**HDFS掉电恢复指南**

HDFS在掉电场景下可能出现两种故障：**文件数据块损坏**，**元数据损坏**

## **元数据损坏**

### 问题原因

数据块损坏是比较严重的故障，需要手工介入恢复。

判断是否元数据损坏：

1. 有无磁盘损坏（大部分场景元数据损坏是由于磁盘损坏引起的）；
2. /metadata/dfs/nn目录是否正常（主备节点）；
3. NN能够启动；

### 恢复手段

假设rjbdnode1上的元数据损坏。按一下手段从rjbdnode2上恢复。

1. 停止CM上的HDFS服务；
2. 备份HDFS上rjbdnode1元数据目录，并copy rjbdnode2上的元数据目录到该节点；

mv /metadata/dfs/nn /metadata/dfs/nn\_bak

cd /metadata/dfs && scp rjbdnode2: /metadata/dfs/nn ./

1. 登录HDFS单独启动NameNode实例
2. 删除Journal上的编辑日志，在rjbdnode1~3三个节点执行以下命令；

mv /metadata/dfs/jn /metadata/dfs/jn\_bak

1. 重建HDFS的编辑共享目录

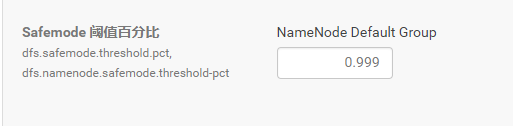


1. 重启HDFS服务

## 数据块损坏

数据块损坏是更常见的故障，当HDFS上损坏的数据块达到00.1%时，会使NN启动后陷入安全模式无法退出。

通过配置项**dfs.safemode.threshold.pct**或**dfs.namenode.safemode.threshold-pct**可以控制HDFS的安全模式阈值，该值默认大小为0.999。



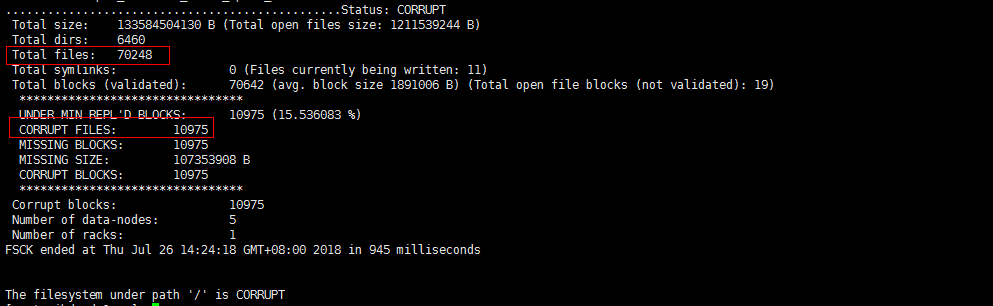
### 恢复手段

当HDFS上出现数据块损坏时，按下面的手段恢复。

1. 执行查看坏块和受影响的文件

sudo –u hdfs hdfs fsck / | egrep -v '^\.+$' | grep -v eplica

sudo –u hdfs hdfs fsck -list-corruptfileblocks -files列出损坏的文件



2、 登录CM页面，强制两个NameNode离开安全模式。



3、执行以下命令删除受损的文件。（**目前HDFS没有自动修复坏块文件的功能，只能删除有坏块的文件，或者手工定位到坏块用其他手段修复。**）

sudo –u hdfs hdfs fsck <dir\_path / file\_path> -delete #删除指定目录下的损坏文件