	The storage where snapshots are stored. The value oss indicates that snapshots are stored in Object Storage Service (OSS).
compress	Indicates whether compression is used. The value true indicates that the metadata of indexes is compressed during snapshot creation.
base_path	The location of snapshots in OSS.
endpoint	The endpoint of the OSS bucket that stores the snapshots.

- Query all snapshots

  Run the GET _snapshot/aliyun_auto_snapshot/_all command to query all snapshots in the aliyun_auto_snapshot repository.

  If the command is successfully executed, the following result is returned:

```
{
 "snapshots": [
  {
   "snapshot": "es-cn-abcdefghij****_20180627091600",
   "uuid": "MMRniVLPRAiawSCm8D****",
   "version_id": 5050399,
   "version": "5.5.3",
   "indices": [
    "index_1",
    ".security",
    ".kibana"
   ],
   "state": "SUCCESS",
   "start_time": "2018-06-27T01:16:01.009Z",
   "start_time_in_millis": 1530062161009,
   "end_time": "2018-06-27T01:16:05.632Z",
   "end_time_in_millis": 1530062165632,
   "duration_in_millis": 4623,
   "failures": [],
   "shards": {
    "total": 12,
    "failed": 0,
    "successful": 12
   }
  }
 ]
}
```


Notice The system uses the system time of the current region to create snapshots. However, the time in the returned result is in UTC. A time zone difference exists between the system time and the time in the returned result. You can convert the time in the returned result based on the time zone difference. For example, the time zone difference between the system time of the China (Beijing) region and UTC is 8 hours. In this case, the system time of the China (Beijing) region is UTC+0080.

The auto snapshot feature also supports the following default parameters that are not displayed.

Parameter	Description
`max_snapshot_bytes_per_sec:40mb`	The maximum speed for snapshot creation on a single node is 40 MB/s.

Parameter	Description
max_restore_bytes_per_sec: 40mb	The maximum speed for data restoration on a single node is 40 MB/s.
chunk_size: Max 1Gb	During snapshot creation, a large index is divided into multiple parts. The maximum size of each part is 1 GB.

- Restore indexes from a snapshot

  You can call the `_restore` operation to restore indexes from snapshots.

  - Run the following command to restore all indexes from a specific snapshot that is stored in the `aliyun_auto_snapshot` repository. The restoration task is executed at the backend.

    ```
    POST _snapshot/aliyun_auto_snapshot/<snapshot>/_restore
    ```

    Replace `<snapshot>` with the name of the specific snapshot, such as `es-cn-abcdefghij****_20180627091600`.

  - Run the following command to restore all indexes from a specific snapshot that is stored in the `aliyun_auto_snapshot` repository. Then, wait until the restoration task is completed.

    The `restore` operation asynchronously runs restoration tasks. An Elasticsearch cluster will immediately return a response if the restore operation is executable. The restoration task is executed at the backend. You can set the `wait_for_completion` parameter to enable the cluster to return a response only after the restoration task is completed.

    ```
    POST _snapshot/aliyun_auto_snapshot/<snapshot>/_restore?wait_for_completion=true
    ```

    Replace `<snapshot>` with the name of the specific snapshot, such as `es-cn-abcdefghij****_20180627091600`.

- Run the following command to restore specified indexes from a specific snapshot that is stored in the `aliyun_auto_snapshot` repository, and rename the restored indexes. The restoration task is executed at the backend.

```
POST _snapshot/aliyun_auto_snapshot/<snapshot>/_restore
{
"indices": "index_1",
"rename_pattern": "index_(.+)",
"rename_replacement": "restored_index_$1"
}
```

Parameter	Description
`<snapshot>`	Replace it with the name of the specific snapshot, such as `es-cn-abcdefghijklmn_20180627091600`.
`indices`	The name of the index you want to restore.
`rename_pattern`	Optional. This parameter specifies the regular expression that is used to match the name of the index you want to restore.
`rename_replacement`	Optional. This parameter specifies the regular expression that is used to rename a matched index.

11.3. Query snapshot status

Query Elasticsearch snapshot status

This topic describes how to query the status of snapshots that are automatically created on your Alibaba Cloud Elasticsearch cluster to obtain the creation progress of snapshots in real time.

Elasticsearch snapshot status Elasticsearch auto snapshot Elasticsearch snapshot
After you enable the auto snapshot feature, log on to the Kibana console of your Elasticsearch cluster. In the left-side navigation pane, click **Dev Tools**. On the page that appears, click the **Console** tab. On this tab, call the *snapshot* operation to query the status of automatically created snapshots.

> **Note**
> - For more information about how to enable the auto snapshot feature, see Create automatic snapshots and restore data from automatic snapshots.
> - For more information about how to log on to the Kibana console, see Log on to the Kibana console.

Query all snapshots

Run the following command to query information about all snapshots stored in the `aliyun_auto_snapshot` repository.

```
GET _snapshot/aliyun_auto_snapshot/_all
```

If the command is executed successfully, the following result is returned:


```
{
 "snapshots": [
  {
   "snapshot": "es-cn-abxxxxxxxxxlmn_20180628092236",
   "uuid": "n7YxxxxxxxxxxxdA",
   "version_id": 5050399,
   "version": "5.5.3",
   "indices": [
    ".kibana"
   ],
   "state": "SUCCESS",
   "start_time": "2018-06-28T01:22:39.609Z",
   "start_time_in_millis": 1530148959609,
   "end_time": "2018-06-28T01:22:39.923Z",
   "end_time_in_millis": 1530148959923,
   "duration_in_millis": 314,
   "failures": [],
   "shards": {
    "total": 1,
    "failed": 0,
    "successful": 1
   }
  },
  {
   "snapshot": "es-cn-abxxxxxxxxxmn_20180628092500",
   "uuid": "frdxxxxxxxxxxxKLA",
   "version_id": 5050399,
   "version": "5.5.3",
   "indices": [
    ".kibana"
   ],
   "state": "SUCCESS",
   "start_time": "2018-06-28T01:25:00.764Z",
   "start_time_in_millis": 1530149100764,
   "end_time": "2018-06-28T01:25:01.482Z",
   "end_time_in_millis": 1530149101482,
   "duration_in_millis": 718,
   "failures": [],
   "shards": {
```

```
    "total": 1,
    "failed": 0,
    "successful": 1
   }
  }
 ]
}
```

`state` : indicates the status of a snapshot. A snapshot can be in one of the following states.

State	Description
`IN_PROGRESS`	The snapshot is being created.
`SUCCESS`	The snapshot is created, and all shards are stored.
`FAILED`	The snapshot fails to be created because some shards cannot be stored.
`PARTIAL`	The snapshot is created, but at least one shard fails to be stored.
`INCOMPATIBLE`	The snapshot version is incompatible with the cluster version.

Query a specified snapshot

Run the following command to query information about a specified snapshot stored in the `aliyun_auto_snapshot` repository.

```
GET _snapshot/aliyun_auto_snapshot/<snapshot>/_status
```

Replace `<snapshot>` with the name of the snapshot, such as `es-cn-abxxxxxxxxxxlmn_20180628092236` . You can use the command in Query all snapshots to query the name of a snapshot.

If the command is executed successfully, the following result is returned:

```
{
 "snapshots": [
  {
   "snapshot": "es-cn-abxxxxxxxxxxlmn_20180628092236",
   "repository": "aliyun_auto_snapshot",
   "uuid": "n7YxxxxxxxxxxxydA",
   "state": "SUCCESS",
   "shards_stats": {
    "initializing": 0,
    "started": 0,
    "finalizing": 0
```

```
  "finalizing": 0,
  "done": 1,
  "failed": 0,
  "total": 1
 },
 "stats": {
  "number_of_files": 4,
  "processed_files": 4,
  "total_size_in_bytes": 3296,
  "processed_size_in_bytes": 3296,
  "start_time_in_millis": 1530148959688,
  "time_in_millis": 77
 },
 "indices": {
  ".kibana": {
   "shards_stats": {
    "initializing": 0,
    "started": 0,
    "finalizing": 0,
    "done": 1,
    "failed": 0,
    "total": 1
   },
   "stats": {
    "number_of_files": 4,
    "processed_files": 4,
    "total_size_in_bytes": 3296,
    "processed_size_in_bytes": 3296,
    "start_time_in_millis": 1530148959688,
    "time_in_millis": 77
   },
   "shards": {
    "0": {
     "stage": "DONE",
     "stats": {
      "number_of_files": 4,
      "processed_files": 4,
      "total_size_in_bytes": 3296,
      "processed_size_in_bytes": 3296,
      "start_time_in_millis": 1530148959688,
      "time_in_millis": 77
```

```
      }
     }
    }
   }
  }
 }
]
}
```

11.4. Commands for creating snapshots and restoring data

Commands for creating snapshots and restoring data

You can call the snapshot operation to back up or restore data for your Alibaba Cloud Elasticsearch cluster. The snapshot operation retrieves the status and data of your cluster and then stores them to a shared repository.

Elasticsearch data backup snapshot operation Elasticsearch data restoration

Precautions

- Snapshots store only index data. The following information of your Elasticsearch cluster is not stored in snapshots: monitoring data (such as indexes with the prefix .monitoring or .security_audit), metadata, translog files, configurations, software packages, built-in and custom plug-ins, and logs.
- This topic uses the following markers to provide descriptions for code: <1>, <2>, and <3>. Remove these markers when you run the code.
- You can run the code provided in this topic in the Kibana console of your Elasticsearch cluster. For more information, see Log on to the Kibana console.
- Some of the content in this topic is referenced from elasticsearch-repository-oss.

Prerequisites

Object Storage Service (OSS) is activated and an OSS bucket is created.

> **Notice** The storage class of the OSS bucket must be Standard. Elasticsearch does not support the Archive storage class. The OSS bucket must reside in the same region as your Elasticsearch cluster.

For more information, see Activate OSS and Create buckets.

Create a repository

```
PUT _snapshot/my_backup
{
  "type": "oss",
  "settings": {
    "endpoint": "http://oss-cn-hangzhou-internal.aliyuncs.com", <1>
    "access_key_id": "xxxx",
    "secret_access_key": "xxxxxx",
    "bucket": "xxxxxx", <2>
    "compress": true,
    "base_path": "snapshot/" <3>
  }
}
```

- <1>: The `endpoint` parameter specifies the internal endpoint of the OSS bucket. For more information, see Regions and endpoints.
- <2>: The bucket parameter specifies the name of an OSS bucket that has been created.
- <3>: The base_path parameter specifies the path of the repository. The default value is the root directory.

Set the size of each part

When you upload a large amount of data to an OSS bucket, you can use the `chunk_size` parameter to set the size of each part. This allows you to upload the data in multiple parts. Example:

```
POST _snapshot/my_backup/ <1>
{
  "type": "oss",
  "settings": {
    "endpoint": "http://oss-cn-hangzhou-internal.aliyuncs.com",
    "access_key_id": "xxxx",
    "secret_access_key": "xxxxxx",
    "bucket": "xxxxxx",
    "chunk_size": "500mb",
    "base_path": "snapshot/" <2>
  }
}
```

- <1>: Use the POST method instead of the PUT method. The POST method updates repository settings.
- <2>: The base_path parameter specifies the path of the repository. The default value is the root directory.

Query repository information

```
GET _snapshot
```

You can also use the `GET _snapshot/my_backup` command to query the information of a specified repository.

Create snapshots

The following command is a basic command that is used to create snapshots:

```
PUT _snapshot/my_backup/snapshot_1
```

This command creates the `snapshot_1` snapshot for all open indexes. The snapshot is stored in the `my_backup` repository. After you run the command, the system immediately returns a response while the snapshot is created at the backend.

If you want the system to return a response after it creates the snapshot, add the `wait_for_completion` parameter as follows:

```
PUT _snapshot/my_backup/snapshot_1? wait_for_completion=true
```

After you run the command, the system does not return a response until the snapshot is created. If the size of the index is large, the response is returned after a longer period of time.

> **Note** The first snapshot is a full copy of the data in a cluster. Subsequent snapshots only store incremental data. Therefore, when you create subsequent snapshots, the system only needs to add data to or delete data from the snapshots. This means that it requires less time to create a subsequent snapshot than the first snapshot.

Create a snapshot for specified indexes

By default, a snapshot contains all open indexes. For Kibana, you may want to ignore all diagnostic indexes (the `.kibana` indexes) when you create a snapshot because of limited disk space. To create a snapshot for specified indexes, run the following command:

> **Notice** A repository stores multiple snapshots. Each snapshot is a copy of all indexes, specified indexes, or a single index in a cluster. When you create a snapshot, make sure that the snapshot name is unique.

```
PUT _snapshot/my_backup/snapshot_2
{
  "indices": "index_1,index_2"
}
```

The preceding command creates a snapshot only for the `index1` and `index2` indexes.

Query snapshot information

In some cases, you may need to query snapshot information. For example, a snapshot name containing a date is hard to remember, such as `backup_2014_10_28`.

To query the information of a snapshot, send a `GET` request that contains both the repository name and snapshot name. Example:

```
GET _snapshot/my_backup/snapshot_2
```

The following response contains detailed information of the snapshot:

```
{
"snapshots": [
 {
  "snapshot": "snapshot_2",
  "indices": [
   ".marvel_2014_28_10",
   "index1",
   "index2"
  ],
  "state": "SUCCESS",
  "start_time": "2014-09-02T13:01:43.115Z",
  "start_time_in_millis": 1409662903115,
  "end_time": "2014-09-02T13:01:43.439Z",
  "end_time_in_millis": 1409662903439,
  "duration_in_millis": 324,
  "failures": [],
  "shards": {
   "total": 10,
   "failed": 0,
   "successful": 10
  }
 }
]
}
```

You can replace the snapshot name in the preceding command with `_all` to query all snapshots in the repository. Example:

```
GET _snapshot/my_backup/_all
```

Monitor snapshot creation progress

The `wait_for_completion` parameter provides a simple method for you to monitor the progress of a snapshot creation task. However, this parameter is not suitable for snapshot creation tasks of medium-size Elasticsearch clusters. You can use one of the following methods to query detailed information about a snapshot:

- Send a `GET` request with the snapshot name specified. Example:

```
GET _snapshot/my_backup/snapshot_3
```

If the system is still creating the snapshot when you run the preceding command, the information of the creation task is returned, such as the time when the snapshot creation task started and the duration.

> **Notice** The preceding command shares a thread pool with the command used to create a snapshot. Therefore, if you create a snapshot for large shards, the preceding command has to wait until the resources that are used by the snapshot creation command in the thread pool are released.

- Call the _status operation to query the snapshot status.

```
{
"snapshots": [
 {
  "snapshot": "snapshot_3",
  "repository": "my_backup",
  "state": "IN_PROGRESS", <1>
  "shards_stats": {
   "initializing": 0,
   "started": 1, <2>
   "finalizing": 0,
   "done": 4,
   "failed": 0,
   "total": 5
  },
  "stats": {
   "number_of_files": 5,
   "processed_files": 5,
   "total_size_in_bytes": 1792,
   "processed_size_in_bytes": 1792,
   "start_time_in_millis": 1409663054859,
   "time_in_millis": 64
  },
  "indices": {
   "index_3": {
    "shards_stats": {
```

```
      "initializing": 0,
      "started": 0,
      "finalizing": 0,
      "done": 5,
      "failed": 0,
      "total": 5
    },
    "stats": {
      "number_of_files": 5,
      "processed_files": 5,
      "total_size_in_bytes": 1792,
      "processed_size_in_bytes": 1792,
      "start_time_in_millis": 1409663054859,
      "time_in_millis": 64
    },
    "shards": {
      "0": {
        "stage": "DONE",
        "stats": {
          "number_of_files": 1,
          "processed_files": 1,
          "total_size_in_bytes": 514,
          "processed_size_in_bytes": 514,
          "start_time_in_millis": 1409663054862,
          "time_in_millis": 22
        }
      },
      ...
```

- <1>: The status of the snapshot. If a snapshot is being created, the value of the field is `IN_PROGRESS`.

- <2>: The number of shards that are being transmitted. If value 1 is returned, a shard of the snapshot is being transmitted, and the other four shards have been transmitted.

  The value of the `shards_stats` parameter contains the status of the snapshot. It also contains statistics about each index and shard. This parameter allows you to learn the detailed information of the snapshot creation progress. A shard can be in one of the following states:

  - `INITIALIZING`: The shard is verifying the status of the cluster to check whether the shard can be stored in a snapshot. In most cases, this process is fast.
  - `STARTED`: Data is being transmitted to the repository.
  - `FINALIZING`: The data transmission process is complete. The shard is sending snapshot metadata.
  - `DONE`: The snapshot is created.
  - `FAILED`: An error occurred during the snapshot creation. The shard, index, or snapshot cannot be processed. You can view logs for more information.

Use a snapshot to migrate data

To use a snapshot to migrate data from an Elasticsearch cluster to another, follow these steps:

1. Back up a snapshot to OSS.
2. Create a snapshot repository on the destination cluster. The repository must use the OSS bucket that stores the snapshot.
3. Set the `base_path` parameter to the path of the snapshot.
4. Run the data restoration command on the destination cluster.

Cancel a snapshot

To cancel a snapshot, run the following command when the snapshot is being created:

```
DELETE _snapshot/my_backup/snapshot_3
```

This command stops the snapshot creation process and deletes the snapshot that is being created from the repository.

Restore indexes from a snapshot

To restore indexes from a snapshot, run the command that is used in Create a repository on the Elasticsearch cluster that you want to restore the indexes. You can use one of the following methods to restore indexes from a snapshot:

- To restore indexes from a specified snapshot, append the `_restore` parameter to the snapshot name in the command to run. Example:

  ```
  POST _snapshot/my_backup/snapshot_1/_restore
  ```

  After you run this command, the system restores all indexes in the snapshot. For example, if the `snapshot_1` snapshot contains five indexes, all these indexes are restored to the Elasticsearch cluster. You can also reference Create a snapshot for specified indexes and specify the indexes that you want to restore.

- Restore specified indexes and rename the indexes. If you only want to verify or process the data in indexes and do not need to overwrite the data, use this method to restore the indexes.

```
POST /_snapshot/my_backup/snapshot_1/_restore
{
 "indices": "index_1", <1>
 "rename_pattern": "index_(.+)", <2>
 "rename_replacement": "restored_index_$1" <3>
}
```

  In this example, the `index_1` index is restored to your Elasticsearch cluster and renamed `restored_index_1`.

  - <1>: The system only restores the `index_1` index from the snapshot.
  - <2>: The system searches for the index that is being restored and matches the index name with the provided pattern.
  - <3>: The system renames the matched index.

  If you want the system to return a response after it restores the index, add the `wait_for_completion` parameter as follows:

```
POST _snapshot/my_backup/snapshot_1/_restore? wait_for_completion=true
```

  After you call the _restore operation, the system immediately returns a response and restores the index at the backend.

Monitor index restoration progress

> **Note** Restoring data from a repository applies the existing restoration mechanism in Elasticsearch. Restoring shards from a repository is the same as restoring data from a node.

You can call the recovery operation to monitor the progress of an index restoration task.

- Monitor a specified index that is being restored.

```
GET restored_index_3/_recovery
```

  The recovery operation is a general-purpose operation that shows the status of the shards that are being transmitted to your cluster.

- Monitor all indexes on the cluster. This may include shards that are irrelevant to the restoration process.

```
GET /_recovery/
```

  The sample response is as follows:

```
{
"restored_index_3" : {
 "shards" : [ {
```

```
"id" : 0,
"type" : "snapshot", <1>
"stage" : "index",
"primary" : true,
"start_time" : "2014-02-24T12:15:59.716",
"stop_time" : 0,
"total_time_in_millis" : 175576,
"source" : { <2>
 "repository" : "my_backup",
 "snapshot" : "snapshot_3",
 "index" : "restored_index_3"
},
"target" : {
 "id" : "ryqJ5lO5S4-lSFbGntkEkg",
 "hostname" : "my.fqdn",
 "ip" : "10.0. **.**",
 "name" : "my_es_node"
},
"index" : {
 "files" : {
  "total" : 73,
  "reused" : 0,
  "recovered" : 69,
  "percent" : "94.5%" <3>
 },
 "bytes" : {
  "total" : 79063092,
  "reused" : 0,
  "recovered" : 68891939,
  "percent" : "87.1%"
 },
 "total_time_in_millis" : 0
},
"translog" : {
 "recovered" : 0,
 "total_time_in_millis" : 0
},
"start" : {
 "check_index_time" : 0,
 "total_time_in_millis" : 0
}
```

```
}]
}
}
```

- <1>: The `type` parameter indicates the type of restoration. The value `snapshot` indicates that the shard is being restored from a snapshot.
- <2>: The `source` parameter indicates the source snapshot and repository.
- <3>: The `percent` parameter indicates the progress of the restoration task. The value `94.5%` indicates that 94.5% of the shard files have been restored.

The response lists all indexes that are being restored and the shards in these indexes. Each shard has statistics about the start or end time, duration, restoration progress, and bytes transmitted.

Cancel index restoration

To cancel index restoration, you only need to delete the indexes that are being restored. A restoration process is a shard restoration process. You can call the DELETE operation to modify the status of the cluster and cancel the restoration process. Example:

```
DELETE /restored_index_3
```

If you run the preceding command when the `restored_index_3` index is being restored, the system stops the restoration and deletes the data that has been restored to the cluster. For more information, see Snapshot And Restore.

Delete a snapshot

You can specify a repository name and a snapshot name, and send a `DELETE` request to delete the specified snapshot. Example:

```
DELETE _snapshot/my_backup/snapshot_2
```

> **Notice**
>
> - You can only use the DELETE operation to delete snapshots. A snapshot is associated with other backup files. Some of the files may also be used by other snapshots. The DELETE operation does not delete files that are still being used by other snapshots. It only deletes files that are associated with the deleted snapshot and are not used by other snapshots.
> - If you choose to manually delete a snapshot, you may delete files that are associated with snapshots by mistake. This may cause data loss.

11.5. Configure a shared OSS repository

Configure a shared OSS repository

Alibaba Cloud Elasticsearch provides shared OSS repositories. You can restore data from the automatic snapshots of an Elasticsearch cluster that are stored in these repositories to a destination Elasticsearch cluster. For example, you have two Elasticsearch V6.7.0 clusters es-cn-a and es-cn-b. The Auto Snapshot feature is enabled for the es-cn-a cluster, and a snapshot is created for the cluster. If you want to restore the data in the snapshot to the es-cn-b cluster, specify a shared OSS repository in the es-cn-b cluster.

Prerequisites

The source and destination Elasticsearch clusters must meet the following requirements:

- The clusters reside in the same region.
- The clusters belong to the same Alibaba Cloud account.
- The version of the source cluster is earlier than or the same as that of the destination cluster.

  If the versions of the source and destination clusters are both V6.7.0 of the Standard Edition, the latest kernel must be used for the clusters. Alternatively, the kernel version of the destination cluster must be later than that of the source cluster.

  > Notice
  > - An Elasticsearch cluster can use only the repository for an Elasticsearch cluster of the same or an earlier version.
  > - When a cluster uses the repository for a cluster of an earlier version, the cluster may be incompatible with the data format of the cluster of an earlier version. For example, you can restore an index that has only one document type to an Elasticsearch V6.7.0 cluster from snapshots in the repository of an Elasticsearch V5.5.3 cluster. However, if you restore an index that has multiple document types to the Elasticsearch V6.7.0 cluster from snapshots in the repository of the Elasticsearch V5.5.3 cluster, an error may occur. This is because Elasticsearch V6.7.0 clusters do not support indexes that have multiple document types.

Add a shared OSS repository

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Snapshots**.
5. In the **Shared OSS Repositories** section, click **Create Now**.

   > Notice If it is not the first time you add a shared OSS repository, click **Create Shared Repository**.

6. In the **Create Shared Repository** dialog box, select an Elasticsearch cluster.

   > Notice The selected cluster must meet the preceding requirements.

7. Click **OK**.
   After the shared repository is added, the current page shows the cluster that owns the repository

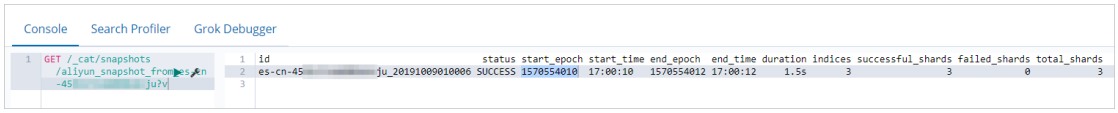
and the repository status.

Notice The system uses your Elasticsearch cluster to retrieve the repository list. If the cluster is updating its configuration, is abnormal, or encounters heavy workloads, the system may fail to retrieve the repository list. If this error occurs, you can log on to the Kibana console and run the `GET _snapshot` command to retrieve the endpoints of all repositories. For more information about how to log on to the Kibana console, see Log on to the Kibana console.

8. Restore an index.**Shared OSS repositories** are only used to share data between Elasticsearch clusters. The system cannot directly restore data for clusters from the shared OSS repositories. If you want to restore an index to an Elasticsearch cluster, run the related command in the Kibana console of the cluster. For example, to restore the file-2019-08-25 index from the es-cn-a cluster, perform the following steps:
   i. Log on to the Kibana console of the Elasticsearch cluster to which you want to restore the index.For more information, see Log on to the Kibana console.
   ii. In the left-side navigation pane, click **Dev Tools**.
   iii. On the **Console** tab of the page that appears, run the following command to query the information of all snapshots in the repository of the es-cn-a cluster:

```
GET /_cat/snapshots/aliyun_snapshot_from_es-cn-a?v
```

The command returns information about all the snapshots in the repository.



Note `aliyun_snapshot_from_es-cn-a` is the name of the **shared repository** that is added in Add a shared OSS repository.

iv. Restore indexes from the snapshot based on the returned information.

> **Notice**
>
> - Before you restore an index, make sure that the destination cluster does not have an index with the same name. If the destination cluster has an index with the same name, make sure that the index is disabled. If the index is enabled, an error occurs during index restoration.
> - Indexes whose names start with `.` are system indexes. We recommend that you do not restore these indexes. If you restore these indexes, you may fail to access the Kibana console.

- Restore a single index

```
POST _snapshot/aliyun_snapshot_from_es-cn-a/es-cn-a_20190705220000/_restore
{"indices": "file-2019-08-25"}
```

- Restore multiple indexes

```
POST _snapshot/aliyun_snapshot_from_es-cn-a/es-cn-a_20190705220000/_restore
{"indices": "kibana_sample_data_ecommerce,kibana_sample_data_logs"}
```

- Restore all indexes other than indexes whose names start with `.`

```
POST _snapshot/aliyun_snapshot_from_es-cn-a/es-cn-a_20190705220000/_restore
{"indices":"*,-.monitoring*,-.security*,-.kibana*","ignore_unavailable":"true"}
```

12.Data visualization

12.1. Kibana

12.1.1. Log on to the Kibana console

Log on to the Kibana console

When you purchase an Alibaba Cloud Elasticsearch cluster, Alibaba Cloud provides a free Kibana node with one vCPU and 2 GiB of memory. You can also choose to purchase a Kibana node with higher specifications. You can use the Kibana console to perform operations such as data queries and visualization.

log on to the Kibana console

Prerequisites

- An Alibaba Cloud Elasticsearch cluster is created.

  For more information, see Create an Elasticsearch cluster.
- The Public Network Access feature is enabled for the cluster. This feature is enabled by default.

  For more information, see Configure a whitelist for access to the Kibana console over the Internet or an internal network.
- The language of the Kibana console is English. The default language is English. If the language is not English, change the language.

  For more information, see Configure the language of the Kibana console.

Context

Elasticsearch provides the Kibana console for business expansion. The Kibana console is a part of the Elastic ecosystem and is seamlessly integrated into Elasticsearch. The Kibana console allows you to monitor the status of your Elasticsearch clusters and manage these clusters.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Data Visualization**.
5. In the **Kibana** section of the page that appears, click **Console**.
6. On the Kibana logon page, enter the username and password, and click Log in.
   - Username: The default value is elastic.
   - Password: Enter the password that is specified when you purchase your Elasticsearch cluster.

What's next

After you log on to the Kibana console, you can perform operations such as data queries and dashboard creation. For more information, see Kibana Guide.

12.1.2. Configure the language of the Kibana console

Configure the language of the Kibana console

This topic describes how to configure the language of the Kibana console.

language of the Kibana console configure the language of the Kibana console

Prerequisites

An Alibaba Cloud Elasticsearch cluster is created. For more information, see Create an Elasticsearch cluster.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Data Visualization**.
5. In the **Kibana** section of the page that appears, click **Edit Configuration**.
6. Click **Edit Configuration** on the right side of **Basic Configuration**.

   Warning The system will restart the Kibana node for the change to take effect. Before you perform the following steps, make sure that the restart does not affect your operations on the Kibana console.

7. In the **Edit Basic Configuration** pane, select a language and click **OK**.

   Note The Kibana console is available in **English** and **Chinese**. The default language is **English**.

   Then, the system automatically restarts the Kibana node.

What's next

After the Kibana node is restarted, log on to the Kibana console and verify that the console is switched to the selected language. Then you can perform operations such as data queries and dashboard creation. For more information, see Kibana Guide.

Note For more information about how to log on to the Kibana console, see Log on to the Kibana console.

12.1.3. Configure a whitelist for access to the Kibana console over the Internet or an internal network

Configure a Kibana whitelist

To access the Kibana console over the Internet or an internal network, you need to add the IP address of your host to a whitelist.

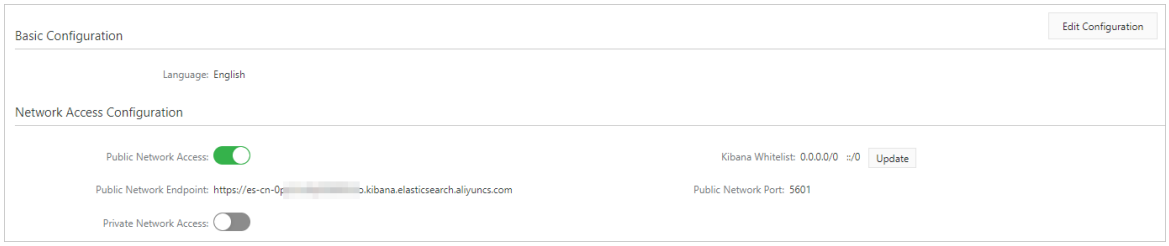
whitelist for access to the Kibana console over the Internet whitelist for access to the Kibana console over an internal network Kibana whitelist

Prerequisites

An Alibaba Cloud Elasticsearch cluster is created. For more information, see Create an Elasticsearch cluster.

Go to the network access configuration page

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Data Visualization**.
5. In the **Kibana** section of the page that appears, click **Edit Configuration**. You can then view the **Network Access Configuration** section on the **Kibana Configuration** page.
6. In the **Network Access Configuration** section, you can perform the following operations:



   - Configure a whitelist for access to the Kibana console over the Internet.
   - Configure a whitelist for access to the Kibana console over an internal network.

Configure a whitelist for access to the Kibana console over the Internet

1. In the **Network Access Configuration** section of the **Kibana Configuration** page, check whether **Public Network Access** is turned on (indicated by the color green).

> **Notice**
>
> By default, **Public Network Access** is turned on.
>
> If **Public Network Access** is turned off, you cannot log on to the Kibana console over the Internet.

- If yes, go to the next step.
- If no, click **Public Network Access** to turn it on.

2. Click **Update** next to **Kibana Whitelist**.
3. Enter the IP address you want to add in the text box.

   The Kibana console supports both IP addresses and Classless Inter-Domain Routing (CIDR) blocks. Enter IP addresses in the format of `192.168.0.1` and CIDR blocks in the format of `192.168.0.0/24`. Separate multiple IP addresses and CIDR blocks with commas (,). You can enter `127.0.0.1` to block all IPv4 addresses or enter `0.0.0.0/0` to allow all IPv4 addresses.

   If your Elasticsearch cluster is deployed in the China (Hangzhou) region, you can add IPv6 addresses to the whitelist in the format of `2401:b180:1000:24::5` or CIDR blocks in the format of `2401:b180:1000::/48`. Enter `::1` to block all IPv6 addresses or `::/0` to allow all IPv6 addresses.
4. Click **OK**.

Configure a whitelist for access to the Kibana console over an internal network

1. In the **Network Access Configuration** section of the **Kibana Configuration** page, check whether **Private Network Access** is turned on (indicated by the color green).

   > **Notice**
   >
   > By default, **Private Network Access** is turned off (indicated by the color gray).
   >
   > If **Private Network Access** is turned off, you cannot log on to the Kibana console over an internal network.

   - If yes, go to the next step.
   - If no, click **Private Network Access** to turn it on.
2. Click **Update** next to **Private Network Whitelist**.
3. Enter the IP address you want to add in the text box.

   The Kibana console supports both IP addresses and CIDR blocks. Enter IP addresses in the format of `192.168.0.1` and CIDR blocks in the format of `192.168.0.0/24`. Separate multiple IP addresses and CIDR blocks with commas (,). You can enter `127.0.0.1` to block all IPv4 addresses or enter `0.0.0.0/0` to allow all IPv4 addresses.
4. Click **OK**.

12.1.4. Install a Kibana plug-in

Install a Kibana plug-in

In addition to open-source community plug-ins, Alibaba Cloud Kibana provides a variety of plug-ins. This topic describes how to install a Kibana plug-in and relevant precautions.

Kibana plug-in install a Kibana plug-in

Prerequisites

- The Kibana node offers at least two vCPUs and 4 GiB of memory. Plug-ins consume resources.

  If the specifications of the Kibana node do not meet requirements, upgrade the Kibana node. For more information, see Upgrade the configuration of a cluster.

- An Alibaba Cloud Elasticsearch cluster of a version earlier than V7.0 is created.

  For more information, see Create an Elasticsearch cluster.

  > **Notice** Alibaba Cloud Elasticsearch clusters of V7.0 or later do not support Kibana plug-ins.

Procedure

1. Log on to the Alibaba Cloud Elasticsearch console.
2. In the top navigation bar, select the region where your cluster resides.
3. In the left-side navigation pane, click **Elasticsearch Clusters**. On the page that appears, find the target cluster and click its ID in the **Cluster ID/Name** column.
4. In the left-side navigation pane, click **Data Visualization**.
5. In the **Kibana** section of the page that appears, click **Edit Configuration**.
6. In the **Plug-in Configuration** section, find the plug-in that you want to install and click **Install** in the **Actions** column. If the specifications of your Kibana node do not meet the preceding requirements, the system prompts you to upgrade the configuration of your cluster. Follow the instructions to upgrade the Kibana node.

   > **Warning** After you confirm the plug-in installation, the system restarts the Kibana node. During the restart, Kibana cannot provide services. Therefore, before you confirm the installation, make sure that the restart does not affect your operations on the Kibana console.

7. In the **Install Plug-in** message, click **OK**. Then, the system restarts the Kibana node. After the node is restarted, the plug-in is installed.
   After the plug-in is installed, the state of the plug-in changes to **Installed**.

What's next

If you no longer require an installed plug-in, you can click **Remove** in the **Actions** column that corresponds to the plug-in in the **Plug-in Configuration** section to remove the plug-in.

> **Warning** After you confirm the plug-in removal, the system restarts the Kibana node. During the restart, Kibana cannot provide services. Therefore, before you confirm the removal, make sure that the restart does not affect your operations on the Kibana console.

12.1.5. Use the bsearch_querybuilder plug-in to query data

Use the bsearch_querybuilder plug-in

bsearch_querybuilder is also known as an advanced query. It is a frontend plug-in. This plug-in provides a visual interface in which you can create complex queries without the need to write complex domain-specific language (DSL) statements.

bsearch_querybuilder plug-in advanced query

Prerequisites

- An Alibaba Cloud Elasticsearch V6.3 or V6.7 cluster is created.

  For more information, see Create an Elasticsearch cluster. This topic uses an Alibaba Cloud Elasticsearch V6.3 cluster as an example.

- The bsearch_querybuilder plug-in is installed.

  For more information, see Install a Kibana plug-in.

- An index is created, and data is imported into the index.

  For more information, see Create an index and Create a document and insert data.

Context

Query DSL is an open-source Java framework that is used to define SQL type-safe queries. It allows you to call API operations to send queries instead of writing statements. Query DSL supports JPA, JDO, SQL, Java Collections, RDF, Lucene, and Hibernate Search.

Elasticsearch provides a complete Query DSL for you based on JSON to define queries. Query DSL provides a number of query expressions. Some queries can wrap other queries, such as boolean queries. Some queries can wrap filters, such as constant_score queries. Some queries can wrap both other queries and filters, such as filtered queries. You can combine any query expression and filter supported by Elasticsearch to create a complex query and filter the returned result. DSL is a complex language and is hard to master. In most cases, users often make mistakes or spend too much time writing DSL statements. The bsearch_querybuilder plug-in simplifies the writing of DSL statements and improves efficiency.

bsearch_querybuilder has the following features:

- Easy to learn: bsearch_querybuilder is a graphical tool. It allows you to create DSL queries with simple click and drag operations. You can customize search conditions without the need for complex coding, which reduces the cost of learning to write DSL statements. It also helps developers write and verify DSL statements.
- Easy to use: All queries that you have defined are stored in Kibana. These queries are ready for use at all times.
- Compact: bsearch_querybuilder only consumes about 14 MiB of disk space and does not stay resident in the memory. This means that the plug-in does not affect the performance of Kibana and Elasticsearch.
- Secure and reliable: bsearch_querybuilder does not rewrite, store, or forward user data. The source code of bsearch_querybuilder has passed the security auditing of Alibaba Cloud.

> **Note** bsearch_querybuilder only supports Elasticsearch V6.3 or V6.7 clusters.

Procedure

1. Log on to the Kibana console of your Alibaba Cloud Elasticsearch cluster. For more information, see

Log on to the Kibana console.

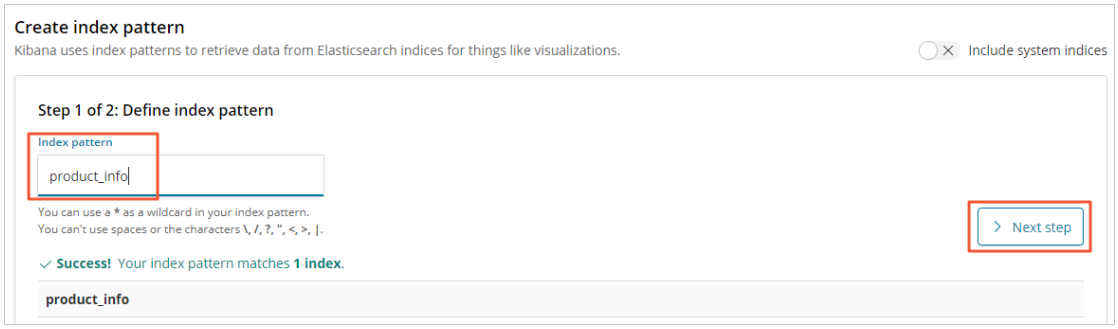
2. In the left-side navigation pane, click **Management** and follow these steps to create an index pattern:

> **Notice** If you have created an index pattern, skip this step.

   i. In the **Kibana** section of the **Management** page, click **Index Patterns**.

   ii. In the **Create index pattern** section, enter an index pattern name (the name of the index that you want to query).

   iii. Click **Next step**.

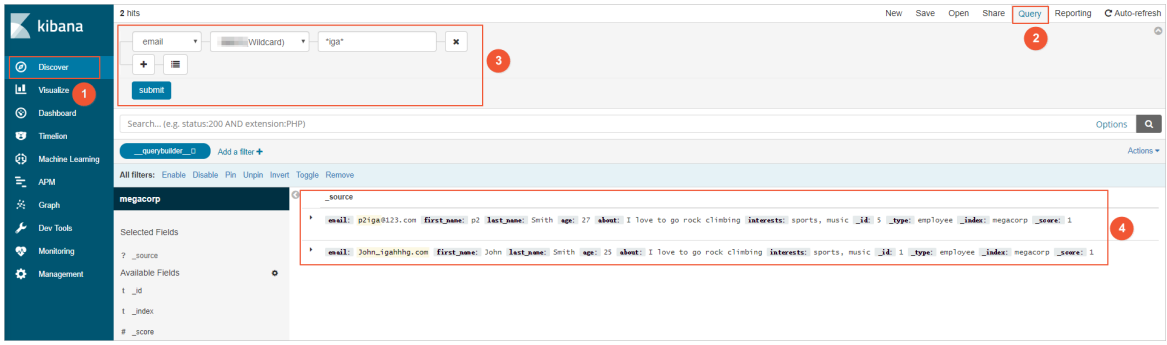


   iv. Click **Create index pattern**.

3. In the left-side navigation pane, click **Discover**.

4. In the top navigation bar of the **Discover** page, click **Query**.

5. In the section that appears, select a search condition and a filter, and click **Submit**.

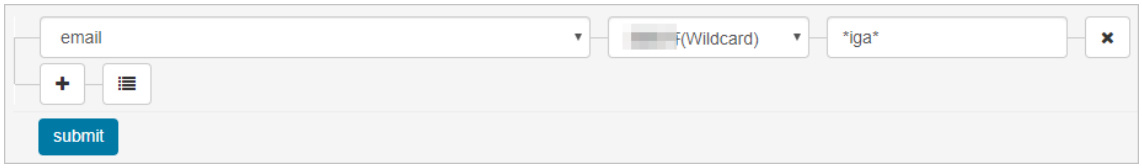


   - Click the + button to add a search condition.
   - Click the ☰ button to add a filter for the search condition.
   - Click the × button to delete a search condition or filter.

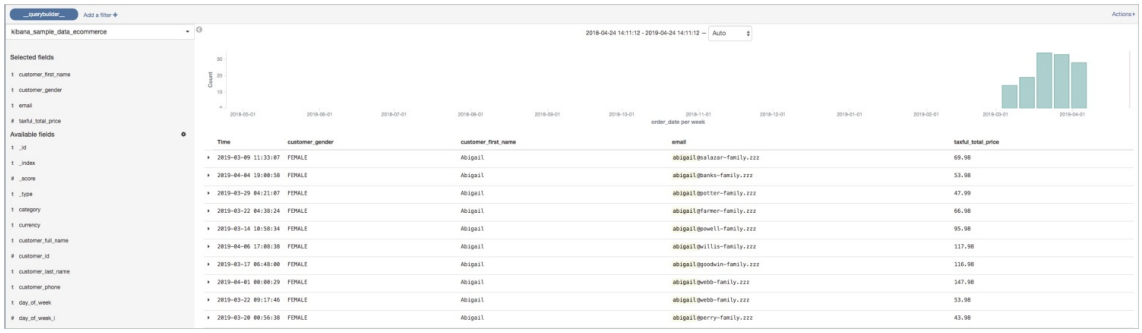
   The bsearch_querybuilder plug-in allows you to create a variety of queries, such as fuzzy queries, boolean queries, and range queries. Examples:

   - Fuzzy query

As shown in the following figure, the **email** condition is added for a fuzzy match. The **email** condition matches all email addresses that contain the **iga** keyword.

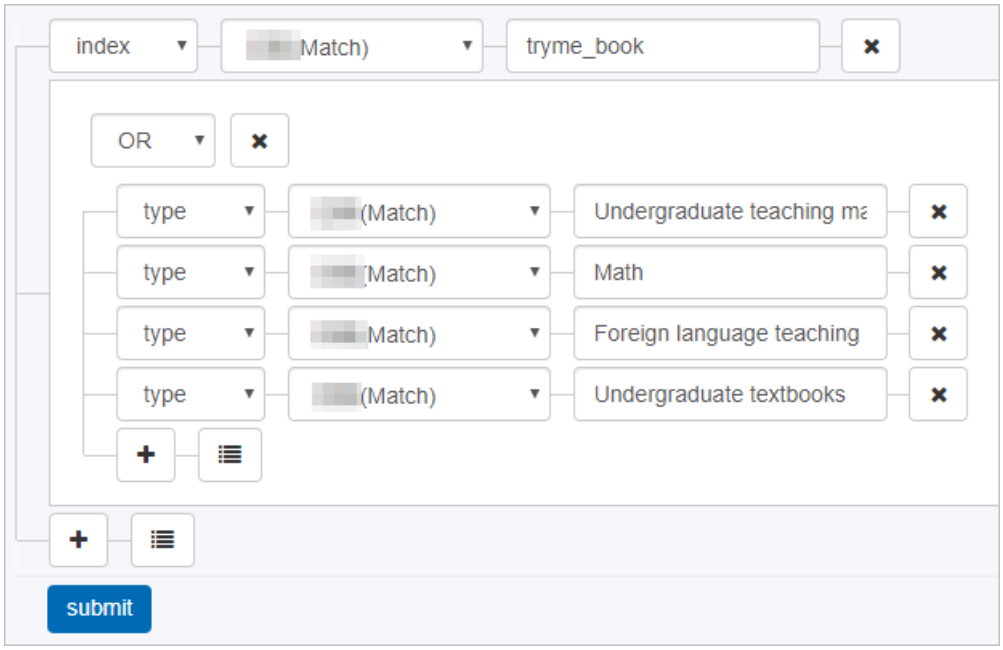


The following figure shows the returned result.

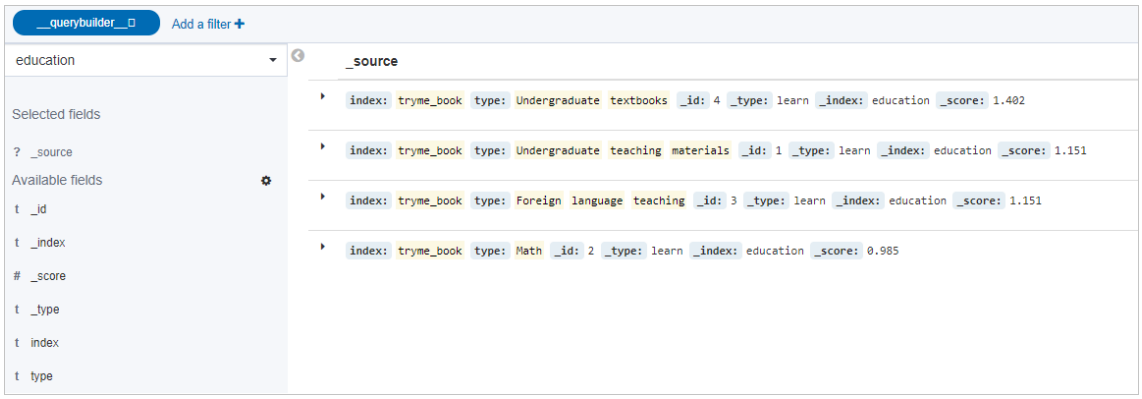


- Boolean query

As shown in the following figure, the **index** condition is set to **tryme_book**. An OR condition that contains multiple filters is also added to filter data by **type**. The **type** filters are set to **Undergraduate teaching materials**, **Math**, **Foreign language teaching**, and **Undergraduate textbooks**.



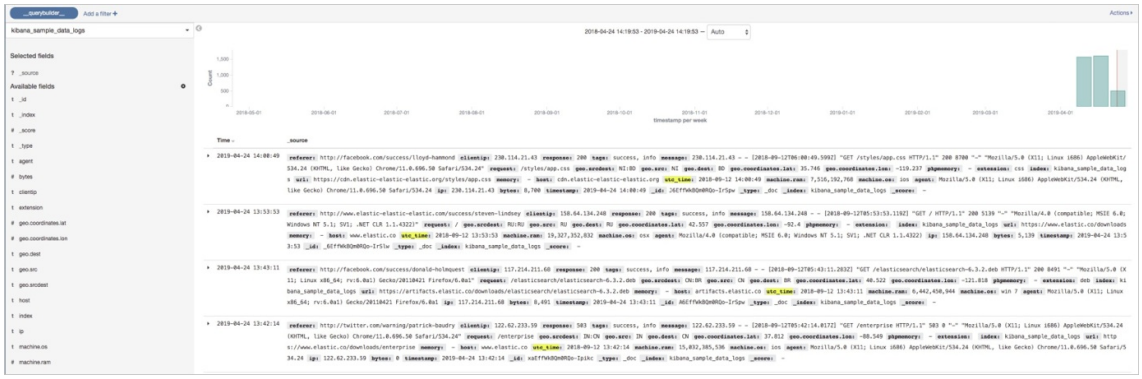
The following figure shows the returned result.



- Range query

  Range queries allow you to search data by date. As shown in the following figure, the range condition is used to filter data based on the **utc_time** field. Only data entries created within the specified time range are returned. The specified time range is `[Current time - 240 days, Current time]` .

The following figure shows the returned result.



12.1.6. Use the bsearch_label plug-in to label data

Use the bsearch_label plug-in to label data

bsearch_label is a frontend data labeling plug-in. It supports visualized data labeling. This way, you do not need to write complex domain-specific language (DSL) statements.

bsearch_label plug-in data labeling

Prerequisites

- An Alibaba Cloud Elasticsearch V6.3 or V6.7 cluster is created.

For more information, see Create an Elasticsearch cluster. This topic uses an Elasticsearch V6.3 cluster as an example.

- The bsearch_label plug-in is installed.

  For more information, see Install a Kibana plug-in.

- An index is created, and data is added to the index.

  For more information, see Create an index and Create a document and insert data.

- The language of the Kibana console is English. The default language is English. If the language is not English, change the language.

  For more information, see Configure the language of the Kibana console.

Context

In most cases, when you analyze data, you may want to use query conditions to filter the data in addition to viewing the data and use tags to classify fields. This procedure is known as data labeling. After you add tags to data, you can use the tags to aggregate and classify the data, perform statistical analysis, and filter data by tag. The labeled data can be used in subsequent procedures.

Procedure

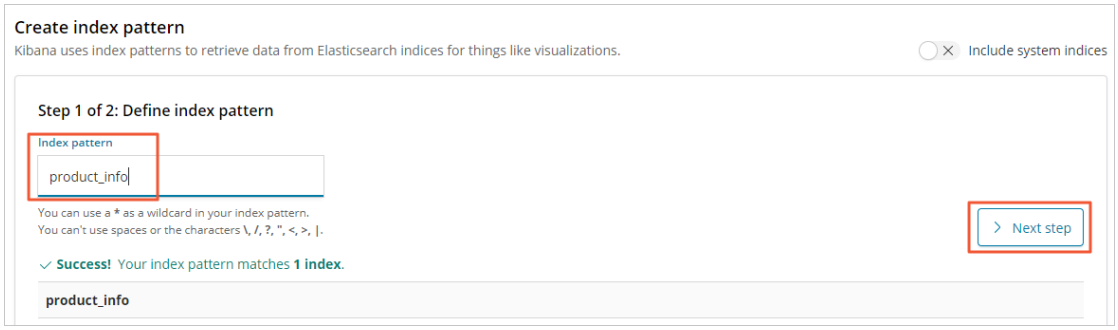
1. Log on to the Kibana console of your Alibaba Cloud Elasticsearch cluster. For more information, see Log on to the Kibana console.
2. In the left-side navigation pane, click **Management** and follow these steps to create an index pattern:

   > **Notice** If you have created an index pattern, skip this step.

   i. In the **Kibana** section of the **Management** page, click **Index Patterns**.

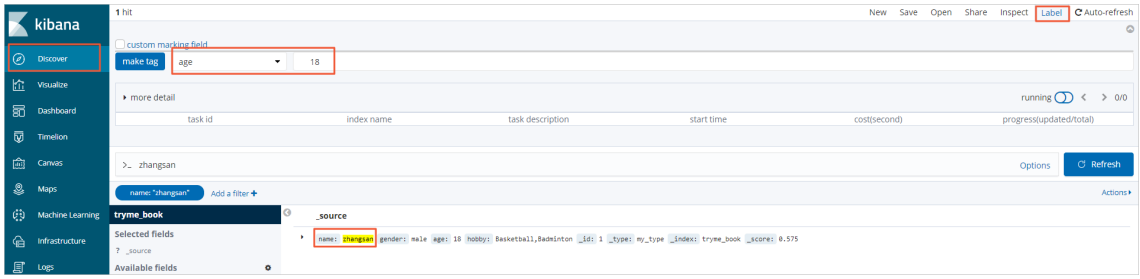
   ii. In the **Create index pattern** section, enter an index pattern name (the name of the index that you want to query).

   iii. Click **Next step**.

   iv. Click **Create index pattern**.

3. In the left-side navigation pane, click **Discover**.
4. In the top navigation bar of the **Discover** page, click **Label**.
5. Use one of the following methods to label the data.
   - Add a tag to an existing field.

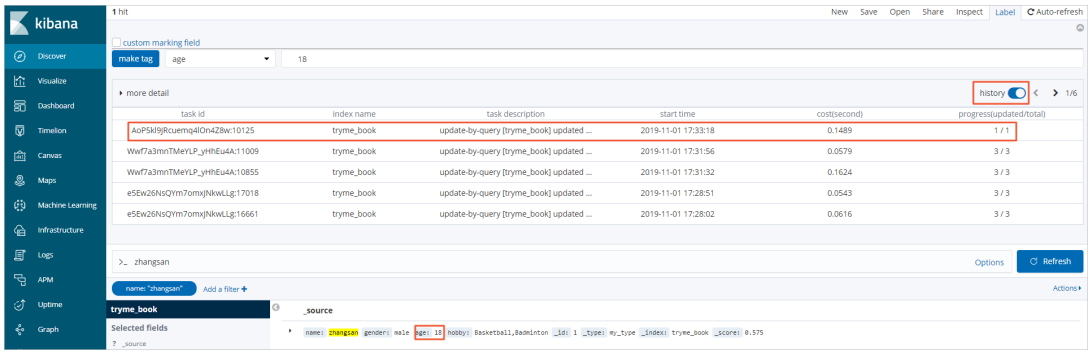


a. As shown in the preceding example, find the record of user zhangsan.

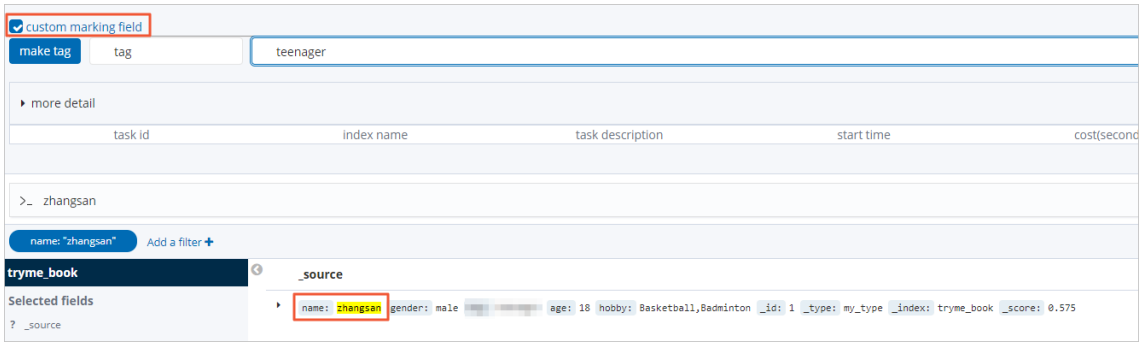
b. Select the **age** field and add tag **18** to this filed.

c. Click **make tag**.

d. Click the **history** switch to view detailed labeling history.



- Add a tag to a new field.



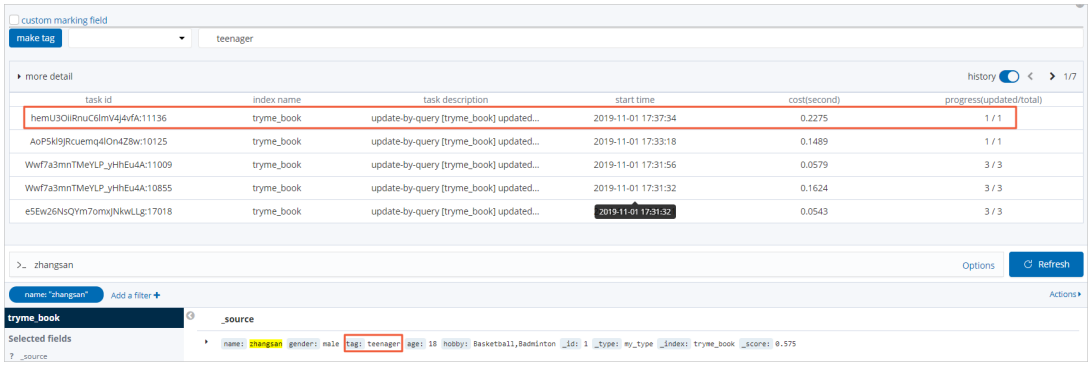
a. As shown in the preceding example, find the record of user zhangsan.

b. Select **custom marking field**.

c. Add the **tag** field and add the **teenager** tag to this field.

d. Click **make tag**.

e. View the labeling result.



13.FAQ

13.1. Incorrect configuration selected for an Alibaba Cloud Elasticsearch cluster

If the configuration of an Alibaba Cloud Elasticsearch cluster does not meet expectations, perform operations based on this topic.

The following table lists the modification method for each configuration item.

Warning If cluster release or subscription cancellation is required, we recommend that you back up your data before the release or subscription cancellation. For more information, see Commands for creating snapshots and restoring data. After the release or subscription cancellation, data on the cluster is deleted and cannot be restored.

Configuration item	Modification method
Billing method	• If you purchased a pay-as-you-go cluster, you can switch it to a subscription cluster. For more information, see Switch the billing method from pay-as-you-go to subscription. • If you purchased a subscription cluster, you cannot switch it to a pay-as-you-go cluster. In this case, we recommend that you cancel the subscription of the cluster and purchase another cluster as needed.
Version	If you purchase an Elasticsearch V6.3.2 cluster but require an Elasticsearch V6.7.0 cluster, you can upgrade the cluster version. For more information, see Upgrade the version of a cluster. In other cases, we recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed.
Region	You cannot modify this configuration item. We recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed.
Zone	You can migrate nodes to the desired zone. For more information, see Migrate nodes in a zone.
Number of zones	You cannot modify this configuration item. We recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed.
Specifications	You can modify this configuration item. For more information, see Upgrade the configuration of a cluster.
Storage type	You cannot modify this configuration item. We recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed.

Configuration item	Modification method
Disk encryption	You cannot modify this configuration item. We recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed.
Storage space per node	You can modify this configuration item. For more information, see Upgrade the configuration of a cluster.
Number of data nodes	You can modify this configuration item. For more information, see Upgrade the configuration of a cluster.
Network type, VPC, and VSwitch	You cannot modify these configuration items. We recommend that you cancel the subscription of or release the cluster and purchase another cluster as needed. **Note** The network type of an Elasticsearch cluster can only be VPC.
Username	The default username is elastic and cannot be changed. You can create a user in the Kibana console and grant the required permissions to the user. For more information, see Create a role and Create a user.
Password	You can modify this configuration item. For more information, see Reset the access password for an Elasticsearch cluster.

For configuration items that are not provided in the preceding table, check whether you can modify the items on the configuration upgrade page. For more information, see Upgrade the configuration of a cluster.

13.2. Access to an Alibaba Cloud Elasticsearch cluster from the classic network

This topic provides answers to some commonly asked questions about access to an Alibaba Cloud Elasticsearch cluster from the classic network.

How do I access an Alibaba Cloud Elasticsearch cluster deployed in a VPC from the classic network?

For network security, your Alibaba Cloud Elasticsearch cluster is deployed in your virtual private cloud (VPC). If your business system is deployed in the classic network, you can use the ClassicLink feature that is supported by VPC to access your Elasticsearch cluster.

What is ClassicLink?

The ClassicLink feature is provided by VPC. It provides a network connection that allows you to access your VPC from the classic network.

What are the limits of ClassicLink?

- Up to 1,000 ECS instances of the classic network can be connected to the same VPC.
- An ECS instance of the classic network can be connected to only one VPC, and the VPC must be under the same account and belong to the same region.

  For cross-account connection such as ones connecting an ECS instance under account A to a VPC under account B, you can transfer the ECS instance from account A to account B.
- To enable the ClassicLink function of a VPC, the following conditions must be met:

VPC CIDR block	Limitations
172.16.0.0/12	There is no custom route entry destined for 10.0.0.0/8 in the VPC.
10.0.0.0/8	○ There is no custom route entry destined for 10.0.0.0/8 in the VPC. ○ Make sure that the CIDR block of the VSwitch to communicate with the ECS instance in the classic network is within 10.111.0.0/16.
192.168.0.0/16	○ There is no custom route entry destined for 10.0.0.0/8 in the VPC. ○ Add a route entry, of which the destination CIDR block is 192.168.0.0/16 and the next hop is the private NIC, to the ECS instance of the classic network. Download the Route script. **Note** Before running the script, read the readme file in the script carefully.

How do I enable ClassicLink?

1. Log on to the VPC console.
2. In the left-side navigation pane, click **VPCs**.
3. In the top navigation bar, select a region.
4. Find your VPC and click **Manage** in the **Actions** column.

   We recommend that you select a VPC that is attached to the Classless Inter-Domain Routing (CIDR) block 172.16.0.0/12.
5. In the upper-right corner of the **VPC Details** page, click **Enable ClassicLink**.

   Note If ClassicLink is enabled, **Disable ClassicLink** appears.
6. In the **Enable ClassicLink** dialog box, click **OK**.

   After ClassicLink is enabled, the value of the ClassicLink parameter changes to **Enabled**.

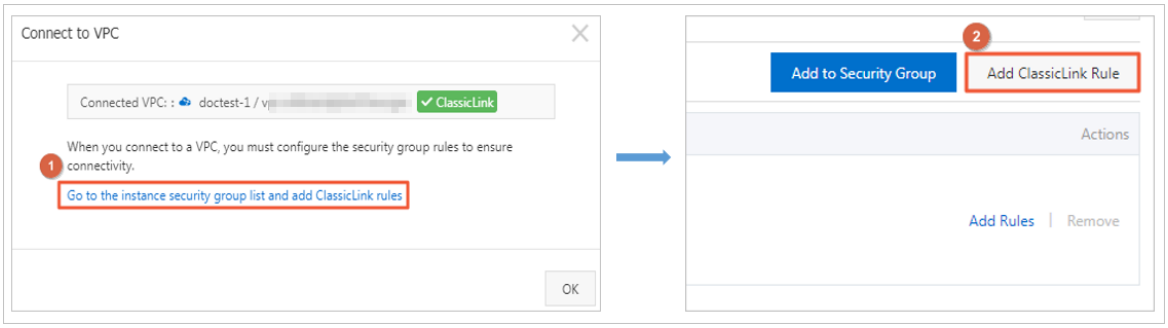
How do I create a ClassicLink?

Before you create a ClassicLink, make sure that you have completed the following operations:

- Read and understand the limits of ClassicLink. For more information, see What are the limits of ClassicLink?

- Enable ClassicLink for the VPC to which you want to establish the ClassicLink. For more information, see How do I enable ClassicLink?

1. Log on to the ECS console.
2. In the left-side navigation pane, choose **Instances & Images > Instances**.
3. In the top navigation bar, select a region.
4. Find your ECS instance. Then, choose **More > Network and Security Group > Set classic link** in the **Actions** column.
5. In the dialog box that appears, select your VPC, click **OK**, and then click **Go to the instance security group list and add ClassicLink rules**.

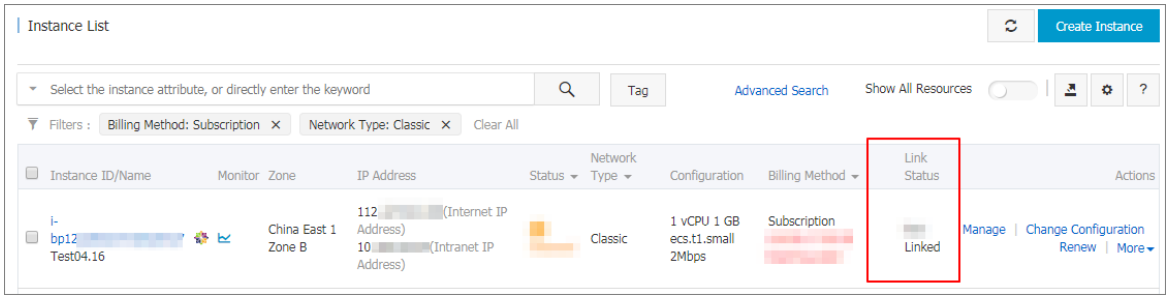


6. Click **Add ClassicLink Rule**. In the dialog box that appears, configure the following parameters and click **OK**.

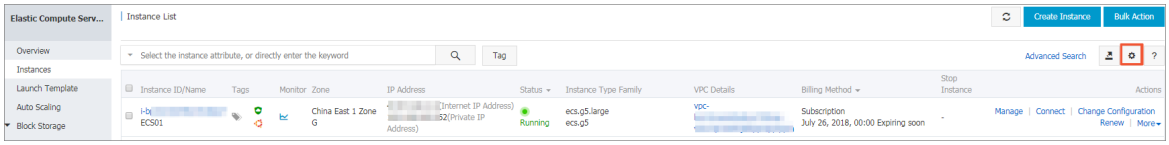
Parameter	Description
Classic Security Group	The name of the security group for the classic network.
Select VPC Security Group	Select a security group for the VPC.
Mode	Select one of the following authorization modes: ○ **Classic <=> VPC**: This mode allows ECS instances in the classic network and cloud resources in a VPC to access each other. We recommend that you select this mode. ○ **Classic => VPC**: This mode allows ECS instances in the classic network to access cloud resources in a VPC. ○ **VPC => Classic**: This mode allows cloud resources in a VPC to access ECS instances in the classic network.
Protocol	Select a communication protocol, such as **Custom TCP**.
Port Range	The ports that are used for communication. Specify the ports in the xx/xx format. For example, to specify port 80, enter 80/80.
Priority	The priority of the rule. A small value indicates a high priority. For example, if you set this parameter to 1, the rule has the highest priority.
Description	Enter a description for the security group.

How do I test the connectivity between the classic network and a VPC?

1. Go to the ECS console and click the



icon in the upper-right corner of the Instances page. In the dialog box that appears, select **Connection Status** and click **OK**. Then, view the connection status of the ECS instance.



2. Log on to the ECS instance from which the ClassicLink is established and run the curl command to access your Elasticsearch cluster in the VPC.

> **Note** If the system displays "curl command not found", run the `yum install curl` command to install cURL on the ECS instance.

```
curl -u <username>:<password> http://<host>:<port>
```

Variable	Description

Variable	Description
<username>	The account that is used to access your Elasticsearch cluster. We recommend that you do not use the elastic account. **Notice** ◦ If you reset the password of the elastic account when you use the account to access your Elasticsearch cluster, it may require some time for the new password to take effect. During this period, you cannot use the account to access your Elasticsearch cluster. Therefore, we recommend that you do not use the elastic account to access your Elasticsearch cluster. ◦ If the version of your Elasticsearch cluster contains "with_X-Pack", you must specify both the username and password to access the cluster.
<password>	The password that is used to access your Elasticsearch cluster. The password is specified when you create the cluster or initialize Kibana.
<host>	The **internal endpoint** of your Elasticsearch cluster. You can obtain the internal endpoint from the **Basic Information** page of the cluster.
<port>	The port number of your Elasticsearch cluster. In most cases, **9200** is used. You can obtain the port number from the **Basic Information** page of the cluster.

Example command:

```
curl -u elastic:es_password http://es-cn-vxxxxxxxxxxxxmedp.elasticsearch.aliyuncs.com:9200
```

If the connection is established, the result shown in the following figure is returned.

13.3. Kibana console password

This topic describes the frequently asked questions about the Kibana console password.

What is a Kibana console password?

Notice The elastic account is the root account in Alibaba Cloud Elasticsearch. It has full permissions to manage clusters. You must keep the password of this account confidential.

When you use one of the following methods to connect to your Elasticsearch instance, you must specify the password of the Kibana console in `elastic/your_password` for permission verification.

- Use the API or SDK to connect to the Elasticsearch instance.
- Use Kibana to connect to the Elasticsearch instance.

What is the better way to manage permissions in the Kibana console of Alibaba Cloud Elasticsearch?

- We recommend that you create user accounts and assign roles in the Kibana console of your Elasticsearch instance instead of using the elastic account to manage the instance. The elastic account has the root permission. For more information about how to create a user account in Kibana, see Kibana.
- We recommend that you do not use the elastic account in searches. The elastic account has the root permission. If the password of the elastic account is disclosed, the security of your cluster may be threatened.
- Use caution when you change the password of the elastic account. The elastic account has the root permission. If you use the elastic account to manage your workloads, after you change the password of your Elasticsearch instance, your requests are rejected due to authentication failures and your workloads are interrupted.

13.4. Installation errors of a custom plug-in

This topic describes how to troubleshoot problems that occur during the installation of a custom plug-in. These problems can include console-reported errors, modification suspension, and verification failures.

Problem description

When you upload and install a custom plug-in, problems occur, such as console-reported errors, modification suspension, and verification failures.

Common solutions

Before you upload and install a custom plug-in, you must put the plug-in in the plugins folder under the installation path of a user-created Elasticsearch cluster. Then, restart the cluster to load the plug-in. After the cluster is restarted, run the `GET /_cat/plugins?v` command to check whether the plug-in is installed. When you install the plug-in, take note of the following items:

- You are not allowed to upload a plug-in that has the same name as a built-in plug-in provided by Alibaba Cloud Elasticsearch.

  If you want to upload such a plug-in, you must change the name of the plug-in first. When you test the installation of a plug-in, such as pinyin analysis or IK analysis, on a user-created Elasticsearch cluster, you must delete the native plug-in that corresponds to the plug-in first. Then, change the name of the plug-in that you want to install, and install it by running the following installation command for a native plug-in. Check whether the installation succeeds. If the installation succeeds, you can upload and install the plug-in to your Alibaba Cloud Elasticsearch cluster.

  ```
  ./bin/elasticsearch-plugin install file:///path-to-your-plugins.zip
  ```

- Custom plug-ins can be uploaded and installed to an Alibaba Cloud Elasticsearch cluster only after they are tested on a user-created Elasticsearch cluster. These custom plug-ins are installed to the user-created cluster by running the following installation command for a native plug-in:

  ```
  ./bin/elasticsearch-plugin install file:///path-to-your-plugins.zip
  ```

- Rolling update is unavailable for custom plug-ins.
- If the security policy of a plug-in defines the permissions on addition, deletion, modification, and query, you must comment out these permissions.

  For example, you can view the security policy of the hanlp plug-in in its configuration file. For this plug-in, you must comment out the following content in the *plugin-security.policy* file:

  ```
  permission java.io.FilePermission "<>", "read,write,delete";
  ```

- You are not allowed to install custom plug-ins for Logstash and Kibana. If you want to install a custom plug-in for Logstash or Kibana, you must install the plug-in in a custom environment by running the installation command for a native plug-in. If the installation is successful, you can submit a ticket to the technical support engineer of Alibaba Cloud Elasticsearch.

  > **Note** After a valid plug-in is installed, you can query the plug-in logs on the Cluster Log tab of the Elasticsearch console.

Console-reported errors

- Cause

The plug-in that you upload is invalid.

- Solution

  Modify the information of the plug-in. For more information, see Common solutions. For example, you can rename the plug-in or modify the configuration file of the plug-in.

Modification suspension

Modification suspension is caused by two reasons. You must perform the following steps to troubleshoot it:

1. Check the size of the plug-in. Make sure that the size is less than 50 MB.

   If the size of the plug-in is greater than or equal to 50 MB, the loading of the plug-in during the installation is slow. In this case, you must terminate the loading and delete the plug-in. Then, you must modify the configuration of the plug-in. For example, you can delete some tokens provided by the plug-in. Make sure that the size of the plug-in is less than 50 MB before you upload and install the plug-in again.

2. Check whether the system is writing data to data nodes.
   - Yes: Wait for a few minutes. The installation process is slow because it is during peak hours of business.
   - No: Terminate the installation. Then, delete the plug-in and test its availability again.

Verification failures

Verification failures are caused by various reasons. For example, the version of the plug-in is not compatible with your Alibaba Cloud Elasticsearch cluster, the *plugin-descriptor.properties* file contains invalid parameter settings, the plug-in is not compressed, or the installation path of the plug-in is invalid. The plug-in is usually compressed in a recursive manner. You can run the zip -r command to compress the plug-in. The preceding problems can occur during the installation of open source plug-ins. You need to troubleshoot them on your own.

13.5. FAQ about Alibaba Cloud Elasticsearch clusters

This topic provides answers to some commonly asked questions about Alibaba Cloud Elasticsearch clusters.

- FAQ about the purchase, subscription cancellation, or release of clusters
  - When I purchase an Elasticsearch cluster, I selected an incorrect zone. How do I change the zone?
  - What are the mappings between versions on the Elasticsearch buy page and specific Elasticsearch versions?
  - After I cancel the subscription of or release an Elasticsearch cluster, I purchase another cluster. Does the endpoint of the new cluster remain the same as that of the original cluster?
  - How do I access an Elasticsearch cluster from the classic network?
  - How do I release an Elasticsearch cluster?
  - When is an Elasticsearch cluster released after it is suspended?
  - Can I purchase an Elasticsearch cluster that has only one node?

- When I purchase an Elasticsearch cluster, resources of a specific category are sold out. What do I do?

- FAQ about features
  - Can I upgrade or downgrade the version of an Elasticsearch cluster?
  - Can I log on to an Elasticsearch cluster over SSH and modify the configuration of the cluster?
  - Is Logstash V6.7 compatible with Elasticsearch V6.3?
  - Can Elasticsearch be used as a data source of Quick BI?
  - Does Elasticsearch support scoring plug-ins?
  - Does Elasticsearch support LDAP?
  - Does Alibaba Cloud provide Elasticsearch SDK for Java?
  - How do I view the kernel version of an Elasticsearch cluster?
- FAQ about cluster restarts
  - How long is required to restart an Elasticsearch cluster or node?
  - Does the system restart an Elasticsearch cluster after I enable or disable the Public Network Access feature for the cluster?
- FAQ about data queries or import

  The CPU utilization and loads of some nodes in an Elasticsearch cluster are normal, whereas other nodes are in the idle state. What do I do?
- FAQ about cluster configuration and configuration updates
  - How do I plan resources before I use Elasticsearch, such as cluster specifications, the number of shards, and the size of each shard?
  - How do I view the configuration of an Elasticsearch cluster?
  - Are services affected when I modify the configuration of an Elasticsearch cluster?
  - Can I change the cloud disk type of an Elasticsearch cluster?
  - Can I convert other types of nodes in an Elasticsearch cluster to warm nodes?
  - Can I downgrade the specifications of an Elasticsearch cluster? If yes, how do I do?
  - In the event of a temporary business surge, how do I modify the configuration of an Elasticsearch cluster to ensure that services run as expected?
  - When I upgrade the configuration of an Elasticsearch cluster, the system displays the "UpgradeVersionMustFromConsole" error message. What do I do?
  - How long is required to upgrade the version of an Elasticsearch cluster?
  - Are services affected when I upgrade the version of an Elasticsearch cluster?
  - Can I change the JVM parameter settings of an Elasticsearch cluster?
  - Can I use the YML configuration file of an Elasticsearch cluster to change the settings of the http.max_content_length and discovery.zen.ping_timeout parameters?
  - Can I switch the VPC of an Elasticsearch cluster?
- FAQ about plug-ins, tokens, and synonyms
  - How do I update dictionary content when I use the IK analysis plug-in?
  - When I use the IK analysis plug-in, the system displays the "ik startOffset" error message. What do I do?

- The IK dictionary files on my on-premises machine are lost. Can I retrieve them on the cluster management page?
- After I update IK dictionaries, how do I apply the new dictionaries to existing data?
- Is a threshold specified for full GC?
- Can I remove plug-ins that are not used?
- Are the dictionaries provided by the IK analysis plug-in of Alibaba Cloud Elasticsearch the same as those provided by the IK analysis plug-in of open source Elasticsearch?
- Can a custom plug-in access an external network, such as reading dictionary files on GitHub?
- Does a custom plug-in support the rolling update method?
- How do I configure the analysis-aliws plug-in? What is the format of the dictionary file for this plug-in?
- What are the differences among Elasticsearch synonyms, IK tokens, and AliNLP tokens?

- FAQ about logs
  - Can I specify a retention period for the .security indexes of an Elasticsearch cluster?
  - I can view Elasticsearch cluster logs that are generated only over the last seven days. How do I view more logs?
  - I cannot view the search and update logs of an Elasticsearch cluster. What do I do?
  - How do I configure and view the slow logs of an Elasticsearch cluster?
  - How do I obtain the slow logs of an Elasticsearch cluster on a regular basis?
- FAQ about data backup and restoration
  - Can I restore data from the snapshots of an Elasticsearch cluster to an Elasticsearch cluster of a different version?
  - When I back up data for an Elasticsearch cluster, the system displays a message indicating that the cluster is unhealthy. What do I do?
  - I enable the Auto Snapshot feature but do not specify shared OSS repositories for an Elasticsearch cluster. Are snapshots created?
  - When I restore data from snapshots, the destination Elasticsearch cluster displays a message. This message indicates that shards are abnormal. After I run the POST /_cluster/reroute?retry_failed=true command to reroute the shards, the issue persists. What do I do?
  - Can I export data from an Elasticsearch cluster to my on-premises machine?
- FAQ about cluster monitoring and alerting
  - How do I use the email notification feature of X-Pack Watcher?
  - What do I do if the system reports an alert indicating that memory cannot be allocated to the garbage collector?
- FAQ about access to clusters
  - How do I use a client to access an Alibaba Cloud Elasticsearch cluster? What is the difference between access to an Alibaba Cloud Elasticsearch cluster and access to an open source Elasticsearch cluster?
  - When I use a client to access an Elasticsearch cluster, can I disable the basic authentication feature?
  - I purchase an ECS instance that resides in the same VPC as but different zone from an Elasticsearch cluster. Can I use the ECS instance to access the Elasticsearch cluster from an internal network?

- How do I access an Elasticsearch cluster from the Internet?

When I purchase an Elasticsearch cluster, I selected an incorrect zone. How do I change the zone?

After the cluster is created and in the Active state, migrate nodes from the incorrect zone to your desired zone. For more information, see Migrate nodes in a zone.

What are the mappings between versions on the Elasticsearch buy page and specific Elasticsearch versions?

Version on the buy page	Specific version
7.7	7.7.1
7.4	7.4.0
6.8	6.8.6
6.7	6.7.0
6.3	6.3.2
5.6	5.6.16
5.5	5.5.3

After I cancel the subscription of or release an Elasticsearch cluster, I purchase another cluster. Does the endpoint of the new cluster remain the same as that of the original cluster?

No, after you purchase the new cluster, we recommend that you modify the client code and cancel the subscription of or release the original cluster to avoid service interruptions.

How do I access an Elasticsearch cluster from the classic network?

You can use the ClassicLink feature to access an Elasticsearch cluster from the classic network. For more information, see Access to an Alibaba Cloud Elasticsearch cluster from the classic network.

How do I release an Elasticsearch cluster?

On the Elasticsearch Clusters page, find the cluster that you want to release. Then, in the Actions column, choose **More > Release**. For more information, see Release a cluster.

When is an Elasticsearch cluster released after it is suspended?

The cluster is released 24 hours after it is suspended. After it is released, all data in the cluster is permanently deleted and cannot be recovered. For more information, see Overdue payments.

Can I purchase an Elasticsearch cluster that has only one node?

No, an Elasticsearch cluster must have a minimum of two data nodes. For more information, see Parameters on the buy page.

When I purchase an Elasticsearch cluster, resources of a specific category are sold out. What do I do?

Take one of the following measures:

- Select another region.
- Select another zone.
- Select another category.

If the resources that you want to purchase are still unavailable after you take all of the preceding measures, try again later. Resources are dynamic. If resources are insufficient, Alibaba Cloud replenishes them as soon as possible.

Can I upgrade or downgrade the version of an Elasticsearch cluster?

Upgrades are supported, whereas downgrades are not supported. You can upgrade Elasticsearch clusters only from V6.3.2 to V6.7.0. For more information, see Upgrade the version of a cluster.

If you want to perform upgrades between other versions or downgrades, purchase an Elasticsearch cluster of the desired version. Then, migrate data from the original cluster to the new cluster and cancel the subscription of or release the original cluster.

Can I log on to an Elasticsearch cluster over SSH and modify the configuration of the cluster?

No, for security purposes, you are not allowed to log on to your Elasticsearch cluster over SSH. If you want to modify the configuration of your cluster, use the cluster configuration feature of Elasticsearch. For more information, see Overview.

Is Logstash V6.7 compatible with Elasticsearch V6.3?

Yes, for more information, see Compatibility matrixes.

Can Elasticsearch be used as a data source of Quick BI?

No, you can use Kibana to analyze and present analysis results.

Does Elasticsearch support scoring plug-ins?

Yes, when you create an index, Elasticsearch allows you to create a tokenizer. When you search for data, Elasticsearch uses a scoring plug-in to sort search results by score. For more information, see Search for data.

Does Elasticsearch support LDAP?

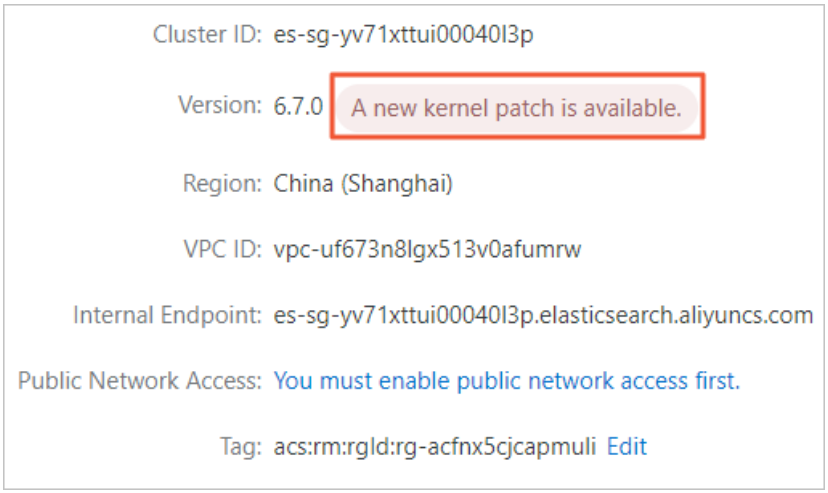
No, if you want to use Lightweight Directory Access Protocol (LDAP) to authenticate requests sent to your Elasticsearch cluster, you must deploy an on-premises Elasticsearch cluster of the same version. Then, use the on-premises Elasticsearch cluster to conduct an authentication test. If LDAP runs as expected, send related configurations to Alibaba Cloud Elasticsearch technical engineers. Then, the engineers can configure your cluster to support LDAP based on the configurations. For more information, see Best practice of integrating X-Pack with LDAP.

Does Alibaba Cloud provide Elasticsearch SDK for Java?

Yes, different Elasticsearch versions use different SDKs. For more information, see Java API.

How do I view the kernel version of an Elasticsearch cluster?

By default, Elasticsearch clusters use the kernel of the latest version. For more information about kernel versions, see AliES release notes. If your cluster does not use the kernel of the latest version, the **A new kernel patch is available** message appears on the Basic Information page of your cluster. You can click the message to view the current kernel version of your cluster.



How long is required to restart an Elasticsearch cluster or node?

When you restart an Elasticsearch cluster or node, the system displays the required time. The time is estimated based on your cluster specifications, data structure, and data volume. Typically, it requires a few hours to restart a cluster. For more information, see Restart a cluster or node.

Does the system restart an Elasticsearch cluster after I enable or disable the Public Network Access feature for the cluster?

No, only the status of the Public Network Access feature changes, which does not affect your cluster.

The CPU utilization and loads of some nodes in an Elasticsearch cluster are normal, whereas other nodes are in the idle state. What do I do?

This issue is caused by unbalanced loads on the cluster. Unbalanced loads may be caused by several reasons. These reasons include inappropriate shard settings, uneven segment sizes, unseparated hot and cold data, and persistent connections that are used for Service Load Balancer (SLB) instances and multi-zone architecture. Resolve the issue based on the actual situation. For more information, see Imbalanced loads on a cluster.

> **Notice** Before you resolve the issue, check the specifications of your cluster. If the specifications of your cluster are 1 vCPU and 2 GiB of memory, upgrade the specifications to 2 vCPUs and 4 GiB of memory or higher. The specifications of 1 vCPU and 2 GiB of memory are used only for tests. For more information about how to upgrade the specifications, see Upgrade the configuration of a cluster.

How do I plan resources before I use Elasticsearch, such as cluster specifications, the number of shards, and the size of each shard?

Evaluate the specifications and storage capacity of your Elasticsearch cluster. For more information, see Evaluate specifications and storage capacity. You can purchase an Elasticsearch cluster or upgrade the configuration of the cluster based on the evaluation results.

How do I view the configuration of an Elasticsearch cluster?

You can view the configuration of your Elasticsearch cluster on the Basic Information page of the cluster. For more information, see View basic information of a cluster.

When you use Transport Client to access an Elasticsearch cluster, set the `cluster.name` parameter to the ID of your cluster. For more information, see Transport Client (5.x).

Are services affected when I modify the configuration of an Elasticsearch cluster?

The system restarts the cluster after you modify its configuration. The system uses the rolling restart method to restart a cluster. Before the restart, make sure that the cluster is in the Active state (indicated by the color green), each index has at least one replica shard for each primary shard, and resource usage is not high. For example, the value of NodeCPUUtilization(%) is about 80%, that of NodeHeapMemoryUtilization is about 50%, and that of NodeLoad_1m is less than the number of vCPUs of the current node. If all the conditions are met, the cluster can still provide services during the restart. You can view the resource usage on the Cluster Monitoring page. However, we recommend that you modify the configuration of your cluster during off-peak hours.

Can I change the cloud disk type of an Elasticsearch cluster?

No, if you want to change the cloud disk type of your cluster, purchase another cluster based on your requirements and migrate data from the original cluster to the new cluster. Then, cancel the subscription of or release the original cluster. For more information about how to migrate data, see Configure a shared OSS repository.

Can I convert other types of nodes in an Elasticsearch cluster to warm nodes?

No, the conversion can cause your cluster to be unstable. For more information, see "Hot-Warm" Architecture in Elasticsearch 5.x.

Can I downgrade the specifications of an Elasticsearch cluster? If yes, how do I do?

No, you can scale in your cluster. For more information, see Scale in an Elasticsearch cluster.

In the event of a temporary business surge, how do I modify the configuration of an Elasticsearch cluster to ensure that services run as expected?

We recommend that you add nodes to the cluster when the temporary business surge occurs and remove the nodes after the business surge. For more information, see Upgrade the configuration of a cluster and Scale in an Elasticsearch cluster. For the changes to take effect, the system restarts the cluster. Before the restart, take note of the following items:

- The cluster is in the Active state (indicated by the color green).
- Each index of the cluster has at least one replica shard for each primary shard, and the resource

usage of the cluster is not high. For example, the value of NodeCPUUtilization(%) is about 80%, that of NodeHeapMemoryUtilization is about 50%, and that of NodeLoad_1m is less than the number of vCPUs of the current node. You can view the resource usage on the Cluster Monitoring page of the cluster.

When I upgrade the configuration of an Elasticsearch cluster, the system displays the "UpgradeVersionMustFromConsole" error message. What do I do?

The error message returned because the version change does not meet requirements. You can upgrade clusters only from V6.3.2 to V6.7.0.

How long is required to upgrade the version of an Elasticsearch cluster?

The required time is determined by the data volume, data structure, and cluster specifications. The version upgrade requires about one hour.

Are services affected when I upgrade the version of an Elasticsearch cluster?

When you upgrade the version of an Elasticsearch cluster, you can still read data from or write data to the cluster but cannot make other changes. We recommend that you perform a version upgrade during off-peak hours. For more information about the precautions and procedure for a version upgrade, see Upgrade the version of a cluster.

Can I use the YML configuration file of an Elasticsearch cluster to change the settings of the http.max_content_length and discovery.zen.ping_timeout parameters?

You are not allowed to configure the two parameters. If you want to add these parameters to the configuration file, contact Alibaba Cloud Elasticsearch technical engineers. Before you add the parameters, make sure that the parameter settings are correct and you accept the impact caused by parameter modifications. If the parameter settings are incorrect, the system fails to perform a rolling restart for the cluster.

> **Note** In most cases, you do not need to change the settings of the discovery.zen.ping_timeout, discovery.zen.fd.ping_timeout, discovery.zen.fd.ping_interval, and discovery.zen.fd.ping_retries parameters.

Can I switch the VPC of an Elasticsearch cluster?

No, you can purchase an Elasticsearch cluster in the desired virtual private cloud (VPC) and migrate data from the original cluster to the new cluster. Then, cancel the subscription of or release the original cluster.

Can I change the JVM parameter settings of an Elasticsearch cluster?

Alibaba Cloud Elasticsearch clusters use JVM parameter settings that are recommended by open source Elasticsearch. We recommend that you do not change the settings. By default, JVM heap memory is half of cluster memory.

How do I update dictionary content when I use the IK analysis plug-in?

You can use the standard update or rolling update feature of the IK analysis plug-in to update dictionary content. For more information, see Use the analysis-ik plug-in.

When I use the IK analysis plug-in, the system displays the "ik startOffset" error message. What do I do?

The error message returned because of an Elasticsearch V6.7 bug. You must restart your cluster. For more information, see Restart a cluster or node. We will fix the bug as soon as possible.

The IK dictionary files on my on-premises machine are lost. Can I retrieve them on the cluster management page?

No, you can only delete or update dictionary files on the cluster management page. We recommend that you download the official main and stopword dictionary files. Then, change the tokens in the files to those in your system dictionary file and upload the files to your cluster.

After I update IK dictionaries, how do I apply the new dictionaries to existing data?

You must perform a reindex operation. If indexes are configured with IK tokens, the new dictionaries apply only to new data in these indexes. If you want to apply the new dictionaries to all the data in these indexes, you must perform a reindex operation. For more information, see Recreate indexes by calling the Reindex operation.

Is there a specific threshold for full GC?

Full garbage collection (GC) is used to clean the entire heap memory. Whether full GC is correctly performed needs to be analyzed based on the service latency, heap memory size before full GC, and heap memory size after full GC. The CMS collector starts to collect garbage when the memory usage is 75%. This is because some space is reserved for burst traffic.

Can I remove plug-ins that are not used?

You can remove only some plug-ins. On the Built-in Plug-ins tab of the Plug-ins page of your Elasticsearch cluster, you can view plug-ins that can be removed. If the system displays Remove in the Actions column of a plug-in, the plug-in can be removed. For more information about how to remove a plug-in, see Install and remove a built-in plug-in.

Are the dictionaries provided by the IK analysis plug-in of Alibaba Cloud Elasticsearch the same as those provided by the IK analysis plug-in of open source Elasticsearch?

Yes, for more information, see IK Analysis for Elasticsearch.

Can a custom plug-in access an external network, such as reading dictionary files on GitHub?

No, if you want your Elasticsearch cluster to access external files, upload the files to Alibaba Cloud Object Storage Service (OSS) and connect your Elasticsearch cluster to OSS.

Does a custom plug-in support the rolling update method?

No, if you want a custom plug-in to support this method, configure the plug-in based on the rolling update method of the IK analysis plug-in. For more information, see IK Analysis for Elasticsearch.

How do I configure the analysis-aliws plug-in? What is the format of the dictionary file for this plug-in?

For more information about how to configure the plug-in, see Use the analysis-aliws plug-in.

The dictionary file must meet the following requirements:

- Name: aliws_ext_dict.txt.
- Encoding format: UTF-8.
- Content: Each row contains one word and ends with \n (line break in UNIX or Linux). No whitespace characters are used before or after this word. If the dictionary file is generated in Windows, you must use the dos2unix tool to convert the file and upload the file to your cluster.

What are the differences among Elasticsearch synonyms, IK tokens, and AliNLP tokens?

Token type	Usage	Description	Supported file type	Tokenizer and analyzer
Synonym	You can upload a synonym dictionary file on the Cluster Configuration page of your cluster to enable the cluster to use it.	After you write several synonyms in the file, the system displays all the synonyms when you query one of them.	The synonym dictionary file must be a TXT file encoded in UTF-8.	Custom tokenizer and analyzer
IK token	The IK tokens are used based on the analysis-ik plug-in.	The system splits a paragraph based on the *main.dic* file. If you send a query request that contains one or more split words, the system returns the entire paragraph in the query result. The analysis-ik plug-in also provides a stopword file named *stop.dic*. The query result does not include the stopwords in the *stop.dic* file You can view the dictionary file from the official documentation.	The main and stopword dictionary files must be DIC files encoded in UTF-8.	Tokenizer: • ik_smart • ik_max_word

Token type	Usage	Description	Supported file type	Tokenizer and analyzer
AliNLP token	The AliNLP tokens are used based on the analysis-aliws plug-in.	The analysis-aliws plug-in works in a similar way as the analysis-ik plug-in, but the analysis-aliws plug-in does not provide a separate stopword dictionary file. Stopwords are integrated into the main dictionary file *aliws_ext_dict.txt*. The file is invisible to you. In addition, you are not allowed to customize stopwords.	The dictionary file name must be aliws_ext_dict.txt. The file must be encoded in UTF-8.	• Analyzer: aliws, which does not return function words, function phrases, or symbols • Tokenizer: aliws_tokenizer

Can I specify a retention period for the .security indexes of an Elasticsearch cluster?

No, Elasticsearch does not automatically delete expired indexes. You must manually delete the expired .security indexes. For more information, see Delete an index.

I can view Elasticsearch cluster logs that are generated only over the last seven days. How do I view more logs?

You can call the ListSearchLog operation to obtain all logs that you require. For more information, see ListSearchLog.

I cannot view the search and update logs of an Elasticsearch cluster. What do I do?

You can configure slow logs and reduce the timestamp precision of log entries. For more information, see Configure slow logs.

How do I configure and view the slow logs of an Elasticsearch cluster?

By default, Elasticsearch logs only read and write operations that require 5 to 10 seconds to complete as slow logs. You can log on to the Kibana console of the cluster and run the related command to reduce the timestamp precision of log entries. This helps capture more logs. For more information, see Configure slow logs.

> **Note** You are not allowed to change the format of slow logs.

How do I obtain the slow logs of an Elasticsearch cluster on a regular basis?

You can call the ListSearchLog operation on a regular basis to obtain the slow logs of your cluster. For more information, see ListSearchLog.

Can I restore data from the snapshots of an Elasticsearch cluster to an Elasticsearch cluster of a different version?

For automatic snapshots, you can restore data from the snapshots only to the original cluster. For more information, see Create automatic snapshots and restore data from automatic snapshots.

For manual snapshots, you can restore data from the snapshots to a cluster other than the original cluster. We recommend that the versions of the destination cluster and original cluster must be the same. If the versions are different, compatibility issues may occur. For more information, see Commands for creating snapshots and restoring data.

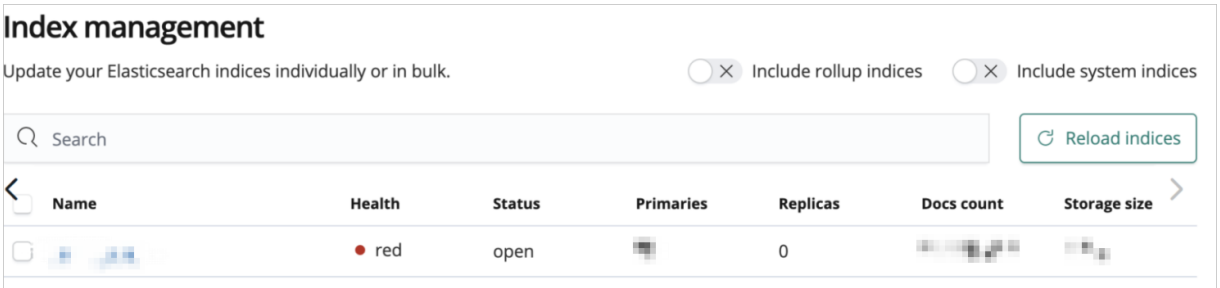
When I back up data for an Elasticsearch cluster, the system displays a message indicating that the cluster is unhealthy. What do I do?

When an Elasticsearch cluster is unhealthy, you cannot use the Auto Snapshot feature and specify shared OSS repositories. You can purchase an OSS bucket that resides in the same region as your Elasticsearch cluster. Then, create an OSS repository and manually create snapshots. For more information, see Commands for creating snapshots and restoring data.

I enable the Auto Snapshot feature but do not specify shared OSS repositories for an Elasticsearch cluster. Are snapshots created?

Elasticsearch provides an OSS bucket for your cluster by default. You can log on to the Kibana console of your cluster and run the `GET _snapshot/aliyun_auto_snapshot/_all` command to obtain automatic snapshots. For more information about how to log on to the Kibana console, see Log on to the Kibana console.

When I restore data from snapshots, the destination Elasticsearch cluster displays a message. This message indicates that shards are abnormal. After I run the `POST /_cluster/reroute?retry_failed=true` command to reroute the shards, the issue persists. What do I do?



Delete the problematic index and call the _restore operation to restore it. You must add the max_restore_bytes_per_sec parameter to the request. This parameter is used to limit the restoration rate. The default value of this parameter is 40mb. This value indicates that the index is restored at a speed of 40 MB per second.

```
POST /_snapshot/aliyun_snapshot_from_instanceId/es-cn-instanceId_datetime/_restore
{
  "indices": "myIndex",
  "settings": {
  "max_restore_bytes_per_sec" : "150mb"
  }
}
```

Note You can also add the following parameters:

- compress: specifies whether to enable data compression. Default value: true.
- max_snapshot_bytes_per_sec: specifies the snapshot creation rate of each node. Default value: 40mb.

Can I export data from an Elasticsearch cluster to my on-premises machine?

Yes, Elasticsearch provides the snapshot feature. For more information, see View the snapshot feature. You can create and store snapshots in OSS and download objects from OSS. For more information, see Download objects.

How do I use the email notification feature of X-Pack Watcher?

You can configure specific actions for X-Pack Watcher. For more information, see Watcher settings in Elasticsearch.

Notice X-Pack Watcher of Elasticsearch cannot directly access the Internet. You must use the internal endpoint of an Elasticsearch cluster to access the Internet. Therefore, you must create an ECS instance that can access both the Internet and the Elasticsearch cluster. Then, use the ECS instance as a proxy to perform actions. For more information, see Configure X-Pack Watcher.

What do I do if the system reports an alert indicating that memory cannot be allocated to the garbage collector?

Possible causes include heavy loads, high query QPS, or large amounts of data to write. Troubleshoot the issue based on the following instructions:

- Heavy loads: For more information, see High disk usage and read-only indexes.
- High query QPS or large amounts of data to write: We recommend that you install the aliyun-qos plug-in on your Elasticsearch cluster to implement read/write throttling. For more information, see Use the aliyun-qos plug-in.

  Note For image searches, we recommend that you install the aliyun-knn plug-in on your cluster and plan your cluster and indexes. For more information, see Use the aliyun-knn plug-in for vector search.

How do I use a client to access an Alibaba Cloud Elasticsearch cluster? What is the difference between access to an Alibaba Cloud Elasticsearch cluster and access to an open source Elasticsearch cluster?

Access an Alibaba Cloud Elasticsearch cluster by using its internal or public endpoint. Access an open source Elasticsearch cluster by using its address. For more information, see Access Alibaba Cloud Elasticsearch by using a client.

When I use a client to access an Elasticsearch cluster, can I disable the basic authentication feature?

No, the basic authentication feature is a Kibana authentication mechanism provided by the built-in Elasticsearch plug-in X-Pack. Therefore, you cannot disable the feature.

I purchase an ECS instance that resides in the same VPC as but different zone from an Elasticsearch cluster. Can I use the ECS instance to access the Elasticsearch cluster from an internal network?

Yes, you can use an ECS instance to access an Elasticsearch cluster from an internal network if they reside in the same VPC.

How do I access an Elasticsearch cluster from the Internet?

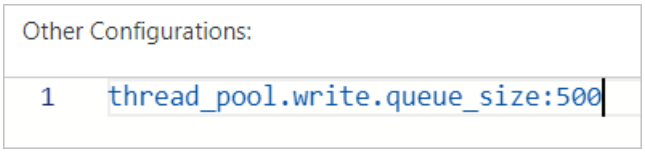
You can access the cluster from the Internet by using its public endpoint and configuring a public IP address whitelist. For more information, see Configure a whitelist to access an Elasticsearch cluster over the Internet or a VPC. When you access the cluster, you must configure parameters, such as the domain name, username, and password. For more information, see Access Alibaba Cloud Elasticsearch by using a client.

13.6. FAQ about open source Elasticsearch

This topic provides answers to some commonly asked questions about open source Elasticsearch.

How do I configure the thread pool size for indexes?

In the YML configuration file of your Elasticsearch cluster, specify the thread_pool.write.queue_size parameter to configure the thread pool size. For more information, see Modify the YML configuration.



Notice If the version of your Elasticsearch cluster is earlier than 6.X, use the thread_pool.index.queue_size parameter to configure the thread pool size.

What do I do if OOM occurs?

Run the following command to clear the cache. Then, analyze the cause, and upgrade the configuration of your Elasticsearch cluster or adjust your business. For more information about how to upgrade the cluster configuration, see Upgrade the configuration of a cluster.

```
curl -u elastic:passwd -XPOST "localhost:9200/<index-name>/_cache/clear?pretty"
```

How do I manually manage a shard?

Use the reroute operation or Cerebro. For more information, see Cluster reroute API and Cerebro.

What are the cache clearing policies for Elasticsearch?

Elasticsearch supports the following cache clearing policies:

- Clear the cache of all indexes

  ```
  curl localhost:9200/_cache/clear?pretty
  ```

- Clear the cache of a single index

  ```
  curl localhost:9200/<index_name>/_cache/clear?pretty
  ```

- Clear the cache of multiple indexes

  ```
  curl localhost:9200/<index_name1>,<index_name2>,<index_name3>/_cache/clear?pretty
  ```

How do I reroute index shards?

If some shards are lost or inappropriately allocated, you can run the following command to reroute the shards:

```
curl -XPOST 'localhost:9200/_cluster/reroute' -d '{
  "commands" : [ {
    "move" :
      {
       "index" : "test", "shard" : 0,
       "from_node" : "node1", "to_node" : "node2"
      }
    },
    {
     "allocate" : {
       "index" : "test", "shard" : 1, "node" : "node3"
     }
    }
  ]
}'
```

When I query an index, the system displays the `statusCode: 500` error message. What do I do?

We recommend that you use a third-party plug-in, such as Cerebro, to query the index.

- If the query succeeds, the issue is caused by an invalid index name. In this case, modify the index name. An index name can contain only letters, underscores (_), and digits.
- If the query fails, the issue is caused by an error in the index or your cluster. In this case, check whether your cluster stores the index and runs in a normal state.

How do I modify the `auto_create_index` parameter?

Run the following command:

```
PUT /_cluster/settings
{
  "persistent" : {
    "action": {
     "auto_create_index": "false"
    }
  }
}
```

Notice The default value of the `auto_create_index` parameter is false. This value indicates that the system does not automatically create indexes. In most cases, we recommend that you do not modify this parameter. Otherwise, excessive indexes are created, and the mappings or settings of the indexes do not meet requirements.

How long is required to create a snapshot that will be stored in OSS?

If the number of shards, memory usage, disk usage, and CPU utilization of your cluster are normal, about 30 minutes are required to create a snapshot for 80 GB of index data.

How do I specify the number of shards when I create an index?

You can divide the total data size by the data size of each shard to obtain the number of shards. We recommend that you limit the data size of each shard to 30 GB. If the data size of each shard exceeds 50 GB, query performance is severely affected.

You can appropriately increase the number of shards to speed up index creation. The query performance is affected no matter whether the number of shards is too small or too large.

- Shards are stored on different nodes. If a large number of shards are configured, more files need to be opened, and more interactions are required among the nodes. This decreases query performance.
- If a small number of shards are configured, each shard stores more data. This also decreases query performance.

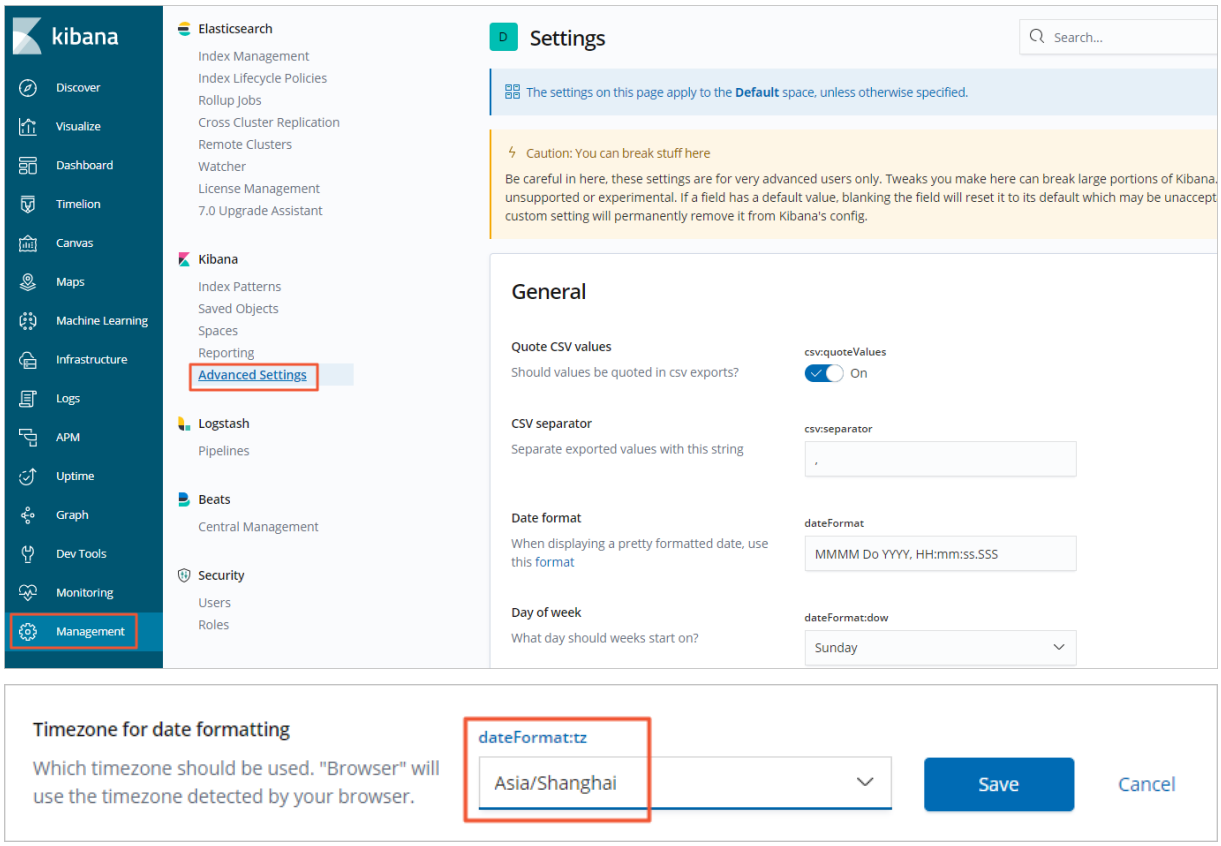
When I use the elasticsearch-repository-oss plug-in to migrate data from a self-managed Elasticsearch cluster, the system displays the following error message. What do I do?

Error message: `ERROR: This plugin was built with an older plugin structure. Contact the plugin author to remove the intermediate "elasticsearch" directory within the plugin zip` .

Change the name of the ZIP plug-in package from elasticsearch to elasticsearch-repository-oss, and copy the package to the plugins directory.

How do I adjust the Elasticsearch server time?

You can change the time zone in the Kibana console, as shown in the following figure. In this example, an Elasticsearch 6.7.0 cluster is used.



What type of data can I perform Elasticsearch term queries on?

Term queries are word-level queries that can be performed on structured data, such as numbers, dates, and keywords other than text.

> **Note** When you perform a full-text query, the system splits words in the text. When you perform a word-level query, the system directly searches the inverted indexes that contain the related fields. Word-level queries are generally performed on fields of a numeric data type or the DATE data type.

What are the precautions for using aliases in Elasticsearch?

The total number of shards for indexes that have the same alias must be less than 1,024.

What do I do if the following error message is returned during a query?

Error message: `"type": "too_many_buckets_exception", "reason": "Trying to create too many buckets. Must be less than or equal to: [10000] but was [10001]`

You can change the value of the size parameter for bucket aggregations. For more information, see Limit the number of buckets that can be created in an aggregation. You can also resolve this issue based on the instructions provided in Increasing max_buckets for specific Visualizations.

How do I delete multiple indexes at a time?

By default, Elasticsearch does not allow you to delete multiple indexes at a time. You can run the following command to enable the deletion. Then, use a wildcard to delete multiple indexes at a time.

```
PUT /_cluster/settings
{
 "persistent": {
   "action.destructive_requires_name": false
 }
}
```