

# Alibaba Cloud E-MapReduce

## Errors and Troubleshooting

Issue: 20190219

## 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 due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall 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 the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. 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 bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, 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 content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, 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 the Alibaba Cloud website, product programs, or content 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 contact Alibaba Cloud directly if you discover any errors in this document.



## Generic conventions

Table -1: Style conventions

| Style   | Description  | Example  |
|---|--|--|
|  | This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results. |  <b>Danger:</b><br>Resetting will result in the loss of user configuration data.                                    |
|  | This warning information indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.  |  <b>Warning:</b><br>Restarting will cause business interruption. About 10 minutes are required to restore business. |
|  | This indicates warning information, supplementary instructions, and other content that the user must understand.                           |  <b>Notice:</b><br>Take the necessary precautions to save exported data containing sensitive information.           |
|   | This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.                       |  <b>Note:</b><br>You can use Ctrl + A to select all files.  |
| >   | Multi-level menu cascade.  | Settings > Network > Set network type  |
| <b>Bold</b>   | It is used for buttons, menus, page names, and other UI elements.  | Click <b>OK</b> .  |
| Courier font  | It is used for commands.   | Run the <code>cd /d C:/windows</code> command to enter the Windows system folder.  |
| <i>Italics</i>  | It is used for parameters and variables.   | <code>bae log list --instanceid <i>Instance_ID</i></code>  |
| [ ] or [a b]  | It indicates that it is an optional value, and only one item can be selected.  | <code>ipconfig [-all -t]</code>  |

| Style                                 | Description  | Example                            |
|---------------------------------------|--|------------------------------------|
| <code>{}</code> or <code>{a b}</code> | It indicates that it is a required value, and only one item can be selected. | <code>swich {stand   slave}</code> |



# Contents

---

|  |    |
|--|----|
| Legal disclaimer.....                    | I  |
| Generic conventions.....                 | I  |
| 1 Configure cluster ports.....           | 1  |
| 2 Use execution plans.....               | 4  |
| 3 Appendix.....                          | 6  |
| 3.1 Cluster Service Management page..... | 6  |
| 3.2 Notes on E-MapReduce versions.....   | 7  |
| 3.3 Error code list.....                 | 8  |
| 3.4 Status list.....                     | 10 |

# 1 Configure cluster ports

---

## Hadoop HDFS

| Service  | Limit | Port number | Access requirement | Parameter                                     | Description                                      |
|----------|-------|-------------|--------------------|---|--|
| NameNode | -     | 9000        | External           | fs.default.name or fs.defaultFS               | fs.default.name is expired but still available   |
| NameNode | -     | 50070       | External           | dfs.http.address or dfs.namenode.http-address | dfs.http.address is expired but still available. |

## Hadoop YARN (MRv2)

| Service           | Limit | Port number | Access requirement | Parameter                                     | Description |
|-------------------|-------|-------------|--------------------|---|-------------|
| JobHistory Server | -     | 10020       | Internal           | mapreduce.jobhistory.address                  | -           |
| JobHistory Server | -     | 19888       | External           | mapreduce.jobhistory.webapp.address           | -           |
| ResourceMa nager  | -     | 8025        | Internal           | yarn.resourcemanager.resource-tracker.address | -           |
| ResourceMa nager  | -     | 8032        | Internal           | yarn.resourcemanager.address                  | -           |

| Service             | Limit | Port number | Access requirement | Parameter  | Description |
|---------------------|-------|-------------|--------------------|--|-------------|
| ResourceMa<br>nager | -     | 8030        | Internal           | yarn.<br>resourcema<br>nager.<br>scheduler.<br>address | -           |
| ResourceMa<br>nager | -     | 8088        | Internal           | yarn.<br>resourcema<br>nager.<br>webapp.<br>address    | -           |

### Hadoop MapReduce (MRv1)

| Service    | Limit | Port number | Access requirement | Parameter                            | Description |
|------------|-------|-------------|--------------------|--------------------------------------|-------------|
| JobTracker | -     | 8021        | External           | mapreduce<br>.jobtracker.<br>address | -           |

### Hadoop HBase

| Service           | Limit | Port number | Access requirement | Parameter                            | Description |
|-------------------|-------|-------------|--------------------|--------------------------------------|-------------|
| HMaster           | -     | 16000       | Internal           | hbase.<br>master.port                | -           |
| HMaster           | -     | 16010       | External           | hbase.<br>master.info.<br>port       | -           |
| HRegionSer<br>ver | -     | 16020       | Internal           | hbase.<br>regionserver<br>.port      | -           |
| HRegionSer<br>ver | -     | 16030       | External           | hbase.<br>regionserver<br>.info.port | -           |
| ThriftServer      | -     | 9099        | External           | -                                    | -           |

**Hadoop Spark**

| Service      | Limit | Port number | Access requirement | Parameter | Description |
|--------------|-------|-------------|--------------------|-----------|-------------|
| SparkHistory | -     | 18080       | External           | -         | -           |

**Storm**

| Service  | Port number | Parameter |
|----------|-------------|-----------|
| Storm UI | 9999        | ui_port   |

**Druid**

| Service       | Port number | Parameter                           |
|---------------|-------------|-------------------------------------|
| overlord      | 18090       | overlord.runtime -> druid.port      |
| coordinator   | 18081       | coordinator.runtime -> druid.port   |
| middleManager | 18091       | middleManager.runtime -> druid.port |
| historical    | 18083       | historical.runtime -> druid.port    |

## 2 Use execution plans

---

### Apply for high-configuration instances

If you have not activated a high-configuration instance, an error will occur when you use high-configuration instances to create a cluster, and the following error message appears:

The specified InstanceType is not authorized for usage.

Click [here](#) to submit a ticket and activate high-configuration instances.

### Use security groups

You need to use security groups that are created in EMR when creating clusters in EMR. This is because only port 22 of the cluster in EMR is accessible. We recommend that you sort your existing instances into different security groups based on their functions. For example, the security group of EMR is "EMR-security group" and you can name your existing security group "User-security group." Each security group applies its own access control based on your needs. If it is necessary to bind the security groups with the cluster that has been created, follow these steps:

- Add an EMR cluster to the existing security group

Click Details. Security groups related to all ECS instances are displayed. In the ECS console, click the Security Group tab in the lower-left corner, find the security group "EMR-security group". Click Manage Instance. ECS instance names starting with emr-xxx are displayed. These are the corresponding ECS instances in the EMR cluster. Select all of these instances, and click Move to Security Group in the upper-right corner to move these instances to another security group.

- Add the existing cluster into the "EMR-security group"

Find the security group in which the existing cluster is located. Repeat the preceding operations, and move the cluster to the "EMR-security group." Select the instances that are not used by the cluster in the ECS console and move them to the "EMR-security group" by using the batch operations.

- Rules of security groups

The security group rules are subject to the OR relationship when an ECS instance is in several different security groups. For example, only port 22 of EMR security

is accessible while all ports of "User-security group" are accessible. When an EMR cluster is added into "User-security group", all ports of instances in EMR open are accessible. Note the following rule:



**Notice:**

When setting up security group rules, make sure that you restrict access by IP address range. Do not set the IP range to 0.0.0.0 to avoid attacks.

### Execution plan FAQs

- Edit an execution plan.

You can edit execution plans that are not in the running or scheduling status. If you cannot click the edit button, confirm the status of the execution plan and try again

- Run an execution plan.

If you set the scheduling mode to Execute immediately when creating an execution plan, the plan is automatically executed after it is created. If it is an existing execution plan, you need to manually run the execution plan. The execution plan is not immediately run after creation.

- Periodical execution time.

The start time of a periodical execution cluster indicates the time when the execution plan starts to run. The time is accurate to minutes. The schedule cycle indicates the interval between two executions since the start time. As shown in the following example:

The first run is at 14:30:00, December 01, 2015 and the second run is at 14:30:00, December 02, 2015. The execution plan is run once a day.

If the current time is later than the time you have scheduled, then the latest time for scheduling is 14:30:00, December 01, 2015.

**Example:**

If the current time is 09:30, December 02, 2015, then the latest time for scheduling is 10:00:00, December 02, 2015, which is based on the scheduling rule. The first run starts at this time.

## 3 Appendix

---

### 3.1 Cluster Service Management page

You can purchase a one vCPU 2 GB ECS instance that runs the Ubuntu system and deploy the instance in a VPC network. You can use this instance as a management client to access the management pages.

The following table lists the endpoints of services in the cluster.

| Software        | Service              | Endpoint  |
|-----------------|----------------------|---|
| <b>Hadoop</b>   |                      |   |
|                 | yarn resourcemanager | masternode1_private_ip:8088,masternode2_private_ip:8088   |
|                 | jobhistory           | masternode1_private_ip:19888                              |
|                 | timeline server      | masternode1_private_ip:8188                               |
|                 | hdfs                 | masternode1_private_ip:50070,masternode2_private_ip:50070 |
| <b>Spark</b>    |                      |   |
|                 | spark ui             | masternode1_private_ip:4040                               |
|                 | history              | masternode1_private_ip:18080                              |
| <b>Tez</b>      |                      |   |
|                 | tez-ui               | masternode1_private_ip:8090/tez-ui2                       |
| <b>Hue</b>      |                      |   |
|                 | hue                  | masternode1_private_ip:8888                               |
| <b>Zeppelin</b> |                      |   |

| Software | Service  | Endpoint                                |
|----------|----------|---|
|          | zeppelin | masternode1_private_ip:<br>8080         |
| Hbase    |          |   |
|          | hbase    | masternode1_private_ip:<br>16010        |
| Presto   |          |   |
|          | Presto   | masternode1_private_ip:<br>9090         |
| Oozie    |          |   |
|          | Oozie    | masternode1_private_ip:<br>11000        |
| Ganglia  |          |   |
|          | Ganglia  | masternode1_private_ip:<br>8085/ganglia |

## 3.2 Notes on E-MapReduce versions

- EMR is regularly updated.
- Versions of software installed within each EMR version are fixed. Currently, EMR does not support choosing different versions of software. We recommend that you do not manually change the versions of software. For example, Hadoop 3.6.0 and Spark 1.4.1 are installed in EMR V1.0.
- If you have selected a version and created a cluster, the version that the cluster uses will not be automatically updated. If you select V1.0, the Hadoop remains at V2.6.0 and Spark remains at V1.4.1. When EMR is updated to V1.1, Hadoop is updated to V2.7.0 and Spark is updated to V1.5.0. These updates do no effect the clusters that you have created. Only new clusters use new mirror images.
- When updating the cluster version, for example, from V1.0 to V1.1, test the jobs in the news software environment, to see if they run successfully and to avoid exceptions caused by incompatibility and a change of software environment.

### 3.3 Error code list

#### Common error codes

| Error code | Error message  |
|------------|--|
| 4001       | The error message returned when the request parameter is invalid, for example , the parameter is missing or the format of the parameter is invalid.  |
| 4005       | The error message returned when you are not authorized to access resources of other users.   |
| 4006       | The error message returned when the cluster is in the Abnormal status and the job cannot be submitted. Check whether the cluster associated with the execution plan has been released.   |
| 4007       | The error message returned when the name of the security group is empty.   |
| 4009       | The error message returned when your account has an overdue payment or is suspended. Check the status of your account.   |
| 4011       | The error message returned when the cluster is in the Abnormal status and cannot be scheduled. Check whether the cluster associated with the execution plan has been released.   |
| 5012       | The error message returned when the number of security groups you can create has exceeded the upper limit. Go to the <a href="#">security group page</a> and delete security groups that are not in use.   |
| 5038       | The error message returned when the job is in a running or pending execution plan and cannot be modified. You can modify the job only after the associated execution plan has been successfully executed. You can clone the job, then modify and use the cloned job. |

| Error code | Error message   |
|------------|---|
| 5039       | The error message returned when you fail to lock the cluster role. You must have certain permissions to use EMR. For <a href="#">role authorization</a> , click <a href="#">here</a> to create cluster roles. |
| 5050       | The error message returned when accessing the database. Try again later.  |
| 6002       | The error message returned when cluster updating failure occurs.  |
| 8002       | The error message returned when you are not authorized to perform the specified operation. Click <a href="#">RAM</a> to apply for authorization.  |
| 8003       | The error message returned when you are not authorized to perform the PassRole operation. Click <a href="#">RAM</a> to apply for authorization.   |
| 9006       | The error message returned when the ID of the cluster does not exist. You need to verify the ID.  |
| 9007       | The error message returned when the password used to log on to the master node is invalid. The password must be 8-30 characters in length and can contain uppercase letters, lowercase letters, and numbers.  |

## ECS-related errors

| Error message   | Description   |
|---|---|
| The specified InstanceType is not authorized for use. | You have not applied for the specified types of instances that are used to create clusters. You can <a href="#">apply for high-configuration instances</a> on the ECS buy page. |
| No zone or cluster resources are available .          | No ECS resources are available in this zone.  |

## 3.4 Status list

### Cluster status list



**Note:**

You can view the cluster status in the cluster list or on the cluster details page.

| Status    | Status code   | Description  |
|-----------|---------------|--|
| Creating  | CREATING      | The cluster is being created. The creation task includes two stages : creating physical ECS machines and activating Spark clusters. It takes a moment for the clusters to start running.                                   |
| Failed    | CREATE_FAILED | An exception occurred during creation. The ECS instance that you have created automatically rolls back. You can click the question mark (?) to the right of the status on the cluster list page to view exception details. |
| Running   | RUNNING       | The computing cluster is running.  |
| Idle      | IDLE          | The cluster is not running any execution plan.   |
| Releasing | RELEASING     | Click Release in the status list to set the cluster to this status. This status indicates that the cluster is in the releasing process . It may take a moment to complete this process.                                    |

| Status         | Status code    | Description  |
|----------------|----------------|--|
| Release Failed | RELEASE_FAILED | An exception occurred when releasing the cluster . You can click the question mark (?) to the right of the status on the cluster list page to view the exception details. When the cluster is in this status, click Release to release the cluster again . |
| Released       | RELEASED       | The computing cluster and the ECS instance that hosts the computing cluster have been released.  |
| Abnormal       | ABNORMAL       | Unrecoverable errors occurred on one or more nodes in the computing cluster. Click Release to release the cluster.   |

### Job status list



Note:

View job status in the job status list

| Status     | Description   |
|------------|---|
| Ready      | The creation information is complete, correct, and successfully saved. The job is ready to be added to the submission queue. It may take a moment for the job to change its status to Submitting. |
| Submitting | The job is in the submission queue of the computing cluster. It has not been submitted to the computing cluster.  |
| Failed     | An exception occurred when submitting the job to the computing cluster. You need to clone and submit the job if you want to submit this job again.  |

| Status    | Description  |
|-----------|--|
| Running   | The job is running in the cluster. Wait a moment and click the corresponding log button in the job list to view output log entries in real time. |
| Succeeded | The job has been successfully executed in the cluster. Click the corresponding log button to view the log entry.                                 |
| Failed    | An exception occurred when executing the job. Click the corresponding log button to view the log entry.  |