# Hadoop Upgrade Guide

**By Ivan**

|  |  |
| --- | --- |
| 升级前版本 | 升级后版本 |
| 2.6.5 | 2.8.0 |

1. 创建新数据工作目录
2. 目录结构调整

/bigdata/

----------/backup

----------/data(新加)

----------/tools

----------/workspace

1. 在各Node创建以下目录

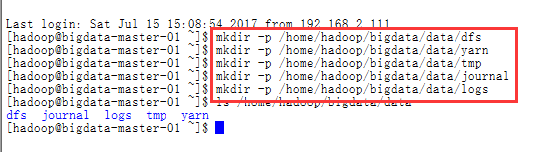
mkdir -p /home/hadoop/bigdata/data/dfs

mkdir -p /home/hadoop/bigdata/data/yarn

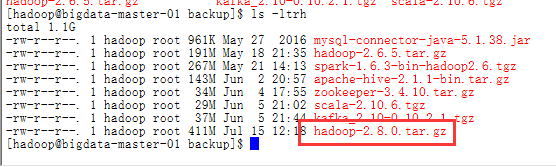
mkdir -p /home/hadoop/bigdata/data/tmp

mkdir -p /home/hadoop/bigdata/data/journal

mkdir -p /home/hadoop/bigdata/data/logs



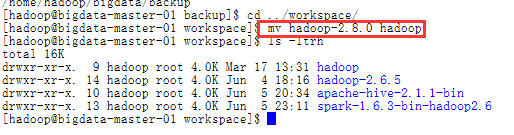
1. 上传2.8.0的hadoop到master-01



1. 解压并修改目录名为hadoop

tar -zxvf /home/hadoop/bigdata/backup/hadoop-2.8.0.tar.gz -C /home/hadoop/bigdata/workspace

mv hadoop-2.8.0 hadoop



1. 修改老版本hadoop配置文件中的工作目录
2. 待修改信息说明

core-site.xml

hadoop.tmp.dir

hdfs-site.xml

dfs.journalnode.edits.dir <!--指定JournalNode集群在对NameNode的目录进行共享时，自己存储数据的磁盘路径-->

dfs.namenode.name.dir <!--指定namenode名称空间的存储地址, 可以是按逗号分隔的目录列表-->

dfs.datanode.data.dir <!--指定datanode数据存储地址, 可以是按逗号分隔的目录列表-->

mapred-site.xml

yarn-site.xml

yarn.nodemanager.local-dirs<!-- nodemanager本地文件存储目录-->

yarn.nodemanager.log-dirs <!--存储container日志的地方-->

1. 修改各节以上配置信息如下：

core-site.xml

hadoop.tmp.dir

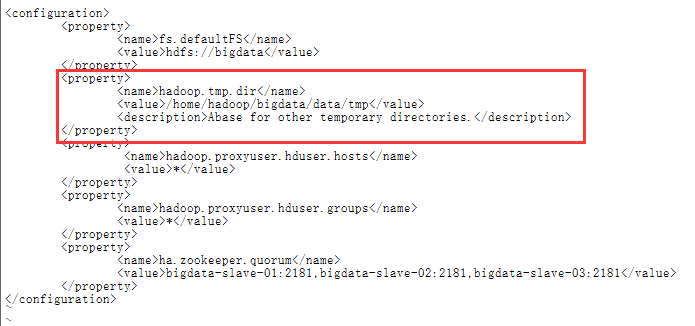
<property>

<name>hadoop.tmp.dir</name>

<value>/home/hadoop/bigdata/data/tmp</value>

<description>Abase for other temporary directories.</description>

</property>



hdfs-site.xml

dfs.journalnode.edits.dir

<property>

<name>dfs.journalnode.edits.dir</name>

<value>/home/hadoop/bigdata/data/journal</value>

</property>



dfs.namenode.name.dir

<property>

<name>dfs.namenode.name.dir</name>

<value>file:/home/hadoop/bigdata/data/dfs/name</value>

</property>

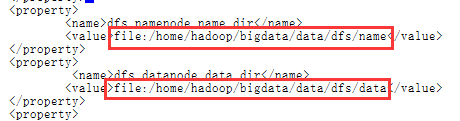
dfs.datanode.data.dir

<property>

<name>dfs.datanode.data.dir</name>

<value>file:/home/hadoop/bigdata/data/dfs/data</value>

</property>



yarn-site.xml

yarn.nodemanager.local-dirs

<property>

<name>yarn.nodemanager.local-dirs</name>

<value>/home/hadoop/bigdata/data/yarn/local</value>

</property>

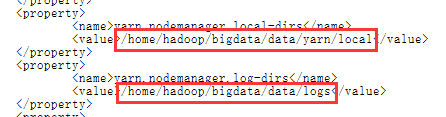
yarn.nodemanager.log-dirs

<property>

<name>yarn.nodemanager.log-dirs</name>

<value>/home/hadoop/bigdata/data/logs</value>

</property>

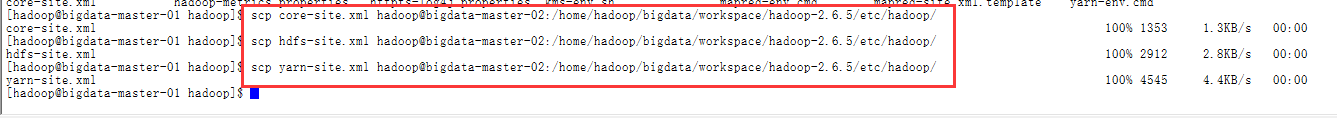


3．将master-01修改好的配置同步到其它node上去（注意master-02 yarn-site.xml中的yarn.resourcemanager.ha.id配置特殊性master-01为rm1, master-02为rm2）。

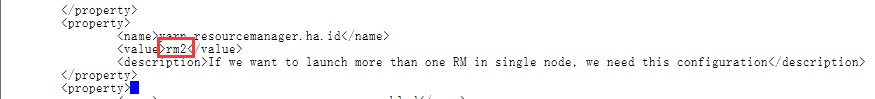
scp core-site.xml [hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/](mailto:hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/)

scp hdfs-site.xml [hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/](mailto:hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/)

scp yarn-site.xml [hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/](mailto:hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/)



4．将master-02 节点上yarn-site.xml中的yarn.resourcemanager.ha.id改回rm2

5.将各节中老版本hadoop 的data工作目录的数据copy到新data目录

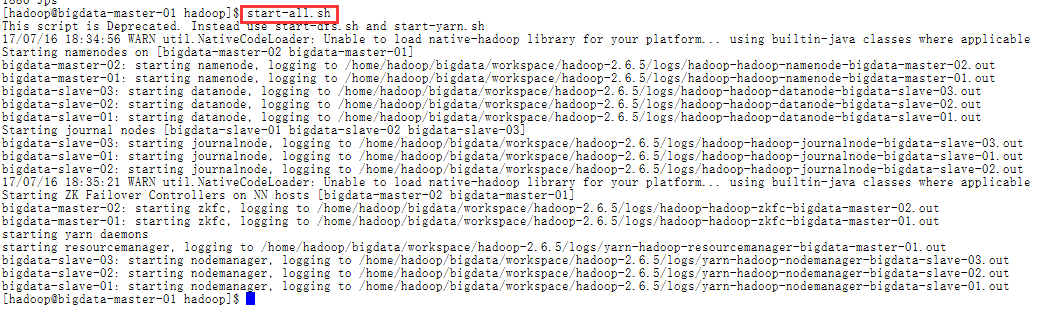
cp -r /home/hadoop/bigdata/workspace/hadoop-2.6.5/dfs /home/hadoop/bigdata/data/

cp -r /home/hadoop/bigdata/workspace/hadoop-2.6.5/yarn /home/hadoop/bigdata/data/

cp -r /home/hadoop/bigdata/workspace/hadoop-2.6.5/tmp /home/hadoop/bigdata/data/

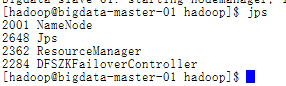
cp -r /home/hadoop/bigdata/workspace/hadoop-2.6.5/journal /home/hadoop/bigdata/data/

1. 启动hadoop，验证修改配置后工作是否正常（注意先启动zookeeper）

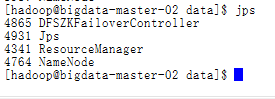


1. 验证各节点进程

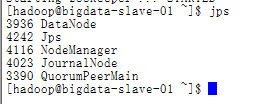
Master-01



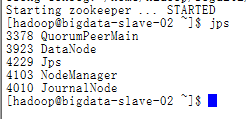
Master-02



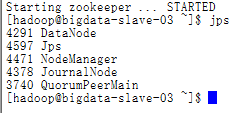
Slave-01



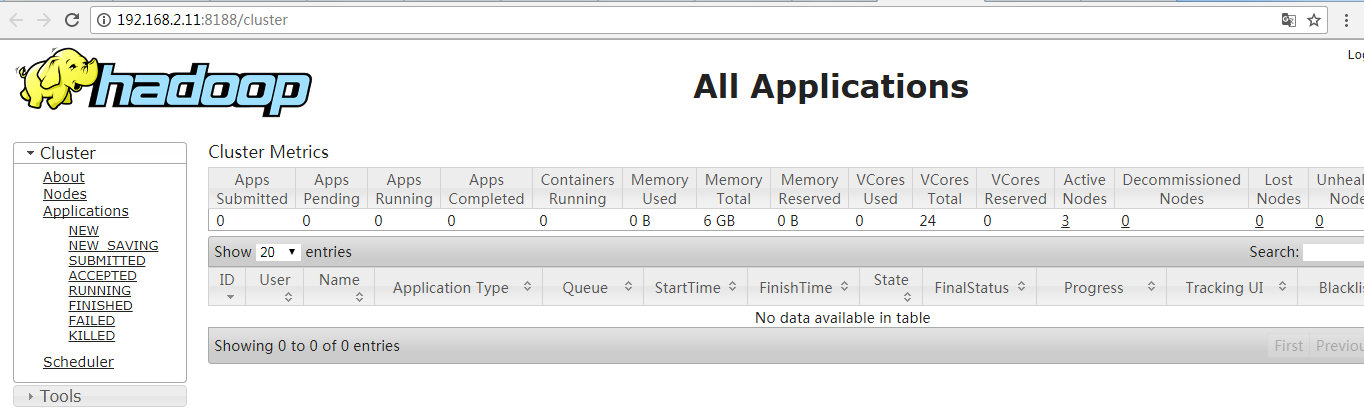
Slave-02

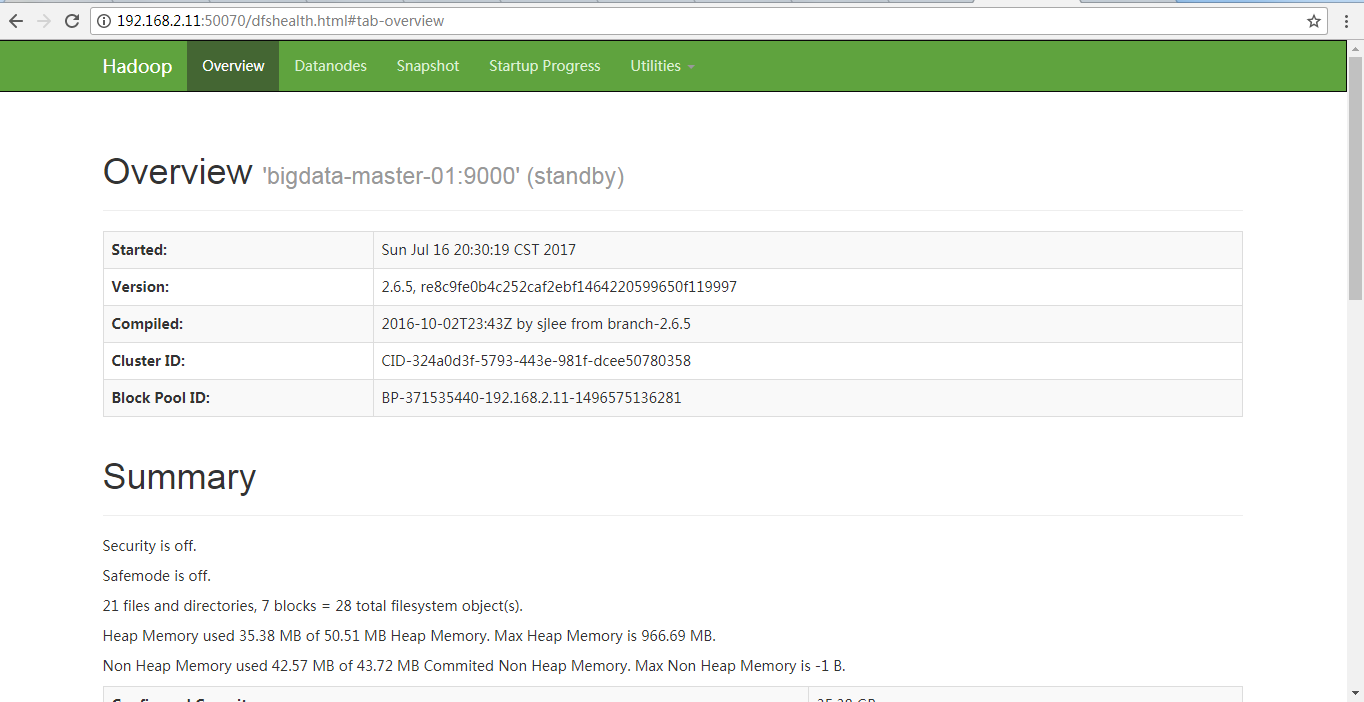


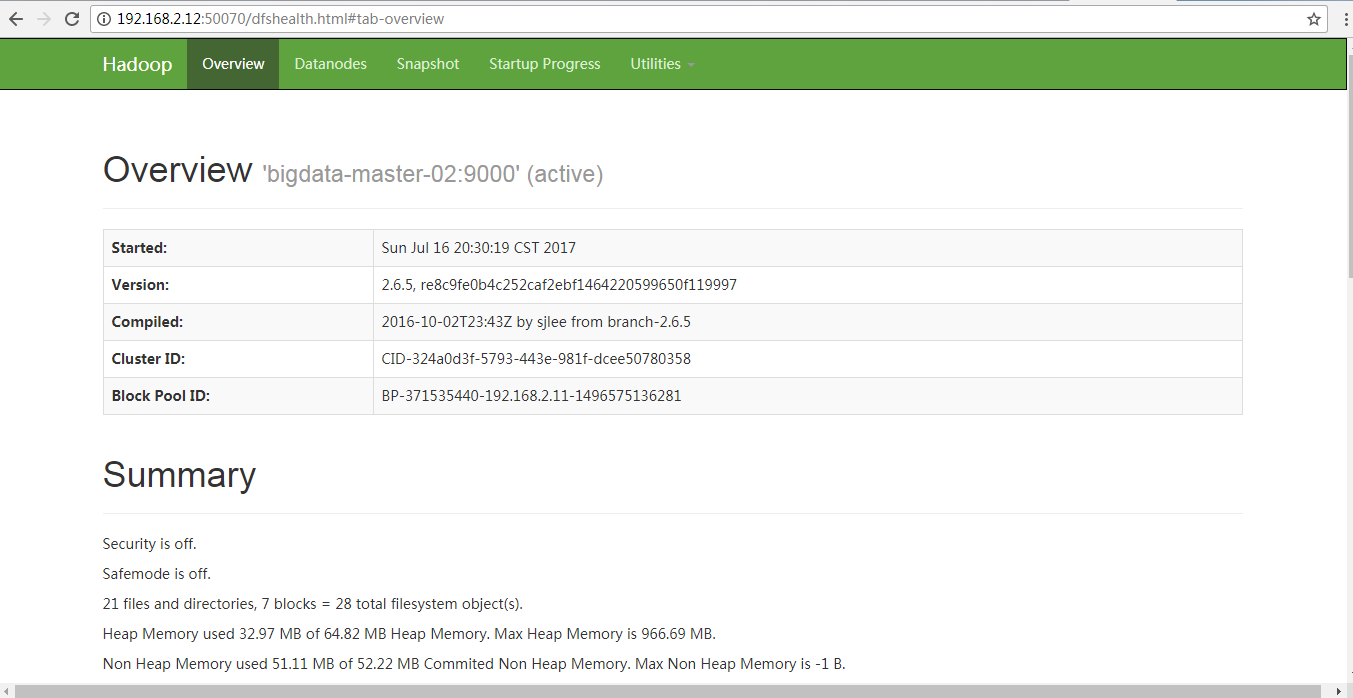
Slave-03

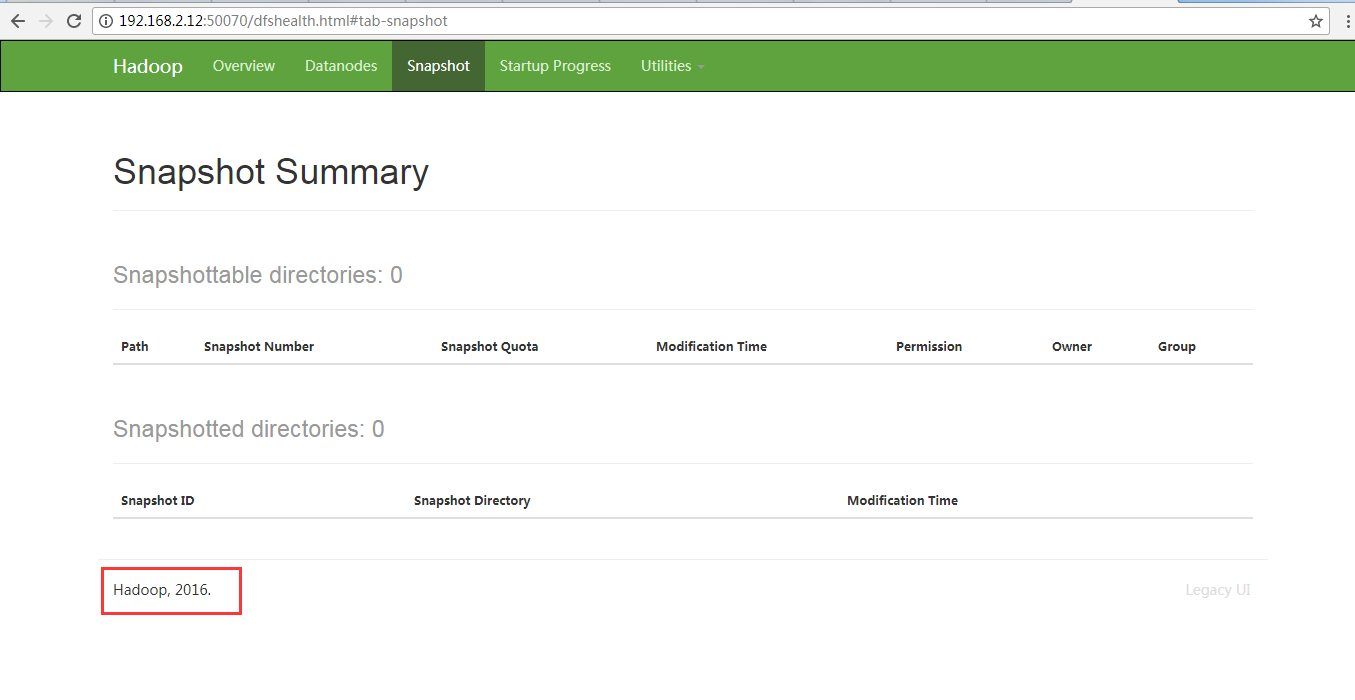


1. 验证ui









1. 准备hive数据
2. 创建表

create table wyp

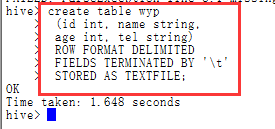
(id int, name string,

age int, tel string)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

STORED AS TEXTFILE;



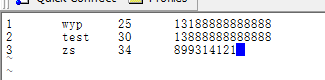
1. 数据文件

wyp.txt

1 wyp 25 13188888888888

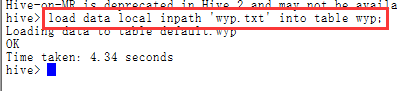
2 test 30 13888888888888

3 zs 34 899314121



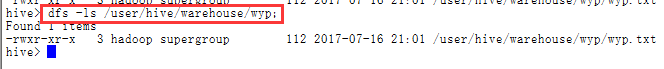
1. 数据导入到wyp表里面

load data local inpath 'wyp.txt' into table wyp;



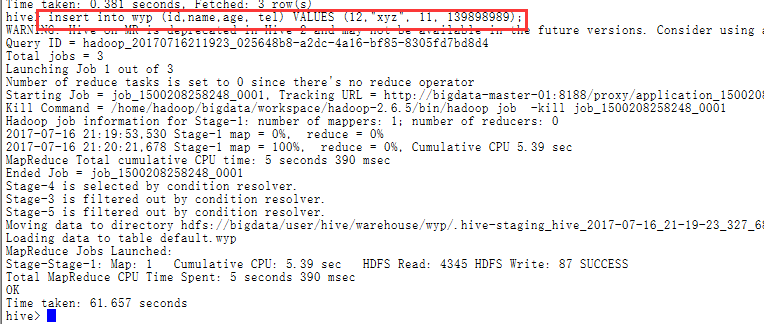
1. 查看文件

dfs -ls /user/hive/warehouse/wyp;

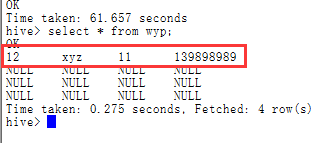


1. Insert 数据

insert into wyp (id,name,age, tel) VALUES (12,"xyz", 11, 139898989);



1. 查询数据



1. 升级hadoop
2. 先关闭集群， 再将新版本的hadoop安装包复制到各个node

scp -r /home/hadoop/bigdata/workspace/hadoop hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/

scp -r /home/hadoop/bigdata/workspace/hadoop hadoop@bigdata-slave-01:/home/hadoop/bigdata/workspace/

scp -r /home/hadoop/bigdata/workspace/hadoop hadoop@bigdata-slave-02:/home/hadoop/bigdata/workspace/

scp -r /home/hadoop/bigdata/workspace/hadoop hadoop@bigdata-slave-03:/home/hadoop/bigdata/workspace/

1. 在各节点执行以下命令将老版本hadoop的配置文件复制到新包中

cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/hadoop-env.sh /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/hdfs-site.xml /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/yarn-site.xml /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

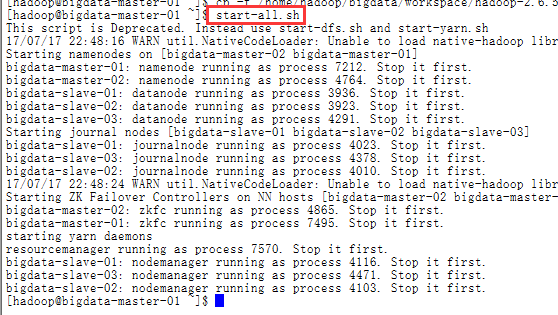
cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/slaves /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/core-site.xml /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

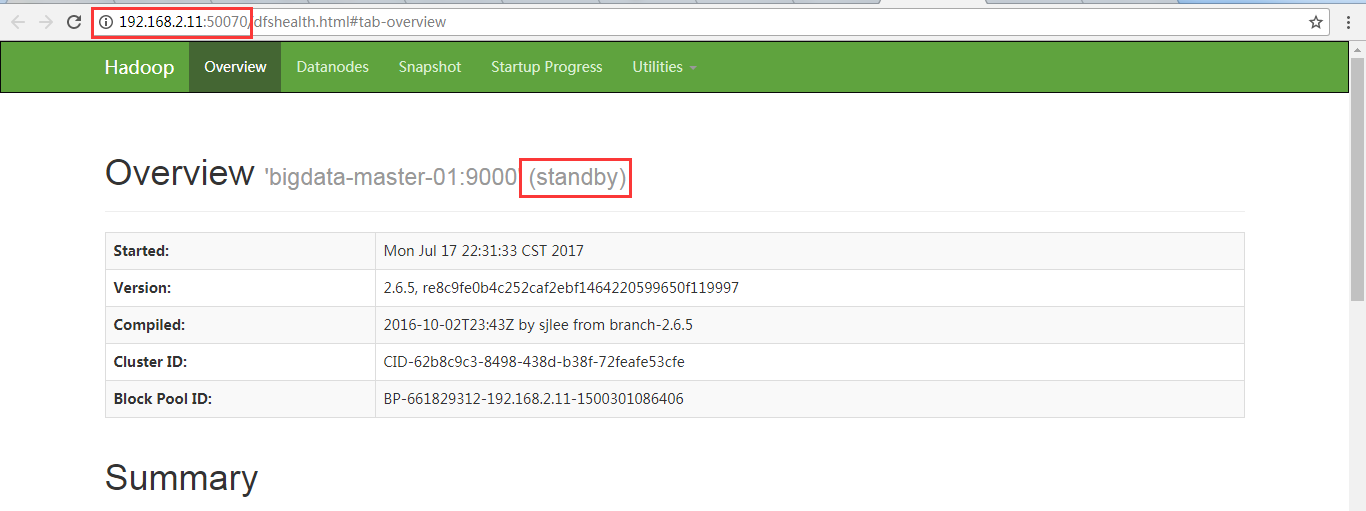
cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/mapred-site.xml /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

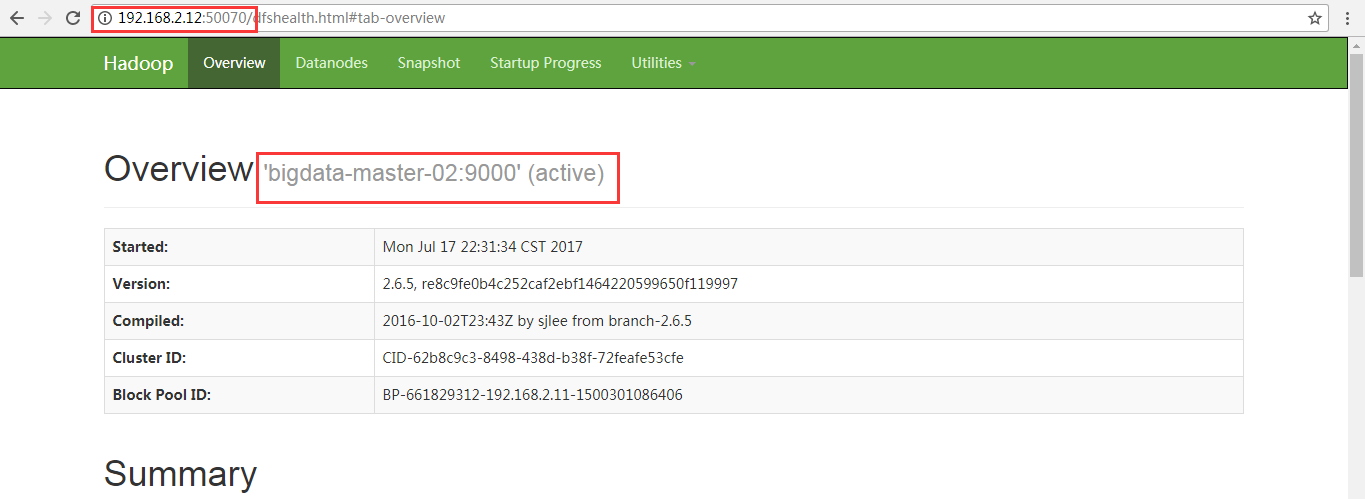
cp -f /home/hadoop/bigdata/workspace/hadoop-2.6.5/etc/hadoop/capacity-scheduler.xml /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

1. 启动老版本的hadoop集群

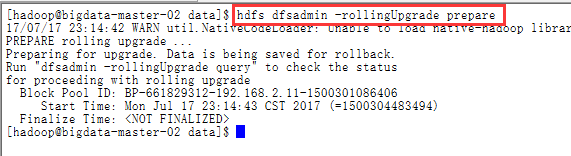


1. 确定active与standby

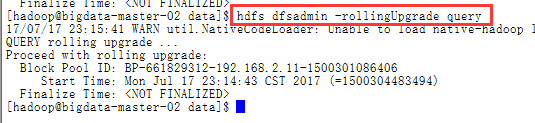




1. Prepare Rolling Upgrade
2. Run “hdfs dfsadmin -rollingUpgrade prepare” to create a fsimage for rollback.

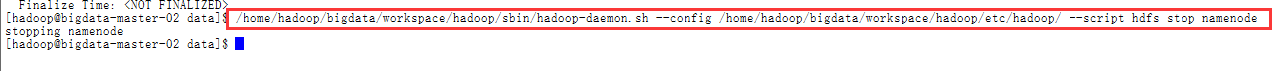


1. Run “hdfs dfsadmin -rollingUpgrade query” to check the status of the rollback image. Wait and re-run the command until the “Proceed with rolling upgrade” message is shown.

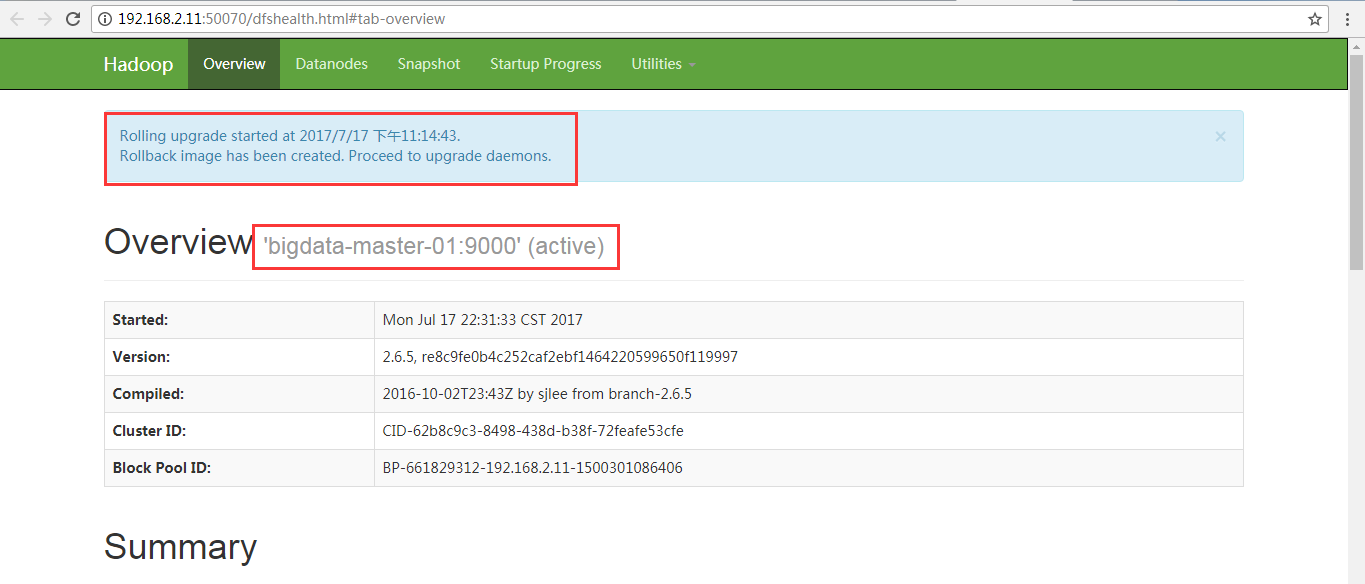


1. Upgrade Active and Standby NameNodes
2. Shutdown and upgrade NameNode2.

/home/hadoop/bigdata/workspace/hadoop/sbin/hadoop-daemon.sh --config /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/ --script hdfs stop namenode

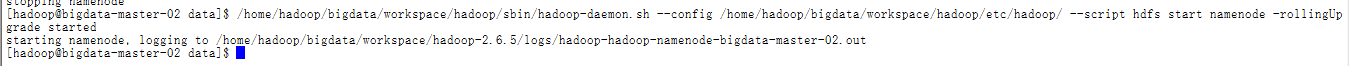


Check current active is NameNode1:

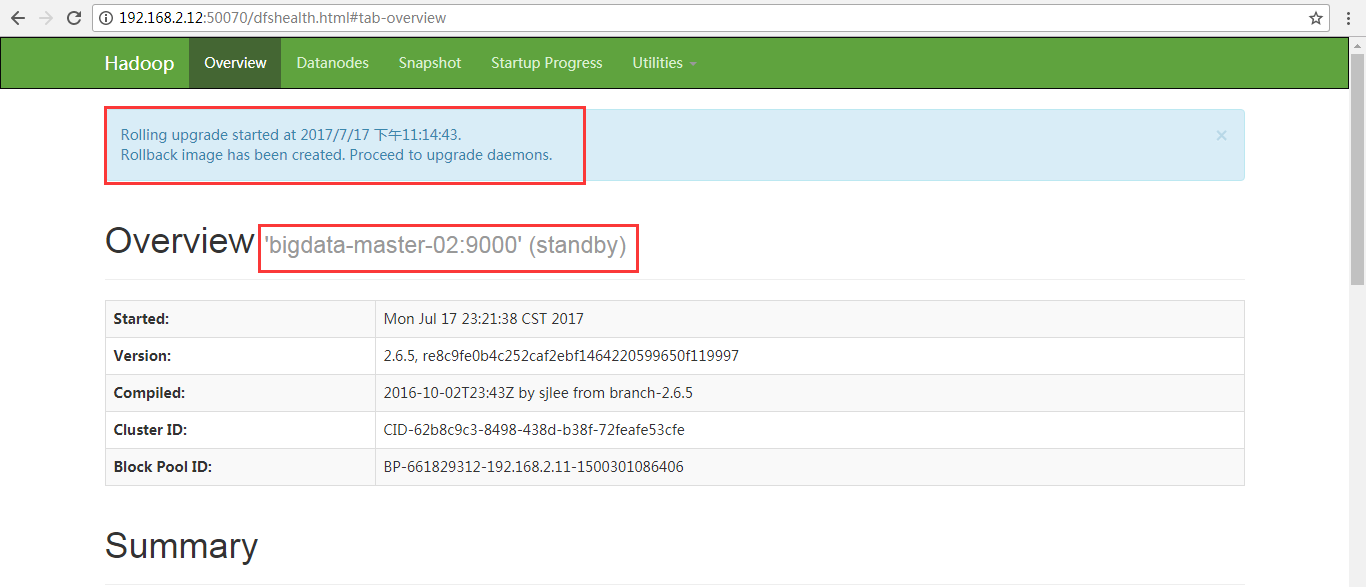


1. Start NN2 as standby with the “-rollingUpgrade started” option.

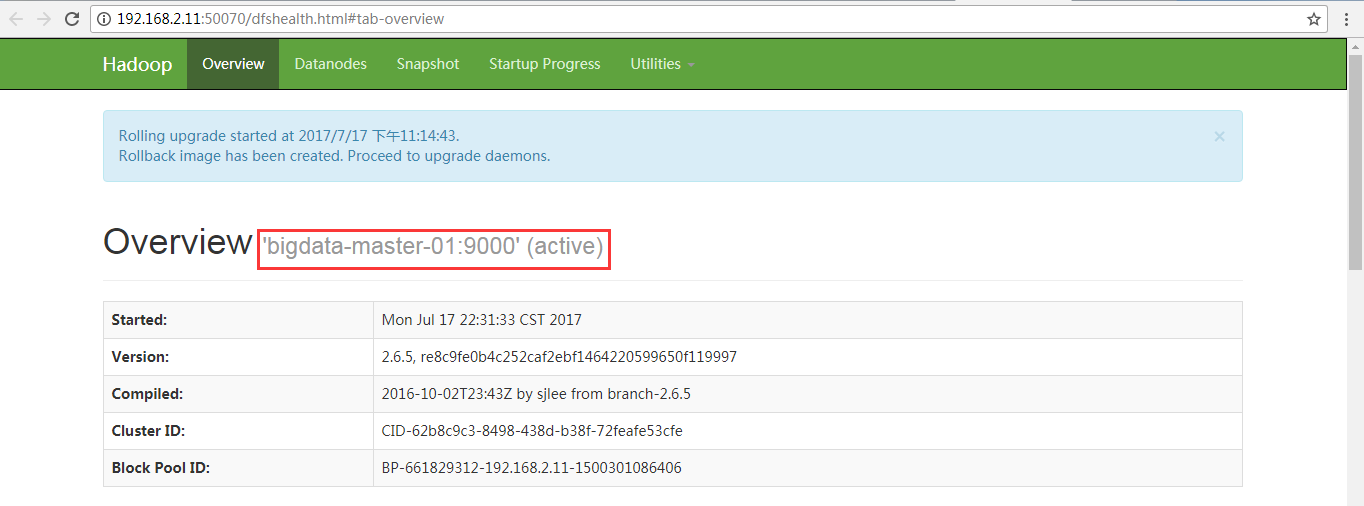
/home/hadoop/bigdata/workspace/hadoop/sbin/hadoop-daemon.sh --config /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/ --script hdfs start namenode -rollingUpgrade started

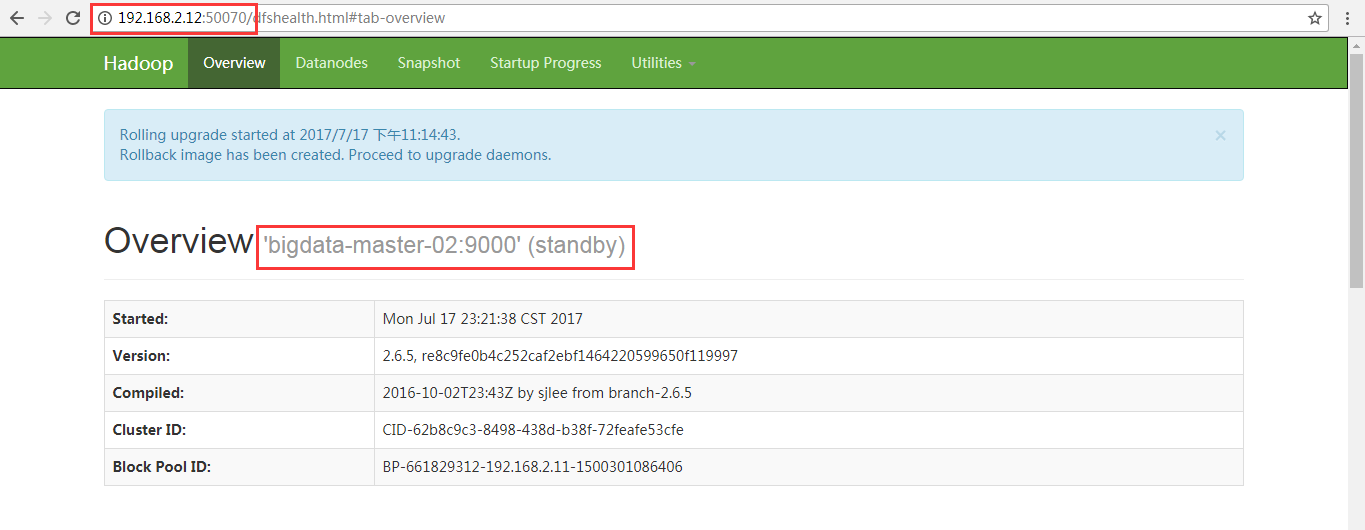


Check current NameNode2 is standby



3) Failover from NameNode1 to NameNode2 so that NN2 becomes active and NN1 becomes standby.



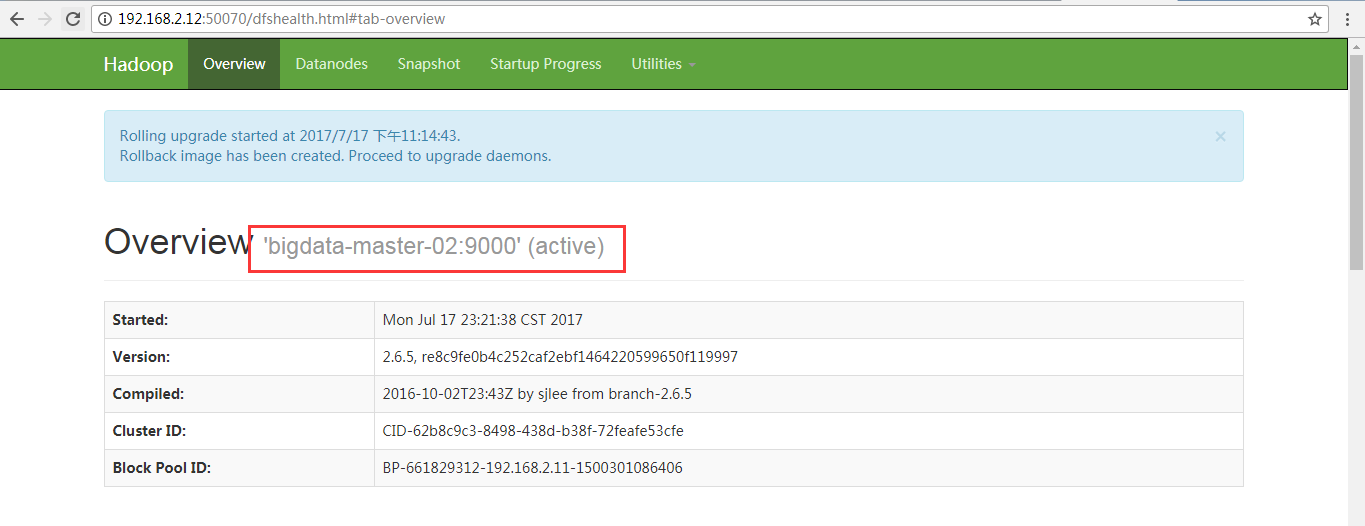


1. Shutdown and upgrade NameNode1.

/home/hadoop/bigdata/workspace/hadoop/sbin/hadoop-daemon.sh --config /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/ --script hdfs stop namenode

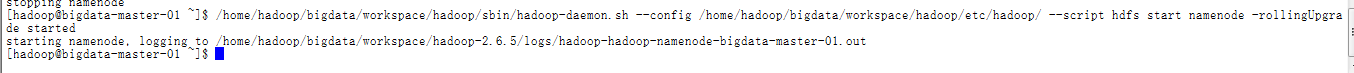


Check current active is NameNode2:

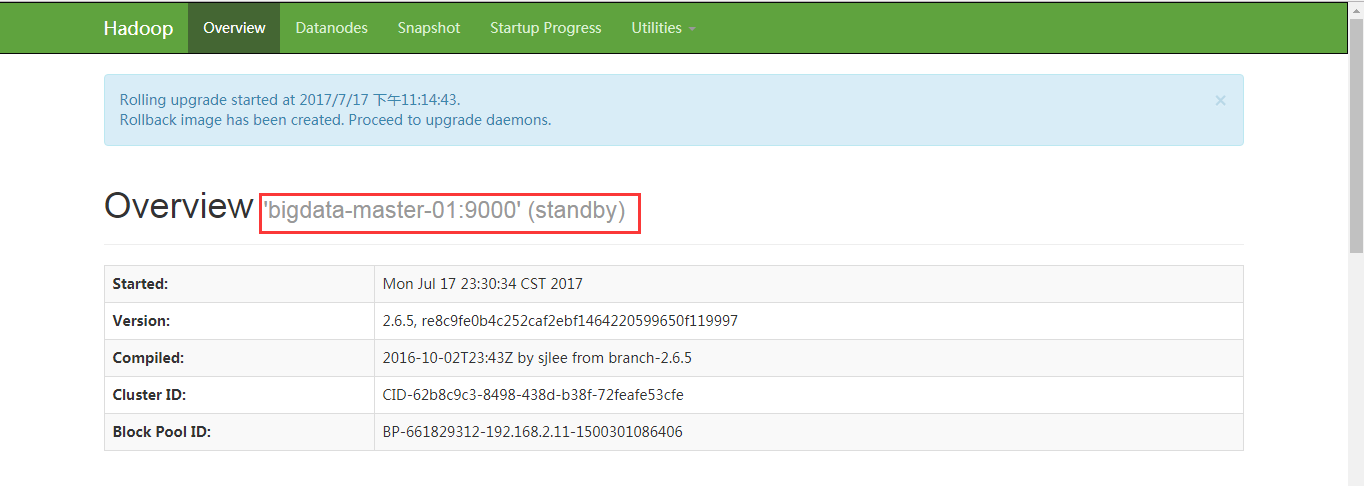


1. Start NameNode1 as standby with the “-rollingUpgrade started” option.

/home/hadoop/bigdata/workspace/hadoop/sbin/hadoop-daemon.sh --config /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/ --script hdfs start namenode -rollingUpgrade started

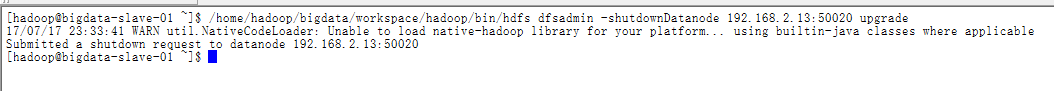


Check NameNode1 is standby



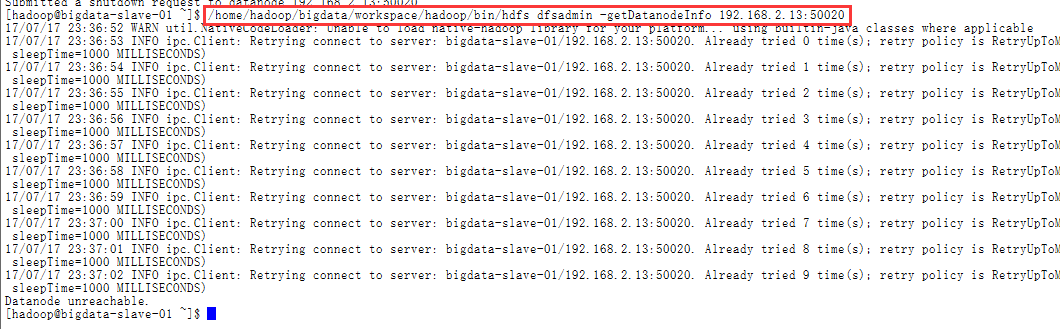
1. Upgrade DateNodes(在每一台DateNode都要如此操作)
2. Run “hdfs dfsadmin -shutdownDatanode <DATANODE\_HOST:IPC\_PORT> upgrade” to shutdown one of the chosen datanodes.

/home/hadoop/bigdata/workspace/hadoop/bin/hdfs dfsadmin -shutdownDatanode 192.168.2.13:50020 upgrade



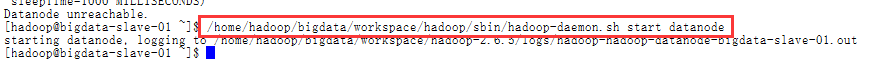
1. Run “hdfs dfsadmin -getDatanodeInfo <DATANODE\_HOST:IPC\_PORT>” to check and wait for the datanode to shutdown.

/home/hadoop/bigdata/workspace/hadoop/bin/hdfs dfsadmin -getDatanodeInfo 192.168.2.13:50020

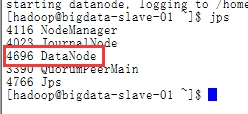


1. Upgrade and restart the datanode.

/home/hadoop/bigdata/workspace/hadoop/sbin/hadoop-daemon.sh start datanode



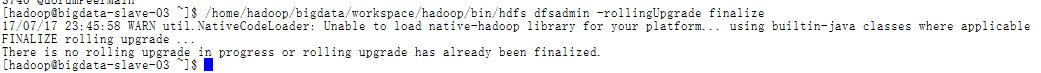
1. 查看节点是否启动



1. Repeat the above steps until all datanodes in the cluster are upgraded.
2. Finalize Rolling Upgrade（此阶段一定要在namenode上操作）

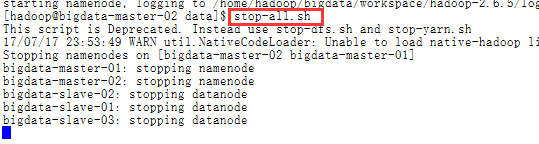
Run “hdfs dfsadmin -rollingUpgrade finalize” to finalize the rolling upgrade.

/home/hadoop/bigdata/workspace/hadoop/bin/hdfs dfsadmin -rollingUpgrade finalize



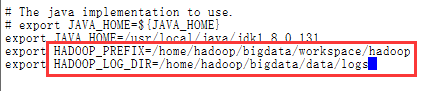
9．关闭集群

stop-all.sh



1. 修改新版本的hadoop-env.sh，配置log文件路径

vi /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/hadoop-env.sh



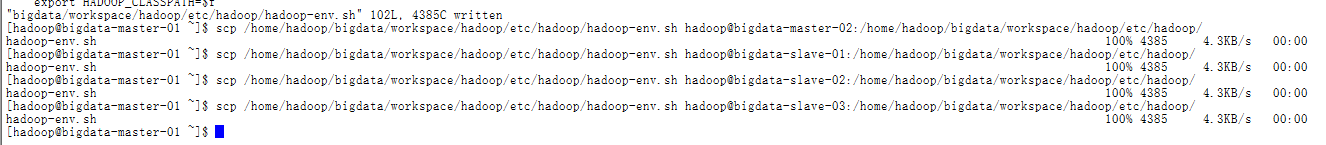
将修改好的配置文件同步到其它节点

scp /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/hadoop-env.sh hadoop@bigdata-master-02:/home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

scp /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/hadoop-env.sh hadoop@bigdata-slave-01:/home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

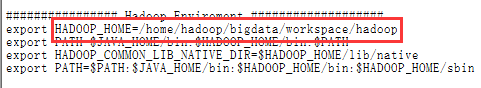
scp /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/hadoop-env.sh hadoop@bigdata-slave-02:/home/hadoop/bigdata/workspace/hadoop/etc/hadoop/

scp /home/hadoop/bigdata/workspace/hadoop/etc/hadoop/hadoop-env.sh hadoop@bigdata-slave-03:/home/hadoop/bigdata/workspace/hadoop/etc/hadoop/



1. 修改各节点的hadoop修改变量

vi ~/.bashrc

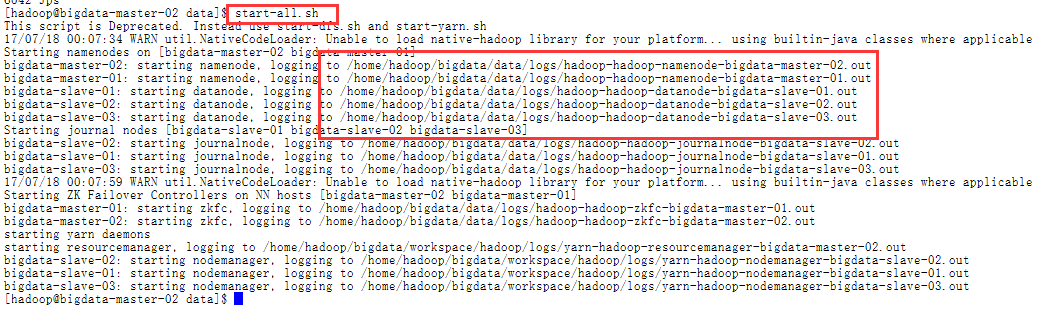


source ~/.bashrc

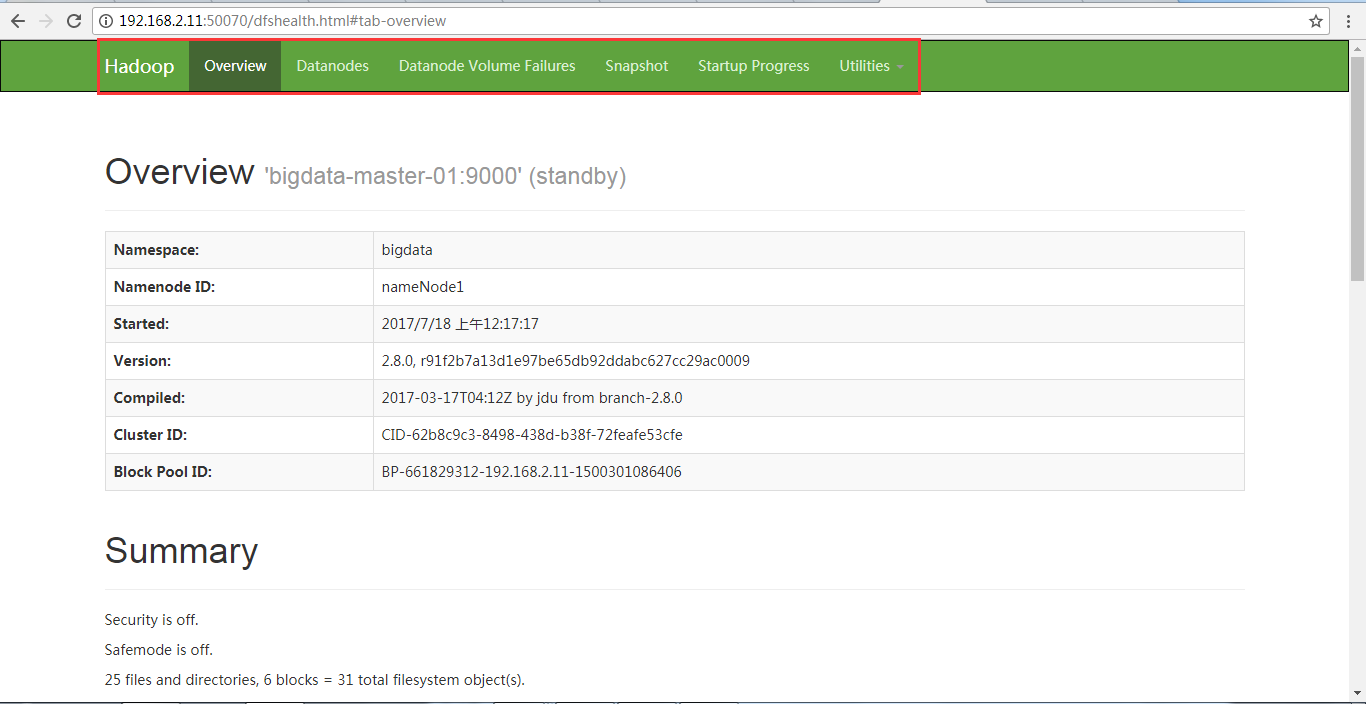
1. 将各节点上老版本hadoop标示为retired

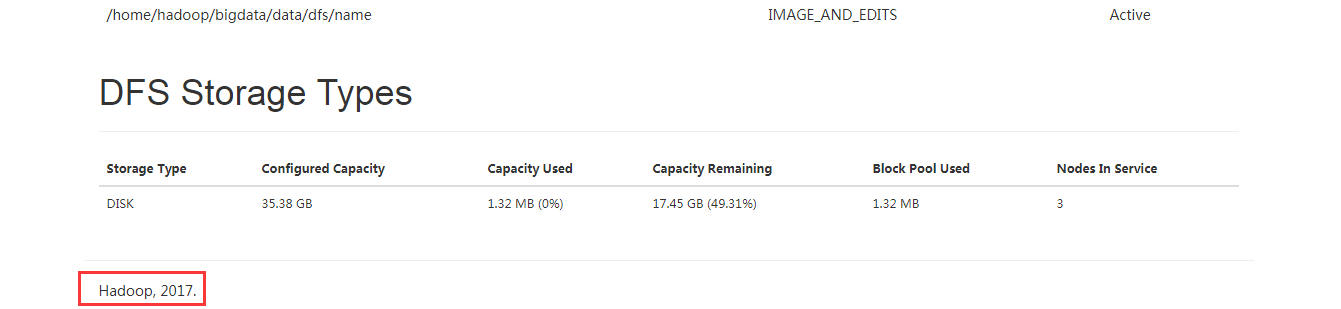
mv /home/hadoop/bigdata/workspace/hadoop-2.6.5 /home/hadoop/bigdata/workspace/hadoop-retired

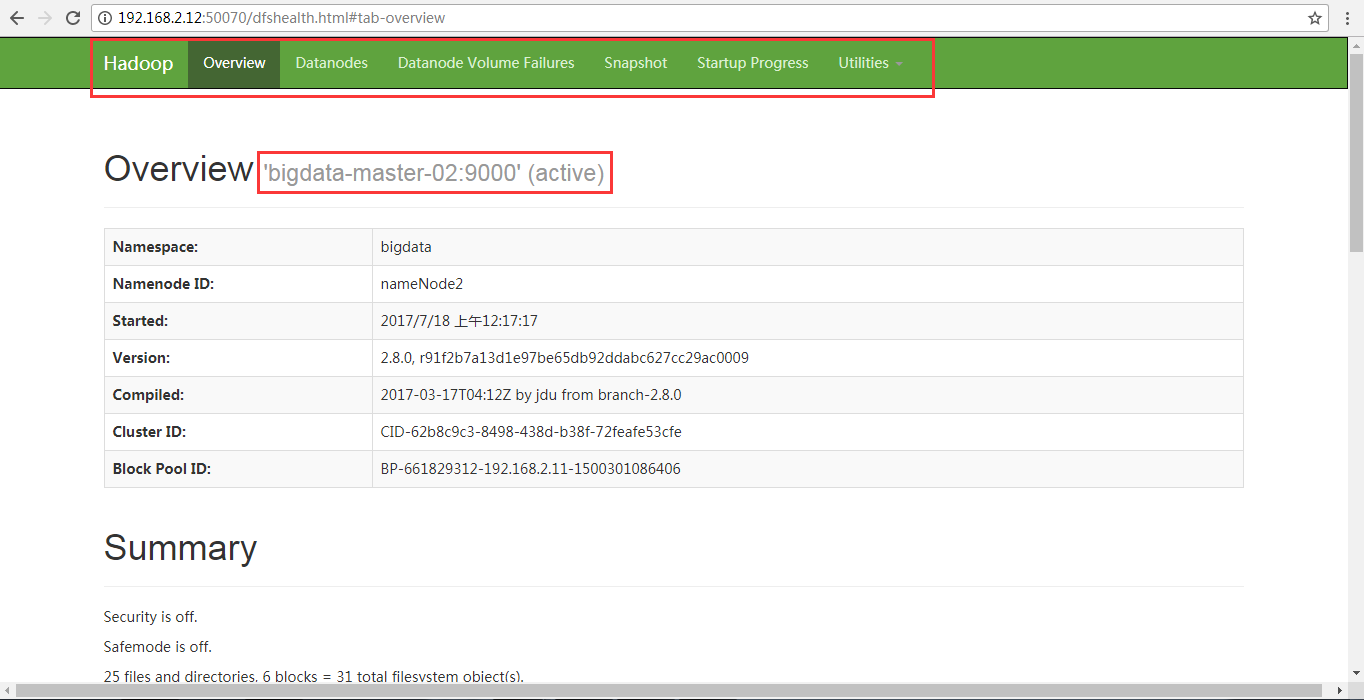
1. 验证hadoop升级是否成功
2. 启动集群

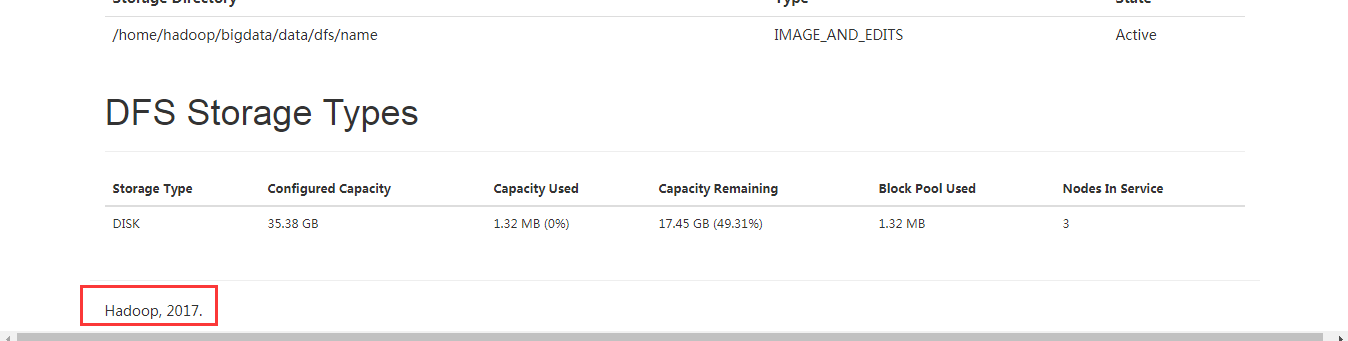


1. 查看hdfs ui版本



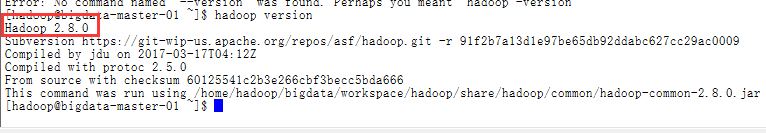




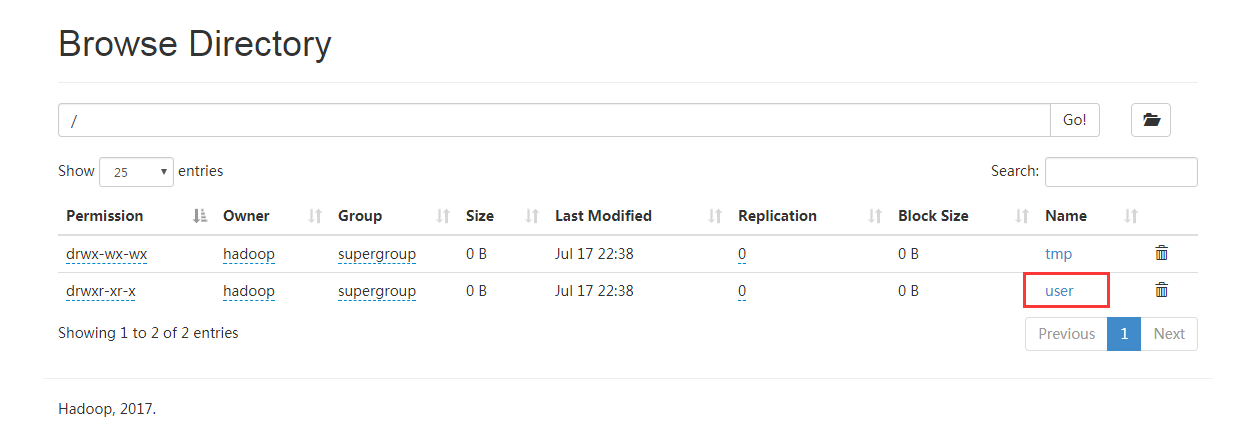


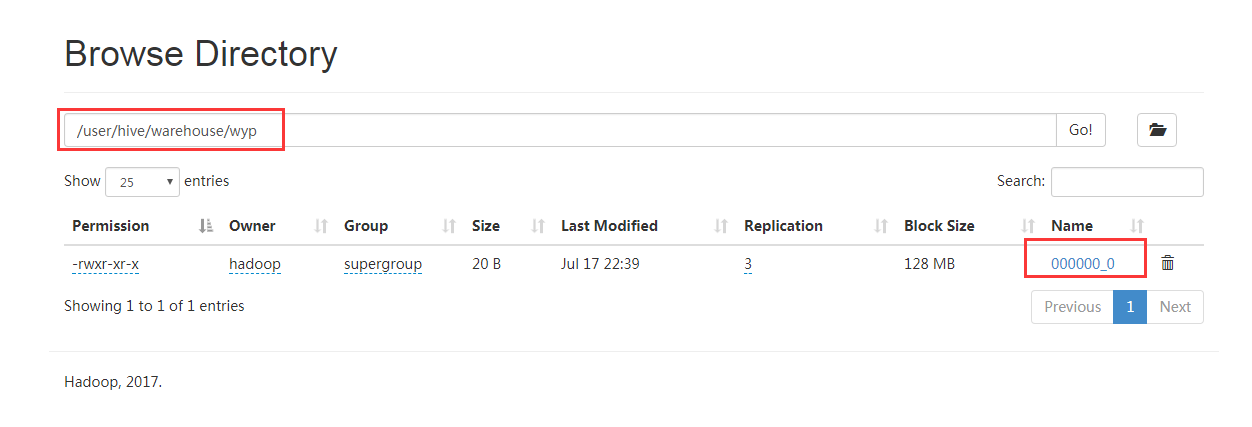
1. 查看当前hadoop版本

hadoop version

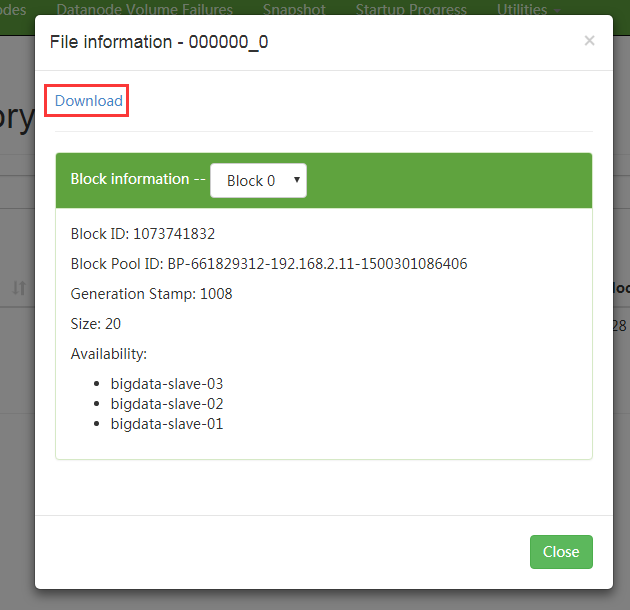


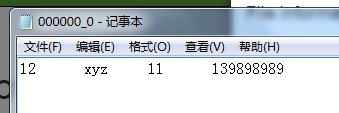
1. 查看hdfs文件





1. 查看文件内容





1. 查看hive

