

HDFS运维管理

www.huawei.com





目标

- 学完本课程后，您将能够：
 - 掌握**HDFS**服务状态检查方法
 - 熟悉**HDFS**服务基本性能调优参数
 - 掌握**HDFS**服务基本故障恢复方法



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用维护命令
5. 常用性能调优参数
6. 常见问题及定位手段

目录介绍

目 录	用 途
/opt/huawei/Bigdata	各组件安装目录
/opt/huawei/Bigdata/fusioninsight/FusionInsight-Hadoop-2.7.2	HDFS 组件安装目录
/srv/BigData/namenode	NameNode 数据存储目录
/srv/BigData/hadoop	DataNode 数据存储目录
/srv/BigData/journalnode	JournalNode 数据存储目录
/var/log/Bigdata/hdfs	HDFS 日志存储目录



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用维护命令
5. 常用性能调优参数
6. 常见问题及定位手段

服务检查

FusionInsight Manager的“服务管理”页面，查看HDFS状态

- HDFS操作状态：已启动
- HDFS健康状态：良好
- HDFS配置状态：已同步



服务	操作状态	健康状态	配置状态
HDFS	✓ 已启动	✓ 良好	✓ 已同步
KrbServer	✓ 已启动	✓ 良好	✓ 已同步
LdapServer	✓ 已启动	✓ 良好	✓ 已同步
Mapreduce	✓ 已启动	✓ 良好	✓ 已同步
SmallFS	✓ 已启动	✓ 良好	✓ 已同步
Yarn	✓ 已启动	✓ 良好	✓ 已同步
ZooKeeper	✓ 已启动	✓ 良好	✓ 已同步

服务检查

“告警管理” 页面中，无和**HDFS**相关的告警：

- **ALM-14000 HDFS**服务不可用
- ...
- **ALM-14011 HDFS DataNode**数据目录配置不合理



服务检查

在“**HDFS**服务状态”页面，执行“更多操作 > 启动服务健康检查”，显示：

- 发送包的平均时间统计： **x.x ns** 小于 **1000000.0 ns**
- 服务状态： 健康
- 告警信息： 无告警信息

服务健康检查

导出报告

项目	开始时间	进度	结束时间
1. HDFS	2016-04-27 09:30:08	100%	2016-04-27 09:30:18
发送包的平均时间统计： 0.0 ns 小于 1000000.0 ns			
服务状态：良好			
告警信息：无告警信息			

操作成功。

单击完成继续。



目录

1. 目录介绍
2. 服务检查
3. 日志分析
4. 常用维护命令
5. 常用性能调优参数
6. 常见问题及定位手段

安装日志

HDFS安装日志的存储路径为：**/var/log/Bigdata/hdfs/角色名**

日志文件名	日志文件描述
cleanupDetail.log	HDFS 服务卸载时候的清理日志。
postinstallDetail.log	HDFS 服务安装前的准备工作日志。
prestartDetail.log	HDFS 服务启动前的准备工作日志。

运行日志

HDFS运行日志的存储路径为：**/var/log/Bigdata/hdfs/角色名**

日志文件名	日志文件描述
hadoop-<SSH_USER>-<process_name>-<hostname>.log	HDFS 系统日志，记录 HDFS 系统运行时候所产生的大部分日志。
hadoop-<SSH_USER>-<process_name>-<hostname>.out	HDFS 运行环境的信息日志。
hadoop.log	Hadoop 客户端操作日志。
namenode-<SSH_USER>-gc.log	NameNode 垃圾回收日志。

审计日志

HDFS审计日志的存储路径为：**/var/log/Bigdata/audit/hdfs**

日志文件名	日志文件描述
hdfs-audit-<process_name>.log	HDFS 操作审计日志（例如：文件增删改查）。
SecurityAuth.audit	HDFS 安全审计日志。

日志过滤

HDFS中提供了下表所示的日志级别，日志级别优先级从高到低分别是**OFF**、**WARN**、**INFO**、**DEBUG**、**TRACE**，程序会打印高于或等于所设置级别的日志，设置的日志等级越高，打印出来的日志就越少。

日志级别	描述
OFF	OFF 表示关闭日志输出。
WARN	WARN 表示当前事件处理存在异常信息。
INFO	INFO 表示系统及各事件正常运行状态信息。
DEBUG	DEBUG 表示系统及系统调试信息。
TRACE	TRACE 表示记录系统及类调用关系的信息。



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用维护命令
5. 常用性能调优参数
6. 常见问题及定位手段

常用维护命令

在**HDFS**客户端，分别执行如下命令进行安全认证登录：

```
source bigdata_env
```

```
kinit <kerberos_user>
```

然后执行维护命令。

维护命令	用途
<code>hdfs fsck <path> [COMMAND_OPTIONS]</code>	运行 HDFS 文件系统检查工具。
<code>hdfs dfsadmin [GENERIC_OPTIONS]</code>	运行一个 HDFS 的 dfsadmin 客户端。
<code>hdfs dfs [COMMAND [COMMAND_OPTIONS]]</code>	在 HDFS 文件系统上运行 filesystem 命令
<code>hdfs getconf [COMMAND_OPTIONS]</code>	从配置目录获取配置信息
...	...

常用维护命令（详细）

维护命令	用途
hdfs dfs -ls <path>	列出<path>路径的目录及文件。
hdfs dfs -mkdir <path>	在<path>路径创建新的目录。
hdfs dfs -rm -r <path>	删除目录。
hdfs dfs -put <LocalFile> <path>	本地文件<LocalFile>上传至HDFS。
hdfs dfs -get <path>	HDFS文件下载到本地。
hdfs dfs -count -q -v <path>	查看目录的Quota和SpaceQuota。
hdfs dfs -getfacl <path>	获取目录的ACL权限。
hdfs fsck <path>	健康状态查看。
hdfs dfsadmin -safemode enter leave get	进入安全模式 离开安全模式 安全模式查看。
...	...

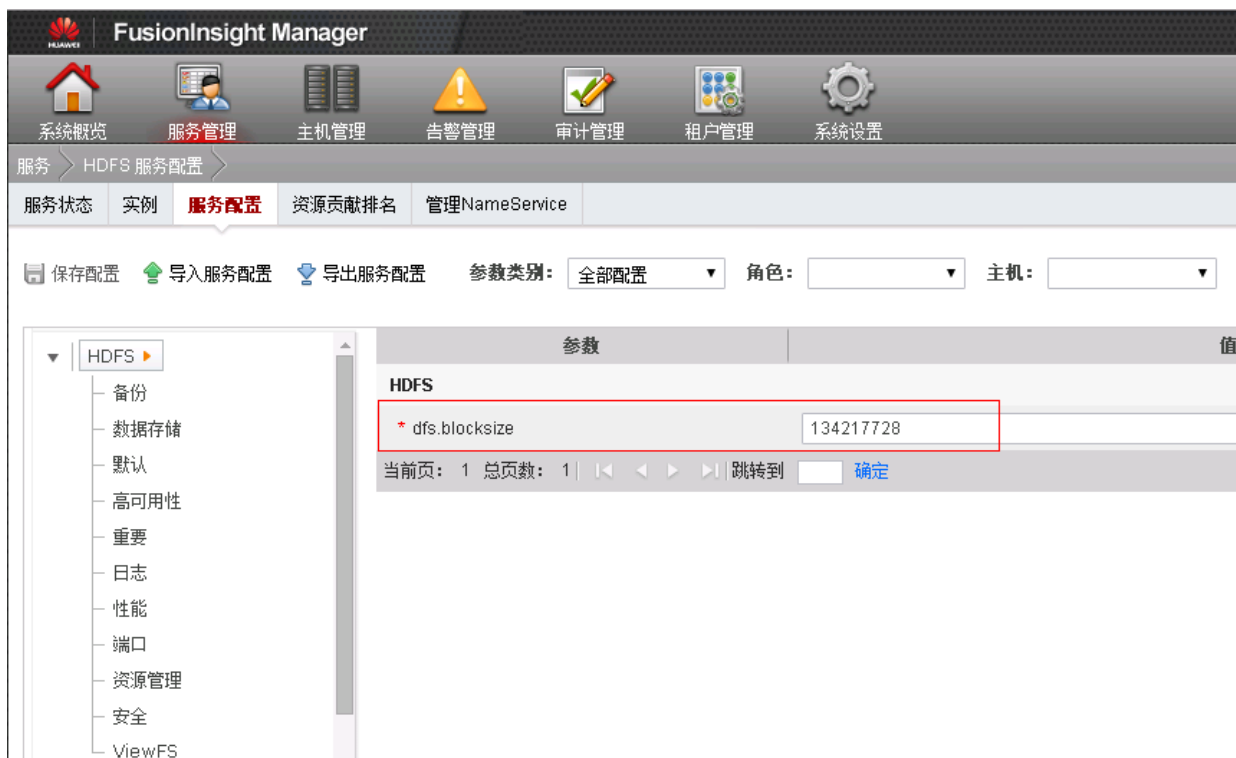


目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用维护命令
- 5. 常用性能调优参数**
6. 常见问题及定位手段

调优参数设置方法

在FusionInsight Manager界面，单击“服务管理 > HDFS > 服务配置”，将“参数类别”设置为“全部配置”。在搜索框里搜索需要调优的参数，会显示所搜索的参数，然后按照需求进行修改。



常用性能调优参数

在HDFS中，通过调整配置的值，使得HDFS集群更适应自身的业务情况，从而提升HDFS的读写性能。

配置项	缺省值	调优场景
dfs.blocksize	134217728 (128M)	表示新建文件的默认块大小。单位：字节。（例如134217728，表示128MB）。 最优值：268435456（256MB） 取值范围：512-1073741824 说明： 参数值必须为512的倍数，否则向HDFS写入文件时会出现错误。
dfs.datanode.drop.cache.behind.reads	true	设置为true表示丢弃缓存的DataNode数据（需在DataNode中配置）。
dfs.client-write-packet-size	262144	当HDFS Client往DataNode写数据时，将数据生成一个包。然后将这个包在网络上传出。此参数指定传输数据包的大小，可以通过各Job来指定。单位：字节。



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用维护命令
5. 常用性能调优参数
6. 常见问题及定位手段

案例1-HDFS服务异常

现象描述

HDFS服务界面显示服务不可用，导致**HBase**、**Hive**、**MapReduce**等服务故障或无法启动

可能原因

- **HDFS**进入安全模式。
- **HDFS**依赖的**ZooKeeper**服务异常。

定位思路

- 检查**HDFS**是否处于安全模式。
- 检查**ZooKeeper**服务是否运行正常。

案例1-HDFS服务异常

处理步骤

1. 确认**HDFS**是否处于安全模式。

是，直接退出**HDFS**的安全模式。

否，执行步骤 2。

2. 登录**FusionInsight Manager**管理界面，在**ZooKeeper**的服务页面，检查服务是否正常，并在“告警管理”页面检查是否有**ZooKeeper**服务的故障告警。

3. 按照**ZooKeeper**的告警指导，恢复服务，并确保**ZooKeeper**服务正常。

4. **ZooKeeper**服务正常后，尝试重启**HDFS**服务，并检查是否正常。

是，操作结束。

否，执行步骤 5。

5. 检查**NameNode**故障，并尝试修复。

案例2-HDFS块丢失异常

现象描述

- **HDFS**进入安全模式，**HDFS**服务不可用。
- **HDFS**服务页面块丢失项的值大于**0**。
- **HDFS**服务启动失败，角色实例启动成功。

可能原因

- 数据节点硬盘故障，或节点故障可能导致数据副本丢失。
- **HDFS**在如下情况进入安全模式：
 1. 当**NameNode**启动且等待**DataNode**上报副本。
 2. **NameNode**所在磁盘空间不足。
 3. 恢复**NameNode**数据后，元数据与业务数据无法匹配。
- **HDFS**对应文件的副本全部丢失

案例2-HDFS块丢失异常

定位思路

1. 检查**FusionInsight Manager**告警界面是否有节点、硬盘故障告警。
2. 检查**NameNode**的块丢失阈值“**dfs.namenode.safemode.threshold-pct**”是否配置过高。

处理步骤

1. 登录**FusionInsight Manager**管理界面，单击“告警管理”，查看是否存在**HDFS**告警、节点故障、硬盘故障等告警。

是，按照告警处理相关章节进行处理。

否，执行步骤 2。

2. 安装客户端程序，以“**admin**”用户使用**HDFS**客户端工具执行**fsck**检查文件系统中文件是否完整。

命令示例：**hdfs fsck /**

客户端程序的安装可参考《**FusionInsight HD V100R002C60SPC200** 管理员指南》的“安装客户端”相关章节。

案例2-HDFS块丢失异常

3. 若**fsck**校验仅显示副本丢失，而不是文件丢失（看已有副本数是否大于**0**），则执行

hdfs dfsadmin -safemode leave退出**HDFS**安全模式，即可修复。

4. 若文件丢失，检查是否执行了恢复**NameNode**数据操作。

是，表示找不到与元数据对应的业务数据块，执行

hdfs dfsadmin -safemode leave退出**HDFS**安全模式，然后执行[步骤 5](#)。

否，表示文件丢失，执行[步骤 5](#)。

5. 若有文件丢失，检查丢失文件的文件路径，并检查文件是否为重要文件。

- **MapReduce**类文件是非重要的，对非重要文件，退出安全模式后，可执行删除操作。

命令示例：

hdfs fsck [path-to-file] -delete

- 恢复**NameNode**数据后可能导致元数据与业务数据无法匹配，需要执行删除操作。

hdfs fsck / -delete

案例2-HDFS块丢失异常

6. 若丢失的文件是重要文件，则需要恢复，必要时联系华为技术支持。

7. 检查**NameNode**的块丢失阈值是否配置过高。

（1）在**FusionInsight Manager**管理界面，单击“服务管理 > **HDFS** > 服务配置”，“参数类别”选为“全部配置”，单击“**NameNode** > 高可用性”，设置“**dfs.namenode.safemode.threshold-pct**”值。

（2）集群中只有三个节点时，建议将**NameNode**的块丢失阈值“**dfs.namenode.safemode.threshold-pct**”参数配置为“0.66”。

8. 重启**HDFS**服务。

总结

- 目录介绍
- 服务检查
- 日志介绍
- 常用维护命令
- 常用性能调优参数
- 常见问题及定位手段

思考题

- **HDFS**审计日志的存储路径是什么？
- 如何查看**HDFS**服务状态？
- **dfs.blocksize** 默认块大小是多少？
- **HDFS**文件系统检查命令是什么？

习题

1. HDFS 默认 **Block Size**的大小是（ ）。

A.32MB

B.64MB

C.128MB

D.256MB

2. 获取**NameNode**安全模式状态的命令是（ ）。

A.hdfs dfs -safemode get

B. hdfs dfsadmin -safemode get

C. hdfs fs -safemode get

D. hdfs fsadmin -safemode get

3. **HDFS**告警故障的故障索引格式是（ ）。

A. 12xxx

B. 14xxx

C. 16xxx

D. 18xxx

Thank you

www.huawei.com