Alibaba Cloud E-MapReduce

FAQ

Issue: 20190115

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 intellectu

al 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.	 **Danger:** 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.	 **Warning:** 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.	 **Notice:** Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructio ns, best practices, tips, and other content that is good to know for the user.	 **Note:** You can use **Ctrl** + **A** to select all files.
>	Multi-level menu cascade.	**Settings** > **Network** > **Set network type**
Bold	It is used for buttons, menus, page names, and other UI elements.	Click **OK**.
`Courier font`	It is used for commands.	Run the `cd /d C:/windows` command to enter the Windows system folder.
`Italics`	It is used for parameters and variables.	`bae log list --instanceid` *`Instance_ID`*
[] or [a|b]	It indicates that it is a optional value, and only one item can be selected.	`ipconfig` *`[-all|-t]`*
{} or {a|b}	It indicates that it is a required value, and only one item can be selected.	`swich` *`{stand | slave}`*

Contents

1 Frequently asked questions about Alibaba Cloud Elastic MapReduce (EMR)

Q: What is the difference between a job and an execution plan?

A: In EMR, two steps are required to run a job:

- Create a job

  In EMR, creating a “job” creates a "job running configuration" which cannot be run directly. If you have created a "job" in EMR, you have created a "configuration about how to run the job." The configuration contains the Job JAR to be run for the job, the input and output addresses of data and certain running parameters. After you create such a configuration and provide a name for it, a job is defined. However, when you want to debug the running job, the execution plan is required.

- Create an execution plan

  The execution plan associates the job with the cluster. Through the execution plan, we can combine multiple jobs into a job sequence and prepare a running cluster for the job. The execution can also automatically create a temporary cluster or associate the job with an existing cluster. The execution plan also helps to configure a periodical execution plan for the job sequence and automatically releases the cluster after the task is completed. You can also view the execution status and log entries on the execution record list.

Q: How do I view the job log entries?

A: In the EMR system, the system has uploaded the job running log entries to OSS by jobid. The path is set by you when you create the cluster. You can view the job log entries on the webpage. If you log on to the master node to submit jobs or run scripts, the logs can be determined by your script.

Q: How do I log on to the core node?

A: Use the following steps:

1. Switch to the hadoop account on the master node.

```
su hadoop
```

2. Log on to the corresponding core node with password-free SSH authentication.

```
ssh emr-worker-1
```

3. Get root privileges through the sudo command.

```
sudo vi /etc/hosts
```

Q: How to view logs on OSS?

A: users can also find all the log files directly from the OSS and download them. However, since OSS does not directly view the logs, it can be much more difficult to use it. If you have enabled logging and specified a log location for OSS, how can you find the job log entries? For example, this path of *OSS://mybucket/emr/spark*.

1. 1. Go to the execution plan page, find the corresponding execution plan and click **View Records** to enter the execution log page.
2. 2. Find the specific execution log entry on the execution log page, such as the last execution log entry. Click the corresponding **Execution Cluster** to view the ID of the execution cluster.
3. 3. Then search for the cluster ID directory *OSS://mybucket/emr/spark/*under *OSS://mybucket/emr/spark directory*.
4. 4. There will be multiple directories under *OSS://mybucket/emr/spark/cluster ID/jobs* based on the execution ID of the job, and each directory stores the running log file of the job.

Q: What is the timing policy of the cluster, execution plan, and running job?

A: Three following timing policies are available:

- The timing policy of the cluster

  In the cluster list, you can see the running time of every cluster. The formula for calculating the running time is: Running time = Time when the cluster is released - Time when the cluster is established. Once a cluster has been created, the timing starts until the end of the life cycle of the cluster.

- The timing policy of the execution plan

  In the running log list of the execution plan, you can see the running time of every execution plan, and the timing policy can be summarized into two situations:

- 1. If the execution plan is executed on demand, the running process of every execution log involves cluster creation, job submission, job running, and cluster release. The formula for calculating the on-demand execution plan is: Running time = The time when the cluster is created + The total time used for running all the jobs in the execution plan + The time when the cluster is released.
- 2. If the execution plan is associated with an existing cluster, the entire execution cycle does not involve the creation and release of a cluster. The running time is the total time used for running all the jobs in the execution plan.

• The timing policy of the running job

The specified job refers to the job assigned to the execution plan. Click the View Job List to the right of the running log of every execution plan to view the job. The formula for calculating the runtime of every job is: Running time = The actual time when the job stops running - The actual time when the job starts running. The specified time when the job starts or stops refers to the time when the job is scheduled to run or stopped running by the Spark or Hadoop cluster.

Q: Why are there no security groups available the first time I run an execution plan?

A: For some security reasons, you cannot directly use an existing security group as an EMR security group. So you are not able to select an available security group if you have not created a security group in EMR We recommend that you create an on-demand cluster for testing. You can create a new EMR security group when creating the cluster. You can set up a scheduling cycle to schedule execution plans after the test is passed. The security groups that you have previously created are also available.

Q: The error message "java.lang.RuntimeException.Parse responsed failed: '<!DOCTYPE html>…'" is returned when reading and writing MaxCompute .

A: Check whether the odps tunnel endpoint is correct. This error occurs if it is wrong.

Q: TPS conflict occurs when multiple consumer IDs consume the same topic.

A: This topic may have been created in the beta test or other environments, causing the consumption data of certain consumer groups conflicted. Report the corresponding topic and consumer ID to ONS for processing by submitting a ticket.

Q: Can I view job log entries on the worker nodes in EMR?

A: Yes. Prerequisite: Click **Save Log** when creating a cluster. How to view the log entries: Choose **Execution plan list** > **More** > **Running log** > **Running log** > **View job list** > **Job list** > **workers log**.

Q: Why do the external tables created in Hive contain no data?

A: For example:

```
CREATE EXTERNAL TABLE storage_log(content STRING) PARTITIONED BY (ds
STRING)
    ROW FORMAT DELIMITED
    FIELDS TERMINATED BY '\t'
    STORED AS TEXTFILE
    LOCATION 'oss://log-124531712/biz-logs/airtake/pro/storage';
    hive> select * from storage_log;
    OK
    Time taken: 0.3 seconds
    The external table that you have created contains no data.
```

Hive does not automatically bind itself with the partition directory of the specified directory. You need to bind them manually, for example:

```
alter table storage_log add partition(ds=123);

                                                OK
    Time taken: 0.137 seconds
    hive> select * from storage_log;
    OK
    abcd    123
    efgh    123
```

Q: Why does the Spark Streaming job stop with no specified reason?

A: First, check whether the current Spark version is earlier than Version 1.6. Spark Version 1.6 repaired a memory leak bug. This bug can cause container memory overuse, which terminates jobs. This error may be one of the causes and this does not mean that Spark 1.6 does not have any issues. In addition, you must check your code to optimize memory usage.

Q: Why is the job still in the running status in the EMR console when the Spark Streaming job has ended?

A: Check whether the running mode of the Spark Streaming job is yarn-client. If yes, set it to yarn -cluster. Errors occur in EMR when it is monitoring the status of Spark Streaming jobs that are in the yarn-client mode. This issue will be repaired as soon as possible.

Q: Why does the error message "error: could not find or load main class" return?

A: Check whether the path protocol header of the Job Jar is `ossref` in the job configuration. If not, change it to `ossref`.

Q: How are machines in a cluster responsible for different tasks?

A: The EMR contains a master node and multiple slave or worker nodes. The master node does not perform data storage or computing tasks. The slave node is used for data storage

and computing tasks. For example, in a cluster with three four-core 8 GB machines, one of the machines serves as the master node and the other two serve as the slave nodes. As a result, the available computing resources of the cluster are two four-core 8 GB machines.

Q: How to use the local shared library in MR jobs?

A: You can use multiple methods, including the following example: Modify the *mapred-site.xml* file, for example:

```
<property>
    <name>mapred.child.java.opts</name>
    <value>-Xmx1024m -Djava.library.path=/usr/local/share/</value>
  </property>
  <property>
    <name>mapreduce.admin.user.env</name>
    <value>LD_LIBRARY_PATH=$HADOOP_COMMON_HOME/lib/native:/usr/local/
lib</value>
  </property>
```

Add the library file as needed and you can use the local shared library.

Q: How can I specify the OSS data source file path in the MR or Spark job?

A: You can use the following `OSS URL: oss://[accessKeyId:accessKeySecret@]bucket[.endpoint]/object/path`.

This URI is used for specifying input/output data sources in the job, and is similar to `hdfs://`. Follow this procedure when you perform operations on OSS data:

- (Recommended) EMR provides MetaService, which allows you to access OSS data without AssessKey, and directly write to the OSS path: *// bucket/Object/path*.
- (Not recommended) You can configure AccessKeyId, AccessKeySecret, and endpoint to Configuration (SparkConf in Spark jobs, Configuration in MR jobs), or you can directly specify AccessKeyId, AccessKeySecret, and endpoint in the URL. For more information, see the *Development preparation* section.

Q: Why does Spark SQL return an error message "Exception in thread "main" java.sql.SQLException: No suitable driver has been found for jdbc:mysql:xxx"?

A:

- 1. Earlier versions of mysql-connector-java may encounter such issues. Update mysql-connector-java to the latest version.
- 2. In the job parameters, use*–driver-class-path ossref://bucket/…/mysql-connector-java-[version].jar* to load *mysql-connector-java* package. This issue may also occur if you directly package *mysql-connector-java*into the Job Jar.

Q: Why is the error message "Invalid authorization specification, message from server: ip not in whitelist" returned when Spark SQL is connected with RDS?

A: Check the RDS whitelist settings and add the internal network IP addresses of the cluster machines to the RDS whitelist.

Q: Notes on creating a cluster based on low-configuration machines.

A:

- If you choose two-core 4 GB machines for the master node, the memory of the master node is heavily loaded. This may cause insufficient memory issues. We recommend that you expand the memory capacity of the master node.
- If you choose two-core 4 GB machines for the slave nodes, adjust the parameters when you run MR or Hive jobs. For MR jobs, add the **`-D yarn.app.mapreduce.am.resource.mb =1024`** parameter. For Hive jobs, add the **`set yarn.app.mapreduce.am.resource.mb= 1024`** parameter. This can prevent jobs to be suspended.

Q: Why is the error message "Failed with exception java.io.IOException:org.apache.parquet.io.ParquetDecodingException: Can not read value at 0 in block -1 in file hdfs://…/…/part-00000-xxx.snappy.parquet" returned when Hive or Impala jobs reads Parquet tables that are imported by Spark SQL ?

A: Hive and Spark SQL use different conversion methods on decimal types to write Parquet. This may cause Hive to fail to correctly read the data imported by Spark SQL. If Hive or Impala needs to use data that has been imported using Spark SQL, we recommend that you add the **`spark. sql.parquet.writeLegacyFormat=true`** parameter and then import data again.

Q: How does Beeline access Kerberos security clusters?

A:

- HA cluster (discovery mode)

```
! connect jdbc:hive2://emr-header-1:2181,emr-header-2:2181,emr-
header-3:2181/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=
hiveserver2;principal=hive/_HOST@EMR.${clusterId). COM
```

- HA cluster (directly connected to a machine)

Connect to emr-header-1.

```
! connect jdbc:hive2://emr-header-1:10000/;principal=hive/emr-header
-1@EMR.${clusterId}. COM
```

Connect to emr-header-2.

```
! connect jdbc:hive2://emr-header-2:10000/;principal=hive/emr-header
-2@EMR.${clusterId}. COM
```

- Non-HA cluster

```
! connect jdbc:hive2://emr-header-1:10000/;principal=hive/emr-header
-1@EMR.${clusterId}. COM
```

2 Notes on EMR 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 Error messages

Error message: Pay-As-You-Go instances are not available in this region.

The error message returned when you cannot purchase Pay-As-You-Go ECS instances in the region that you want to create clusters. We recommend that you switch to another region to purchase instances.

Error message: The request processing has failed due to an unknown error, exception or failure.

This is an unknown error that occurs in the ECS management system. EMR is built on Alibaba Cloud Elastic Compute Service (ECS) and is also affected by this error. You can try later or submit a ticket to troubleshoot the issues.

Error message: The Node Controller is temporarily unavailable

EMR is built on ECS. The error message returned when the ECS management system has temporary issues. Try creating clusters later.

Error message: No quota or zone is available.

The error message returned when there is no ECS quota available in the specified zone. You can manually switch to another zone or the system will automatically select a zone for you.

Error message: The specified InstanceType is not authorized for use.

You need to apply to use Pay-As-You-Go high-configuration instances (instances with more than eight cores). Click *Here* to apply. You can create high-configuration instances after your application is approved. Make sure that you apply for instances that are supported by EMR, including eight-core 16 GB, eight-core 32 GB, and 16-core 64 GB types.

4 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.

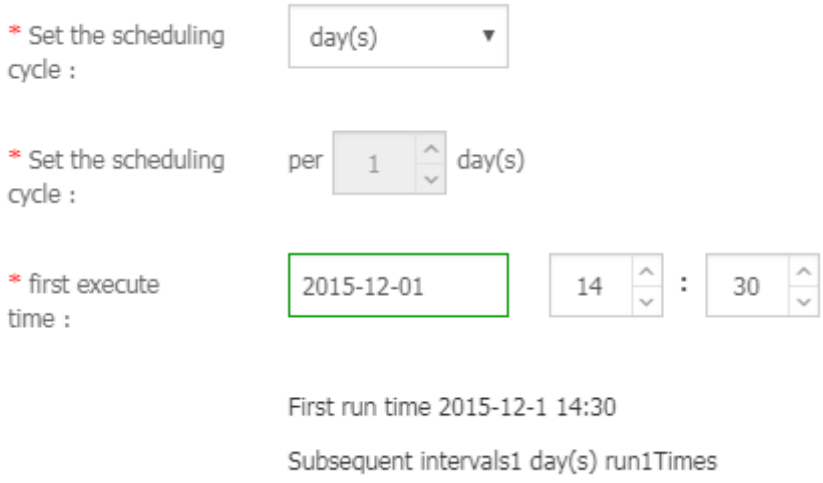
  You can edit execution plans that are not in the running or scheduling status. If you cannot click the edit button, confirm the status of the execution plan and try again.

- 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.

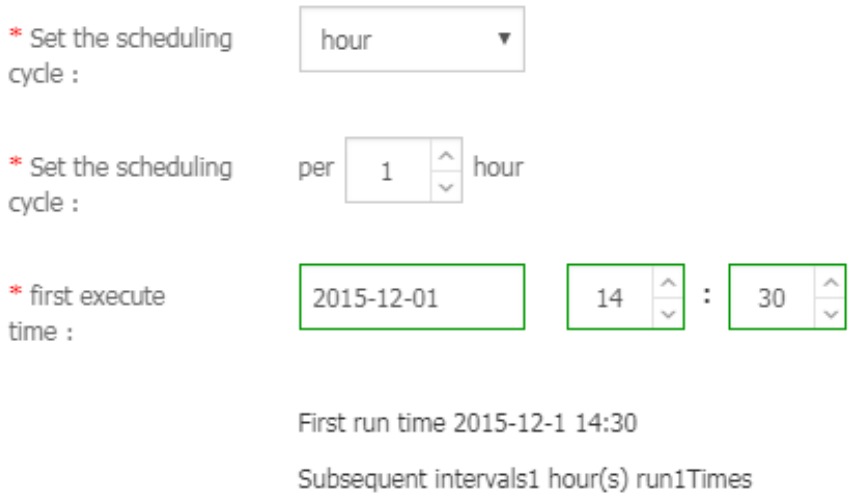
  The start time of a periodical execution cluster indicates the time when the execution plan starts to run. The time is accurate to minutes. The schedule cycle indicates the interval between two executions since the start time. As shown in the following example:



  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.

5 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	-
ResourceManager	-	8025	Internal	yarn.resourcemanager.resource-tracker.address	-
ResourceManager	-	8032	Internal	yarn.resourcemanager.address	-

Service	Limit	Port number.	Access Requirement	Parameter	Description
ResourceManager	-	8030	Internal	yarn.resourcemanager.scheduler.address	-
ResourceManager	-	8088	Internal	yarn.resourcemanager.webapp.address	-

Hadoop MapReduce (MRv1)

Service	Limit	Port number	Access Requirement	Parameter	Description
JobTracker	-	8021	External	mapreduce.jobtracker.address	-

Hadoop HBase

Service	Limit	Port No.	Access Requirement	Parameter	Description
HMaster	-	16000	Internal	hbase.master.port	-
HMaster	-	16010	External	hbase.master.info.port	-
HRegionServer	-	16020	Internal	hbase.regionserver.port	-
HRegionServer	-	16030	External	hbase.regionserver.info.port	-
ThriftServer	-	9099	External	-	-

Hadoop Spark

Service	Limit	Port number	Access Requirement	Parameter	Description
SparkHistory	-	18080	External	-	-

6 Job exception

Q: Why does a Spark job report "Container killed by YARN for exceeding memory limits" or a MapReduce job report "Container is running beyond physical memory limits"?

A: The amount of memory assigned is low when the application is submitted. The JVM consumes too much memory during startup, exceeding the assigned amount. This causes the job to be terminated by NodeManager. This also affects Spark jobs, which may consume more off-heap memory. For Spark jobs, increase the value of **`spark.yarn.driver.memoryOverhead`** or **`spark.yarn.executor.memoryOverhead`**. For MapReduce jobs, increase the value of **`mapreduce.map.memory.mb`** and **`mapreduce.reduce.memory.mb`**.

Q: Why is "Error: Java heap space" returned when I submit a job?

A: The task has large amounts of data in the process but the JVM has insufficient memory. As a result, the OutOfMemoryError error is returned. For Tez jobs, increase the value of **`hive.tez.java.opts`**. For Spark jobs, increase the value of **`spark.executor.memory`** or **`park.driver.memory`**. For MapReduce jobs, increase the value of **`mapreduce.map.java.opts`** or mapreduce.reduce.java.opts.

Q: Why is "No space left on device" returned when I submit a job?

A: Master or worker node has insufficient storage place, which causes a failure of submitting the job. If the disk is full, exceptions in local Hive meta databases such as MySQL Server, or Hive Metastore connection errors may occur. We recommend that you clear enough disk space of the master node, including the system disk and HDFS space.

Q: Why is "ConnectTimeoutException" or "ConnectionException" returned when I use OSS or Log Service?

A: The OSS endpoint is a public network address, but the EMR worker node does not have a public IP address. Therefore, you cannot access OSS or Log Service. For example, the statement `select * from tbl limit 10` can be successfully executed, but `Hive SQL: select count(1) from tbl` fails.

Set the OSS endpoint to an internal network address, such as *`oss-cn-hangzhou-internal.aliyuncs.com`*, or use MetaService provided by EMR. If you choose to use MetaService, you do not need to specify an endpoint.

```
alter table tbl set location "oss://bucket.oss-cn-hangzhou-internal.
aliyuncs.com/xxx"
```

```
alter table tbl partition (pt = 'xxxx-xx-xx') set location "oss://
bucket.oss-cn-hangzhou-internal.aliyuncs.com/xxx"
```

Q: Why is "OutOfMemoryError" returned when I read a Snappy file?

A: The format of standard Snappy files written by Log Service is different from that of the Hadoop Snappy files. By default, EMR processes Hadoop Snappy files.When it processes standard Snappy files, the OutOfMemoryError error is returned. You can set the value of the corresponding parameters to true for troubleshooting. For Hive jobs, configure `set io.compression.codec.snappy.native=true`. For MapReduce jobs, configure `Dio.compression.codec.snappy.native=true`. For Spark jobs, configure `spark.hadoop.io.compression.codec.snappy.native=true`.

Q: Why is "Invalid authorization specification, message from server: "ip not in whitelist or in blacklist, client ip is xxx" returned when I connect the EMR cluster to an RDS instance?

A: You need to configure the whitelist on the RDS instance when you connect the EMR cluster to an RDS instance. If you do not add the IP addresses of the cluster nodes to the whitelist, especially after expanding the cluster, this error occurs.

Q: Why is "Exception in thread “main” java.lang.RuntimeException: java.lang.ClassNotFoundException: Class com.aliyun.fs.oss.nat.NativeOssFileSystem not found" returned when reading or writing OSS data?

A: When reading or writing OSS data in Spark jobs, you need to package the EMR SDK into the job JAR. For more information, see *Prerequisites*.

Q: Why is the available memory of the Spark node exceeded when Spark is connected to Flume?

A: Check whether the data receiving mode is Push-based. If not, set the mode to Push-based. For more information, see *Documentation*.

Q: Why is "Caused by: java.io.IOException: Input stream cannot be reset as 5242880 bytes have been written, exceeding the available buffer size of 524288" returned when I connect OSS to the Internet?

A: This is a bug caused by insufficient space for caching during network connection retries. We recommend that you use the EMR SDK with a version later than V1.1.0.

Q: Why is "Failed to access metastore. This class should not accessed in runtime.org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException:

Unable to instantiate org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClient" returned when Spark is running ?

A: When Spark processes Hive data, you must set the execution mode of Spark to yarn-client or local. Do not set the mode to yarn-cluste. Otherwise, this error occurs. If the JAR package of the job contains third-party files, this error may occur when Spark is running.

Q: Why is "java.lang.NoSuchMethodError:org.apache.http.conn.ssl.SSLConnetionSocketFactory.init(Ljavax/net/ssl/SSLContext;Ljavax/net/ssl/HostnameVerifier)" returned when using the OSS SDK in Spark?

A: The http-core and http-client packages that the OSS SDK is dependent on have version dependency conflicts with the running environments of Spark and Hadoop. We recommend that you do not use the OSS SDK in your code. Otherwise, you must manually resolve this issue. If you need to perform some basic operations to handle OSS files, such as listing objects, click *here* to view the detailed information about how to handle OSS files.

Q: Why is "java.lang.IllegalArgumentException: Wrong FS: oss://xxxxx, expected: hdfs://ip:9000" returned when I use OSS?

A: The default filesystem of HDFS is used when you process OSS data. You must use the OSS path to initialize the filesystem so that it can be used to process data on OSS in the following steps.

```
Path outputPath = new Path(EMapReduceOSSUtil.buildOSSCompleteUri("oss
://bucket/path", conf));        org.apache.hadoop.fs.FileSystem fs =
org.apache.hadoop.fs.FileSystem.get(outputPath.toUri(), conf);
if (fs.exists(outputPath)) {
  fs.delete(outputPath, true);
}
```

Q: Why does garbage collection take a long time and job execution become slower?

A: If the size of the heap memory on the JVM that executes the job is too small, garbage collection may take a longer time and the performance of the job is affected. We recommend that you expand the **`Java Heap Size`**. For Tez jobs, increase the value of the **`hive.tez.java.opts`** Hive parameter. For Spark jobs, increase the value of **`spark.executor.memory`** or **`spark.driver.memory`**. For MapReduce jobs, increase the value of **`mapreduce.map.java.opts or mapreduce.reduce.java.opts`**.

Q: Why does AppMaster take a long time to start a task?

A: If there are too many job tasks or Spark executors, AppMaster may take a long time to start a task. The runtime of a single task is short, and the overhead for scheduling jobs becomes large. We recommend that you use CombinedInputFormat to reduce the number of tasks. You can also increase the block size (**`dfs.blocksize`**) of data that is produced by former jobs, or increase the value of **`mapreduce.input.fileinputformat.split.maxsize`**. For Spark jobs, you can reduce the number of executors (**`spark.executor.instances`**) or reduce the number of concurrent jobs (**`spark.default.parallelism`**).

Q: Why does it take a long time to apply for resources, which causes a job pending issue?

A: After the job is submitted, AppMaster needs to apply for resources to start the task. The cluster is occupied during this period and it may take a long time to apply for resources, causing a job pending issue. We recommend that you check whether the configurations of resource groups are inappropriate, and whether the current resource group is occupied but the cluster still has available resources. If so, you can adjust the configurations of key resource groups or resize the cluster to make full use of the resources .

Q: Why does a small number of tasks take a long time to execute, and the overall runtime of the job become longer (data skew problem)?

A: During a certain stage of the task, data is distributed unevenly. In this circumstance, most tasks are quickly executed, but a small number of tasks takes a long time to execute due to large amounts of data. This makes the overall runtime of the job become longer. We recommend that you use the mapjoin feature of Hive and `set hive.optimize.skewjoin = true`.

Q: Why does a failed task attempt make the job runtime longer?

A: A job has a failed task attempt or failed job attempt. Although the job may end normally, the failed attempt may make the runtime of the job become longer. We recommend that you locate the cause of task failures from this section.

Q: Why is "java.lang.IllegalArgumentException: Size exceeds Integer.MAX_VALUE" returned when the Spark job is running?

A: The block size may become too large if the number of partitions is too small. The maximum value of Integer.MAX_VALUE(2 GB) may then be exceeded when you perform data shuffling. We recommend that you increase the number of partitions, and increase the value of **`spark.default.parallelism`**, spark.sql.shuffle.partitions, or perform the repartition operation before you perform data shuffling.

7 O&M FAQ

7.1 Does EMR support real-time computing?

EMR provides three types of real-time computing services, including Spark Streaming, Storm, and Flink. For more information, see *Developer guide*.

If the issue persists, contact*technical support*.

7.2 Service exception caused by disk exception

A disk exception can occur when the disk is full or when the disk is corrupted.

The following details describe how to resolve these issues:

The disk is full

1. Log on to the corresponding machine, locate the full disk, and delete any unnecessary data to free up some of the disk space. Before you delete any data, note the following:
   - Do NOT delete Kafka data directories. Otherwise, you will lose all of your data.
   - We recommend that you review the oldest log data in your selected partitions (that is, the oldest segments and corresponding index files) and delete data you no longer require.
   - We recommend you do not clean up Kafka topics, such as **consumer_offsets** or schema.
2. Restart the Kafka broker service.

The disk is corrupted

If more than 25% of the disk is corrupted, the machine migration mode can be used for operation and maintenance. To access the machine maintenance mode, submit a ticket to Alibaba Cloud technical support.

8 Cluster 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
Hadoop		
	yarn resourcemanager	k 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
spark		
	spark ui	masternode1_private_ip:4040
	history	masternode1_private_ip:18080
tez		
	tez-ui	masternode1_private_ip:8090/tez-ui2
hue		
	hue	masternode1_private_ip:8888
zeppelin		
	zeppelin	masternode1_private_ip:8080
hbase		
	hbase	masternode1_private_ip:16010
presto		
	presto	masternode1_private_ip:9090
oozie		
	oozie	masternode1_private_ip:11000

Software	Service	Endpoint
ganglia		
	ganglia	masternode1_private_ip:8085/ganglia

9 Appendix

9.1 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 *security group page* and delete security groups that are not in use.

Error code	Error message
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.
5039	The error message returned when you fail to lock the cluster role. You must have certain permissions to use EMR. For *role authorization*, click *here* 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 *RAM* to apply for authorization.
8003	The error message returned when you are not authorized to perform the PassRole operation. Click *RAM* 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 *apply for high-configuration instances* on the ECS buy page.
No zone or cluster resources are available.	No ECS resources are available in this zone.

9.2 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.