Alibaba Cloud E-MapReduce

FAQ

Issue: 20191101

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, consequent**

ial, 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 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 contact Alibaba Cloud directly if you discover any errors in this document.

Document conventions

Style	Description	Example
	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.
	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.
	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.
	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]`

Style	Description	Example
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Contents

1 Q&A about E-MapReduce

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

A: You need to perform two steps to run an EMR job:

- **Create a job**

  An EMR job is essentially a set of configurations for running a job. An EMR job cannot be run directly. Instead, you need to specify the job JAR file, data input and output paths, and some run parameters in the configurations for running an EMR job. Provide a name for the set of configurations to complete the creation of a job. When you need to debug the job, an execution plan is required.

- **Create an execution plan**

  An execution plan associates a job with a cluster. You can use an execution plan to run a sequence of jobs manually or schedule an execution plan to run the jobs periodically. You can choose a cluster for jobs to run using an execution plan. The cluster can be an on-demand cluster or an existing cluster. An on-demand cluster is released automatically after the execution of all jobs is completed. You can view the running status of each execution of an execution plan on the corresponding running log page.

Q: How do I view the logs of jobs?

A: The EMR system uploads the logs of jobs to the OSS log path that you set when creating the cluster based on the job IDs. You can view logs of jobs in the EMR console. If you submit and run jobs on the master node, you need to go to the log path you set to view the logs.

Q: How do I connect to a core node?

A: You need to perform the following steps.

1. **On the master node, switch to the hadoop user using the su command**

   ```
   su hadoop
   ```

2. **Then you can connect to a core node using SSH without entering a password.**

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

3. **Gain root privileges using the sudo command.**

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

Q: Can I view logs in the OSS console?

A: You can search for logs in the OSS console and download them. Viewing logs in the OSS console is not supported. The following example describes how to locate the logs of a job, assuming that you have already enabled the running logs feature and specified the log path in OSS. Assume that the log path is set to *OSS://mybucket/emr/spark*.

1. **Go to the execution plan page, click Running log for the execution plan, which contains the logs you want to view.**
2. **On the running log page, find the execution record that you want. Click the corresponding cluster name in the Execute cluster column to view the cluster ID on the Details page.**
3. **Locate** *OSS://mybucket/emr/spark/cluserID* **under the** *OSS://mybucket/emr/spark* **directory.**
4. **Log files of jobs are stored in the corresponding folders that are created based on the job execution order ID under the** *OSS://mybucket/emr/spark/clusterID/jobs* **directory.**

Q: How is the running time calculated for a cluster, an execution plan, and a job?

A: The corresponding running time calculation policies are as follows.

- **For a cluster**

  You can view the running time of each cluster on the cluster list. Formula: running time = completion time of cluster release - start time of cluster creation. Calculation of the running time starts when a cluster has been created and ends when the cluster has been released.

- **For an execution plan**

  You can view the running time of each execution plan on the running log page. The running time calculation policy is based on the cluster that the execution plan runs on.

  - **If you choose a create-as-needed cluster for your execution plan, then each run of the execution plan involves creating a cluster, submitting jobs, and releasing the cluster. Formula: running time = consumption time for cluster creation + consumption time for running all jobs + consumption time for cluster release.**
  - **If you choose an existing cluster for your execution plan, then cluster creation and cluster release will not be involved in the run cycle. Therefore, running time = consumption time for running all jobs.**

- **For a job**

  Specifically, we are referring to jobs that are included in an execution plan . For each running log of an execution plan, you can click View job list in the Operation column to see all jobs that are included in the corresponding execution plan. The running time calculation formula for each job is: running time = job execution completion time - start time of job execution. Job execution refers to jobs being scheduled to run on a Spark or Hadoop cluster.

Q: Why are there no security groups available when running an execution plan for the first time?

A: For security reasons, you cannot use an existing ECS security group as an EMR security group. Therefore, if you have not created any EMR security groups, no security groups are available for an execution plan. We recommend that you manually create an on-demand cluster for job testing. Create an EMR security group when you manually create a cluster. After testing all jobs, create an execution plan to schedule jobs. At this point, existing ECS security groups are available for your execution plan to choose from.

Q: Why "java.lang.RuntimeException.Parse responsed failed: ‘<!DOCTYPE html>…’ " is returned when I upload data to or download data from MaxCompute using Tunnel ?

A: Check whether the tunnel endpoint is correct. This error occurs when the tunnel endpoint is incorrect.

Q: Can I view the logs of jobs, which are stored in worker nodes, in the EMR console?

A: Yes. Prerequisites: You have enabled the Running log feature when creating the cluster. Path to job logs: Execution plan list > More > Running log > Running record > View job list > job list > job instance

Q: Why data cannot be retrieved using the external table created by Hive?

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
    No data has been retrieved using the external table.
```

This issue occurs because no partition directory is available for Hive to locate. To solve this problem, you can use ALTER TABLE ADD PARTITION to add partitions to the table. 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 a Spark Streaming job stop running unexpectedly?

A: First, check whether the Spark version is earlier than v 1.6. Spark v 1.6 has fixed a memory leak bug. This bug may cause a container to be terminated for exceeding memory limits, which is one probable cause of a Spark Streaming job terminating unexpectedly. Second, check whether your code has been optimized for effective memory usage.

Q: Why does the EMR console show that a Sparking Streaming job is running when the job has already stopped?

A: We recommend that you change the running mode of the Spark Streaming job from yarn-client to yarn-cluster. EMR has problems monitoring the status of a Sparking Streaming job that runs in yarn-client mode. We will fix the problem as soon as we can.

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

A: Check whether the Class-Path header of the job JAR file is `ossref`. If not, modify it to `ossref`.

Q: How do the master node and slave nodes work together?

A: An EMR cluster consists of a single master node and multiple slave (worker) nodes. Only slave (worker) nodes store and process data. For example, a cluster consists of three instances. Each instance has four vCPUs and 8 GB of memory. One instance serves as the master node and the other two serve as slave nodes. Therefore, the available computing resources of this cluster are two instances (the two slave nodes), each with four vCPUs and 8 GB of memory.

Q: How do I include local shared libraries in a MapReduce job?

A: You have multiple ways to achieve this. The following example describes one way. 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>
```

You only need to specify the path of the library that you want.

Q: How can I specify the OSS data source file path for a MapReduce or Spark job?

A: The OSS data source path format is shown as follows: `oss://[accessKeyId:accessKeySecret@]bucket[.endpoint]/object/path`

You can use the URI format to specify input and output OSS data sources for a job. Similarly, when the data sources are in HDFS, the corresponding URI starts with `hdfs://`. You can access OSS data with or without the AccessKey.

- **(Recommended) EMR provides MetaService, which allows you to access OSS data without an AccessKey so that you can specify the data source using the *oss://bucket/object/path* path format.**
- **(Not recommended) You can set the AccessKeyId, AccessKeySecret, and endpoint parameters on the Configuration object for a MapReduce job (for a Spark job, set these parameters on the SparkConf object). Or you can include the AccessKeyId,**

the AccessKeySecret, and the endpoint in a URI directly. 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:

- **This error may occur when you use earlier versions of mysql-connector-java. Use the latest version of mysql-connector-java.**
- **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 file.**

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

A: Include the internal IP addresses of the cluster nodes in the whitelist of ApsaraDB for RDS.

Q: What do I need to consider when creating a cluster of low-specification nodes?

A:

- **If you choose an instance with two vCPUs and 4 GB of memory as the master node , then the master node is prone to running out of memory. We recommend that you increase the memory of the master node.**
- **If you choose an instance with two vCPUs and 4 GB of memory as a slave (worker) node, set the parameters as follows when running a MapReduce or Hive job. For a MapReduce job, set the** `yarn.app.mapreduce.am.resource.mb` **parameter to 1024. For a Hive job, set the** `yarn.app.mapreduce.am.resource.mb` **parameter to 1024. This step is to prevent a job from being suspended due to an OOM error.**

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 reading Parquet data (including columns of the decimal type) written by Spark SQL using Hive or Impala ?

A: The decimal type has different representations in the different Parquet conventions used in Hive/Impala and Spark SQL. Therefore, Parquet data (including columns of the decimal type) written by Spark SQL cannot be read

properly using Hive or Impala. To solve this issue, we recommend that you set the `spark.sql.parquet.writeLegacyFormat` **parameter to true (this setting makes Spark use the same convention as Hive/Impala for writing the Parquet data) before importing the Parquet data written by Spark SQL to Hive or Impala.**

Q: How do I connect to Kerberos-authenticated clusters using Beeline?

A:

- **High-availability cluster (service 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
```

- **High-availability cluster (directly connecting to a node)**

  Connect to the emr-header-1 node.

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

  Connect to the emr-header-2 node.

```
! 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
```

Q: Why do I receive a "Connection refused telnet emr-header-1 10001" error message?

A:

You can view logs in the *`/mnt/disk1/log/spark`* **directory.**

This issue is caused by the Thrift Server running out of memory (OOM). You need to increase memory by raising the value of the `spark.driver.memory` **parameter**

.

2 Error messages

If the creation of an E-MapReduce cluster fails, you can find the corresponding solution based on the error message.

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.

3 The FAQ about product use

Q: How do I create a Pay-As-You- Go high-specification instance for an EMR cluster?

A Pay-As-You-Go instance that has more than eight vCPUs (high-specification) is not shown in the EMR console. You need to submit a ticket to apply for a high-specification instance in the ECS console. We recommend that you use a Subscription cluster, which saves your time for applying for high-specification instances.

Q: What is a high-security cluster?

A high-security cluster is a Kerberos-authenticated cluster. You can turn on the High Security Mode switch to create a high-security cluster on the Create Cluster page. See *Introduction to Kerberos* for more information. You cannot turn off the High Security Mode switch for a cluster of a version earlier than V3.12. If you want to use a non-high security cluster, create a new cluster. V3.12 and later versions support turning off the High Security Mode.

Q: Why do I fail to create a cluster?

A: This issue occurs mainly because the maximum number of Pay-As-You-Go instances that you can create is exceeded. The quota of Pay-As-You-Go instances depends on the initial purchase of a user. You can submit a ticket to apply for a higher quota of Pay-As-You-Go instances. Another cause is that you do not have permission to create instances of the type you want. You need to enable the corresponding permission for the instance type in the ECS console.

Q: How do I renew a cluster?

For more information, see *Cluster renewal*. Many users fail to renew the clusters after renewing the subscription for ECS. This issue occurs because you have not renewed the subscription for EMR, which is also required for renewing a cluster. You can view the expiration dates of ECS and EMR on the cluster renewal page.

Q: How do I adopt automatic renewal?

You can enable the Automatic Renewal feature in the EMR console to renew the subscription for EMR and ECS automatically.

Q: Can I add an existing ECS instance to an EMR cluster?

Currently, you can only create an ECS instance in the EMR console and then add it to an EMR cluster.

Q: Can I install software on the master node of an EMR cluster?

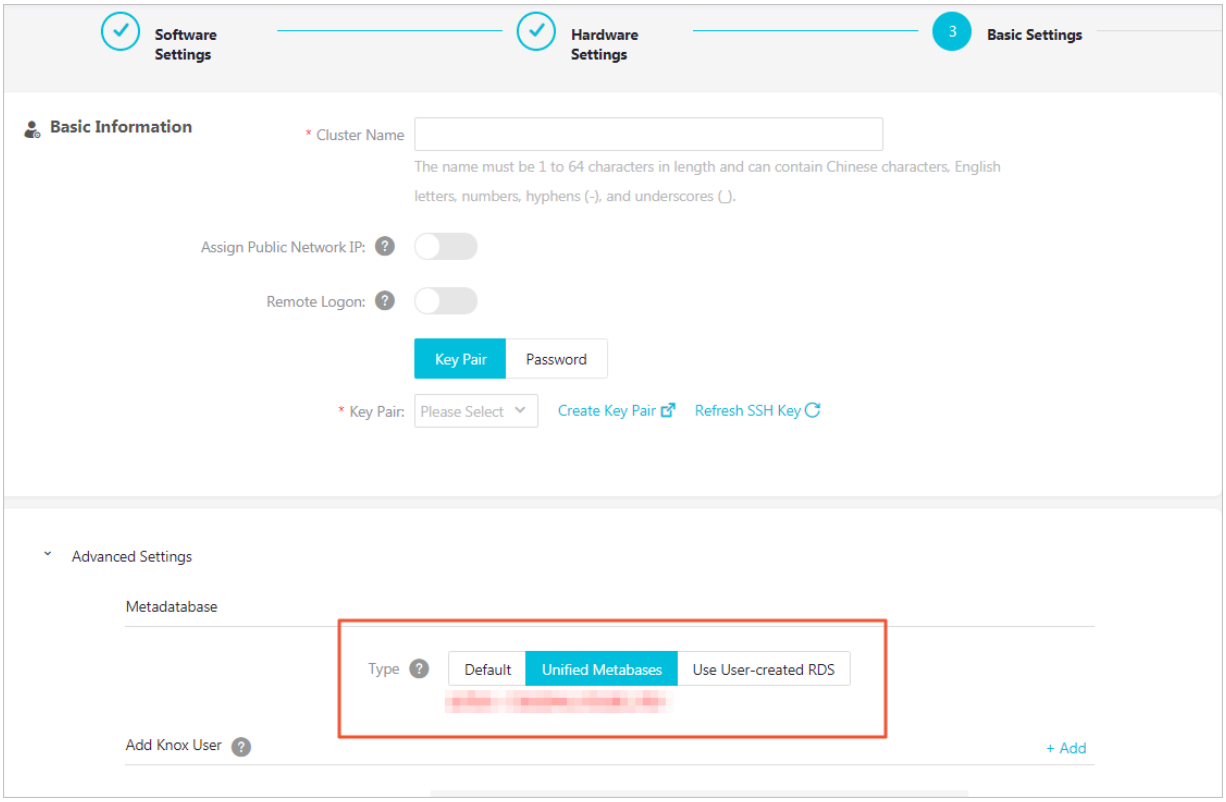
A: Technically, you are allowed to install software on the master node as long as the cluster environment is not affected. However, we recommend that you do not perform this operation because the running of software may impact the stability of a cluster.

Q: How do I connect to a core node and gain root privileges?

For more information, see the Connecting to a core node section in the *Create a cluster*.

Q: Can I release the insecure EIP of a header node in the ECS console? Does the operation affect EMR services?

A: You need to use an EIP to connect to the uniform meta database. If you have not enabled the uniform meta database feature, it is safe to release the EIP.



Q: Do services start automatically when nodes are powered on? Do services restart automatically after stopping unexpectedly?

A: Services start and restart automatically. A service retries starting three times after it fails to start.

Q: How do I access open source components?

A: Currently, the latest clusters are accessed using Knox. For more information, see *#unique_10*. **This issue may occur when the EIP has been released. Check whether the access links can be resolved to IP addresses by connecting to the links using Telnet on port 8443.**

Q: What is required for a RAM user to operate EMR?

A: The corresponding Alibaba Cloud account needs to enable the EMRdefaultRole and ECSdefaultRole policies first. The RAM user needs to adopt the EMRfullaccessrole policy.

Q: How to use Hue?

A: For the initial username and password of Hue, see *How to use Hue*. **You can also refer to this document when you forget your username or password. If Hue fails to access HDFS, set the value of the** `dfs.webhdfs.enabled` **property to true in the EMR console and restart HDFS.**

Q: Anonymous users are allowed to access Zeppelin by default . How do I disable this feature? Do I need to modify the configuration file on the master node if I cannot find the configuration option I want on the configuration page of Zeppelin in the EMR console?

A: You need to modify the configurations manually and restart Zeppelin. For more information, see XXX. *http://blog.csdn.net/mergerly/article/details/53196918*

Q: Does EMR support spot instances?

A: Currently, spot instances are not supported.

Q: Why cannot I run Hive queries with the uniform meta database feature enabled? The error message is as follows. "FAILED: SemanticException org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClient"

A: This issue occurs mainly because you do not have an available EIP to assign to your cluster. EIPs are required for connecting to uniform metadatabases. You need

to assign an EIP to your cluster manually and submit a ticket to the Alibaba Cloud product team for adding the EIP to the security group of your databases.

Q: What is a low-specification node?

A: A low-specification node is configured with two vCPUs and 8 GB of memory. You can create a low-specification node for some users in your whitelist to use. Note: Issues are prone to occurring. You need to resolve issues by yourself.

Q: Why there is no instance type available for the master node?

A: No instance type is available for the master node in your zone. You can switch to another zone.

Q: How can I make other ECS instances submit jobs and return results?

A: You can use Gateway. See *Gateway instance* **for purchasing Gateway in the EMR console.**

Q: How do I expand the disk capacity of an EMR cluster?

A: For more information, see *Disk capacity expansion***. Expanding the disk capacity in the EMR console will be supported soon.**

Q: Can I enable logging in OSS for an existing cluster?

A: Currently, it is not supported. We recommend that you create a workflow instance to run jobs. Workflow is a new feature that you can find on the Data Platform page. By using Workflow, you can view running logs without enabling the running logs feature .

Q: Can a Zookeeper/Kafka/Storm cluster created using EMR communicate with HBase?

A: Yes. The EMR cluster and the HBase cluster need to be in the same VPC. Make sure to add the IP address of the EMR cluster to the whitelist of HBase.

Q: How do I remove core nodes and task nodes from an EMR cluster? For example, I have a cluster of four core nodes and two task nodes. Is it feasible to make the cluster contain two core nodes?

A: Currently, removing core nodes in the console is not supported. If you want to scale in your cluster, submit a ticket to Alibaba Cloud Customer Services to request to remove core nodes. The sequence of removing nodes that run the ZooKeeper services is based on the worker IDs. The latest node is removed first. For example, the sequence of removing worker1, worker 2, worker3, and worker4 is worker4 > worker3 > worker2 > worker1. Currently, removing Subscription core nodes and

tasks nodes in the console is not supported. Removing Pay-As-You-Go task nodes in the console is supported.

Q: How to get refunds for EMR?

A: Submit a ticket including the reason for refunds to the EMR product team.

Q: What is the difference between EMR and MaxCompute?

Both EMR and MaxCompute are used for big data processing. EMR is a big data platform built completely based on open source technologies. It is 100% compatible with open sources for use and practices. MaxCompute is a proprietar y platform developed by Alibaba Cloud. It is easy to use with encapsulation and saves costs for operation and maintenance. EMR is designed based on the open-source Hadoop ecosystem. It is easy to get started for developers with Hadoop prior knowledge. Using MaxCompute requires a little modification on the code.

Q: What is the password for MySQL in an EMR cluster?

A: On the Clusters and Services page, click Hive and click Configuration to view the password.

```
javax.jdo.option.ConnectionURL
```

```
javax.jdo.option.ConnectionUserName
```

```
javax.jdo.option.ConnectionPassword
```

Q: How do I resolve the issue that Zeppelin 0.71 does not support Spark 2.2?

A: This issue occurs in earlier versions of EMR. EMR V3.11 and later versions have resolved this issue by upgrading Zeppelin to V0.73.

Q: Is automatic storage balancing available? Do I manually rebalance storage?

A: For manual storage balancing, choose console > Cluster Management > Clusters and Services > HDFS. Click Actions and click rebalance.

Q: Does EMR support downgrading configurations? For example, reducing 16 vCPUs and 32 GB of memory to 8 vCPUS and 16 GB of memory for the master node, core nodes, and tasks nodes.

A: Currently, it is not supported.

Q: Why does the " The specified DataDisk Size beyond the permitted range, or the capacity of snapshot exceeds the size limit of the specified disk category" error message appear?

A: The value you set for the DataDisk Size parameter is too small for a ???

Q: Why does the "User real name authenticate failed!" message appear ?

A: Your Alibaba Cloud account has not been real-name authenticated. Complete the authentication in the Alibaba Cloud console. EMR also requires real-name authentication of users. You can contact the EMR product team to complete authentication.

Q: Why does the "Your account does not have enough balance" error message appear?

A: Your account balance is insufficient.

Q: Why does the "The maximum number of Pay-As-You-Go instances is exceeded: create ecs vcpu quota per region limited by user quota [xxx]." error message appear?

A: Your quota for Pay-As-You-Go instances has been exceeded. You need to apply for a higher quota in the ECS console or release instances to create an EMR cluster.

Q: I cannot find Flume in the EMR console. How do I steam data to the OSS path configured for my Hadoop cluster using Flume?

A: Currently, Flume has not been integrated with EMR. You need to install Flume manually.

Q: Does Spark support submitting jobs using the standalone mode?

A: Currently, the Spark On Yarn mode is used by default in the EMR console. The standalone mode is not supported.

Q: How do I modify software configurations?

A: Earlier versions of EMR do not support modifying software configurations in the console. You can perform the following steps to modify software configurations.

1. **Log on to the master node of your cluster.**
2. **Go to the directory of configuration templates.**

```
cd /var/lib/ecm-agent/cache/ecm/service/
```

3. **Locate the directory of the service you want. Assuming the service is Hue, go to the directory of Hue.**

4. **Go to the corresponding directory to the version of Hue. For example,** `/var/lib/ecm-agent/cache/ecm/service/HUE/4.1.0.1.3`**.**
5. **Corresponding configuration files are shown in the** `/package/templates/` **directory.**
6. **Modify the configurations as needed. You can add a configuration or modify an existing configuration.**
   - **If you want to add a configuration, make sure the format is correct. Note: Check whether the line breaks and spaces are used properly.**
7. **After the modification is complete, restart the service for the configurations to take effect.**

4 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.

5 How do I set the authentication mode of HiveServer2 to LDAP?

This topic describes how to set the authentication mode of HiveServer2 to Lightweight Directory Access Protocol (LDAP).

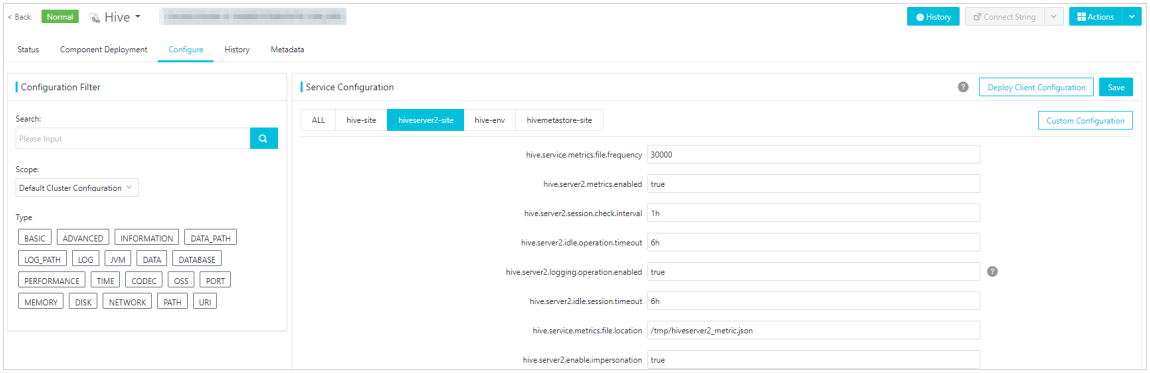
Question

How do I set the authentication mode of HiveServer2 to LDAP?

Answer

In an E-MapReduce cluster, HiveServer2 supports multiple authentication modes, including NOSASL, None, LDAP, Kerberos, PAM, and Custom. You can set the hive.server2.authentication parameter to specify the authentication mode.

1. **Log on to the *Alibaba Cloud E-MapReduce console*.**
2. **Set the authentication mode of HiveServer2 to LDAP and restart HiveServer2.**
   a. **On the Cluster Management page, click the ID of the E-MapReduce cluster for which you want to set the authentication mode of HiveServer2. In the left-side navigation pane, click Hive. Click the Configure tab and then click hiveserver2-site in the Service Configuration section.**



   b. **Click Custom Configuration to add parameters.**

      To set the authentication mode of HiveServer2 to LDAP, you must add the three parameters listed in the following table.

Parameter	Value	Description
hive.server2.authentication	LDAP	The authentication mode.

Parameter	Value	Description
hive.server2.authentication.ldap.url	ldap://${emr-header-1-hostname}:10389	Replace ${emr-header-1-hostname} with the actual hostname. You can use SSH to log on to the E-MapReduce cluster and run the `hostname` command on the emr-header-1 instance of the cluster to obtain the hostname. For more information, see *#unique_17*.
hive.server2.authentication.ldap.baseDN	ou=people,dc=emr	N/A

c. After the parameters are added, click Save in the upper-right corner.

d. In the dialog box that appears, enter the change description and click OK. A message appears, indicating that the parameters are added.

e. In the upper-right corner, choose Operation > Restart HiveServer2.

3. Add an account to the LDAP service.

   In an E-MapReduce cluster, OpenLDAP is an LDAP service, which is used to manage Knox accounts by default. HiveServer2 can reuse the Knox accounts for LDAP authentication. For more information about how to add an account, see *Knox*. In this example, add the emr-test account.

4. Check whether you can use the new account to log on to HiveServer2.

   Use */usr/lib/hive-current/bin/beeline* to log on to HiveServer2 as follows:

```
beeline> ! connect jdbc:hive2://emr-header-1:10000/
Enter username for jdbc:hive2://emr-header-1:10000/: emr-guest
Enter password for jdbc:hive2://emr-header-1:10000/: emr-guest-pwd
Transaction isolation: TRANSACTION_REPEATABLE_READ
```

   If the account or password is incorrect, the following error message appears:

```
Error: Could not open client transport with JDBC Uri: jdbc:hive2://
emr-header-1:10000/: Peer indicated failure: Error validating the
login (state=08S01,code=0)
```

6 FAQ about the HDFS balancer in E-MapReduce

This topic describes how to adjust the settings of related parameters to optimize the performance of the Hadoop Distributed File System (HDFS) balancer in E-MapReduce.

Question:

How does the HDFS balancer work? How can I optimize the performance of the HDFS balancer?

Answer:

HDFS provides a balancer utility that analyzes block placement and balances data across the DataNodes. Besides, HDFS offers the data locality feature for improving the cluster balancing performance. For example, after you add a large amount of empty DataNodes to the Hadoop cluster, you can balance data in time by moving computation close to data. This enhances the performance of the Hadoop cluster. You can follow these steps to run the HDFS balancer:

1. **Log on to a DataNode in the target cluster.**
2. **Run the following commands to switch to the HDFS user and run the HDFS balancer:**

```
su hdfs
/usr/lib/hadoop-current/sbin/start-balancer.sh -threshold 10
```

3. **Run either of the following commands to check whether the HDFS balancer is running:**

```
less /var/log/hadoop-hdfs/hadoop-hdfs-balancer-emr-header-xx.cluster
-xxx.log
```

or

```
tailf /var/log/hadoop-hdfs/hadoop-hdfs-balancer-emr-header-xx.
cluster-xxx.log
```

Note:

If the command output includes `Successfully`, the HDFS balancer is running.

The following table describes the HDFS balancer parameters.

Parameter	Description
threshold	The maximum difference allowed between the storage utilization of a DataNode and the storage utilization of its cluster. The default value is 10, indicating that the storage utilization of each DataNode differs by no more than plus or minus 10% of the average storage utilization of the cluster. Decrease the value when the storage utilization of the cluster is high, and increase the value when the storage utilization of the cluster is low. This improves the cluster performance. After you add a large number of DataNodes to the cluster, you can specify a high threshold first. In this way, the HDFS balancer efficiently moves data from overutilized DataNodes to underutilized DataNodes. After the HDFS balancer runs for a while, decrease the threshold to involve more DataNodes in load balancing. This increases the number of concurrent threads and guarantees efficient load balancing.

Parameter	Description
dfs.datanode.balance.max.concurrent.moves	The maximum number of concurrent threads on a DataNode used by the HDFS balancer for moving blocks. The default value is 5. Generally, set the value based on the number of disks for the DataNode. We recommend that you set the value to quadruple of the disk quantity on the DataNode side, and set the value to the disk quantity on the HDFS balancer side. In this way, the HDFS balancer can schedule the concurrent threads within the limit. Assume that a DataNode has 28 disks. Set the value of this parameter to 28 on the HDFS balancer side, and set the value of this parameter to 112 (28 x 4) on the DataNode side. Adjust the value based on the cluster load. Increase the value when the cluster load is low, and decrease the value when the cluster load is high. If the maximum numbers of concurrent threads set on the DataNode and HDFS balancer sides are different, the smaller one prevails. **Note:** After you set this parameter for a DataNode, restart the DataNode for the parameter setting to take effect.

Parameter	Description
dfs.balancer.dispatcherThreads	The number of dispatcher threads used by the HDFS balancer to decide which blocks to move. Before the HDFS balancer moves a certain amount of data between two DataNodes, it repeatedly gets block lists for moving blocks until the required amount of data is scheduled.  Note: The default value is 200.
dfs.balancer.rpc.per.sec	The number of remote procedure calls (RPCs) sent by dispatcher threads per second. The default value is 20. Before the HDFS balancer moves a certain amount of data between two DataNodes, it uses dispatcher threads to repeatedly send the getBlocks() RPC to the NameNode. This results in a heavy load on the NameNode. To avoid this issue and balance the cluster load, we recommend that you set this parameter to limit the number of RPCs sent per second. For example, you can set this parameter to 10 or 5, which slightly affects block moves in the cluster.

Parameter	Description
dfs.balancer.getBlocks.size	The total data size of the blocks moved each time. Before the HDFS balancer moves a certain amount of data between two DataNodes, it repeatedly gets block lists for moving blocks until the required amount of data is scheduled. By default, the size of blocks in each block list is 2 GB. When the NameNode receives a getBlocks() RPC , the NameNode is locked. If an RPC queries a large amount of blocks, the NameNode is locked for a long time, which slows down data writing. To avoid this issue, we recommend that you set this parameter based on the NameNode load.
dfs.balancer.moverThreads	The number of threads used by the HDFS balancer to move blocks. The default value is 1000. Each block move requires a thread. This parameter limits the number of total concurrent moves for balancing in the entire cluster.

Parameter	Description
dfs.namenode.balancer.request.standby	Specifies whether the HDFS balancer sends RPCs to query the standby NameNode for blocks to be moved. The default value is false. When the NameNode receives a getBlocks() RPC, the NameNode is locked. If an RPC queries a large amount of blocks, the NameNode is locked for a long time, which slows down data writing. If you set high-availability clusters, the HDFS balancer only sends RPCs to the standby NameNode.
dfs.balancer.getBlocks.min-block-size	The minimum size of blocks to be queried by the getBlocks() RPC. The default value is 10, in MB. After you set this parameter, the getBlocks() RPC skips blocks smaller than the minimum size. This improves the query efficiency.
dfs.balancer.max-iteration-time	The maximum duration of each iteration for moving blocks between two DataNodes. The default value is 1200000, in milliseconds. After the duration of an iteration exceeds this limit, the HDFS balancer enters the next iteration.

Parameter	Description
dfs.balancer.block-move.timeout	The timeout duration for each iteration . The default value is 0, in milliseconds , indicating that the iteration does not time out. When the HDFS balancer moves blocks , an iteration may last for a long time because some block moves are not complete. You can set this parameter to avoid such issue.

The following table describes the DataNode parameters.

Parameter	Description
dfs.datanode.balance.bandwidthPerSec	The maximum number of bytes per second that each DataNode can use to balance the cluster. The default bandwidth is 1 Mbit/s. We recommend that you set this parameter to a value greater than 100 Mbit/s so that the cluster can be balanced quickly. You can also set this parameter based on the cluster load. Alternatively, you can set the dfsadmin -setBalancerBandwidth parameter to specify the bandwidth. Modifying this configuration does not require restarting DataNodes. For example, you can increase the bandwidth when the cluster load is low , and decrease the bandwidth when the cluster load is high.

Parameter	Description
dfs.datanode.balance.max.concurrent.moves	**The maximum number of concurrent threads on a DataNode used by the HDFS balancer for moving blocks.**

7 Use execution plans

Apply for high-configuration instances

You must activate a high configuration instance before you use it to create a cluster. If an instance is not activated, the error message The specified InstanceType is not authorized for usage appears when you try to create a cluster.

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.

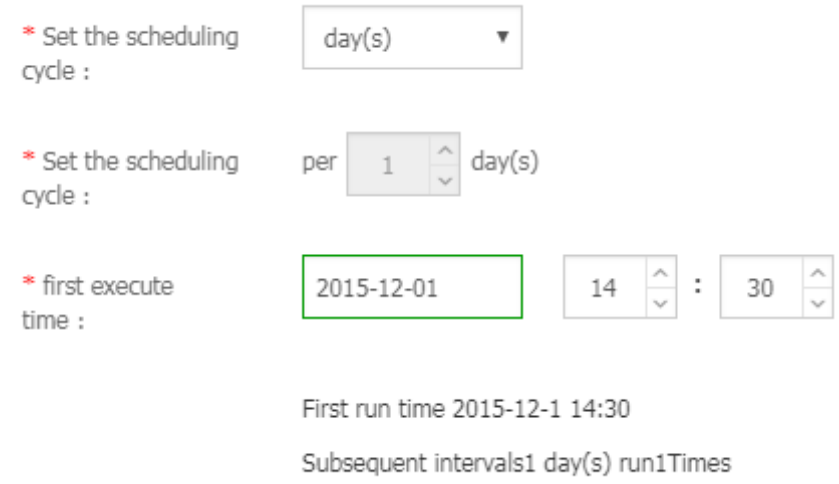
- **Run an execution plan.**

  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

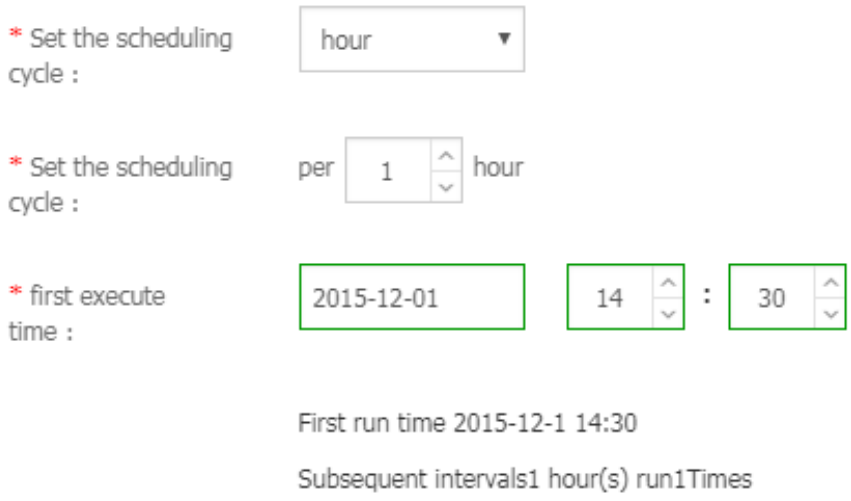
indicates the interval between two executions since the start time. As shown in the following example:



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.

8 O&M FAQ

8.1 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*.

8.2 How do I handle disk exceptions in a Kafka cluster?

A disk exception may occur when a disk is fully occupied or is damaged. This topic describes how to handle disk exceptions.

Fully occupied disk

Follow these steps to handle a fully occupied disk:

1. **Log on to the machine where a disk is fully occupied.**
2. **Find the fully occupied disk and delete unnecessary data to free up some of the disk space. Note the following rules:**
   - **Do not directly delete Kafka data directories. Otherwise, you may lose all of your data.**
   - **Do not delete Kafka topics, such as consumer_offsets and schema.**
   - **Find the topics that occupy a large space or you no longer need. Delete the oldest log data from your selected partitions. Delete segments and correspond ing index and timeindex files.**
3. **Restart the Kafka broker of this machine.**

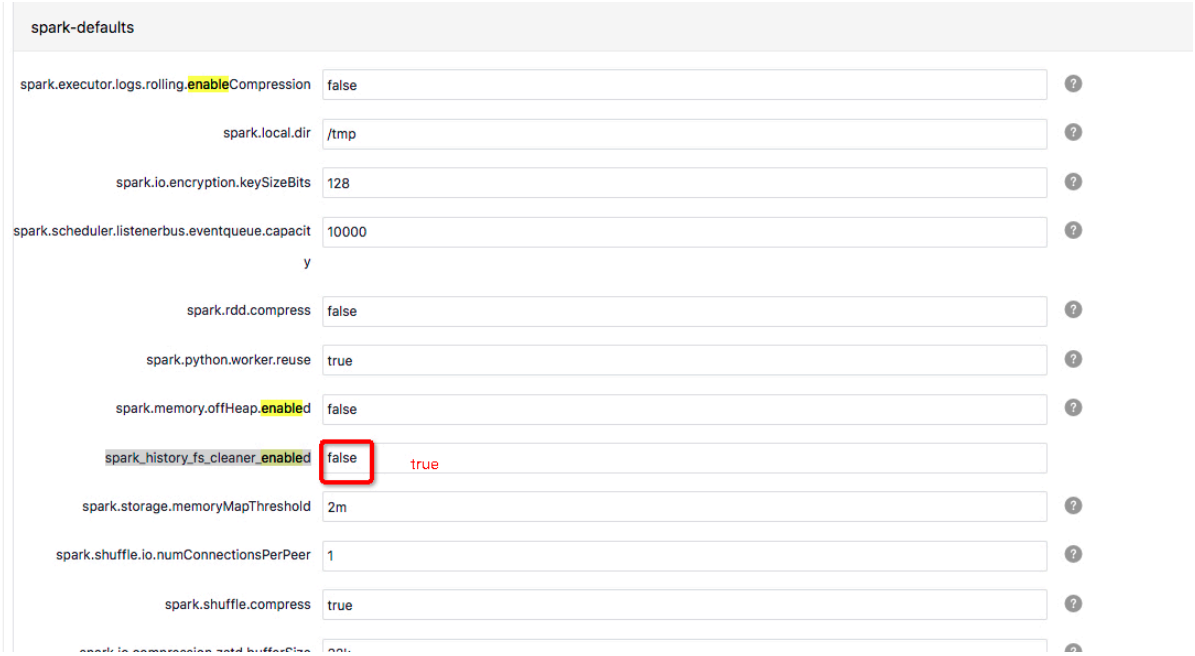
Damaged disk

Use one of the following solutions to handle a damaged disk:

- **If no more than 25% of disks are damaged on a machine, you do not need to take any action.**
- **If more than 25% of disks are damaged on a machine, you can migrate data to another machine. In this case, open a ticket to ask Alibaba Cloud for technical support.**

8.3 HDFS capacity of a cluster is full with large amounts of data stored in the /spark-history directory

You can enable the cleaner (spark.history.fs.cleaner) of Spark history on the Spark configuration management page to clean up event logs for completed jobs periodically.



The event logs to clean up include logs (in-progress logs excluded) that are stored in the /spark-history directory of HDFS.

If your cluster has many long-running Spark Streaming jobs, set the spark.eventLog.enabled property to false to avoid increasing event logs. If you cannot find the spark.eventLog.enabled configuration option on the page, create a custom configuration file on the server. For the configurations to take effect, restart the Spark Streaming jobs.