

Alibaba Cloud E-MapReduce

Kernel Enhancement

Issue: 20190729

Legal disclaimer

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

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, special, or punitive damages, including lost profits arising from the use

or trust in this document, even if Alibaba Cloud has been notified of the possibility of such a loss.

5. By law, all the content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of the Alibaba Cloud website, product programs, or content shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates).
6. Please contact Alibaba Cloud directly if you discover any errors in this document.

Generic conventions

Table -1: Style conventions

| Style | Description | Example |
|---|--|--|
|  | This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results. |  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 instructions, 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 <code>cd / d C :/ windows</code> command to enter the Windows system folder. |
| <i>Italics</i> | It is used for parameters and variables. | <code>bae log list --instanceid Instance_ID</code> |
| [] or [a b] | It indicates that it is an optional value, and only one item can be selected. | <code>ipconfig [-all -t]</code> |

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

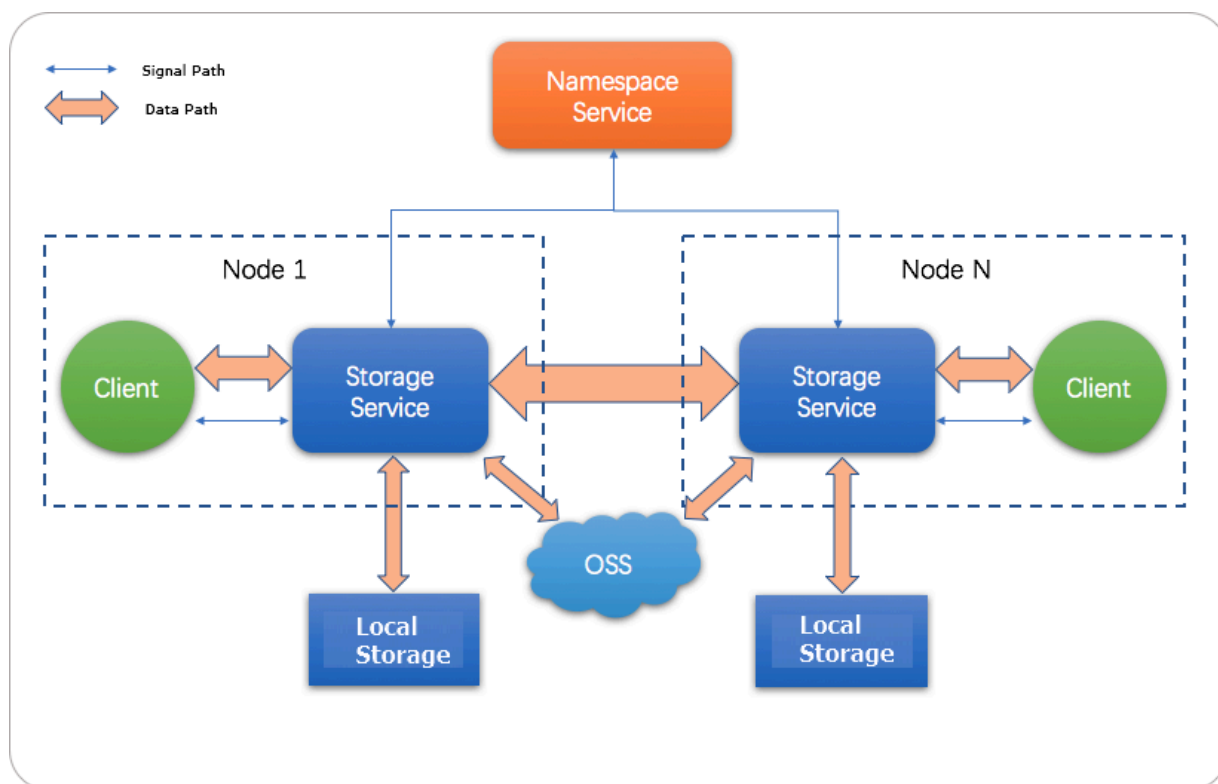
Contents

| | |
|---------------------------|---|
| Legal disclaimer..... | I |
| Generic conventions..... | I |
| 1 JindoFS user guide..... | 1 |

1 JindoFS user guide

JindoFS is a cloud-native file system that combines the advantages of OSS and local storage. JindoFS is also the next-generation storage system that provides efficient and reliable storage services for cloud computing. This topic describes how to configure and use JindoFS, and its scenarios.

Overview

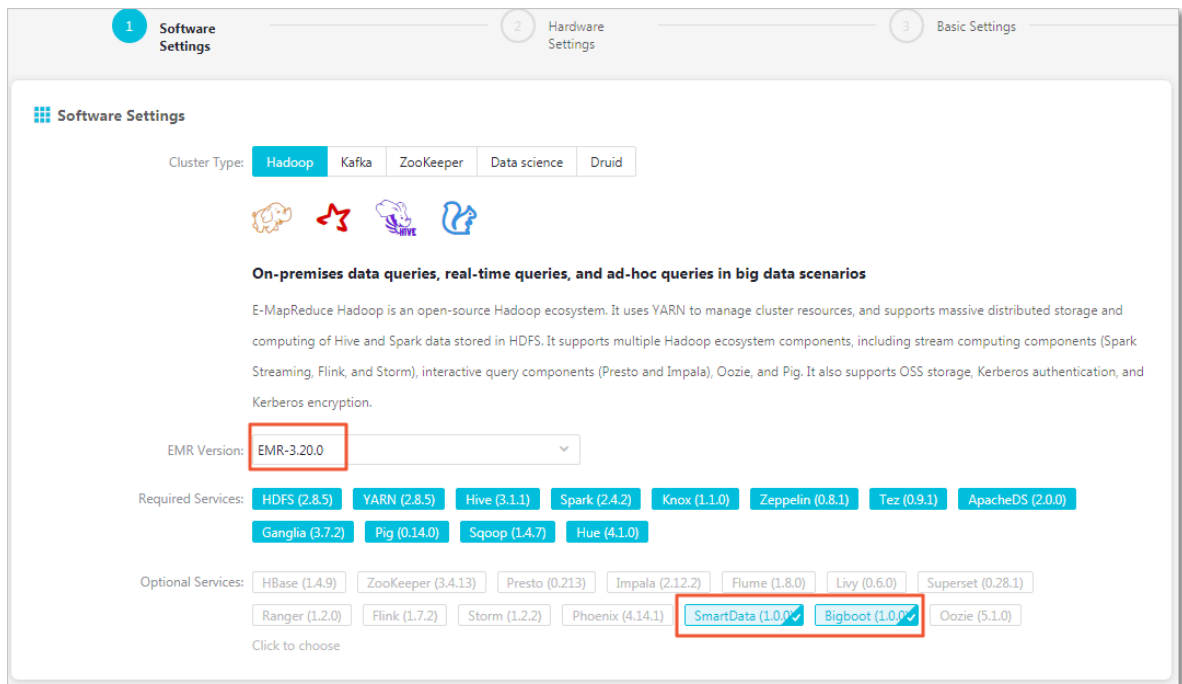


JindoFS adopts a heterogeneous multi-backup mechanism. Storage Service provides data storage capacity. OSS is used as the back-end storage to ensure high reliability of data. Local storage is used to implement redundant backup. Backups stored in the local storage can accelerate read operations. Namespace Service serves as the local service that manages metadata of JindoFS. This way, metadata is queried from Namespace Service instead of OSS, which improves query performance. This query method of JindoFS is the same as that of HDFS. EMR V3.20.0 and later support Jindo FS. To use JindoFS, select the related services when creating an E-MapReduce cluster.

Environment preparation

- Create E-MapReduce clusters

Select EMR V3.20.0 or later. Select SmartData and Bigboot. For more information about how to create an E-MapReduce cluster, see [Create a cluster](#). Bigboot provides distributed data management in the E-MapReduce console. SmartData is based on Bigboot to provide JindoFS for the application layer.



- Configure E-MapReduce clusters

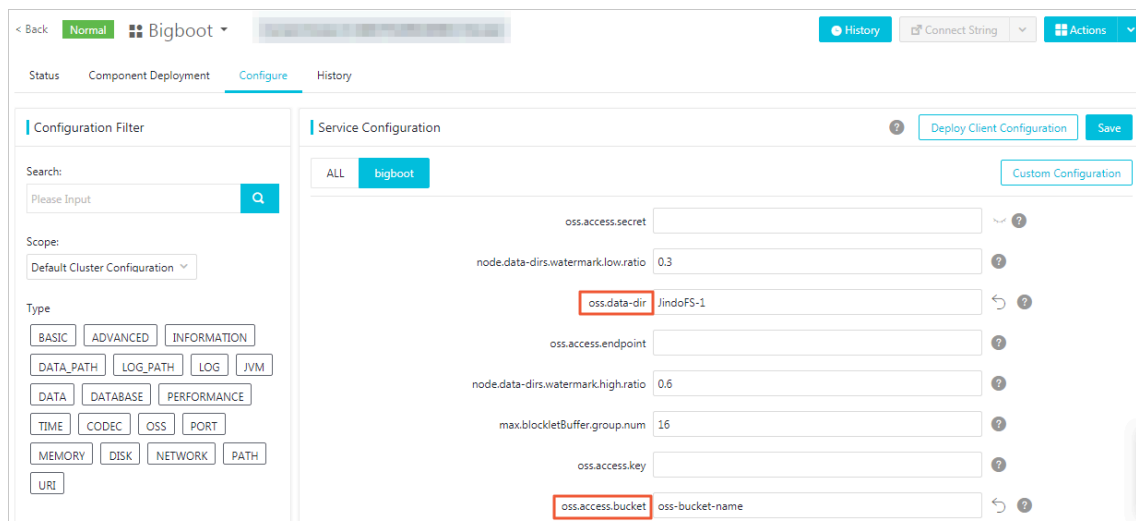
JindoFS provided by SmartData uses OSS as the storage back end. Therefore, you need to configure OSS-related parameters before using JindoFS. The following section provides two configuration methods. If you use the first configuration method, you need to create an E-MapReduce cluster, modify Bigboot-related parameters, and restart SmartData for the configurations take effect. If you use the second configuration method, you need to add custom configurations when you create an E-MapReduce cluster. This method ensures that the related services are restarted based on custom parameters after the E-MapReduce cluster is created.

- Initialize parameters after an E-MapReduce cluster is created

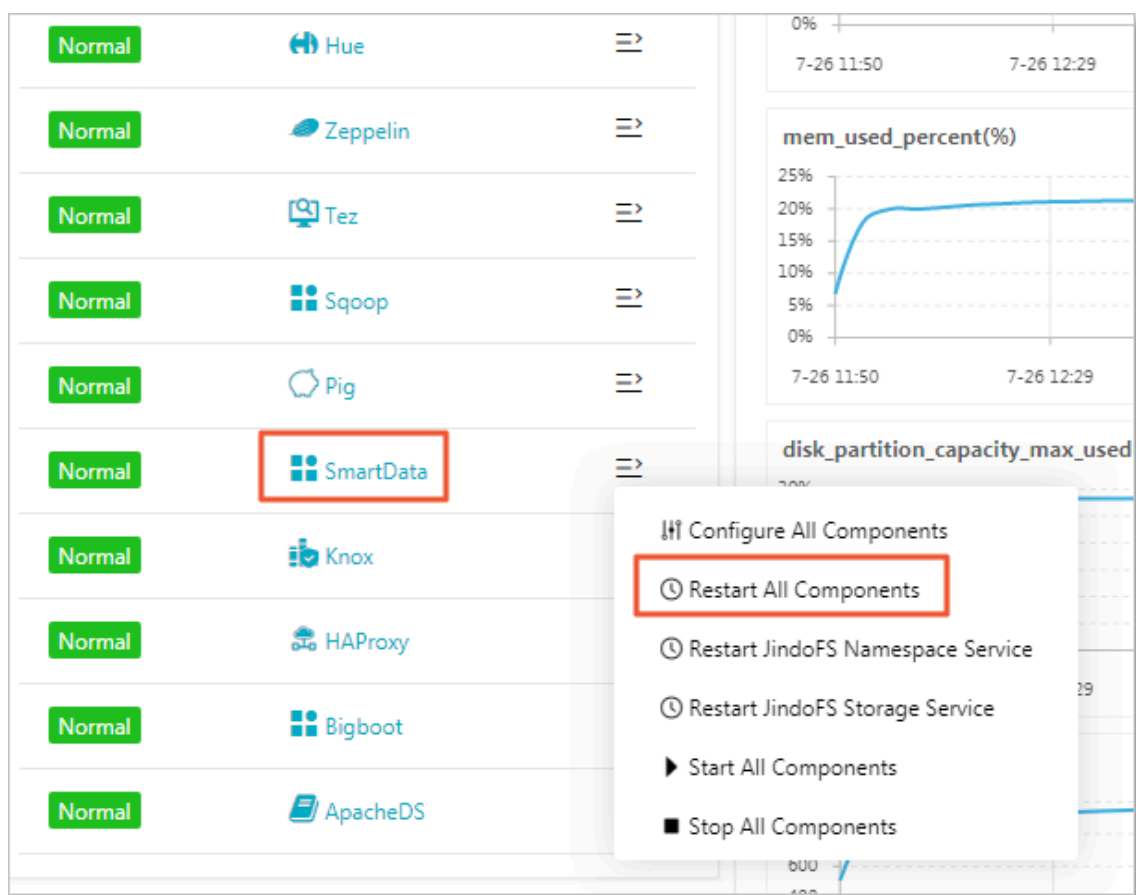
All JindoFS-related configurations are contained in Bigboot. As shown in the following figure, parameters in the red box are required. `oss.access.bucket` specifies the name of the bucket in OSS. `oss.data-dir` specifies the directory used by JindoFS in the OSS bucket. Note that the directory is the back-end storage

directory of JindoFS. Data stored in this directory cannot be modified or deleted. Ensure that the directory can be used only as the back-end storage of JindoFS. The directory configured by users are automatically created when data is written to JindoFS. You can move the pointer over the information icons to view instructions of the `oss.access.endpoint`, `oss.access.key`, and `oss.access.secret` parameters. To maintain performance and stability, we recommend that you select an OSS bucket that is located in the same region as the E-MapReduce cluster when you configure the storage back end. This way, the E-MapReduce

cluster can access OSS without using a password. In this case, you do not need to configure these parameters.



Save and deploy the configurations. Restart all components in SmartData to start using JindoFS.



- Add custom configurations when creating an E-MapReduce cluster

You can add custom configurations when you create an E-MapReduce cluster. The following section uses password-free access to OSS within the same

region as an example. Turn on Custom Software Settings. Add the following configurations to the field in the Advanced Settings section.

```
[
  {
    "ServiceName": "BIGBOOT ",
    "FileName": "bigboot ",
    "ConfigKey": "oss . data - dir ",
    "ConfigValue": "jindoFS - 1 "
  },
  {
    "ServiceName": "BIGBOOT ",
    "FileName": "bigboot ",
    "ConfigKey": "oss . access . bucket ",
    "ConfigValue": "oss - bucket - name "
  }
]
```

EMR Version: EMR-3.20.0

Required Services: HDFS (2.8.5) YARN (2.8.5) Hive (3.1.1) Spark (2.4.2) Knox (1.1.0) Zeppelin (0.8.1) Tez (0.9.1) ApacheDS (2.0.0) Ganglia (3.7.2) Pig (0.14.0) Sqoop (1.4.7) Hue (4.1.0)

Optional Services: HBase (1.4.9) ZooKeeper (3.4.13) Presto (0.213) Impala (2.12.2) Flume (1.8.0) Livy (0.6.0) Superset (0.28.1) Ranger (1.2.0) Flink (1.7.2) Storm (1.2.2) Phoenix (4.14.1) SmartData (1.0.0) Bigboot (1.0.0) Oozie (5.1.0)

Click to choose

Advanced Settings

Kerberos Mode: ☐ All components in a cluster that is enabled with the Kerberos mode will use Kerberos authentication to validate requests. For more information, see [Kerberos introduction](#)

Custom Software Settings: ☒ Before creating a cluster, you can use a JSON file to customize parameters for the cluster components. For more information, see [Software Settings](#)

```
[{"ServiceName": "BIGBOOT", "FileName": "bigboot", "ConfigKey": "oss.data-dir", "ConfigValue": "jindoFS-1"}, {"ServiceName": "BIGBOOT", "FileName": "bigboot", "ConfigKey": "oss.access.bucket", "ConfigValue": "oss-bucket-name"}]
```

Use JindoFS

The use of JindoFS is similar to HDFS. JindoFS also provides a prefix. To use JindoFS, you need only to replace `hdfs` with `jfs`. Example:

```
hadoop fs -ls jfs:/// hadoop fs -mkdir jfs:/// test -
dirhadoop fs -put test.log jfs:/// test - dir /
```

Data can be read from JindoFS only when jobs of MapReduce, Hive, and Spark components are running in the E-MapReduce cluster.

Scenarios

E-MapReduce provides the following big data-based storage systems: EMR OssFileSystem, EMR HDFS, and EMR JindoFS. OssFileSystem and JindoFS are cloud-based storage solutions. The following table compares features of these storage systems and the Hadoop-Aliyun module.

| Feature | Hadoop-Aliyun module | EMR OssFileSystem | EMR HDFS | EMR JindoFS |
|-----------------------------|--|---|---|---|
| Storage space | Capable of storing large amounts of data | Capable of storing large amounts of data | Dependent on the cluster scale | Capable of storing large amounts of data |
| Reliability | High | High | High | High |
| Factor affecting throughput | Server | I/O performance of disks in the E-MapReduce cluster | I/O performance of disks in the E-MapReduce cluster | I/O performance of disks in the E-MapReduce cluster |
| Metadata query efficiency | Slow | Medium | Quick | Quick |
| Scale-out operation | Easy | Easy | Easy | Easy |
| Scale-in operation | Easy | Easy | Decommissioning required | Easy |
| Data locality | None | Weak | Strong | Relatively strong |

JindoFS has the following features:

- Infinitely scalable storage capable of storing large amounts of data. OSS is used as the storage back end that is independent of the local cluster. The local cluster can be scaled as needed.
- Read operations can be accelerated if data can be read from the storage resources in the local cluster. This capacity applies to the local cluster that is capable of storing a certain amount of data. Based on this capacity, the throughput efficiency of the local cluster can be maximized. The throughput efficiency can be greatly improved for Write Once Read Many (WORM)-compliant storage solutions.

- High efficiency of metadata-related queries. This Namespace Service-based query method of JindoFS is the same as that of HDFS. This method helps resolve the problems of inefficient queries from OSS and instability caused by frequent access.
- Maximal data locality provided when jobs are executed in the E-MapReduce cluster . This method eases network transmission pressure and further improves read performance.

Disk space usage-based control

The back end of JindoFS is based on OSS that is capable of storing large amounts of data. However, the storage capacity of local disks is limited. Therefore, JindoFS releases data backups that are less frequently accessed. Alibaba Cloud uses `node . data - dirs . watermark . high . ratio` and `node . data - dirs . watermark . low . ratio` to adjust the usage of local disk storage. The average value ranging from 0 to 1 indicates the storage usage. JindoFS uses the total storage of all data disks by default. The `node.data-dirs.watermark.high.ratio` parameter specifies the upper limit of storage usage of each data disk. Less frequently accessed data stored in disks is released if their storage used by JindoFS reaches the upper limit. The `node.data-dirs.watermark.low.ratio` parameter specifies the lower limit of storage usage. After storage usage of data disks reaches the upper limit, less frequently accessed data will be released until the storage usage of the disks reaches a value that is smaller than the lower limit. You can set the upper limit and lower limit to adjust and assign disk space to JindoFS. Ensure that the upper limit is greater than the lower limit.