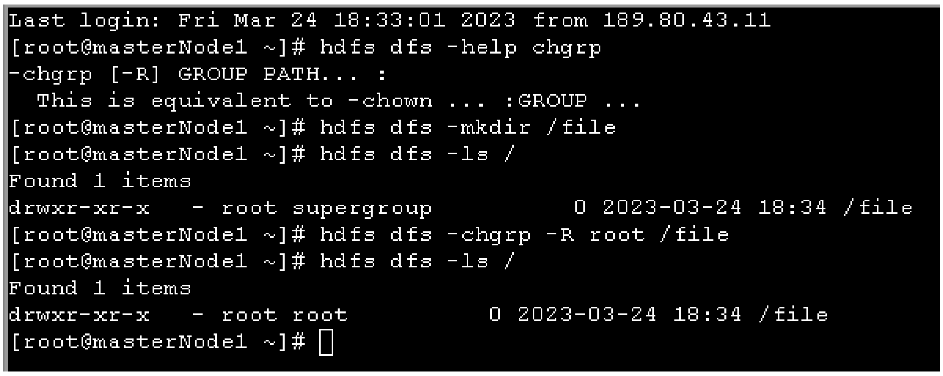
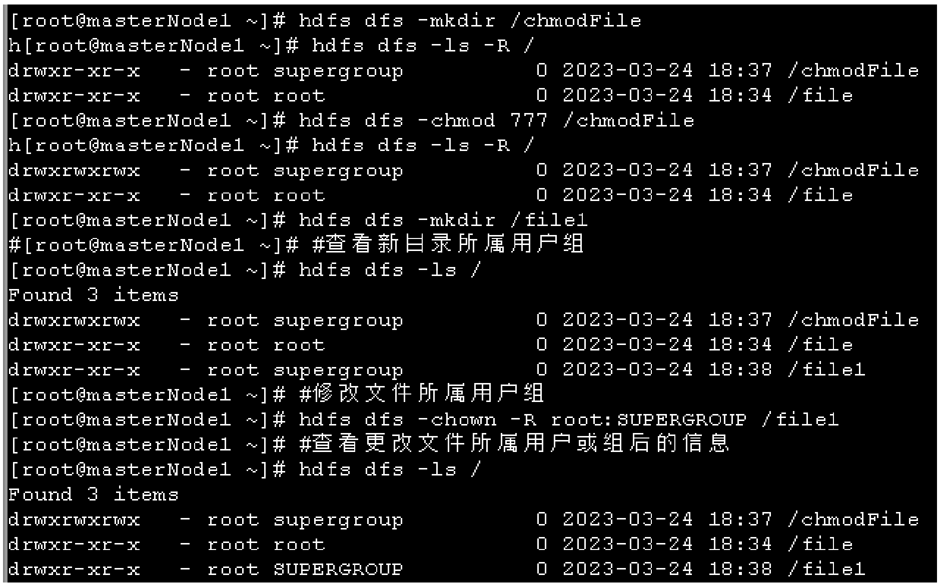
第一次作业

查看帮助命令



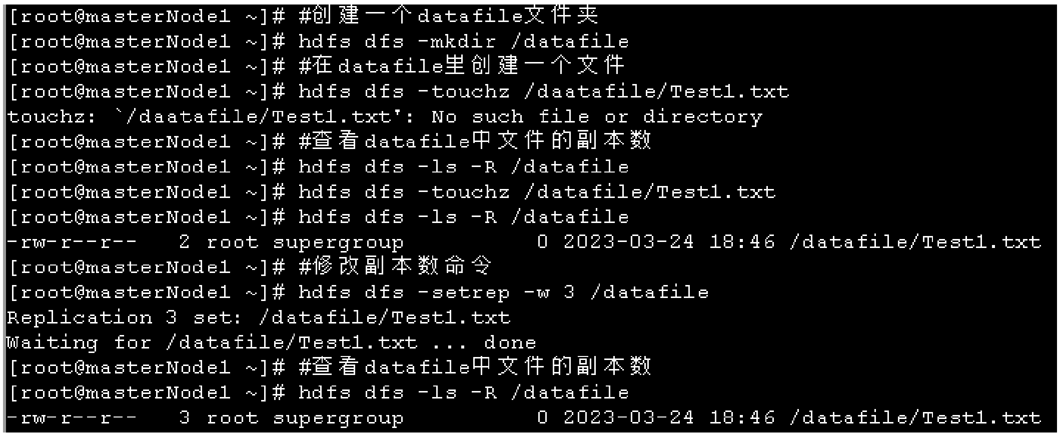
```
Last login: Fri Mar 24 18:33:01 2023 from 189.80.43.11
[root@masterNode1 ~]# hdfs dfs -help chgrp
-chgrp [-R] GROUP PATH... :
  This is equivalent to -chown ... :GROUP ...
[root@masterNode1 ~]# hdfs dfs -mkdir /file
[root@masterNode1 ~]# hdfs dfs -ls /
Found 1 items
drwxr-xr-x   - root supergroup          0 2023-03-24 18:34 /file
[root@masterNode1 ~]# hdfs dfs -chgrp -R root /file
[root@masterNode1 ~]# hdfs dfs -ls /
Found 1 items
drwxr-xr-x   - root root          0 2023-03-24 18:34 /file
[root@masterNode1 ~]# 
```

查看新目录所属用户组、修改文件所属用户组



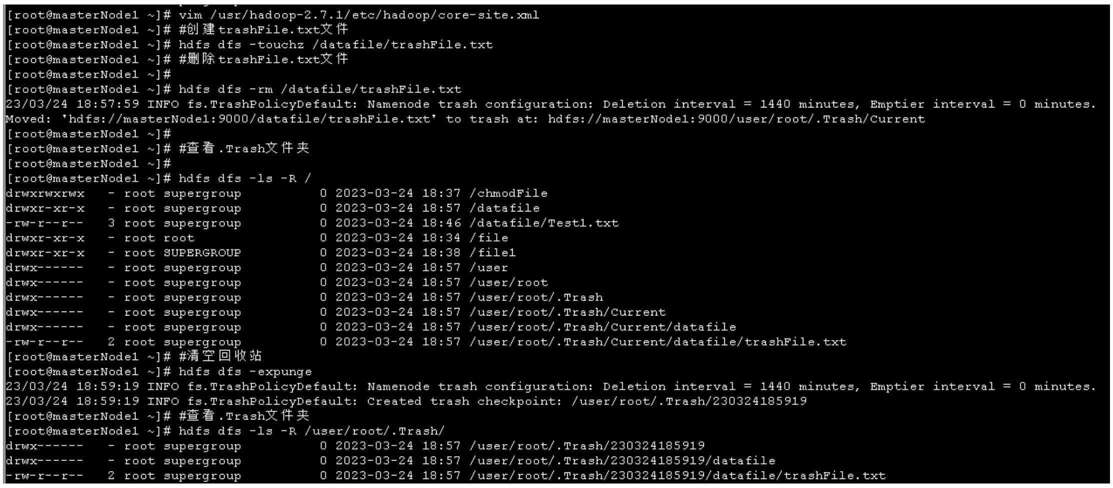
```
[root@masterNode1 ~]# hdfs dfs -mkdir /chmodFile
h[root@masterNode1 ~]# hdfs dfs -ls -R /
drwxr-xr-x   - root supergroup          0 2023-03-24 18:37 /chmodFile
drwxr-xr-x   - root root                0 2023-03-24 18:34 /file
[root@masterNode1 ~]# hdfs dfs -chmod 777 /chmodFile
h[root@masterNode1 ~]# hdfs dfs -ls -R /
drwxrwxrwx   - root supergroup          0 2023-03-24 18:37 /chmodFile
drwxr-xr-x   - root root                0 2023-03-24 18:34 /file
[root@masterNode1 ~]# hdfs dfs -mkdir /file1
#[root@masterNode1 ~]# #查看新目录所属用户组
[root@masterNode1 ~]# hdfs dfs -ls /
Found 3 items
drwxrwxrwx   - root supergroup          0 2023-03-24 18:37 /chmodFile
drwxr-xr-x   - root root                0 2023-03-24 18:34 /file
drwxr-xr-x   - root supergroup          0 2023-03-24 18:38 /file1
[root@masterNode1 ~]# #修改文件所属用户组
[root@masterNode1 ~]# hdfs dfs -chown -R root:SUPERGROUP /file1
[root@masterNode1 ~]# #查看更改文件所属用户或组后的信息
[root@masterNode1 ~]# hdfs dfs -ls /
Found 3 items
drwxrwxrwx   - root supergroup          0 2023-03-24 18:37 /chmodFile
drwxr-xr-x   - root root                0 2023-03-24 18:34 /file
drwxr-xr-x   - root SUPERGROUP          0 2023-03-24 18:38 /file1
```

创建一个 datafile 文件夹、在 datafile 里创建一个文件、查看 datafile 中文件的副本数、修改副本数命令



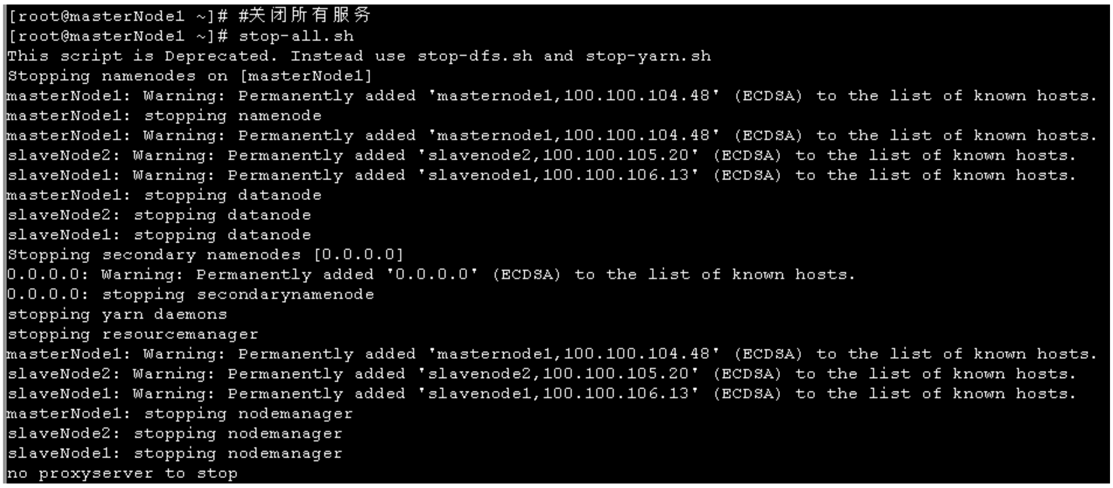
```
[root@masterNode1 ~]# #创建一个datafile文件夹
[root@masterNode1 ~]# hdfs dfs -mkdir /datafile
[root@masterNode1 ~]# #在datafile里创建一个文件
[root@masterNode1 ~]# hdfs dfs -touchz /daatafile/Test1.txt
touchz: `/daatafile/Test1.txt': No such file or directory
[root@masterNode1 ~]# #查看datafile中文件的副本数
[root@masterNode1 ~]# hdfs dfs -ls -R /datafile
[root@masterNode1 ~]# hdfs dfs -touchz /datafile/Test1.txt
[root@masterNode1 ~]# hdfs dfs -ls -R /datafile
-rw-r--r--   2 root supergroup          0 2023-03-24 18:46 /datafile/Test1.txt
[root@masterNode1 ~]# #修改副本数命令
[root@masterNode1 ~]# hdfs dfs -setrep -w 3 /datafile
Replication 3 set: /datafile/Test1.txt
Waiting for /datafile/Test1.txt ... done
[root@masterNode1 ~]# #查看datafile中文件的副本数
[root@masterNode1 ~]# hdfs dfs -ls -R /datafile
-rw-r--r--   3 root supergroup          0 2023-03-24 18:46 /datafile/Test1.txt
```

创建 trashFile.txt 文件、查看 Trash 文件、清空回收站



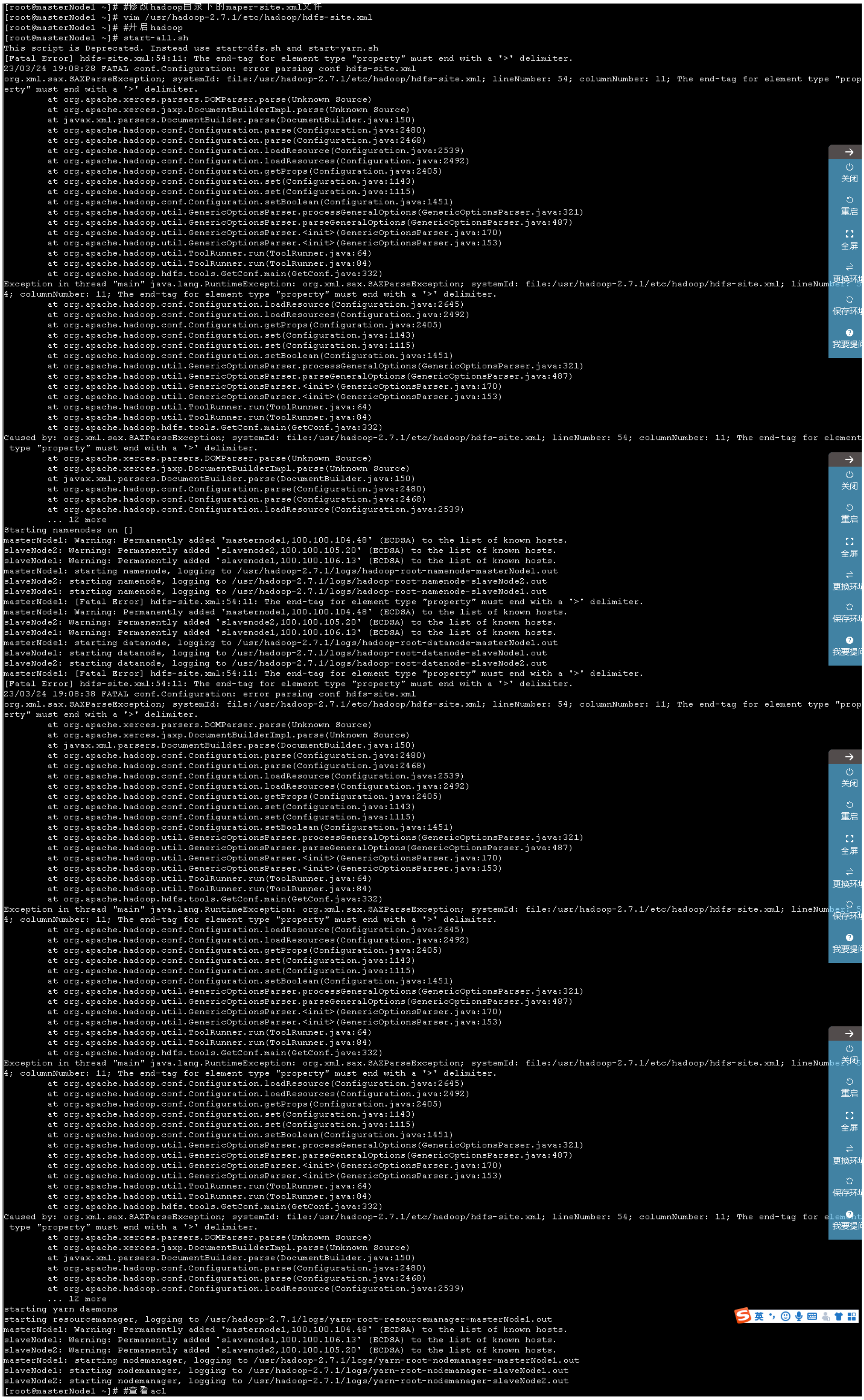
```
[root@masterNode1 ~]# vim /usr/hadoop-2.7.1/etc/hadoop/core-site.xml
[root@masterNode1 ~]# #创建trashFile.txt文件
[root@masterNode1 ~]# hdfs dfs -touchz /datafile/trashFile.txt
[root@masterNode1 ~]# #删除trashFile.txt文件
[root@masterNode1 ~]#
[root@masterNode1 ~]# hdfs dfs -rm /datafile/trashFile.txt
23/03/24 18:57:59 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 1440 minutes, Emptier interval = 0 minutes.
Moved: 'hdfs://masterNode1:9000/datafile/trashFile.txt' to trash at: hdfs://masterNode1:9000/user/root/.Trash/Current
[root@masterNode1 ~]#
[root@masterNode1 ~]# #查看.Trash文件夹
[root@masterNode1 ~]#
[root@masterNode1 ~]# hdfs dfs -ls -R /
drwxrwxrwx   - root supergroup          0 2023-03-24 18:37 /chmodFile
drwxr-xr-x   - root supergroup          0 2023-03-24 18:57 /datafile
-rw-r--r--   3 root supergroup          0 2023-03-24 18:46 /datafile/Test1.txt
drwxr-xr-x   - root root                0 2023-03-24 18:34 /file
drwxr-xr-x   - root SUPERGROUP          0 2023-03-24 18:38 /file1
drwx------   - root supergroup          0 2023-03-24 18:57 /user
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root/.Trash
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root/.Trash/Current
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root/.Trash/Current/datafile
-rw-r--r--   2 root supergroup          0 2023-03-24 18:57 /user/root/.Trash/Current/datafile/trashFile.txt
[root@masterNode1 ~]# #清空回收站
[root@masterNode1 ~]# hdfs dfs -expunge
23/03/24 18:59:19 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 1440 minutes, Emptier interval = 0 minutes.
23/03/24 18:59:19 INFO fs.TrashPolicyDefault: Created trash checkpoint: /user/root/.Trash/230324185919
[root@masterNode1 ~]# #查看.Trash文件夹
[root@masterNode1 ~]# hdfs dfs -ls -R /user/root/.Trash/
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root/.Trash/230324185919
drwx------   - root supergroup          0 2023-03-24 18:57 /user/root/.Trash/230324185919/datafile
-rw-r--r--   2 root supergroup          0 2023-03-24 18:57 /user/root/.Trash/230324185919/datafile/trashFile.txt
```

关闭所有服务



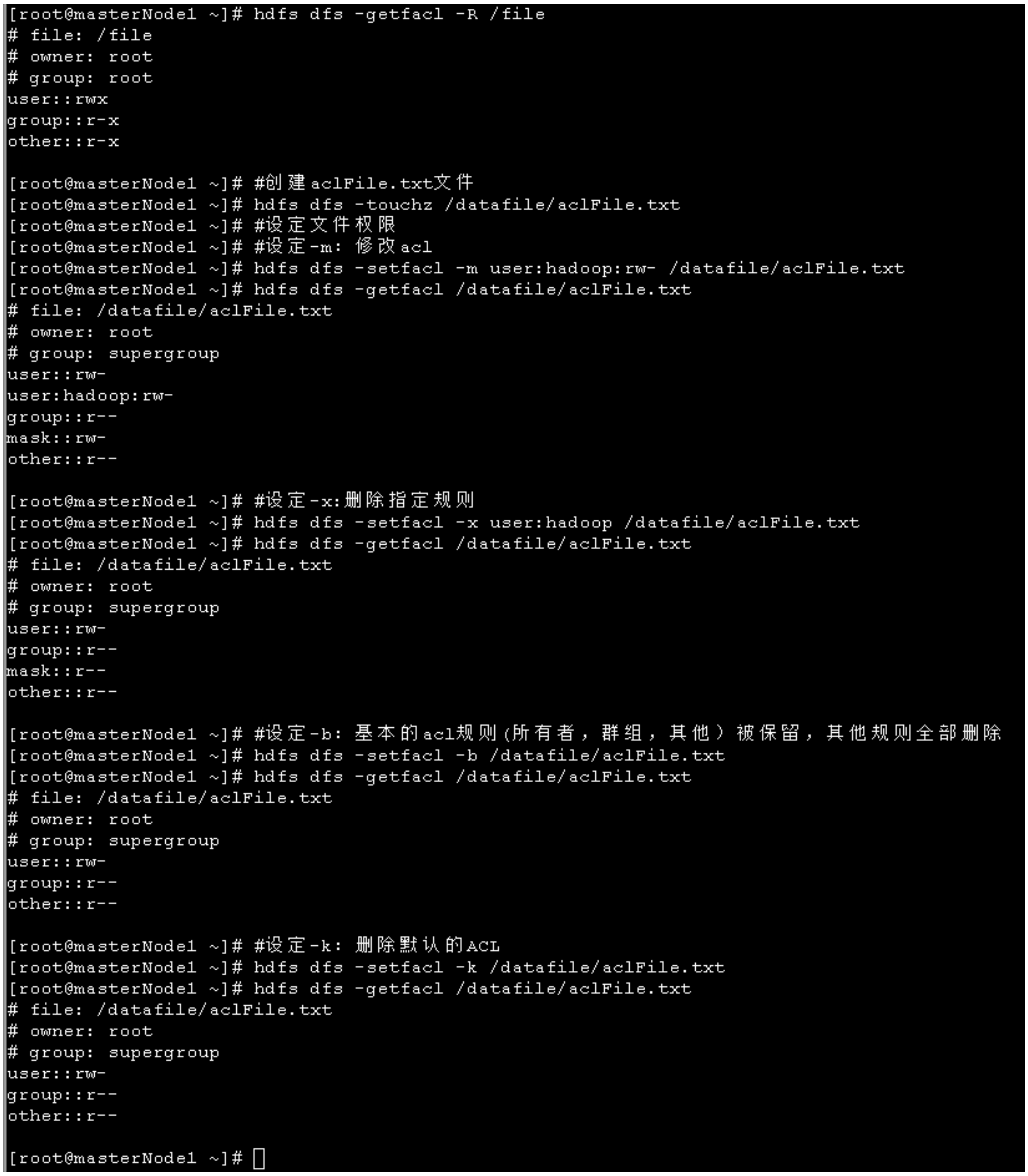
```
[root@masterNode1 ~]# #关闭所有服务
[root@masterNode1 ~]# stop-all.sh
This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh
Stopping namenodes on [masterNode1]
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
masterNode1: stopping namenode
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
slaveNode2: Warning: Permanently added 'slavenode2,100.100.105.20' (ECDSA) to the list of known hosts.
slaveNode1: Warning: Permanently added 'slavenode1,100.100.106.13' (ECDSA) to the list of known hosts.
masterNode1: stopping datanode
slaveNode2: stopping datanode
slaveNode1: stopping datanode
Stopping secondary namenodes [0.0.0.0]
0.0.0.0: Warning: Permanently added '0.0.0.0' (ECDSA) to the list of known hosts.
0.0.0.0: stopping secondarynamenode
stopping yarn daemons
stopping resourcemanager
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
slaveNode2: Warning: Permanently added 'slavenode2,100.100.105.20' (ECDSA) to the list of known hosts.
slaveNode1: Warning: Permanently added 'slavenode1,100.100.106.13' (ECDSA) to the list of known hosts.
masterNode1: stopping nodemanager
slaveNode2: stopping nodemanager
slaveNode1: stopping nodemanager
no proxyserver to stop
```

开启 hadoop

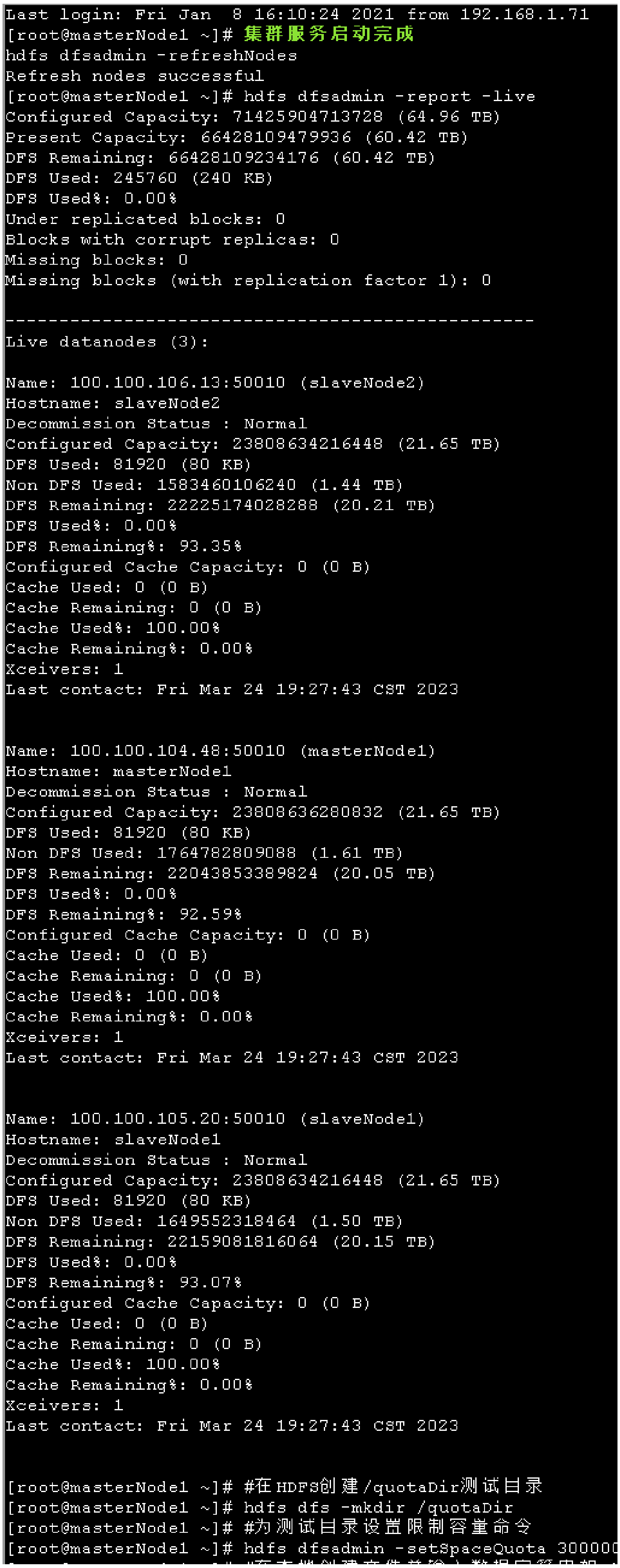


```
[root@masterNode1 ~]# #修改hadoop目录下的maper-site.xml文件
[root@masterNode1 ~]# vim /usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml
[root@masterNode1 ~]# #开启hadoop
[root@masterNode1 ~]# start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
[Fatal Error] hdfs-site.xml:54:11: The end-tag for element type "property" must end with a '>' delimiter.
23/03/24 19:08:28 FATAL conf.Configuration: error parsing conf hdfs-site.xml
org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 54; columnNumber: 11; The end-tag for element type "prop
erty" must end with a '>' delimiter.
        at org.apache.xerces.parsers.DOMParser.parse(Unknown Source)
        at org.apache.xerces.jaxp.DocumentBuilderImpl.parse(Unknown Source)
        at javax.xml.parsers.DocumentBuilder.parse(DocumentBuilder.java:150)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2480)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2468)
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2539)
        at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2492)
        at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2405)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1143)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1115)
        at org.apache.hadoop.conf.Configuration.setBoolean(Configuration.java:1451)
        at org.apache.hadoop.util.GenericOptionsParser.processGeneralOptions(GenericOptionsParser.java:321)
        at org.apache.hadoop.util.GenericOptionsParser.parseGeneralOptions(GenericOptionsParser.java:487)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:170)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:153)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:64)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
        at org.apache.hadoop.hdfs.tools.GetConf.main(GetConf.java:332)
Exception in thread "main" java.lang.RuntimeException: org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 5
4; columnNumber: 11; The end-tag for element type "property" must end with a '>' delimiter.
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2645)
        at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2492)
        at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2405)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1143)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1115)
        at org.apache.hadoop.conf.Configuration.setBoolean(Configuration.java:1451)
        at org.apache.hadoop.util.GenericOptionsParser.processGeneralOptions(GenericOptionsParser.java:321)
        at org.apache.hadoop.util.GenericOptionsParser.parseGeneralOptions(GenericOptionsParser.java:487)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:170)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:153)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:64)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
        at org.apache.hadoop.hdfs.tools.GetConf.main(GetConf.java:332)
Caused by: org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 54; columnNumber: 11; The end-tag for element
 type "property" must end with a '>' delimiter.
        at org.apache.xerces.parsers.DOMParser.parse(Unknown Source)
        at org.apache.xerces.jaxp.DocumentBuilderImpl.parse(Unknown Source)
        at javax.xml.parsers.DocumentBuilder.parse(DocumentBuilder.java:150)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2480)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2468)
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2539)
        ... 12 more
Starting namenodes on []
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
slaveNode2: Warning: Permanently added 'slavenode2,100.100.105.20' (ECDSA) to the list of known hosts.
slaveNode1: Warning: Permanently added 'slavenode1,100.100.106.13' (ECDSA) to the list of known hosts.
masterNode1: starting namenode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-namenode-masterNode1.out
slaveNode2: starting namenode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-namenode-slaveNode2.out
slaveNode1: starting namenode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-namenode-slaveNode1.out
masterNode1: [Fatal Error] hdfs-site.xml:54:11: The end-tag for element type "property" must end with a '>' delimiter.
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
slaveNode2: Warning: Permanently added 'slavenode2,100.100.105.20' (ECDSA) to the list of known hosts.
slaveNode1: Warning: Permanently added 'slavenode1,100.100.106.13' (ECDSA) to the list of known hosts.
masterNode1: starting datanode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-datanode-masterNode1.out
slaveNode1: starting datanode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-datanode-slaveNode1.out
slaveNode2: starting datanode, logging to /usr/hadoop-2.7.1/logs/hadoop-root-datanode-slaveNode2.out
masterNode1: [Fatal Error] hdfs-site.xml:54:11: The end-tag for element type "property" must end with a '>' delimiter.
[Fatal Error] hdfs-site.xml:54:11: The end-tag for element type "property" must end with a '>' delimiter.
23/03/24 19:08:38 FATAL conf.Configuration: error parsing conf hdfs-site.xml
org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 54; columnNumber: 11; The end-tag for element type "prop
erty" must end with a '>' delimiter.
        at org.apache.xerces.parsers.DOMParser.parse(Unknown Source)
        at org.apache.xerces.jaxp.DocumentBuilderImpl.parse(Unknown Source)
        at javax.xml.parsers.DocumentBuilder.parse(DocumentBuilder.java:150)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2480)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2468)
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2539)
        at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2492)
        at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2405)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1143)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1115)
        at org.apache.hadoop.conf.Configuration.setBoolean(Configuration.java:1451)
        at org.apache.hadoop.util.GenericOptionsParser.processGeneralOptions(GenericOptionsParser.java:321)
        at org.apache.hadoop.util.GenericOptionsParser.parseGeneralOptions(GenericOptionsParser.java:487)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:170)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:153)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:64)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
        at org.apache.hadoop.hdfs.tools.GetConf.main(GetConf.java:332)
Exception in thread "main" java.lang.RuntimeException: org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 5
4; columnNumber: 11; The end-tag for element type "property" must end with a '>' delimiter.
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2645)
        at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2492)
        at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2405)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1143)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1115)
        at org.apache.hadoop.conf.Configuration.setBoolean(Configuration.java:1451)
        at org.apache.hadoop.util.GenericOptionsParser.processGeneralOptions(GenericOptionsParser.java:321)
        at org.apache.hadoop.util.GenericOptionsParser.parseGeneralOptions(GenericOptionsParser.java:487)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:170)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:153)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:64)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
        at org.apache.hadoop.hdfs.tools.GetConf.main(GetConf.java:332)
Exception in thread "main" java.lang.RuntimeException: org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 5
4; columnNumber: 11; The end-tag for element type "property" must end with a '>' delimiter.
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2645)
        at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2492)
        at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2405)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1143)
        at org.apache.hadoop.conf.Configuration.set(Configuration.java:1115)
        at org.apache.hadoop.conf.Configuration.setBoolean(Configuration.java:1451)
        at org.apache.hadoop.util.GenericOptionsParser.processGeneralOptions(GenericOptionsParser.java:321)
        at org.apache.hadoop.util.GenericOptionsParser.parseGeneralOptions(GenericOptionsParser.java:487)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:170)
        at org.apache.hadoop.util.GenericOptionsParser.<init>(GenericOptionsParser.java:153)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:64)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
        at org.apache.hadoop.hdfs.tools.GetConf.main(GetConf.java:332)
Caused by: org.xml.sax.SAXParseException; systemId: file:/usr/hadoop-2.7.1/etc/hadoop/hdfs-site.xml; lineNumber: 54; columnNumber: 11; The end-tag for element
 type "property" must end with a '>' delimiter.
        at org.apache.xerces.parsers.DOMParser.parse(Unknown Source)
        at org.apache.xerces.jaxp.DocumentBuilderImpl.parse(Unknown Source)
        at javax.xml.parsers.DocumentBuilder.parse(DocumentBuilder.java:150)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2480)
        at org.apache.hadoop.conf.Configuration.parse(Configuration.java:2468)
        at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2539)
        ... 12 more
starting yarn daemons
starting resourcemanager, logging to /usr/hadoop-2.7.1/logs/yarn-root-resourcemanager-masterNode1.out
masterNode1: Warning: Permanently added 'masternode1,100.100.104.48' (ECDSA) to the list of known hosts.
slaveNode1: Warning: Permanently added 'slavenode1,100.100.106.13' (ECDSA) to the list of known hosts.
slaveNode2: Warning: Permanently added 'slavenode2,100.100.105.20' (ECDSA) to the list of known hosts.
masterNode1: starting nodemanager, logging to /usr/hadoop-2.7.1/logs/yarn-root-nodemanager-masterNode1.out
slaveNode1: starting nodemanager, logging to /usr/hadoop-2.7.1/logs/yarn-root-nodemanager-slaveNode1.out
slaveNode2: starting nodemanager, logging to /usr/hadoop-2.7.1/logs/yarn-root-nodemanager-slaveNode2.out
[root@masterNode1 ~]# #查看acl
```

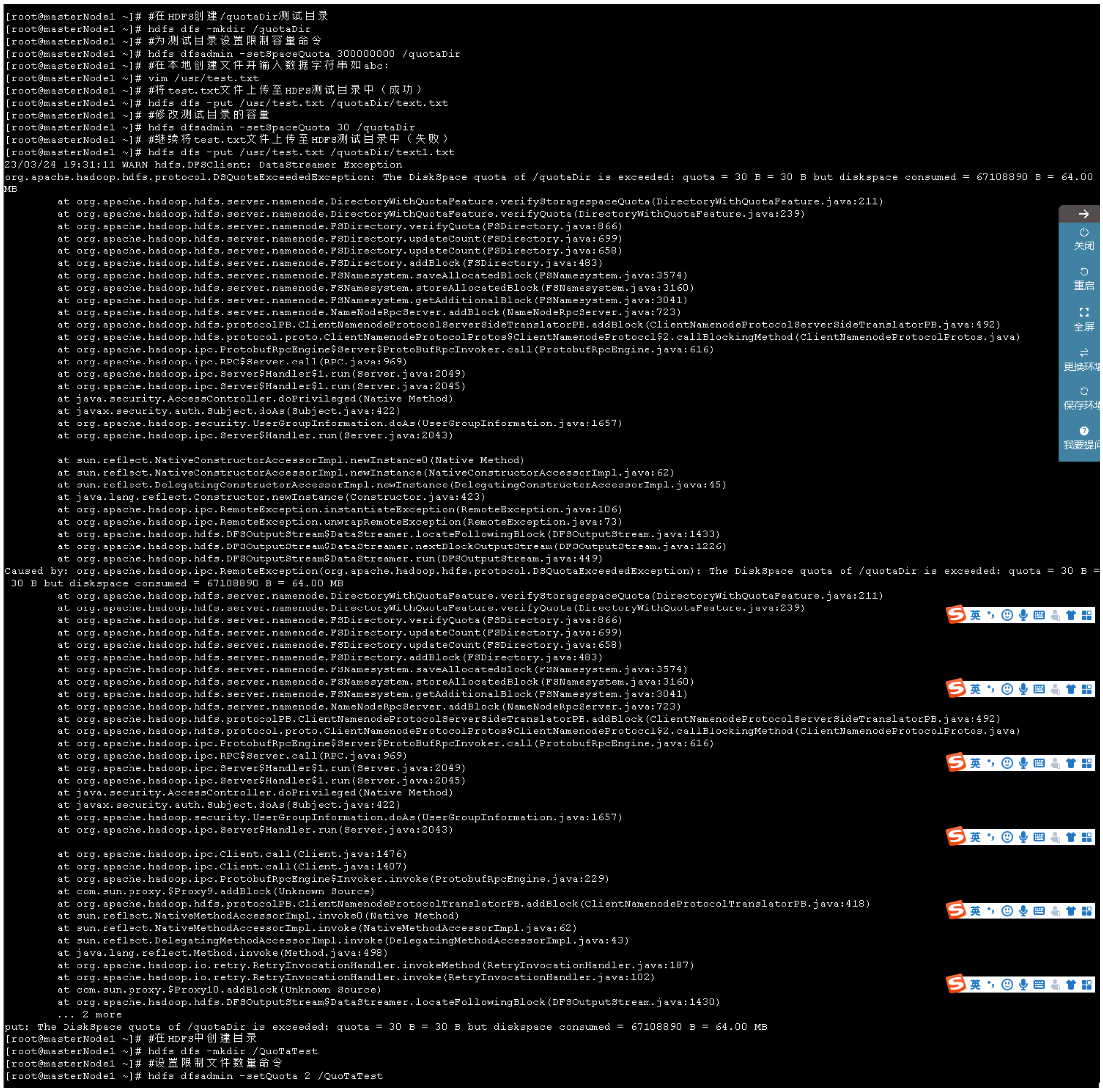
创建 aclFile.txt 文件、设定文件权限



```
[root@masterNode1 ~]# hdfs dfs -getfacl -R /file
# file: /file
# owner: root
# group: root
user::rwx
group::r-x
other::r-x

[root@masterNode1 ~]# #创建aclFile.txt文件
[root@masterNode1 ~]# hdfs dfs -touchz /datafile/aclFile.txt
[root@masterNode1 ~]# #设定文件权限
[root@masterNode1 ~]# #设定-m: 修改acl
[root@masterNode1 ~]# hdfs dfs -setfacl -m user:hadoop:rw- /datafile/aclFile.txt
[root@masterNode1 ~]# hdfs dfs -getfacl /datafile/aclFile.txt
# file: /datafile/aclFile.txt
# owner: root
# group: supergroup
user::rw-
user:hadoop:rw-
group::r--
mask::rw-
other::r--

[root@masterNode1 ~]# #设定-x:删除指定规则
[root@masterNode1 ~]# hdfs dfs -setfacl -x user:hadoop /datafile/aclFile.txt
[root@masterNode1 ~]# hdfs dfs -getfacl /datafile/aclFile.txt
# file: /datafile/aclFile.txt
# owner: root
# group: supergroup
user::rw-
group::r--
mask::r--
other::r--

[root@masterNode1 ~]# #设定-b: 基本的acl规则(所有者，群组，其他）被保留，其他规则全部删除
[root@masterNode1 ~]# hdfs dfs -setfacl -b /datafile/aclFile.txt
[root@masterNode1 ~]# hdfs dfs -getfacl /datafile/aclFile.txt
# file: /datafile/aclFile.txt
# owner: root
# group: supergroup
user::rw-
group::r--
other::r--

[root@masterNode1 ~]# #设定-k: 删除默认的ACL
[root@masterNode1 ~]# hdfs dfs -setfacl -k /datafile/aclFile.txt
[root@masterNode1 ~]# hdfs dfs -getfacl /datafile/aclFile.txt
# file: /datafile/aclFile.txt
# owner: root
# group: supergroup
user::rw-
group::r--
other::r--

[root@masterNode1 ~]# 
```

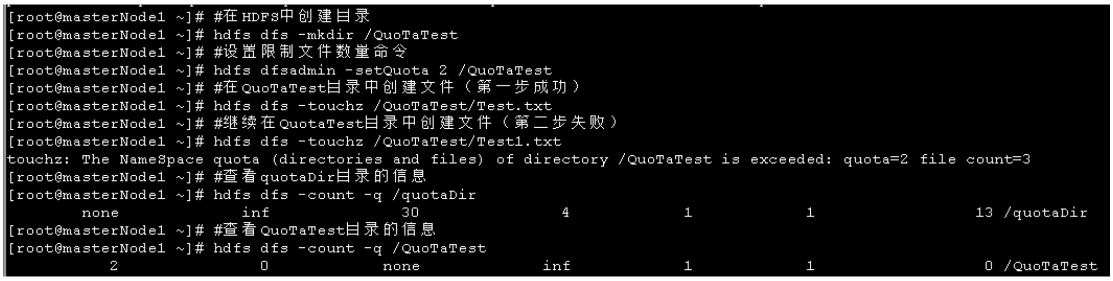


```
Last login: Fri Jan  8 16:10:24 2021 from 192.168.1.71
[root@masterNode1 ~]# 集群服务启动完成
hdfs dfsadmin -refreshNodes
Refresh nodes successful
[root@masterNode1 ~]# hdfs dfsadmin -report -live
Configured Capacity: 71425904713728 (64.96 TB)
Present Capacity: 66428109479936 (60.42 TB)
DFS Remaining: 66428109234176 (60.42 TB)
DFS Used: 245760 (240 KB)
DFS Used%: 0.00%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0

-------------------------------------------------
Live datanodes (3):

Name: 100.100.106.13:50010 (slaveNode2)
Hostname: slaveNode2
Decommission Status : Normal
Configured Capacity: 23808634216448 (21.65 TB)
DFS Used: 81920 (80 KB)
Non DFS Used: 1583460106240 (1.44 TB)
DFS Remaining: 22225174028288 (20.21 TB)
DFS Used%: 0.00%
DFS Remaining%: 93.35%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Fri Mar 24 19:27:43 CST 2023


Name: 100.100.104.48:50010 (masterNode1)
Hostname: masterNode1
Decommission Status : Normal
Configured Capacity: 23808636280832 (21.65 TB)
DFS Used: 81920 (80 KB)
Non DFS Used: 1764782809088 (1.61 TB)
DFS Remaining: 22043853389824 (20.05 TB)
DFS Used%: 0.00%
DFS Remaining%: 92.59%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Fri Mar 24 19:27:43 CST 2023


Name: 100.100.105.20:50010 (slaveNode1)
Hostname: slaveNode1
Decommission Status : Normal
Configured Capacity: 23808634216448 (21.65 TB)
DFS Used: 81920 (80 KB)
Non DFS Used: 1649552318464 (1.50 TB)
DFS Remaining: 22159081816064 (20.15 TB)
DFS Used%: 0.00%
DFS Remaining%: 93.07%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Fri Mar 24 19:27:43 CST 2023


[root@masterNode1 ~]# #在HDFS创建/quotaDir测试目录
[root@masterNode1 ~]# hdfs dfs -mkdir /quotaDir
[root@masterNode1 ~]# #为测试目录设置限制容量命令
[root@masterNode1 ~]# hdfs dfsadmin -setSpaceQuota 300000
```



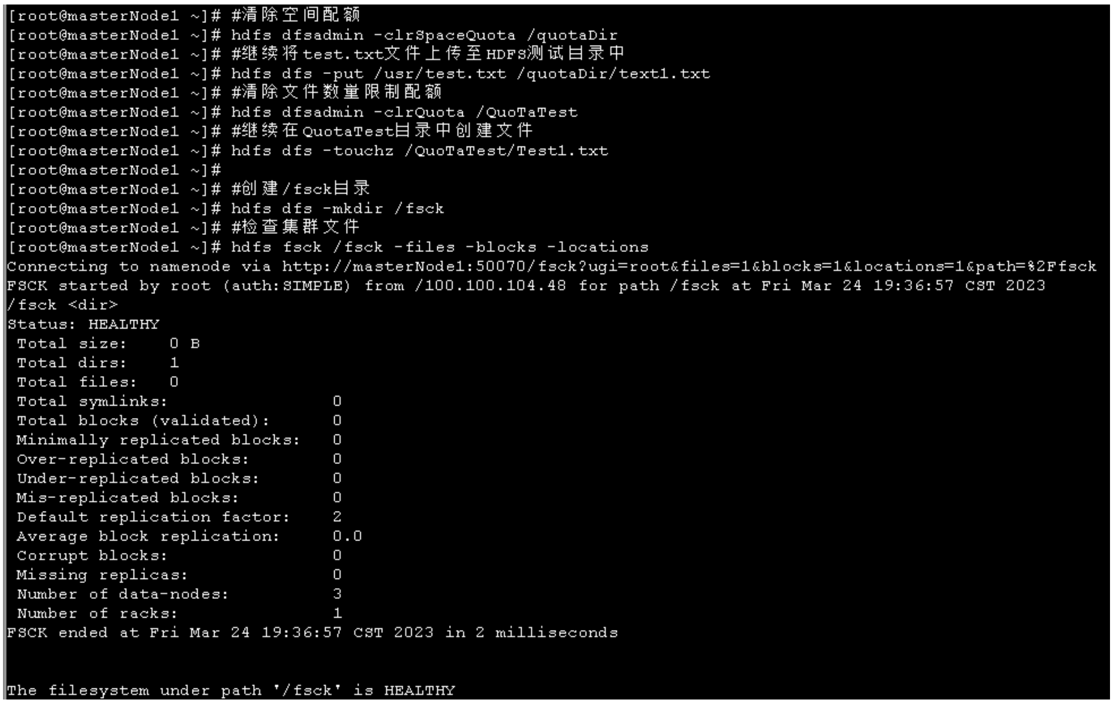
```
[root@masterNode1 ~]# #在HDFS创建/quotaDir测试目录
[root@masterNode1 ~]# hdfs dfs -mkdir /quotaDir
[root@masterNode1 ~]# #为测试目录设置限制容量命令
[root@masterNode1 ~]# hdfs dfsadmin -setSpaceQuota 300000000 /quotaDir
[root@masterNode1 ~]# #在本地创建文件并输入数据字符串如abc:
[root@masterNode1 ~]# vim /usr/test.txt
[root@masterNode1 ~]# #将test.txt文件上传至HDFS测试目录中（成功）
[root@masterNode1 ~]# hdfs dfs -put /usr/test.txt /quotaDir/text.txt
[root@masterNode1 ~]# #修改测试目录的容量
[root@masterNode1 ~]# hdfs dfsadmin -setSpaceQuota 30 /quotaDir
[root@masterNode1 ~]# #继续将test.txt文件上传至HDFS测试目录中（失败）
[root@masterNode1 ~]# hdfs dfs -put /usr/test.txt /quotaDir/text1.txt
23/03/24 19:31:11 WARN hdfs.DFSClient: DataStreamer Exception
org.apache.hadoop.hdfs.protocol.DSQuotaExceededException: The DiskSpace quota of /quotaDir is exceeded: quota = 30 B = 30 B but diskspace consumed = 67108890 B = 64.00
MB
        at org.apache.hadoop.hdfs.server.namenode.DirectoryWithQuotaFeature.verifyStoragespaceQuota(DirectoryWithQuotaFeature.java:211)
        at org.apache.hadoop.hdfs.server.namenode.DirectoryWithQuotaFeature.verifyQuota(DirectoryWithQuotaFeature.java:239)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.verifyQuota(FSDirectory.java:866)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.updateCount(FSDirectory.java:699)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.updateCount(FSDirectory.java:658)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.addBlock(FSDirectory.java:483)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.saveAllocatedBlock(FSNamesystem.java:3574)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.storeAllocatedBlock(FSNamesystem.java:3160)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:3041)
        at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.addBlock(NameNodeRpcServer.java:723)
        at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.addBlock(ClientNamenodeProtocolServerSideTranslatorPB.java:492)
        at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
        at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:616)
        at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:969)
        at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2049)
        at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2045)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1657)
        at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2043)

        at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
        at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)
        at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
        at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
        at org.apache.hadoop.ipc.RemoteException.instantiateException(RemoteException.java:106)
        at org.apache.hadoop.ipc.RemoteException.unwrapRemoteException(RemoteException.java:73)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.locateFollowingBlock(DFSOutputStream.java:1433)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.nextBlockOutputStream(DFSOutputStream.java:1226)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.run(DFSOutputStream.java:449)
Caused by: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.hdfs.protocol.DSQuotaExceededException): The DiskSpace quota of /quotaDir is exceeded: quota = 30 B =
 30 B but diskspace consumed = 67108890 B = 64.00 MB
        at org.apache.hadoop.hdfs.server.namenode.DirectoryWithQuotaFeature.verifyStoragespaceQuota(DirectoryWithQuotaFeature.java:211)
        at org.apache.hadoop.hdfs.server.namenode.DirectoryWithQuotaFeature.verifyQuota(DirectoryWithQuotaFeature.java:239)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.verifyQuota(FSDirectory.java:866)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.updateCount(FSDirectory.java:699)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.updateCount(FSDirectory.java:658)
        at org.apache.hadoop.hdfs.server.namenode.FSDirectory.addBlock(FSDirectory.java:483)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.saveAllocatedBlock(FSNamesystem.java:3574)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.storeAllocatedBlock(FSNamesystem.java:3160)
        at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:3041)
        at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.addBlock(NameNodeRpcServer.java:723)
        at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.addBlock(ClientNamenodeProtocolServerSideTranslatorPB.java:492)
        at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
        at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:616)
        at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:969)
        at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2049)
        at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2045)
        at java.security.AccessController.doPrivileged(Native Method)
        at javax.security.auth.Subject.doAs(Subject.java:422)
        at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1657)
        at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2043)

        at org.apache.hadoop.ipc.Client.call(Client.java:1476)
        at org.apache.hadoop.ipc.Client.call(Client.java:1407)
        at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:229)
        at com.sun.proxy.$Proxy9.addBlock(Unknown Source)
        at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.addBlock(ClientNamenodeProtocolTranslatorPB.java:418)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:187)
        at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:102)
        at com.sun.proxy.$Proxy10.addBlock(Unknown Source)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.locateFollowingBlock(DFSOutputStream.java:1430)
        ... 2 more
put: The DiskSpace quota of /quotaDir is exceeded: quota = 30 B = 30 B but diskspace consumed = 67108890 B = 64.00 MB
[root@masterNode1 ~]# #在HDFS中创建目录
[root@masterNode1 ~]# hdfs dfs -mkdir /QuoTaTest
[root@masterNode1 ~]# #设置限制文件数量命令
[root@masterNode1 ~]# hdfs dfsadmin -setQuota 2 /QuoTaTest
```

在 HDFS 中创建目录

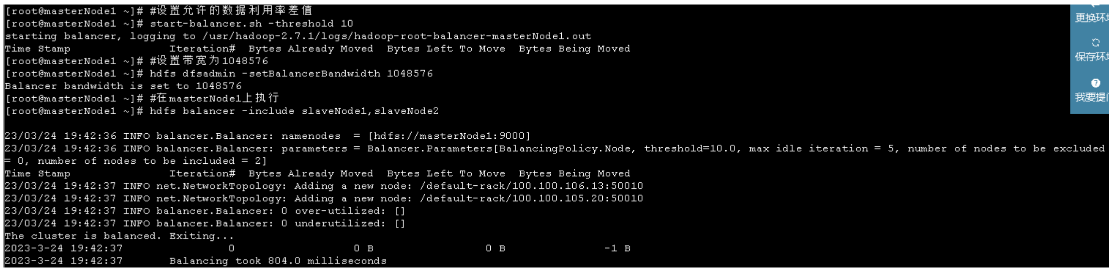


```
[root@masterNode1 ~]# #在HDFS中创建目录
[root@masterNode1 ~]# hdfs dfs -mkdir /QuoTaTest
[root@masterNode1 ~]# #设置限制文件数量命令
[root@masterNode1 ~]# hdfs dfsadmin -setQuota 2 /QuoTaTest
[root@masterNode1 ~]# #在QuoTaTest目录中创建文件（第一步成功）
[root@masterNode1 ~]# hdfs dfs -touchz /QuoTaTest/Test.txt
[root@masterNode1 ~]# #继续在QuoTaTest目录中创建文件（第二步失败）
[root@masterNode1 ~]# hdfs dfs -touchz /QuoTaTest/Test1.txt
touchz: The NameSpace quota (directories and files) of directory /QuoTaTest is exceeded: quota=2 file count=3
[root@masterNode1 ~]# #查看quotaDir目录的信息
[root@masterNode1 ~]# hdfs dfs -count -q /quotaDir
        none             inf              30               4            1            1               13 /quotaDir
[root@masterNode1 ~]# #查看QuoTaTest目录的信息
[root@masterNode1 ~]# hdfs dfs -count -q /QuoTaTest
           2               0            none             inf            1            1                0 /QuoTaTest
```

清除空间配额、检查集群文件



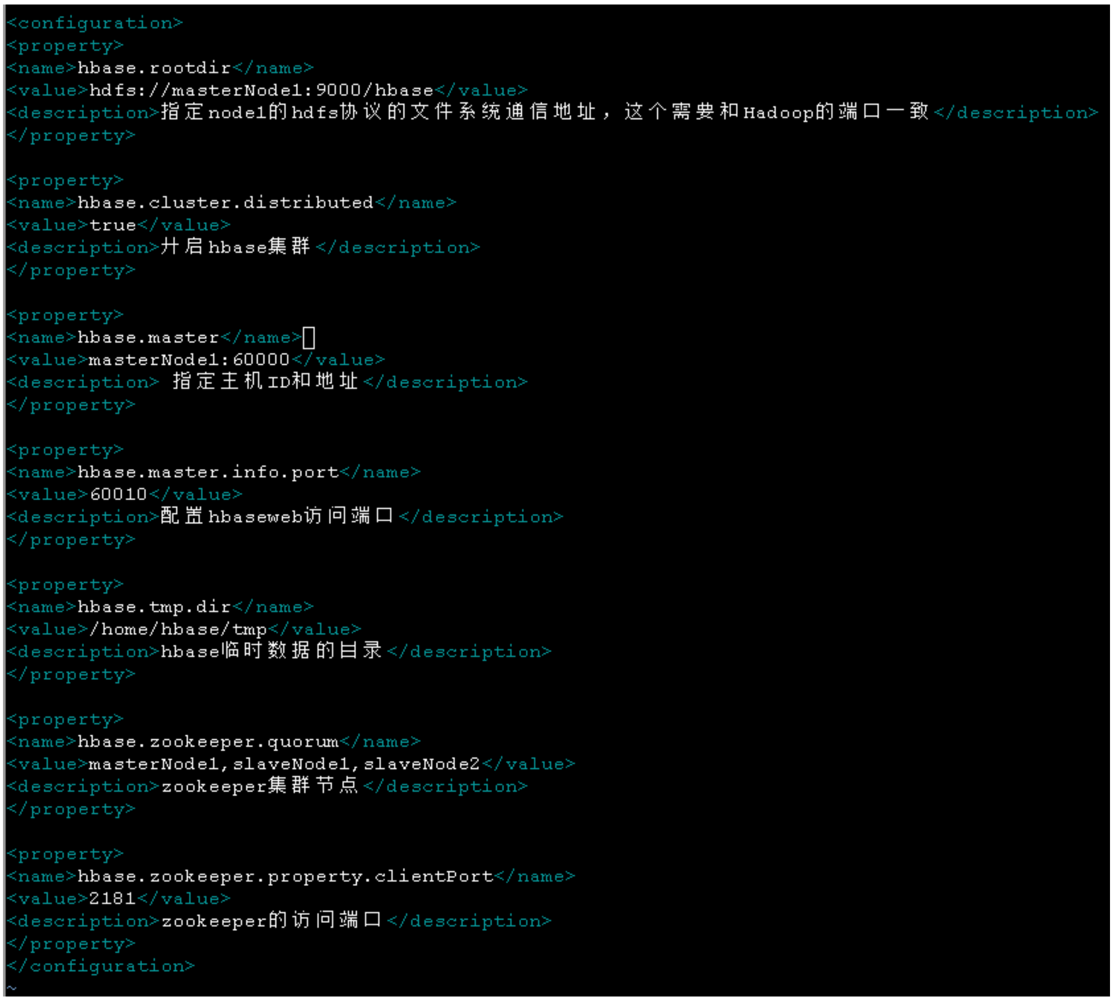
```
[root@masterNode1 ~]# #清除空间配额
[root@masterNode1 ~]# hdfs dfsadmin -clrSpaceQuota /quotaDir
[root@masterNode1 ~]# #继续将test.txt文件上传至HDFS测试目录中
[root@masterNode1 ~]# hdfs dfs -put /usr/test.txt /quotaDir/text1.txt
[root@masterNode1 ~]# #清除文件数量限制配额
[root@masterNode1 ~]# hdfs dfsadmin -clrQuota /QuoTaTest
[root@masterNode1 ~]# #继续在QuotaTest目录中创建文件
[root@masterNode1 ~]# hdfs dfs -touchz /QuoTaTest/Test1.txt
[root@masterNode1 ~]#
[root@masterNode1 ~]# #创建/fsck目录
[root@masterNode1 ~]# hdfs dfs -mkdir /fsck
[root@masterNode1 ~]# #检查集群文件
[root@masterNode1 ~]# hdfs fsck /fsck -files -blocks -locations
Connecting to namenode via http://masterNode1:50070/fsck?ugi=root&files=1&blocks=1&locations=1&path=%2Ffsck
FSCK started by root (auth:SIMPLE) from /100.100.104.48 for path /fsck at Fri Mar 24 19:36:57 CST 2023
/fsck <dir>
Status: HEALTHY
 Total size:    0 B
 Total dirs:    1
 Total files:   0
 Total symlinks:                0
 Total blocks (validated):      0
 Minimally replicated blocks:   0
 Over-replicated blocks:        0
 Under-replicated blocks:       0
 Mis-replicated blocks:         0
 Default replication factor:    2
 Average block replication:     0.0
 Corrupt blocks:                0
 Missing replicas:              0
 Number of data-nodes:          3
 Number of racks:               1
FSCK ended at Fri Mar 24 19:36:57 CST 2023 in 2 milliseconds


The filesystem under path '/fsck' is HEALTHY
```



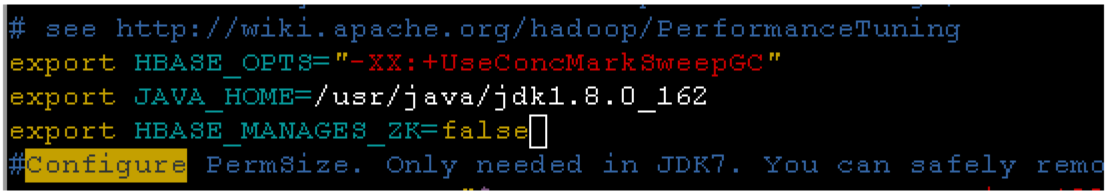
```
[root@masterNode1 ~]# #设置允许的数据利用率差值
[root@masterNode1 ~]# start-balancer.sh -threshold 10
starting balancer, logging to /usr/hadoop-2.7.1/logs/hadoop-root-balancer-masterNode1.out
Time Stamp               Iteration#  Bytes Already Moved  Bytes Left To Move  Bytes Being Moved
[root@masterNode1 ~]# #设置带宽为1048576
[root@masterNode1 ~]# hdfs dfsadmin -setBalancerBandwidth 1048576
Balancer bandwidth is set to 1048576
[root@masterNode1 ~]# #在masterNode1上执行
[root@masterNode1 ~]# hdfs balancer -include slaveNode1,slaveNode2

23/03/24 19:42:36 INFO balancer.Balancer: namenodes  = [hdfs://masterNode1:9000]
23/03/24 19:42:36 INFO balancer.Balancer: parameters = Balancer.Parameters[BalancingPolicy.Node, threshold=10.0, max idle iteration = 5, number of nodes to be excluded
= 0, number of nodes to be included = 2]
Time Stamp               Iteration#  Bytes Already Moved  Bytes Left To Move  Bytes Being Moved
23/03/24 19:42:37 INFO net.NetworkTopology: Adding a new node: /default-rack/100.100.106.13:50010
23/03/24 19:42:37 INFO net.NetworkTopology: Adding a new node: /default-rack/100.100.105.20:50010
23/03/24 19:42:37 INFO balancer.Balancer: 0 over-utilized: []
23/03/24 19:42:37 INFO balancer.Balancer: 0 underutilized: []
The cluster is balanced. Exiting...
2023-3-24 19:42:37                0                  0 B                 0 B               -1 B
2023-3-24 19:42:37       Balancing took 804.0 milliseconds
```

第二次作业

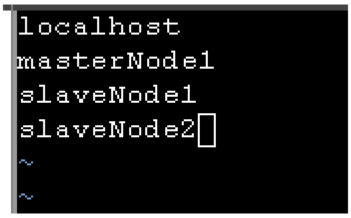
进入 shell 模式



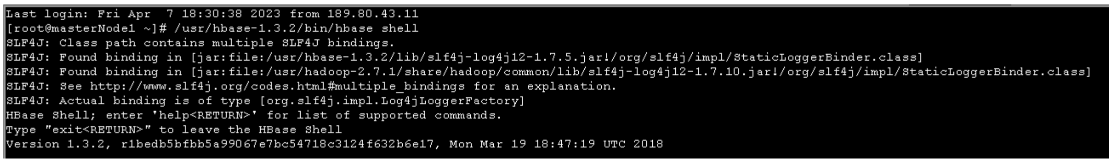
```
<configuration>
<property>
<name>hbase.rootdir</name>
<value>hdfs://masterNode1:9000/hbase</value>
<description>指定node1的hdfs协议的文件系统通信地址，这个需要和Hadoop的端口一致</description>
</property>

<property>
<name>hbase.cluster.distributed</name>
<value>true</value>
<description>开启hbase集群</description>
</property>

<property>
<name>hbase.master</name>
<value>masterNode1:60000</value>
<description> 指定主机ID和地址</description>
</property>

<property>
<name>hbase.master.info.port</name>
<value>60010</value>
<description>配置hbaseweb访问端口</description>
</property>

<property>
<name>hbase.tmp.dir</name>
<value>/home/hbase/tmp</value>
<description>hbase临时数据的目录</description>
</property>

<property>
<name>hbase.zookeeper.quorum</name>
<value>masterNode1,slaveNode1,slaveNode2</value>
<description>zookeeper集群节点</description>
</property>

<property>
<name>hbase.zookeeper.property.clientPort</name>
<value>2181</value>
<description>zookeeper的访问端口</description>
</property>
</configuration>
~
```



```
# see http://wiki.apache.org/hadoop/PerformanceTuning
export HBASE_OPTS="-XX:+UseConcMarkSweepGC"
export JAVA_HOME=/usr/java/jdk1.8.0_162
export HBASE_MANAGES_ZK=false
#Configure PermSize. Only needed in JDK7. You can safely remo
```

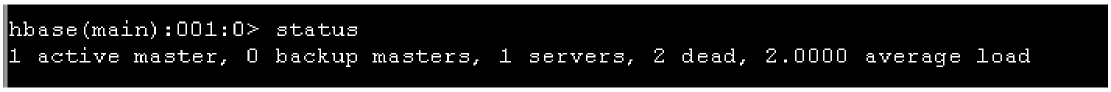


```
localhost
masterNode1
slaveNode1
slaveNode2
~
~
```



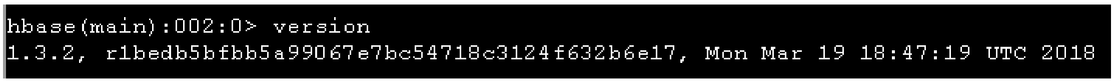
```
Last login: Fri Apr  7 18:30:38 2023 from 189.80.43.11
[root@masterNode1 ~]# /usr/hbase-1.3.2/bin/hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hbase-1.3.2/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hadoop-2.7.1/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.3.2, r1bedb5bfbb5a99067e7bc54718c3124f632b6e17, Mon Mar 19 18:47:19 UTC 2018
```

查看 HBase 状态



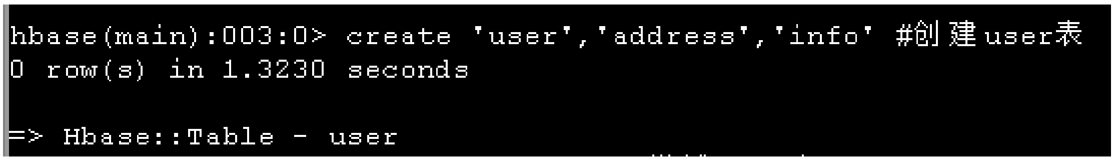
```
hbase(main):001:0> status
1 active master, 0 backup masters, 1 servers, 2 dead, 2.0000 average load
```

查看 HBase 版本信息



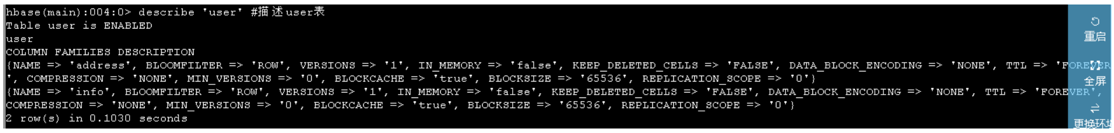
```
hbase(main):002:0> version
1.3.2, r1bedb5bfbb5a99067e7bc54718c3124f632b6e17, Mon Mar 19 18:47:19 UTC 2018
```

创建 user 表



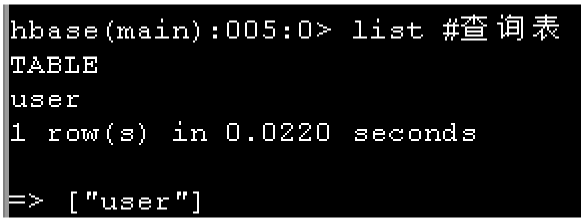
```
hbase(main):003:0> create 'user','address','info' #创建user表
0 row(s) in 1.3230 seconds

=> Hbase::Table - user
```

描述 user 表



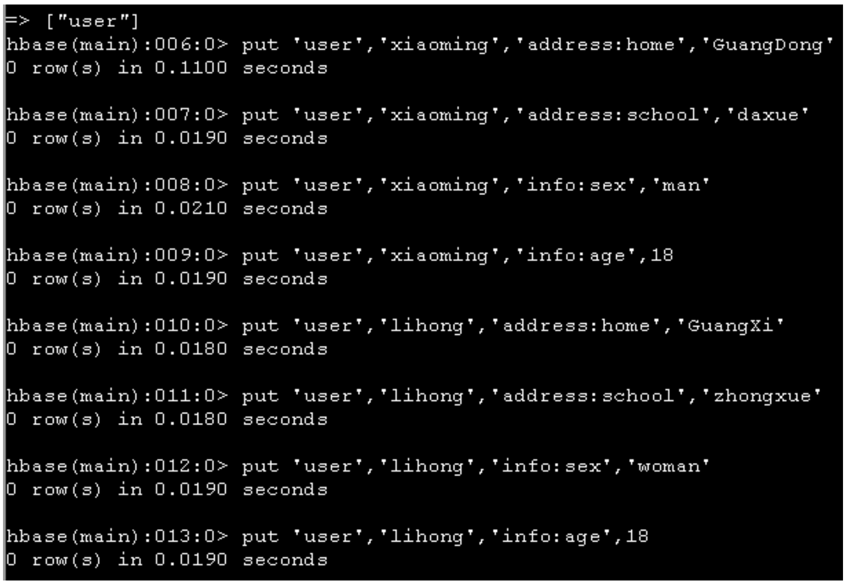
```
hbase(main):004:0> describe 'user' #描述user表
Table user is ENABLED
user
COLUMN FAMILIES DESCRIPTION
{NAME => 'address', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER
', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'info', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER',
COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
2 row(s) in 0.1030 seconds
```

查询表



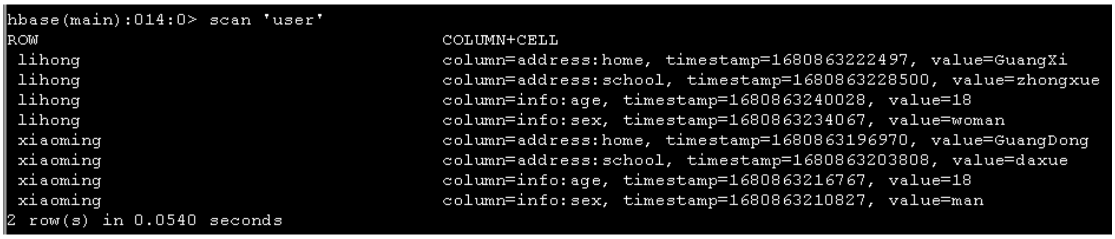
```
hbase(main):005:0> list #查询表
TABLE
user
1 row(s) in 0.0220 seconds

=> ["user"]
```

插入数据

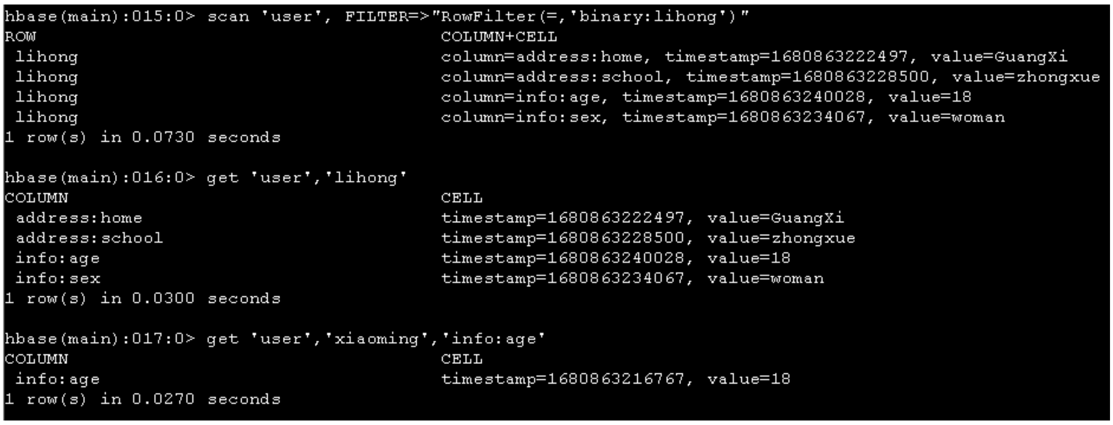


```
=> ["user"]
hbase(main):006:0> put 'user','xiaoming','address:home','GuangDong'
0 row(s) in 0.1100 seconds

hbase(main):007:0> put 'user','xiaoming','address:school','daxue'
0 row(s) in 0.0190 seconds

hbase(main):008:0> put 'user','xiaoming','info:sex','man'
0 row(s) in 0.0210 seconds

hbase(main):009:0> put 'user','xiaoming','info:age',18
0 row(s) in 0.0190 seconds

hbase(main):010:0> put 'user','lihong','address:home','GuangXi'
0 row(s) in 0.0180 seconds

hbase(main):011:0> put 'user','lihong','address:school','zhongxue'
0 row(s) in 0.0180 seconds

hbase(main):012:0> put 'user','lihong','info:sex','woman'
0 row(s) in 0.0190 seconds

hbase(main):013:0> put 'user','lihong','info:age',18
0 row(s) in 0.0190 seconds
```

查询数据

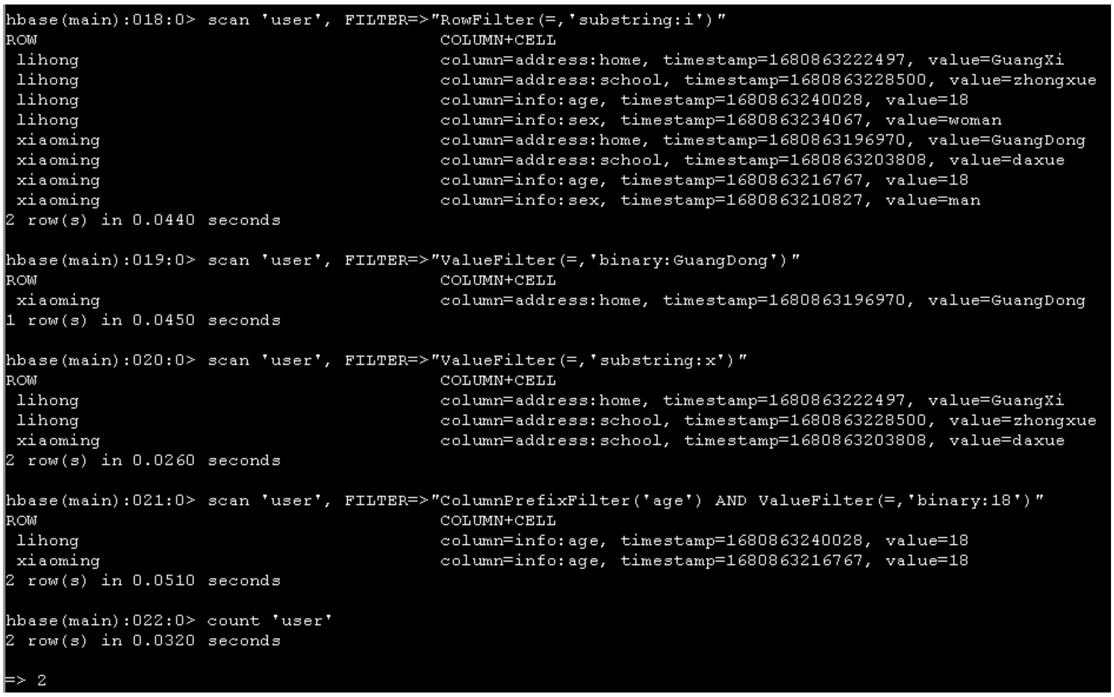


```
hbase(main):014:0> scan 'user'
ROW                                      COLUMN+CELL
 lihong                                  column=address:home, timestamp=1680863222497, value=GuangXi
 lihong                                  column=address:school, timestamp=1680863228500, value=zhongxue
 lihong                                  column=info:age, timestamp=1680863240028, value=18
 lihong                                  column=info:sex, timestamp=1680863234067, value=woman
 xiaoming                                column=address:home, timestamp=1680863196970, value=GuangDong
 xiaoming                                column=address:school, timestamp=1680863203808, value=daxue
 xiaoming                                column=info:age, timestamp=1680863216767, value=18
 xiaoming                                column=info:sex, timestamp=1680863210827, value=man
2 row(s) in 0.0540 seconds
```

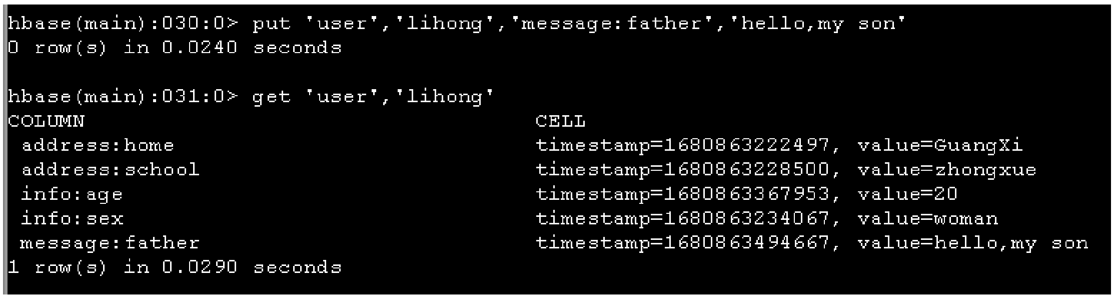


```
hbase(main):015:0> scan 'user', FILTER=>"RowFilter(=,'binary:lihong')"
ROW                                      COLUMN+CELL
 lihong                                  column=address:home, timestamp=1680863222497, value=GuangXi
 lihong                                  column=address:school, timestamp=1680863228500, value=zhongxue
 lihong                                  column=info:age, timestamp=1680863240028, value=18
 lihong                                  column=info:sex, timestamp=1680863234067, value=woman
1 row(s) in 0.0730 seconds

hbase(main):016:0> get 'user','lihong'
COLUMN                                   CELL
 address:home                            timestamp=1680863222497, value=GuangXi
 address:school                          timestamp=1680863228500, value=zhongxue
 info:age                                timestamp=1680863240028, value=18
 info:sex                                timestamp=1680863234067, value=woman
1 row(s) in 0.0300 seconds

hbase(main):017:0> get 'user','xiaoming','info:age'
COLUMN                                   CELL
 info:age                                timestamp=1680863216767, value=18
1 row(s) in 0.0270 seconds
```

查询任何一列中包含“ x “的数据

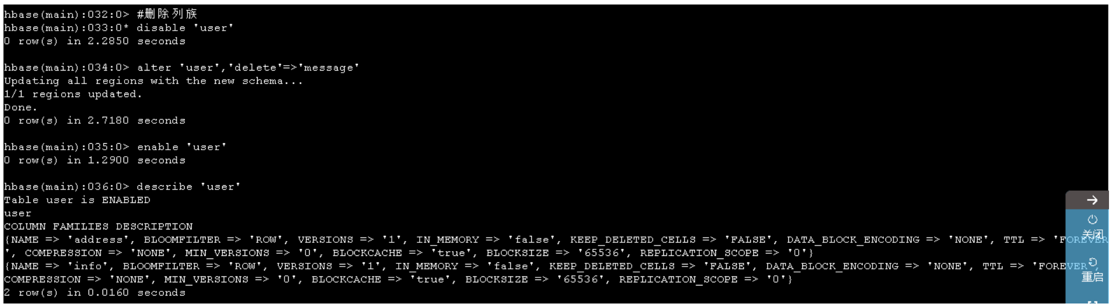


```
hbase(main):018:0> scan 'user', FILTER=>"RowFilter(=,'substring:i')"
ROW                                      COLUMN+CELL
 lihong                                  column=address:home, timestamp=1680863222497, value=GuangXi
 lihong                                  column=address:school, timestamp=1680863228500, value=zhongxue
 lihong                                  column=info:age, timestamp=1680863240028, value=18
 lihong                                  column=info:sex, timestamp=1680863234067, value=woman
 xiaoming                                column=address:home, timestamp=1680863196970, value=GuangDong
 xiaoming                                column=address:school, timestamp=1680863203808, value=daxue
 xiaoming                                column=info:age, timestamp=1680863216767, value=18
 xiaoming                                column=info:sex, timestamp=1680863210827, value=man
2 row(s) in 0.0440 seconds

hbase(main):019:0> scan 'user', FILTER=>"ValueFilter(=,'binary:GuangDong')"
ROW                                      COLUMN+CELL
 xiaoming                                column=address:home, timestamp=1680863196970, value=GuangDong
1 row(s) in 0.0450 seconds

hbase(main):020:0> scan 'user', FILTER=>"ValueFilter(=,'substring:x')"
ROW                                      COLUMN+CELL
 lihong                                  column=address:home, timestamp=1680863222497, value=GuangXi
 lihong                                  column=address:school, timestamp=1680863228500, value=zhongxue
 xiaoming                                column=address:school, timestamp=1680863203808, value=daxue
2 row(s) in 0.0260 seconds

hbase(main):021:0> scan 'user', FILTER=>"ColumnPrefixFilter('age') AND ValueFilter(=,'binary:18')"
ROW                                      COLUMN+CELL
 lihong                                  column=info:age, timestamp=1680863240028, value=18
 xiaoming                                column=info:age, timestamp=1680863216767, value=18
2 row(s) in 0.0510 seconds

hbase(main):022:0> count 'user'
2 row(s) in 0.0320 seconds

=> 2
```



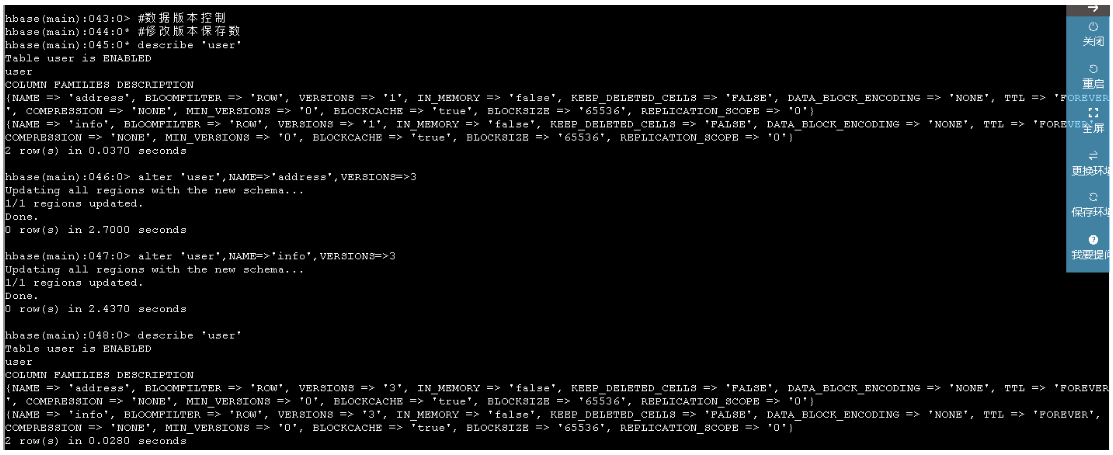
```
hbase(main):030:0> put 'user','lihong','message:father','hello,my son'
0 row(s) in 0.0240 seconds

hbase(main):031:0> get 'user','lihong'
COLUMN                                   CELL
 address:home                            timestamp=1680863222497, value=GuangXi
 address:school                          timestamp=1680863228500, value=zhongxue
 info:age                                timestamp=1680863367953, value=20
 info:sex                                timestamp=1680863234067, value=woman
 message:father                          timestamp=1680863494667, value=hello,my son
1 row(s) in 0.0290 seconds
```

删除列族



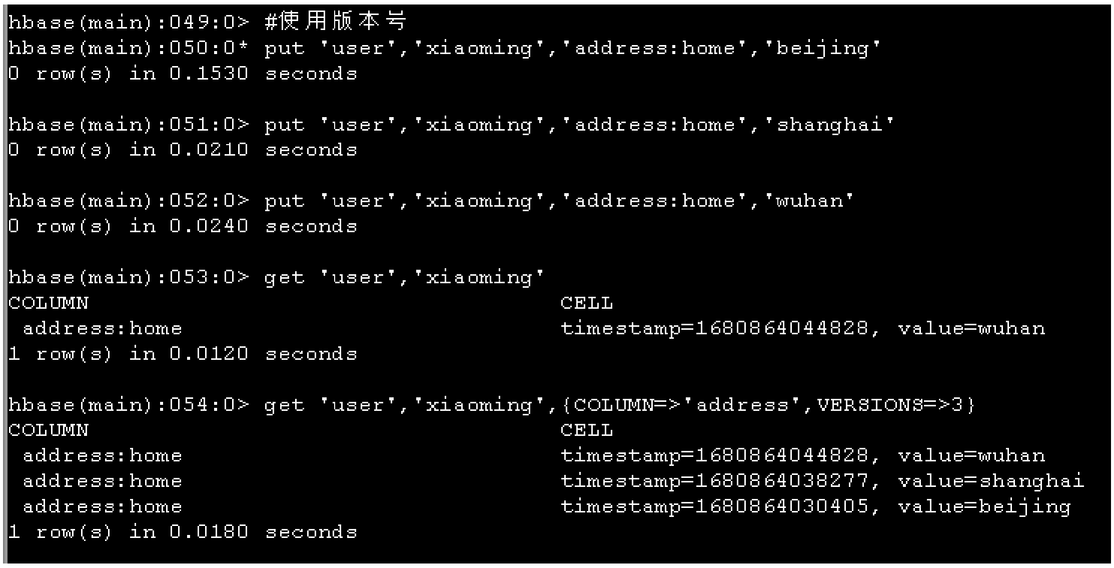
```
hbase(main):032:0> #删除列族
hbase(main):033:0* disable 'user'
0 row(s) in 2.2850 seconds

hbase(main):034:0> alter 'user','delete'=>'message'
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.7180 seconds

hbase(main):035:0> enable 'user'
0 row(s) in 1.2900 seconds

hbase(main):036:0> describe 'user'
Table user is ENABLED
user
COLUMN FAMILIES DESCRIPTION
{NAME => 'address', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER
', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'info', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER',
COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
2 row(s) in 0.0160 seconds
```

数据版本控制



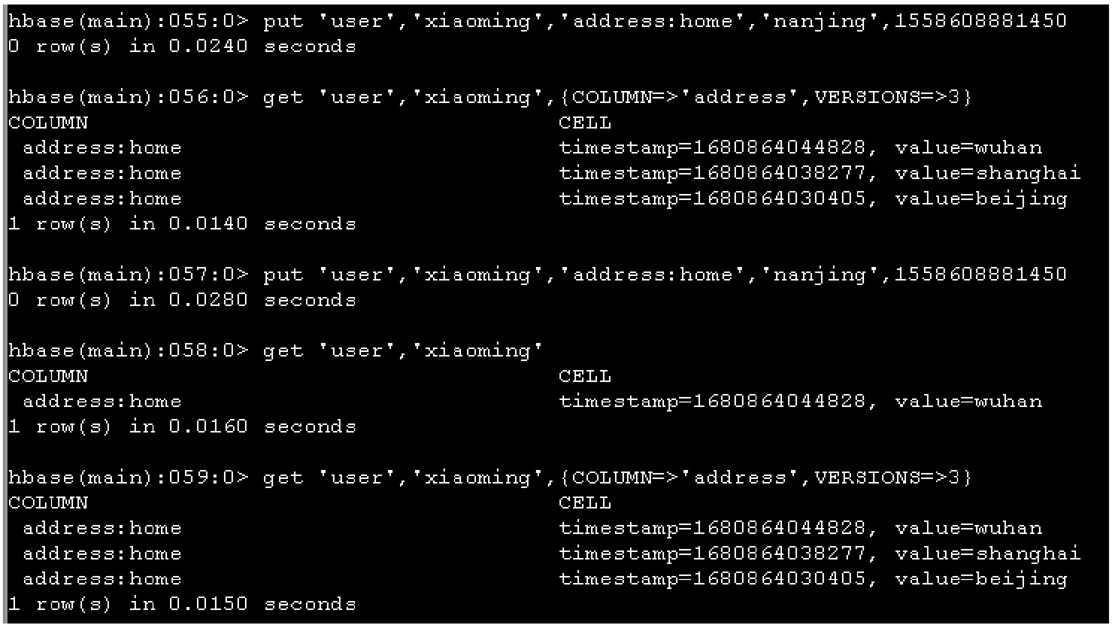
```
hbase(main):043:0> #数据版本控制
hbase(main):044:0* #修改版本保存数
hbase(main):045:0* describe 'user'
Table user is ENABLED
user
COLUMN FAMILIES DESCRIPTION
{NAME => 'address', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER
', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'info', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER',
COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
2 row(s) in 0.0370 seconds

hbase(main):046:0> alter 'user',NAME=>'address',VERSIONS=>3
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.7000 seconds

hbase(main):047:0> alter 'user',NAME=>'info',VERSIONS=>3
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.4370 seconds

hbase(main):048:0> describe 'user'
Table user is ENABLED
user
COLUMN FAMILIES DESCRIPTION
{NAME => 'address', BLOOMFILTER => 'ROW', VERSIONS => '3', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER
', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
{NAME => 'info', BLOOMFILTER => 'ROW', VERSIONS => '3', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER',
COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
2 row(s) in 0.0280 seconds
```

使用版本号

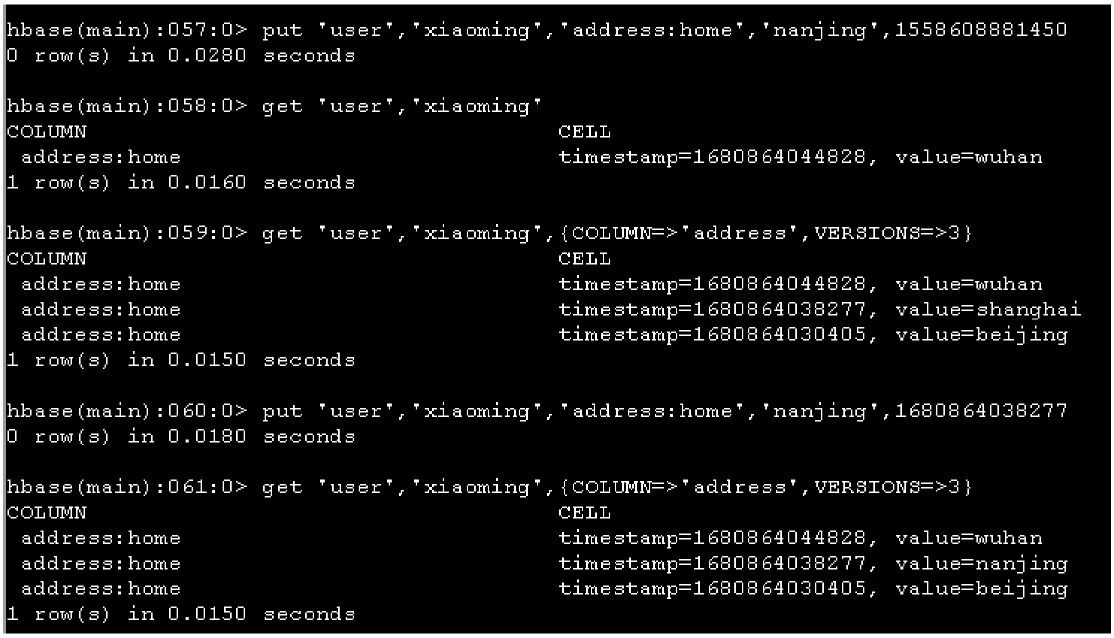


```
hbase(main):049:0> #使用版本号
hbase(main):050:0* put 'user','xiaoming','address:home','beijing'
0 row(s) in 0.1530 seconds

hbase(main):051:0> put 'user','xiaoming','address:home','shanghai'
0 row(s) in 0.0210 seconds

hbase(main):052:0> put 'user','xiaoming','address:home','wuhan'
0 row(s) in 0.0240 seconds

hbase(main):053:0> get 'user','xiaoming'
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
1 row(s) in 0.0120 seconds

hbase(main):054:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864038277, value=shanghai
 address:home                            timestamp=1680864030405, value=beijing
1 row(s) in 0.0180 seconds
```

更新数据



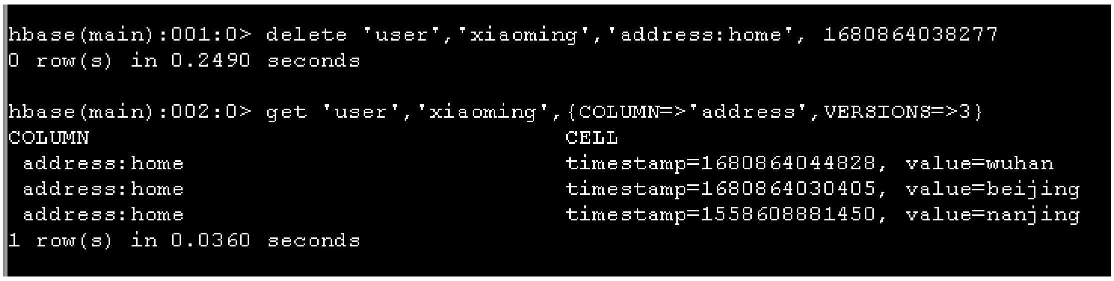
```
hbase(main):055:0> put 'user','xiaoming','address:home','nanjing',1558608881450
0 row(s) in 0.0240 seconds

hbase(main):056:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864038277, value=shanghai
 address:home                            timestamp=1680864030405, value=beijing
1 row(s) in 0.0140 seconds

hbase(main):057:0> put 'user','xiaoming','address:home','nanjing',1558608881450
0 row(s) in 0.0280 seconds

hbase(main):058:0> get 'user','xiaoming'
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
1 row(s) in 0.0160 seconds

hbase(main):059:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864038277, value=shanghai
 address:home                            timestamp=1680864030405, value=beijing
1 row(s) in 0.0150 seconds
```



```
hbase(main):057:0> put 'user','xiaoming','address:home','nanjing',1558608881450
0 row(s) in 0.0280 seconds

hbase(main):058:0> get 'user','xiaoming'
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
1 row(s) in 0.0160 seconds

hbase(main):059:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864038277, value=shanghai
 address:home                            timestamp=1680864030405, value=beijing
1 row(s) in 0.0150 seconds
```

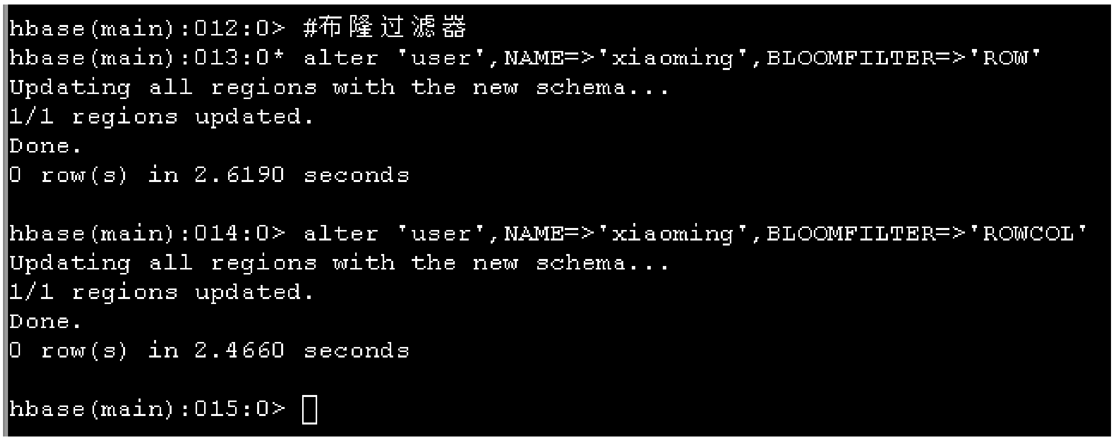
```
hbase(main):060:0> put 'user','xiaoming','address:home','nanjing',1680864038277
0 row(s) in 0.0180 seconds

hbase(main):061:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864038277, value=nanjing
 address:home                            timestamp=1680864030405, value=beijing
1 row(s) in 0.0150 seconds
```

读取最大 VESTONS 个记录

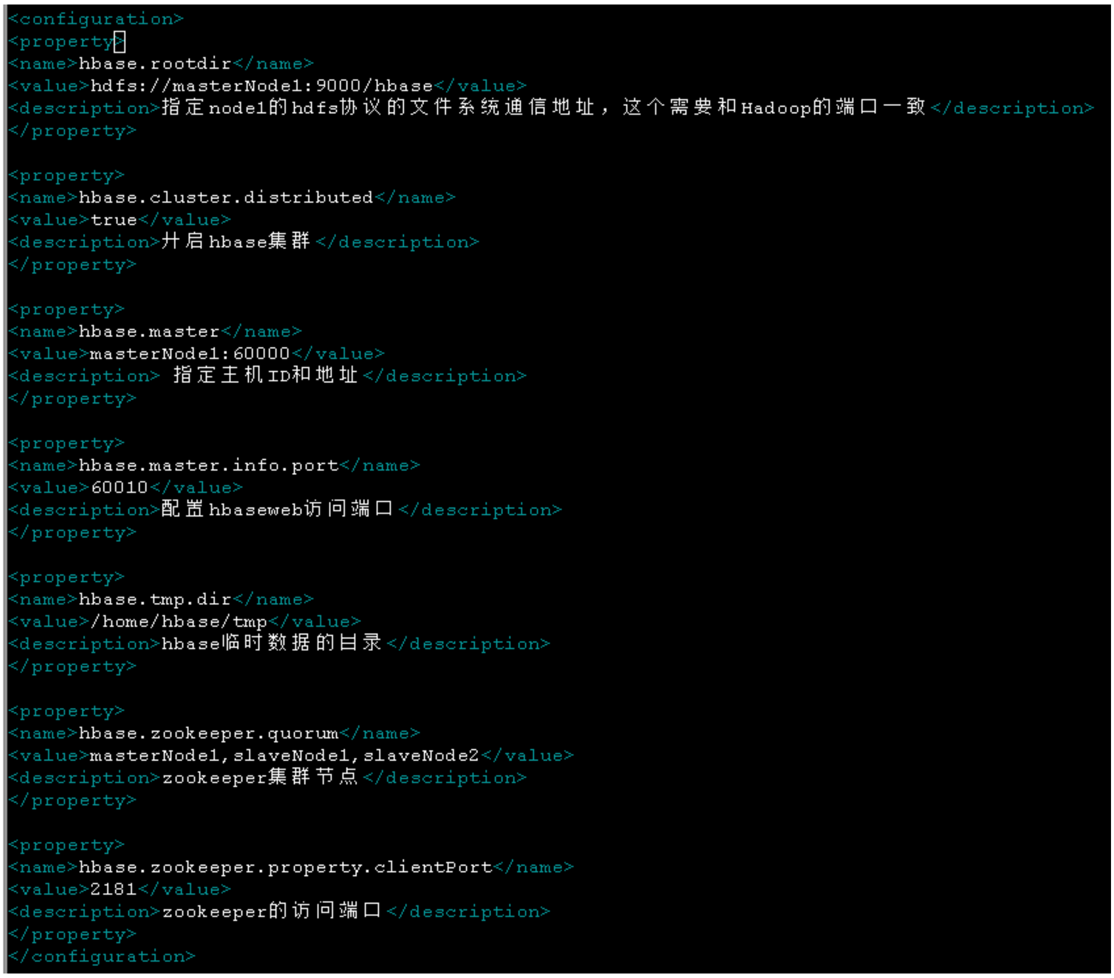


```
hbase(main):001:0> delete 'user','xiaoming','address:home', 1680864038277
0 row(s) in 0.2490 seconds

hbase(main):002:0> get 'user','xiaoming',{COLUMN=>'address',VERSIONS=>3}
COLUMN                                   CELL
 address:home                            timestamp=1680864044828, value=wuhan
 address:home                            timestamp=1680864030405, value=beijing
 address:home                            timestamp=1558608881450, value=nanjing
1 row(s) in 0.0360 seconds
```

布隆过滤器

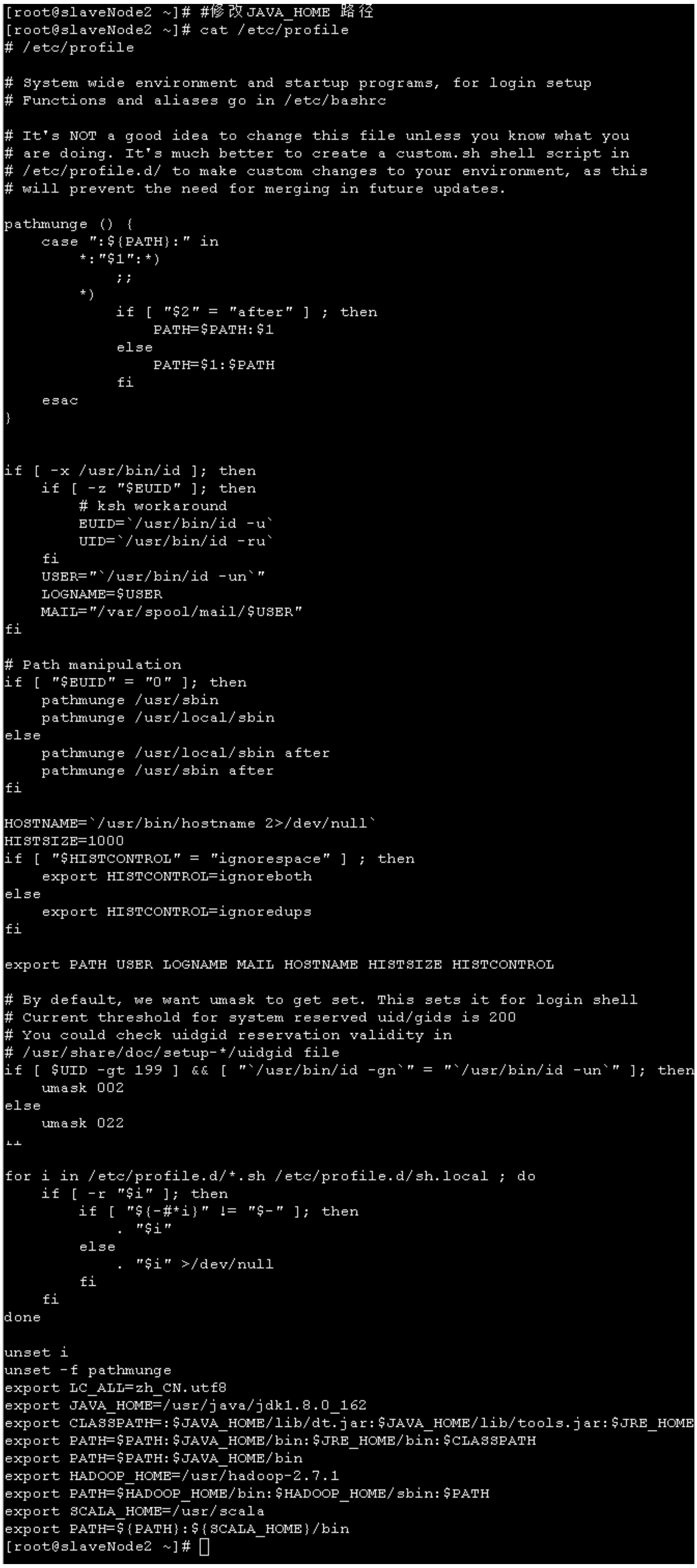


```
hbase(main):012:0> #布隆过滤器
hbase(main):013:0* alter 'user',NAME=>'xiaoming',BLOOMFILTER=>'ROW'
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.6190 seconds

hbase(main):014:0> alter 'user',NAME=>'xiaoming',BLOOMFILTER=>'ROWCOL'
Updating all regions with the new schema...
1/1 regions updated.
Done.
0 row(s) in 2.4660 seconds

hbase(main):015:0> 
```



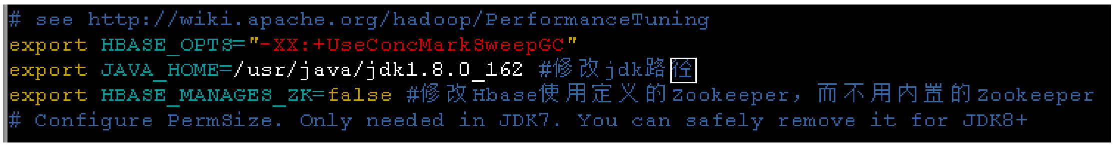
```
<configuration>
<property>
<name>hbase.rootdir</name>
<value>hdfs://masterNode1:9000/hbase</value>
<description>指定node1的hdfs协议的文件系统通信地址，这个需要和Hadoop的端口一致</description>
</property>

<property>
<name>hbase.cluster.distributed</name>
<value>true</value>
<description>开启hbase集群</description>
</property>

<property>
<name>hbase.master</name>
<value>masterNode1:60000</value>
<description> 指定主机ID和地址</description>
</property>

<property>
<name>hbase.master.info.port</name>
<value>60010</value>
<description>配置hbaseweb访问端口</description>
</property>

<property>
<name>hbase.tmp.dir</name>
<value>/home/hbase/tmp</value>
<description>hbase临时数据的目录</description>
</property>

<property>
<name>hbase.zookeeper.quorum</name>
<value>masterNode1,slaveNode1,slaveNode2</value>
<description>zookeeper集群节点</description>
</property>

<property>
<name>hbase.zookeeper.property.clientPort</name>
<value>2181</value>
<description>zookeeper的访问端口</description>
</property>
</configuration>
```

修改 JAVE_HOME 路径



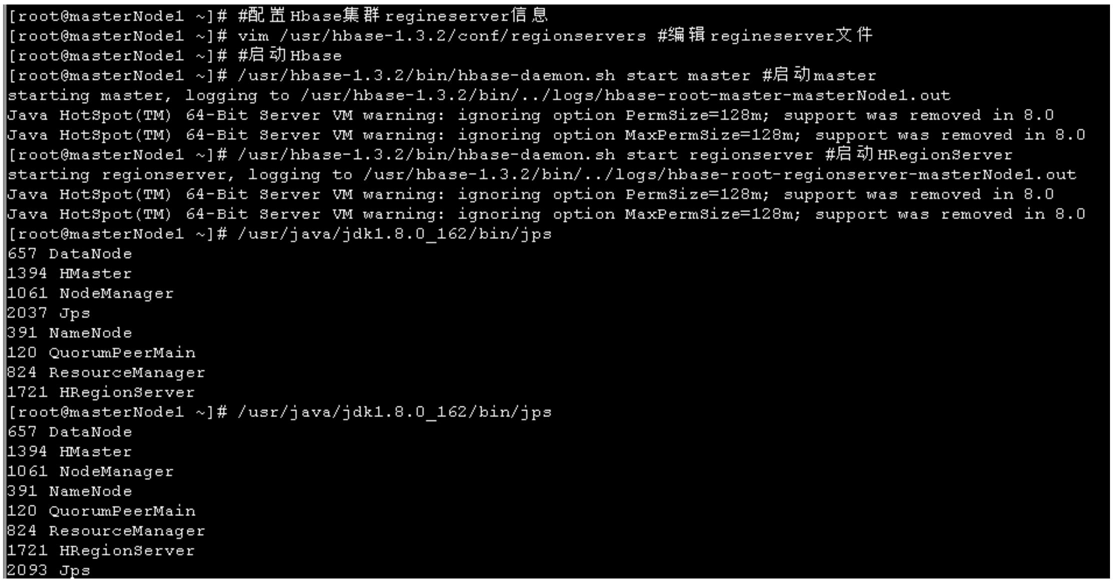
```
[root@slaveNode2 ~]# #修改JAVA_HOME 路径
[root@slaveNode2 ~]# cat /etc/profile
# /etc/profile

# System wide environment and startup programs, for login setup
# Functions and aliases go in /etc/bashrc

# It's NOT a good idea to change this file unless you know what you
# are doing. It's much better to create a custom.sh shell script in
# /etc/profile.d/ to make custom changes to your environment, as this
# will prevent the need for merging in future updates.

pathmunge () {
    case ":${PATH}:" in
        *:"$1":*)
            ;;
        *)
            if [ "$2" = "after" ] ; then
                PATH=$PATH:$1
            else
                PATH=$1:$PATH
            fi
    esac
}


if [ -x /usr/bin/id ]; then
    if [ -z "$EUID" ]; then
        # ksh workaround
        EUID=`/usr/bin/id -u`
        UID=`/usr/bin/id -ru`
    fi
    USER="`/usr/bin/id -un`"
    LOGNAME=$USER
    MAIL="/var/spool/mail/$USER"
fi

# Path manipulation
if [ "$EUID" = "0" ]; then
    pathmunge /usr/sbin
    pathmunge /usr/local/sbin
else
    pathmunge /usr/local/sbin after
    pathmunge /usr/sbin after
fi

HOSTNAME=`/usr/bin/hostname 2>/dev/null`
HISTSIZE=1000
if [ "$HISTCONTROL" = "ignorespace" ] ; then
    export HISTCONTROL=ignoreboth
else
    export HISTCONTROL=ignoredups
fi

export PATH USER LOGNAME MAIL HOSTNAME HISTSIZE HISTCONTROL

# By default, we want umask to get set. This sets it for login shell
# Current threshold for system reserved uid/gids is 200
# You could check uidgid reservation validity in
# /usr/share/doc/setup-*/uidgid file
if [ $UID -gt 199 ] && [ "`/usr/bin/id -gn`" = "`/usr/bin/id -un`" ]; then
    umask 002
else
    umask 022
fi

for i in /etc/profile.d/*.sh /etc/profile.d/sh.local ; do
    if [ -r "$i" ]; then
        if [ "${-#*i}" != "$-" ]; then
            . "$i"
        else
            . "$i" >/dev/null
        fi
    fi
done

unset i
unset -f pathmunge
export LC_ALL=zh_CN.utf8
export JAVA_HOME=/usr/java/jdk1.8.0_162
export CLASSPATH=:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar:$JRE_HOME
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin:$CLASSPATH
export PATH=$PATH:$JAVA_HOME/bin
export HADOOP_HOME=/usr/hadoop-2.7.1
export PATH=$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$PATH
export SCALA_HOME=/usr/scala
export PATH=${PATH}:${SCALA_HOME}/bin
[root@slaveNode2 ~]# 
```

修改 JKS 路径



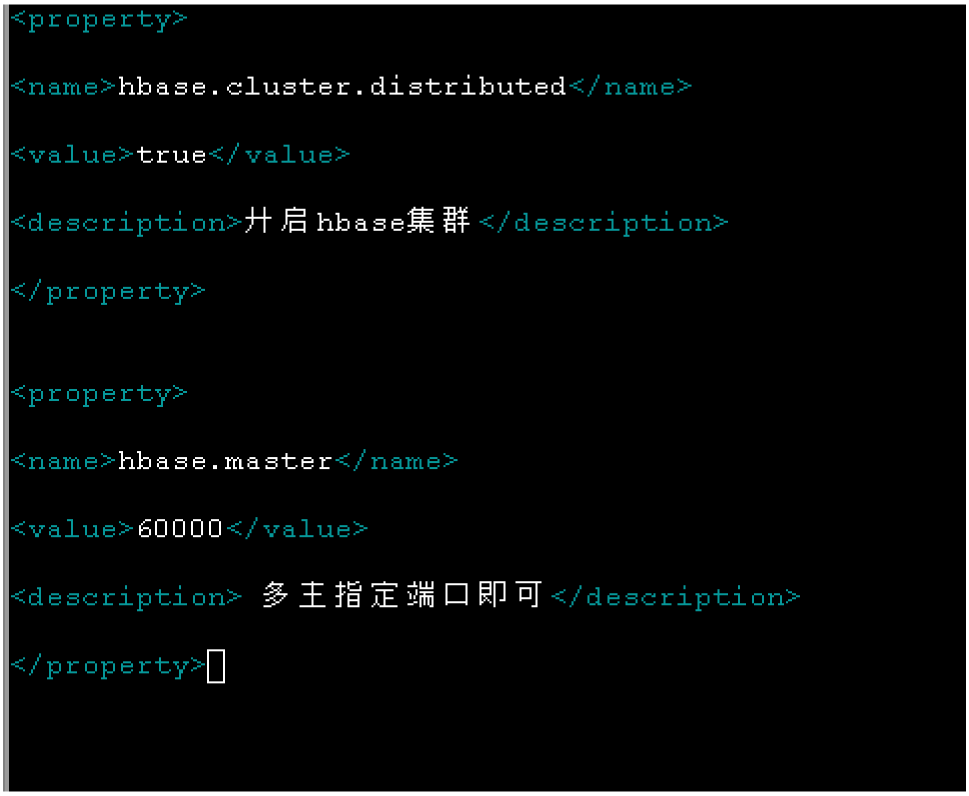
```
# see http://wiki.apache.org/hadoop/PerformanceTuning
export HBASE_OPTS="-XX:+UseConcMarkSweepGC"
export JAVA_HOME=/usr/java/jdk1.8.0_162 #修改jdk路径
export HBASE_MANAGES_ZK=false #修改Hbase使用定义的zookeeper，而不用内置的zookeeper
# Configure PermSize. Only needed in JDK7. You can safely remove it for JDK8+
```

启动 HBase 标准方式启动

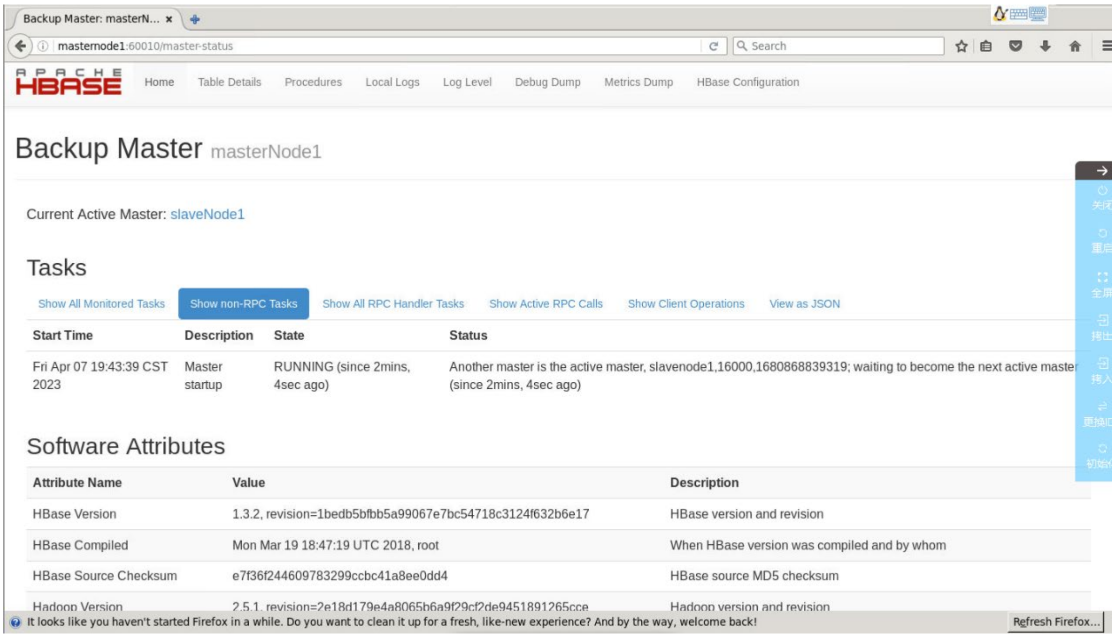


```
[root@masterNode1 ~]# #配置Hbase集群regineserver信息
[root@masterNode1 ~]# vim /usr/hbase-1.3.2/conf/regionservers #编辑regineserver文件
[root@masterNode1 ~]# #启动Hbase
[root@masterNode1 ~]# /usr/hbase-1.3.2/bin/hbase-daemon.sh start master #启动master
starting master, logging to /usr/hbase-1.3.2/bin/../logs/hbase-root-master-masterNode1.out
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
[root@masterNode1 ~]# /usr/hbase-1.3.2/bin/hbase-daemon.sh start regionserver #启动HRegionServer
starting regionserver, logging to /usr/hbase-1.3.2/bin/../logs/hbase-root-regionserver-masterNode1.out
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
[root@masterNode1 ~]# /usr/java/jdk1.8.0_162/bin/jps
657 DataNode
1394 HMaster
1061 NodeManager
2037 Jps
391 NameNode
120 QuorumPeerMain
824 ResourceManager
1721 HRegionServer
[root@masterNode1 ~]# /usr/java/jdk1.8.0_162/bin/jps
657 DataNode
1394 HMaster
1061 NodeManager
391 NameNode
120 QuorumPeerMain
824 ResourceManager
1721 HRegionServer
2093 Jps
```

配置 HBase 集群 regineserver 信息



```
<property>

<name>hbase.cluster.distributed</name>

<value>true</value>

<description>开启hbase集群</description>

</property>


<property>

<name>hbase.master</name>

<value>60000</value>

<description> 多主指定端口即可</description>

</property>
```



Backup Master masterNode1

Current Active Master: slaveNode1

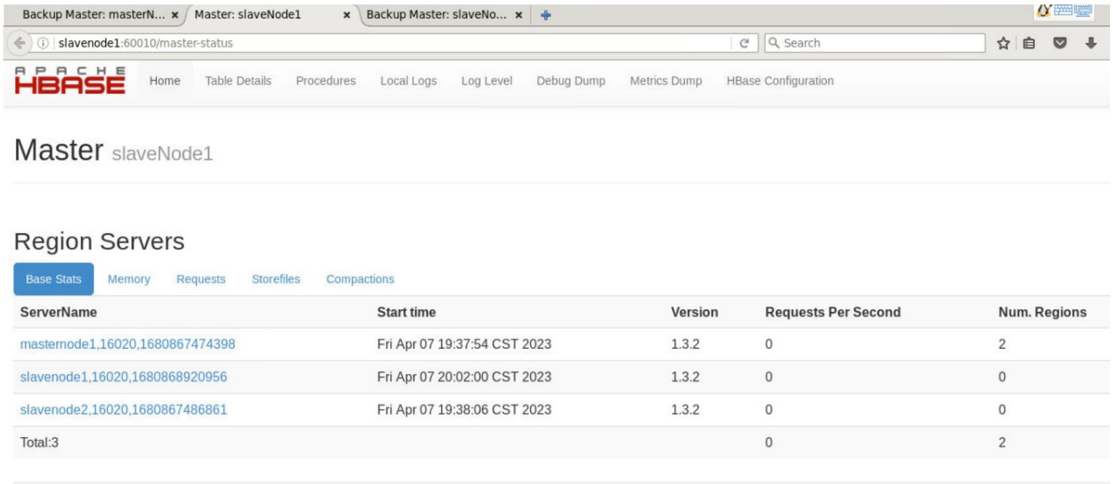
Tasks

Start Time	Description	State	Status
Fri Apr 07 19:43:39 CST 2023	Master startup	RUNNING (since 2mins, 4sec ago)	Another master is the active master, slavenode1,16000,1680868839319; waiting to become the next active master (since 2mins, 4sec ago)

Software Attributes

Attribute Name	Value	Description
HBase Version	1.3.2, revision=1bedb5bfbb5a99067e7bc54718c3124f632b6e17	HBase version and revision
HBase Compiled	Mon Mar 19 18:47:19 UTC 2018, root	When HBase version was compiled and by whom
HBase Source Checksum	e7f36f244609783299ccbc41a8ee0dd4	HBase source MD5 checksum
Hadoop Version	2.5.1, revision=2e18d179e4a8065b6a9f29cf2de9451891265cce	Hadoop version and revision

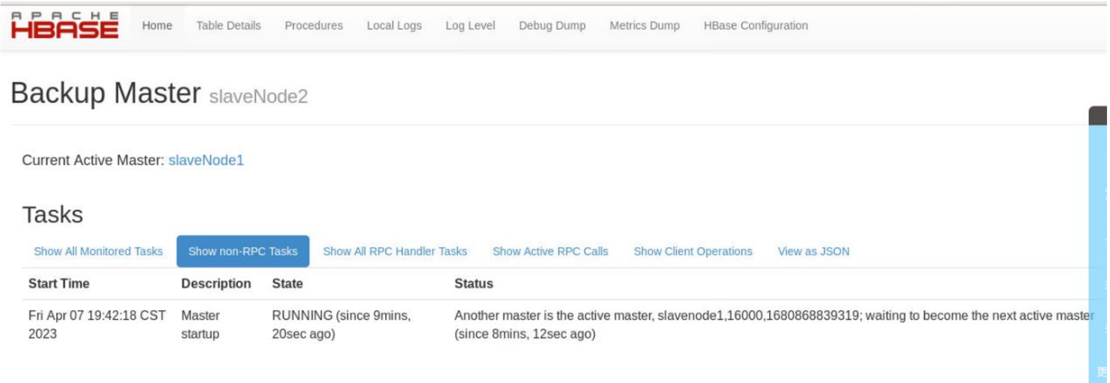
It looks like you haven't started Firefox in a while. Do you want to clean it up for a fresh, like-new experience? And by the way, welcome back!



Master slaveNode1

Region Servers

ServerName	Start time	Version	Requests Per Second	Num. Regions
masternode1,16020,1680867474398	Fri Apr 07 19:37:54 CST 2023	1.3.2	0	2
slavenode1,16020,1680868920956	Fri Apr 07 20:02:00 CST 2023	1.3.2	0	0
slavenode2,16020,1680867486861	Fri Apr 07 19:38:06 CST 2023	1.3.2	0	0
Total:3			0	2

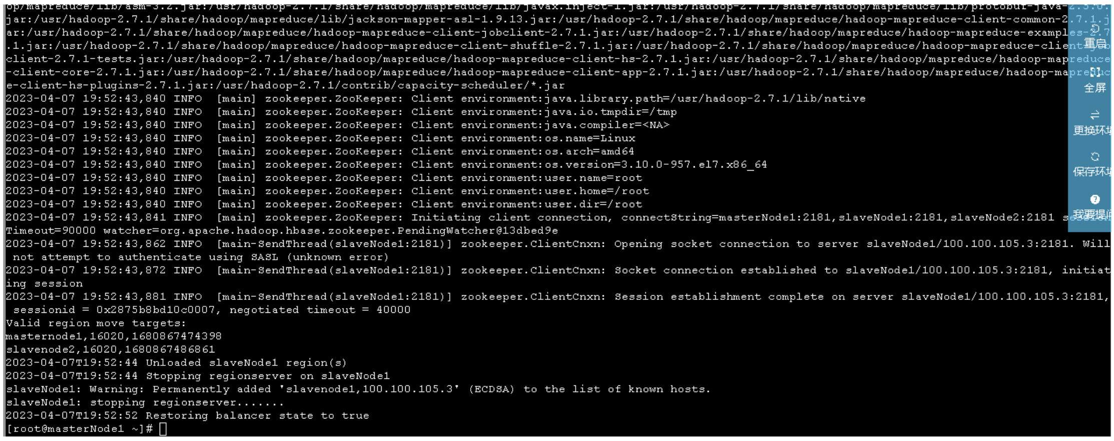


Backup Master slaveNode2

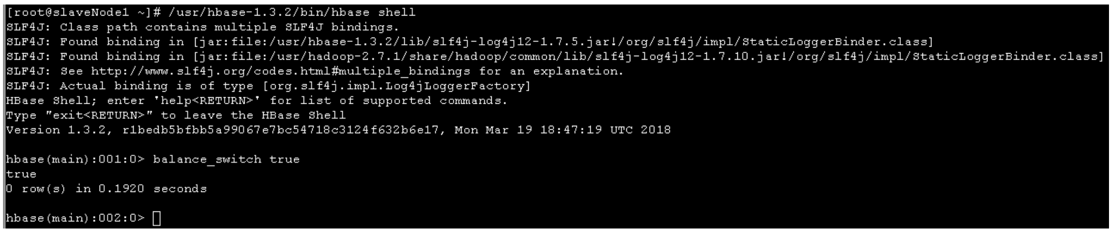
Current Active Master: slaveNode1

Tasks

Start Time	Description	State	Status
Fri Apr 07 19:42:18 CST 2023	Master startup	RUNNING (since 9mins, 20sec ago)	Another master is the active master, slavenode1,16000,1680868839319; waiting to become the next active master (since 8mins, 12sec ago)



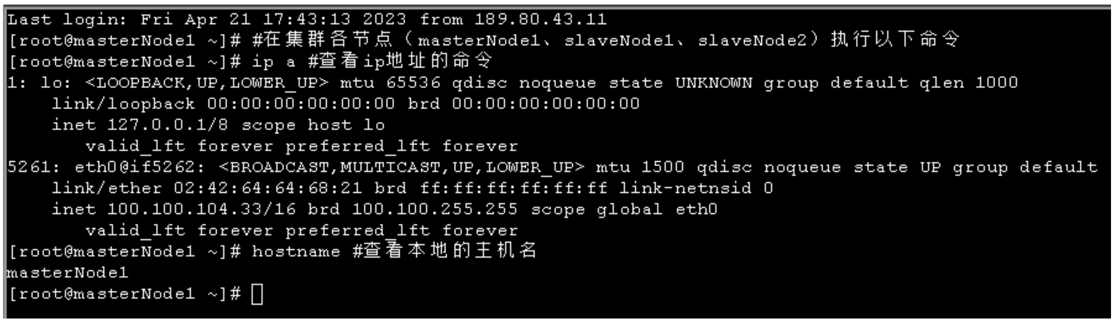
```
op/mapreduce/lib/asm-3.2.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/lib/javax.inject-1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/lib/protobuf-java-2.5.0.
jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/lib/jackson-mapper-asl-1.9.13.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-common-2.7.1.j
ar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-jobclient-2.7.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7
.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-shuffle-2.7.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-
client-2.7.1-tests.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-hs-2.7.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce
-client-core-2.7.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduce-client-app-2.7.1.jar:/usr/hadoop-2.7.1/share/hadoop/mapreduce/hadoop-mapreduc
e-client-hs-plugins-2.7.1.jar:/usr/hadoop-2.7.1/contrib/capacity-scheduler/*.jar
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:java.library.path=/usr/hadoop-2.7.1/lib/native
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:java.io.tmpdir=/tmp
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:java.compiler=<NA>
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:os.name=Linux
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:os.arch=amd64
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:os.version=3.10.0-957.el7.x86_64
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:user.name=root
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:user.home=/root
2023-04-07 19:52:43,840 INFO  [main] zookeeper.ZooKeeper: Client environment:user.dir=/root
2023-04-07 19:52:43,841 INFO  [main] zookeeper.ZooKeeper: Initiating client connection, connectString=masterNode1:2181,slaveNode1:2181,slaveNode2:2181 s
Timeout=90000 watcher=org.apache.hadoop.hbase.zookeeper.PendingWatcher@13dbed9e
2023-04-07 19:52:43,862 INFO  [main-SendThread(slaveNode1:2181)] zookeeper.ClientCnxn: Opening socket connection to server slaveNode1/100.100.105.3:2181. Will
 not attempt to authenticate using SASL (unknown error)
2023-04-07 19:52:43,872 INFO  [main-SendThread(slaveNode1:2181)] zookeeper.ClientCnxn: Socket connection established to slaveNode1/100.100.105.3:2181, initiat
ing session
2023-04-07 19:52:43,881 INFO  [main-SendThread(slaveNode1:2181)] zookeeper.ClientCnxn: Session establishment complete on server slaveNode1/100.100.105.3:2181,
 sessionid = 0x2875b8bd10c0007, negotiated timeout = 40000
Valid region move targets:
masternode1,16020,1680867474398
slavenode2,16020,1680867486861
2023-04-07T19:52:44 Unloaded slaveNode1 region(s)
2023-04-07T19:52:44 Stopping regionserver on slaveNode1
slaveNode1: Warning: Permanently added 'slavenode1,100.100.105.3' (ECDSA) to the list of known hosts.
slaveNode1: stopping regionserver.......
2023-04-07T19:52:52 Restoring balancer state to true
[root@masterNode1 ~]# 
```



```
[root@slaveNode1 ~]# /usr/hbase-1.3.2/bin/hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hbase-1.3.2/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hadoop-2.7.1/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.3.2, r1bedb5bfbb5a99067e7bc54718c3124f632b6e17, Mon Mar 19 18:47:19 UTC 2018

hbase(main):001:0> balance_switch true
true
0 row(s) in 0.1920 seconds

hbase(main):002:0> 
```

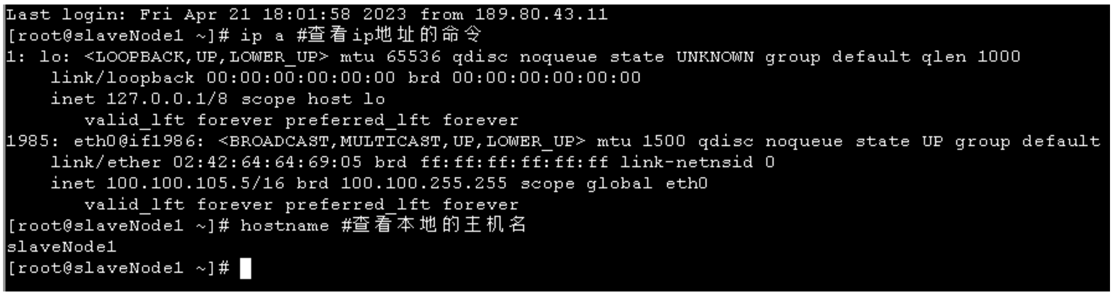
第三次作业

集群环境配置
查看 ip 地址命令
查看本地主机名



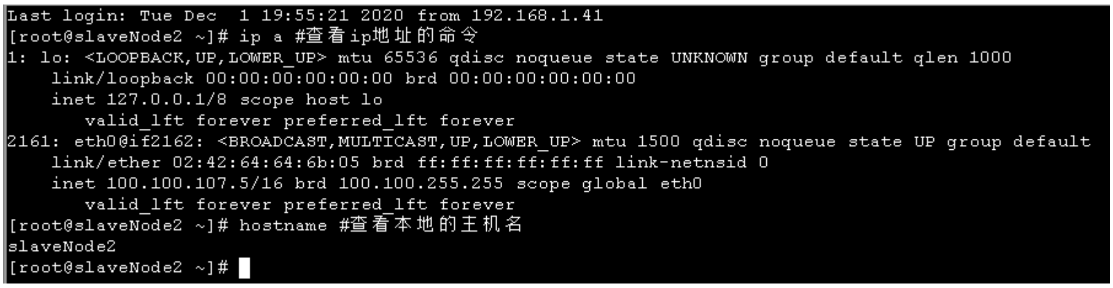
```
Last login: Fri Apr 21 17:43:13 2023 from 189.80.43.11
[root@masterNode1 ~]# #在集群各节点（masterNode1、slaveNode1、slaveNode2）执行以下命令
[root@masterNode1 ~]# ip a #查看ip地址的命令
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
5261: eth0@if5262: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    link/ether 02:42:64:64:68:21 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 100.100.104.33/16 brd 100.100.255.255 scope global eth0
       valid_lft forever preferred_lft forever
[root@masterNode1 ~]# hostname #查看本地的主机名
masterNode1
[root@masterNode1 ~]# 
```

slaveNode1



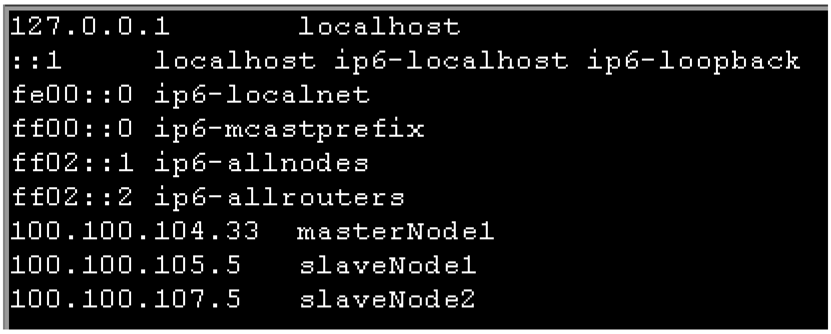
```
Last login: Fri Apr 21 18:01:58 2023 from 189.80.43.11
[root@slaveNode1 ~]# ip a #查看ip地址的命令
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
1985: eth0@if1986: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    link/ether 02:42:64:64:69:05 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 100.100.105.5/16 brd 100.100.255.255 scope global eth0
       valid_lft forever preferred_lft forever
[root@slaveNode1 ~]# hostname #查看本地的主机名
slaveNode1
[root@slaveNode1 ~]# 
```

slaveNode2



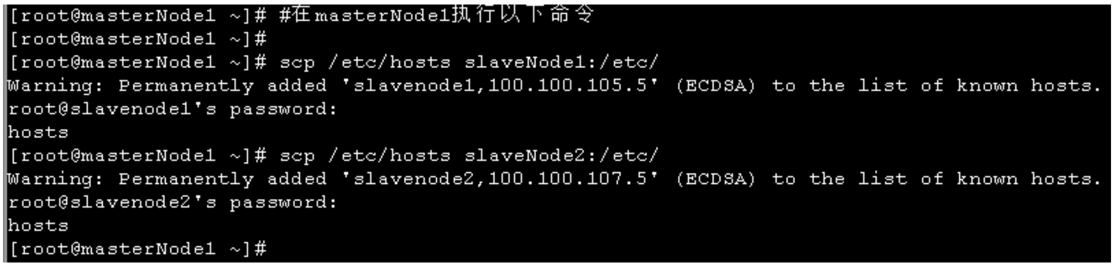
```
Last login: Tue Dec  1 19:55:21 2020 from 192.168.1.41
[root@slaveNode2 ~]# ip a #查看ip地址的命令
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
2161: eth0@if2162: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    link/ether 02:42:64:64:6b:05 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 100.100.107.5/16 brd 100.100.255.255 scope global eth0
       valid_lft forever preferred_lft forever
[root@slaveNode2 ~]# hostname #查看本地的主机名
slaveNode2
[root@slaveNode2 ~]# 
```

向/etc/hosts文件中添加 IP 和 hostname，格式如下图所示，其中间使用制表符隔开。



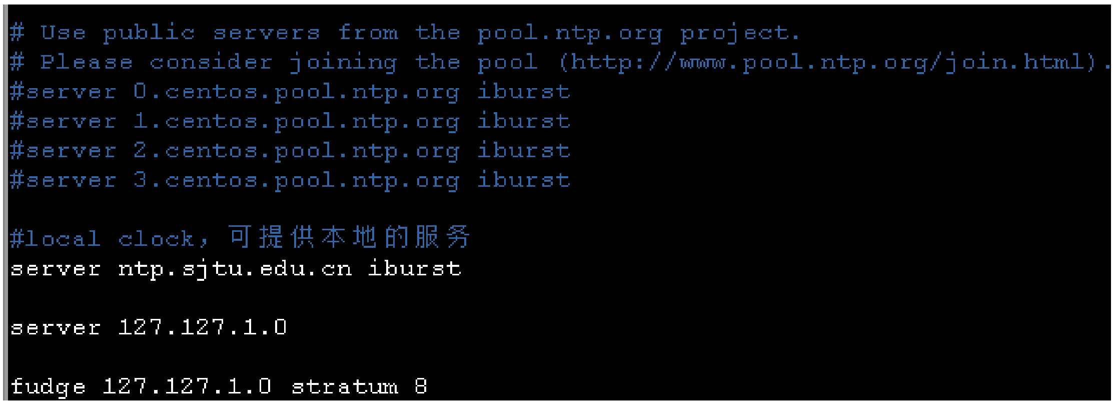
```
127.0.0.1       localhost
::1     localhost ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
100.100.104.33  masterNode1
100.100.105.5   slaveNode1
100.100.107.5   slaveNode2
```

分发 masterNode1 节点的/etc/hosts 文件到其余节点，此时我们可以使用其余节点的 hostname 进行分发。

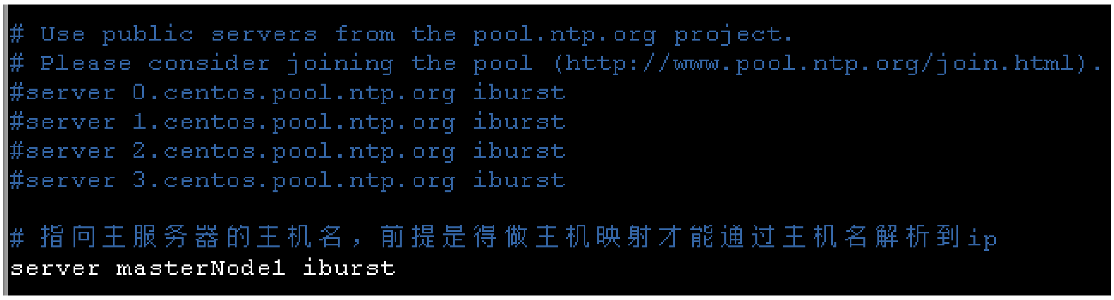


```
[root@masterNode1 ~]# #在masterNode1执行以下命令
[root@masterNode1 ~]#
[root@masterNode1 ~]# scp /etc/hosts slaveNode1:/etc/
Warning: Permanently added 'slavenode1,100.100.105.5' (ECDSA) to the list of known hosts.
root@slavenode1's password:
hosts
[root@masterNode1 ~]# scp /etc/hosts slaveNode2:/etc/
Warning: Permanently added 'slavenode2,100.100.107.5' (ECDSA) to the list of known hosts.
root@slavenode2's password:
hosts
[root@masterNode1 ~]#
```

让 masterNode1 节点作为 NTP 服务器的服务端。注释掉原文件所有以‘server’开头的配置项

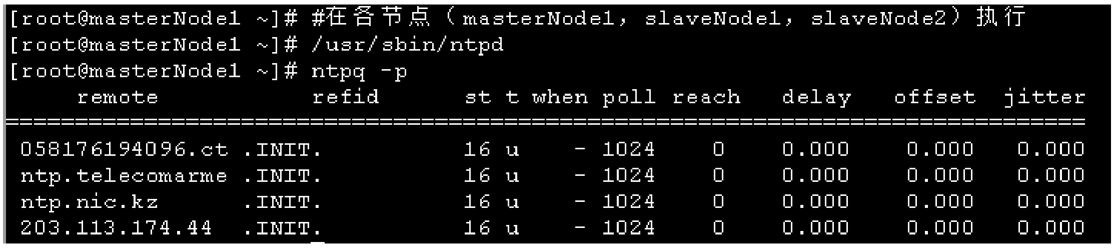


```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
#server 0.centos.pool.ntp.org iburst
#server 1.centos.pool.ntp.org iburst
#server 2.centos.pool.ntp.org iburst
#server 3.centos.pool.ntp.org iburst

#local clock, 可提供本地的服务
server ntp.sjtu.edu.cn iburst

server 127.127.1.0

fudge 127.127.1.0 stratum 8
```

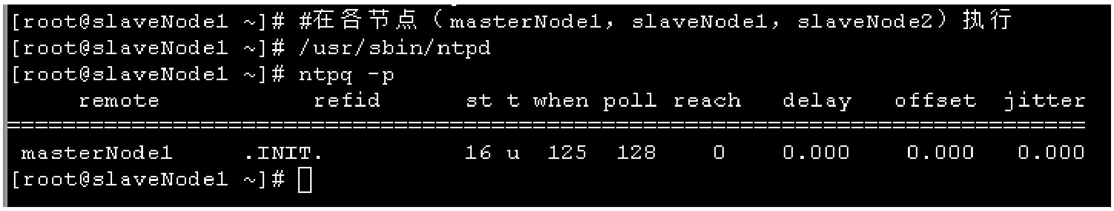


```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
#server 0.centos.pool.ntp.org iburst
#server 1.centos.pool.ntp.org iburst
#server 2.centos.pool.ntp.org iburst
#server 3.centos.pool.ntp.org iburst

# 指向主服务器的主机名，前提是得做主机映射才能通过主机名解析到ip
server masterNode1 iburst
```

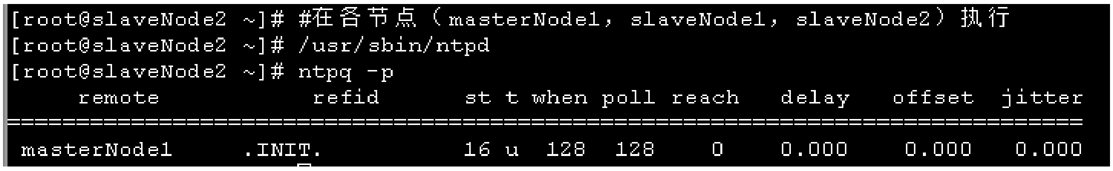
从 yum 源安装 NTP 服务
编辑 NTP 配置文件
查看时间同步是否成功：



```
[root@masterNode1 ~]# #在各节点（masterNode1, slaveNode1, slaveNode2）执行
[root@masterNode1 ~]# /usr/sbin/ntpd
[root@masterNode1 ~]# ntpq -p
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
 058176194096.ct .INIT.          16 u    - 1024    0    0.000    0.000   0.000
 ntp.telecomarme .INIT.          16 u    - 1024    0    0.000    0.000   0.000
 ntp.nic.kz      .INIT.          16 u    - 1024    0    0.000    0.000   0.000
 203.113.174.44  .INIT.          16 u    - 1024    0    0.000    0.000   0.000
```

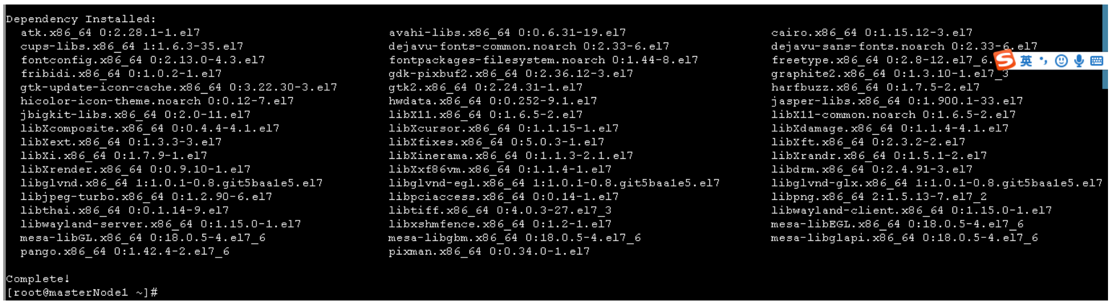


```
[root@slaveNode1 ~]# #在各节点（masterNode1, slaveNode1, slaveNode2）执行
[root@slaveNode1 ~]# /usr/sbin/ntpd
[root@slaveNode1 ~]# ntpq -p
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
 masterNode1     .INIT.          16 u  125  128    0    0.000    0.000   0.000
[root@slaveNode1 ~]# 
```

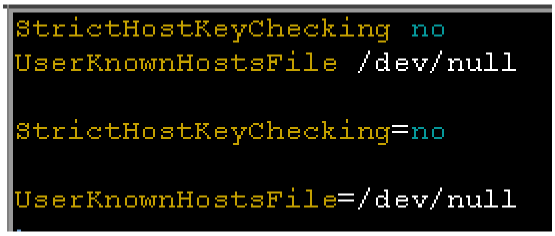


```
[root@slaveNode2 ~]# #在各节点（masterNode1, slaveNode1, slaveNode2）执行
[root@slaveNode2 ~]# /usr/sbin/ntpd
[root@slaveNode2 ~]# ntpq -p
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
 masterNode1     .INIT.          16 u  128  128    0    0.000    0.000   0.000
```

SSH 免密登录



```
Dependency Installed:
  atk.x86_64 0:2.28.1-1.el7                          avahi-libs.x86_64 0:0.6.31-19.el7                          cairo.x86_64 0:1.15.12-3.el7
  cups-libs.x86_64 1:1.6.3-35.el7                    dejavu-fonts-common.noarch 0:2.33-6.el7                    dejavu-sans-fonts.noarch 0:2.33-6.el7
  fontconfig.x86_64 0:2.13.0-4.3.el7                 fontpackages-filesystem.noarch 0:1.44-8.el7                freetype.x86_64 0:2.8-12.el7_6.
  fribidi.x86_64 0:1.0.2-1.el7                       gdk-pixbuf2.x86_64 0:2.36.12-3.el7                         graphite2.x86_64 0:1.3.10-1.el7_3
  gtk-update-icon-cache.x86_64 0:3.22.30-3.el7       gtk2.x86_64 0:2.24.31-1.el7                                harfbuzz.x86_64 0:1.7.5-2.el7
  hicolor-icon-theme.noarch 0:0.12-7.el7             hwdata.x86_64 0:0.252-9.1.el7                              jasper-libs.x86_64 0:1.900.1-33.el7
  jbigkit-libs.x86_64 0:2.0-11.el7                   libX11.x86_64 0:1.6.5-2.el7                                libX11-common.noarch 0:1.6.5-2.el7
  libXcomposite.x86_64 0:0.4.4-4.1.el7               libXcursor.x86_64 0:1.1.15-1.el7                           libXdamage.x86_64 0:1.1.4-4.1.el7
  libXext.x86_64 0:1.3.3-3.el7                       libXfixes.x86_64 0:5.0.3-1.el7                             libXft.x86_64 0:2.3.2-2.el7
  libXi.x86_64 0:1.7.9-1.el7                         libXinerama.x86_64 0:1.1.3-2.1.el7                         libXrandr.x86_64 0:1.5.1-2.el7
  libXrender.x86_64 0:0.9.10-1.el7                   libXxf86vm.x86_64 0:1.1.4-1.el7                            libdrm.x86_64 0:2.4.91-3.el7
  libglvnd.x86_64 1:1.0.1-0.8.git5baa1e5.el7         libglvnd-egl.x86_64 1:1.0.1-0.8.git5baa1e5.el7             libglvnd-glx.x86_64 1:1.0.1-0.8.git5baa1e5.el7
  libjpeg-turbo.x86_64 0:1.2.90-6.el7                libpciaccess.x86_64 0:0.14-1.el7                           libpng.x86_64 2:1.5.13-7.el7_2
  libthai.x86_64 0:0.1.14-9.el7                      libtiff.x86_64 0:4.0.3-27.el7_3                            libwayland-client.x86_64 0:1.15.0-1.el7
  libwayland-server.x86_64 0:1.15.0-1.el7            libxshmfence.x86_64 0:1.2-1.el7                            mesa-libEGL.x86_64 0:18.0.5-4.el7_6
  mesa-libGL.x86_64 0:18.0.5-4.el7_6                 mesa-libgbm.x86_64 0:18.0.5-4.el7_6                        mesa-libglapi.x86_64 0:18.0.5-4.el7_6
  pango.x86_64 0:1.42.4-2.el7_6                      pixman.x86_64 0:0.34.0-1.el7

Complete!
[root@masterNode1 ~]#
```



```
StrictHostKeyChecking no
UserKnownHostsFile /dev/null

StrictHostKeyChecking=no

UserKnownHostsFile=/dev/null
```

#在 masterNode1,slaveNode1,slaveNode2 执行

```
/usr/sbin/sshd
```

#生成 ecdsa 密钥

```
ssh-keygen -t ecdsa -f /etc/ssh/ssh_host_ecdsa_key -N ""
```

#生成 ed25519 密钥

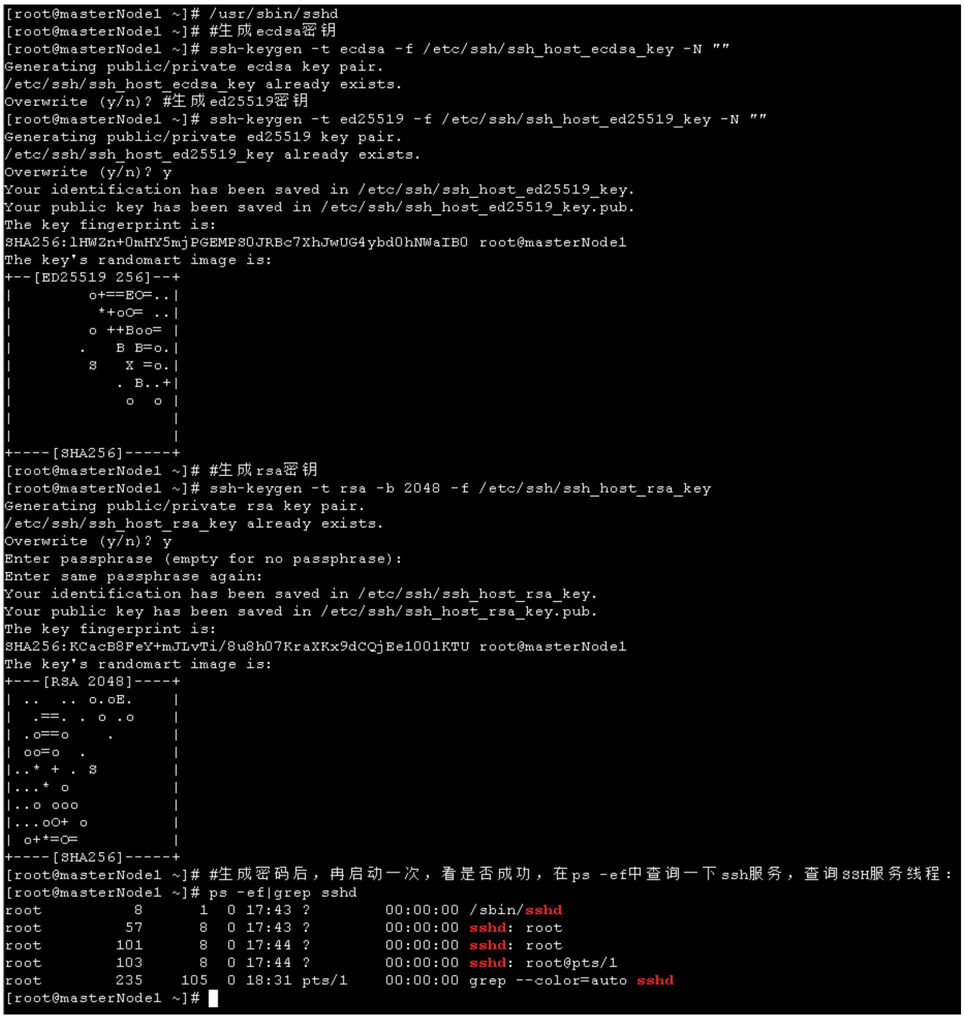
```
ssh-keygen -t ed25519 -f /etc/ssh/ssh_host_ed25519_key -N ""
```

#生成 rsa 密钥

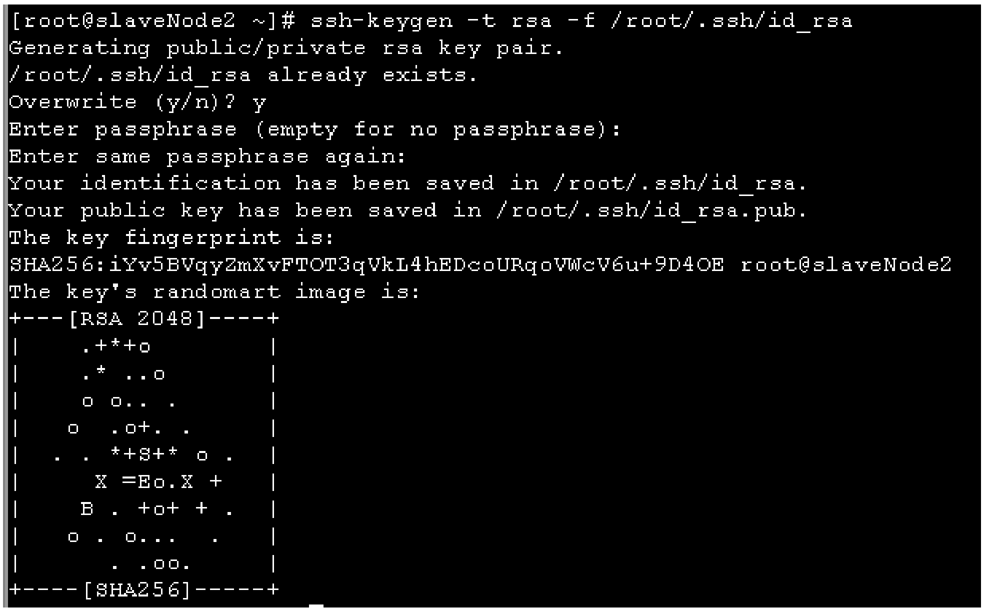
```
ssh-keygen -t rsa -b 2048 -f /etc/ssh/ssh_host_rsa_key
```

#生成密码后，再启动一次，看是否成功，在 ps -ef 中查询一下 ssh 服务，查询 SSH 服务线程：

```
ps -ef|grep sshd
```



```
[root@masterNode1 ~]# /usr/sbin/sshd
[root@masterNode1 ~]# #生成ecdsa密钥
[root@masterNode1 ~]# ssh-keygen -t ecdsa -f /etc/ssh/ssh_host_ecdsa_key -N ""
Generating public/private ecdsa key pair.
/etc/ssh/ssh_host_ecdsa_key already exists.
Overwrite (y/n)? #生成ed25519密钥
[root@masterNode1 ~]# ssh-keygen -t ed25519 -f /etc/ssh/ssh_host_ed25519_key -N ""
Generating public/private ed25519 key pair.
/etc/ssh/ssh_host_ed25519_key already exists.
Overwrite (y/n)? y
Your identification has been saved in /etc/ssh/ssh_host_ed25519_key.
Your public key has been saved in /etc/ssh/ssh_host_ed25519_key.pub.
The key fingerprint is:
SHA256:lHWZn+0mHY5mjPGEMPS0JRBc7XhJwUG4ybd0hNWaIB0 root@masterNode1
The key's randomart image is:
+--[ED25519 256]--+
|        o+==EO=..|
|         *+oO= ..|
|        o ++Boo= |
|       .   B B=o.|
|        S   X =o.|
|           . B..+|
|            o  o |
|                 |
|                 |
+----[SHA256]-----+
[root@masterNode1 ~]# #生成rsa密钥
[root@masterNode1 ~]# ssh-keygen -t rsa -b 2048 -f /etc/ssh/ssh_host_rsa_key
Generating public/private rsa key pair.
/etc/ssh/ssh_host_rsa_key already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /etc/ssh/ssh_host_rsa_key.
Your public key has been saved in /etc/ssh/ssh_host_rsa_key.pub.
The key fingerprint is:
SHA256:KCacB8FeY+mJLvTi/8u8h07KraXKx9dCQjEe1001KTU root@masterNode1
The key's randomart image is:
+---[RSA 2048]----+
| ..  .. o.oE.    |
|  .==. . o .o    |
| .o==o    .      |
| oo=o  .         |
|..* + . S        |
|...* o           |
|..o ooo          |
|...oO+ o         |
| o+*=O=          |
+----[SHA256]-----+
[root@masterNode1 ~]# #生成密码后，再启动一次，看是否成功，在ps -ef中查询一下ssh服务，查询SSH服务线程：
[root@masterNode1 ~]# ps -ef|grep sshd
root          8      1  0 17:43 ?        00:00:00 /sbin/sshd
root         57      8  0 17:43 ?        00:00:00 sshd: root
root        101      8  0 17:44 ?        00:00:00 sshd: root
root        103      8  0 17:44 ?        00:00:00 sshd: root@pts/1
root        235    105  0 18:31 pts/1    00:00:00 grep --color=auto sshd
[root@masterNode1 ~]# 
```

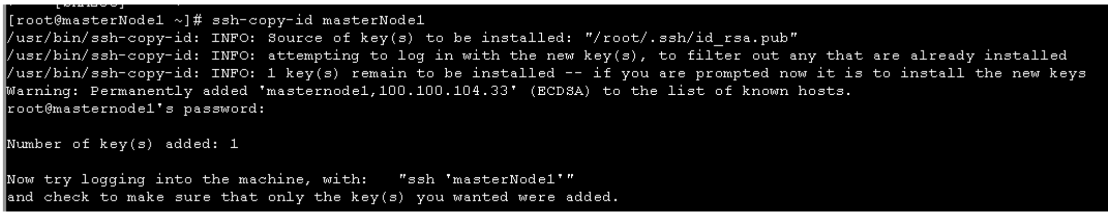


```
[root@slaveNode2 ~]# ssh-keygen -t rsa -f /root/.ssh/id_rsa
Generating public/private rsa key pair.
/root/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:iYv5BVqyZmXvFTOT3qVkL4hEDcoURqoVWcV6u+9D4OE root@slaveNode2
The key's randomart image is:
+---[RSA 2048]----+
|     .+*+o       |
|     .* ..o      |
|     o o.. .     |
|    o  .o+. .    |
|   . .  *+S+* o .|
|      X =Eo.X +  |
|     B . +o+ + . |
|    o . o...   . |
|       . .oo.    |
+----[SHA256]-----+
```

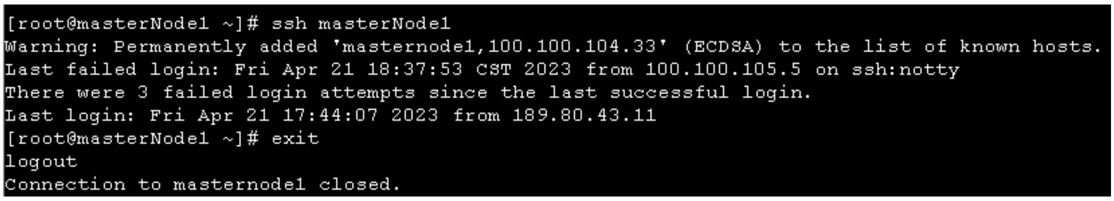
将证书发送到本机及其他机器上，需要输入登录密码 datastudio

```
#在 masterNode1,slaveNode1,slaveNode2 执行

ssh-copy-id  masterNode1

ssh-copy-id  slaveNode1

ssh-copy-id  slaveNode2
```

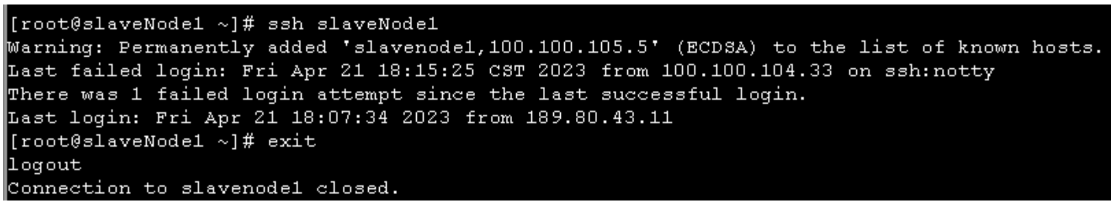
测试验证，使用在各节点中使用 SSH 命令分别进行登录测试，查看能否无密码登入，均可登录即设置成功，可使用 exit 命令退出。



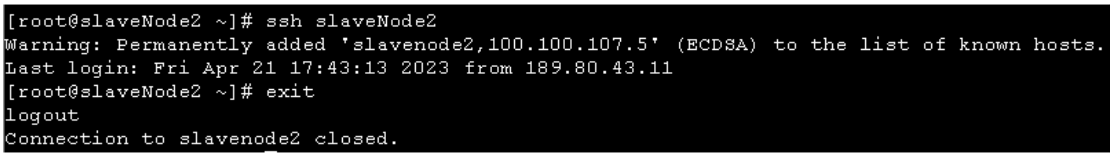
```
[root@masterNode1 ~]# ssh-copy-id masterNode1
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/root/.ssh/id_rsa.pub"
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompted now it is to install the new keys
Warning: Permanently added 'masternode1,100.100.104.33' (ECDSA) to the list of known hosts.
root@masternode1's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'masterNode1'"
and check to make sure that only the key(s) you wanted were added.
```



```
[root@masterNode1 ~]# ssh masterNode1
Warning: Permanently added 'masternode1,100.100.104.33' (ECDSA) to the list of known hosts.
Last failed login: Fri Apr 21 18:37:53 CST 2023 from 100.100.105.5 on ssh:notty
There were 3 failed login attempts since the last successful login.
Last login: Fri Apr 21 17:44:07 2023 from 189.80.43.11
[root@masterNode1 ~]# exit
logout
Connection to masternode1 closed.
```

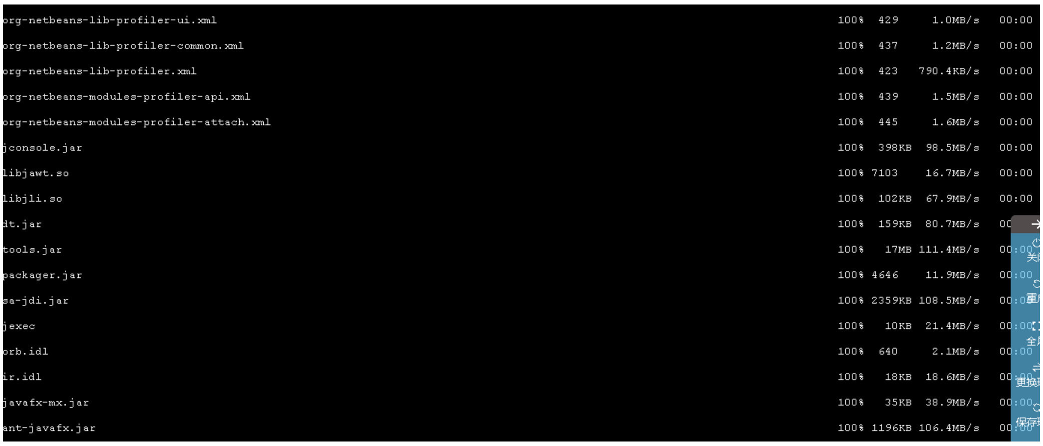


```
[root@slaveNode1 ~]# ssh slaveNode1
Warning: Permanently added 'slavenode1,100.100.105.5' (ECDSA) to the list of known hosts.
Last failed login: Fri Apr 21 18:15:25 CST 2023 from 100.100.104.33 on ssh:notty
There was 1 failed login attempt since the last successful login.
Last login: Fri Apr 21 18:07:34 2023 from 189.80.43.11
[root@slaveNode1 ~]# exit
logout
Connection to slavenode1 closed.
```



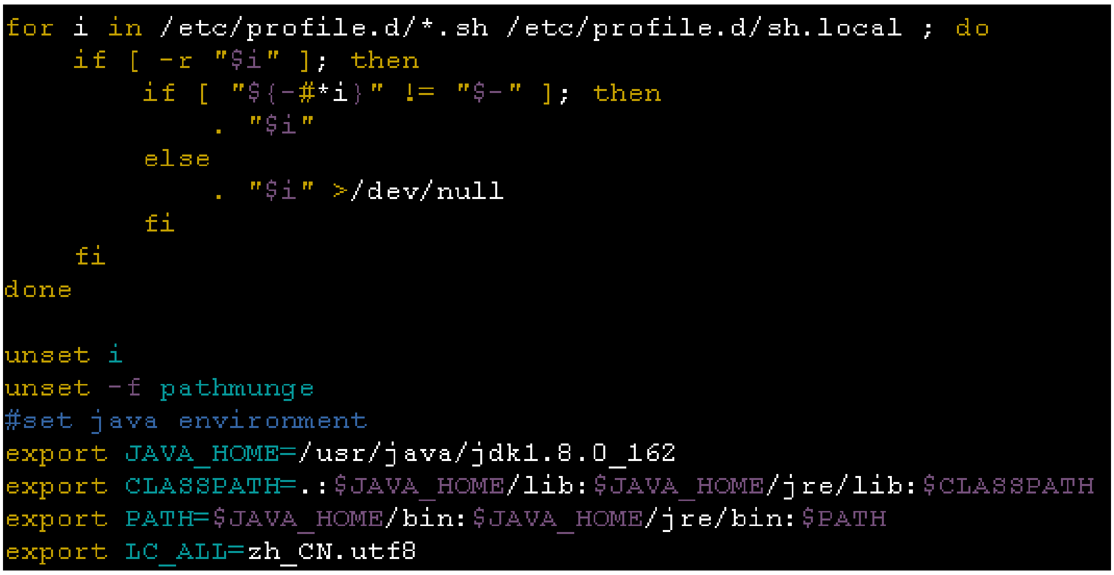
```
[root@slaveNode2 ~]# ssh slaveNode2
Warning: Permanently added 'slavenode2,100.100.107.5' (ECDSA) to the list of known hosts.
Last login: Fri Apr 21 17:43:13 2023 from 189.80.43.11
[root@slaveNode2 ~]# exit
logout
Connection to slavenode2 closed.
```

安装 JDK



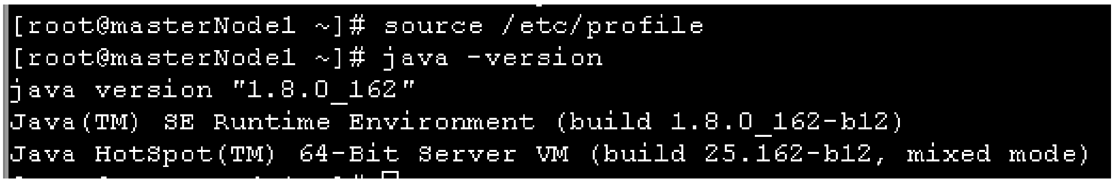
```
org-netbeans-lib-profiler-ui.xml                  100%  429     1.0MB/s   00:00
org-netbeans-lib-profiler-common.xml              100%  437     1.2MB/s   00:00
org-netbeans-lib-profiler.xml                     100%  423   790.4KB/s   00:00
org-netbeans-modules-profiler-api.xml             100%  439     1.5MB/s   00:00
org-netbeans-modules-profiler-attach.xml          100%  445     1.6MB/s   00:00
jconsole.jar                                      100%  398KB  98.5MB/s   00:00
libjawt.so                                        100% 7103    16.7MB/s   00:00
libjli.so                                         100%  102KB  67.9MB/s   00:00
dt.jar                                            100%  159KB  80.7MB/s   00
tools.jar                                         100%   17MB 111.4MB/s   00:00
packager.jar                                      100% 4646    11.9MB/s   00:00
sa-jdi.jar                                        100% 2359KB 108.5MB/s   00:0
jexec                                             100%   10KB  21.4MB/s   00:00
orb.idl                                           100%  640     2.1MB/s   00:00
ir.idl                                            100%   18KB  18.6MB/s   00:00
javafx-mx.jar                                     100%   35KB  38.9MB/s   00:00
ant-javafx.jar                                    100% 1196KB 106.4MB/s   00:00
```

为了使用方便，我们可以为 JDK 添加环境变量，在/usr/profile 文件末尾追加以下内容，注意 JDK 的环境路径。



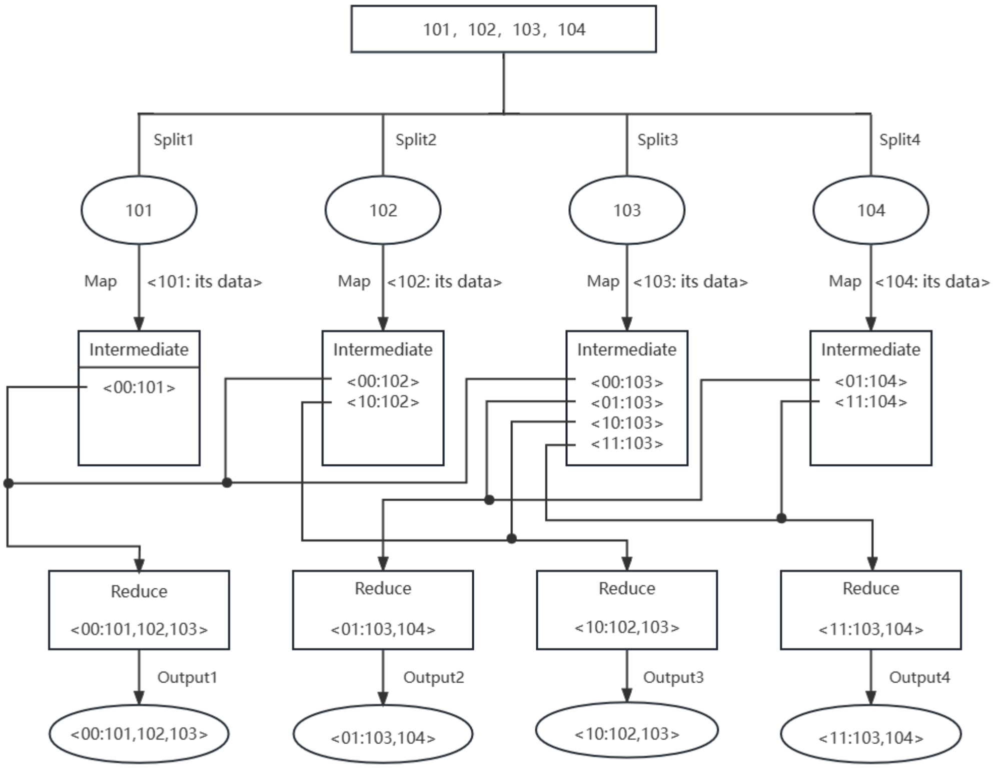
```
for i in /etc/profile.d/*.sh /etc/profile.d/sh.local ; do
    if [ -r "$i" ]; then
        if [ "${-#*i}" != "$-" ]; then
            . "$i"
        else
            . "$i" >/dev/null
        fi
    fi
done

unset i
unset -f pathmunge
#set java environment
export JAVA_HOME=/usr/java/jdk1.8.0_162
export CLASSPATH=.:$JAVA_HOME/lib:$JAVA_HOME/jre/lib:$CLASSPATH
export PATH=$JAVA_HOME/bin:$JAVA_HOME/jre/bin:$PATH
export LC_ALL=zh_CN.utf8
```

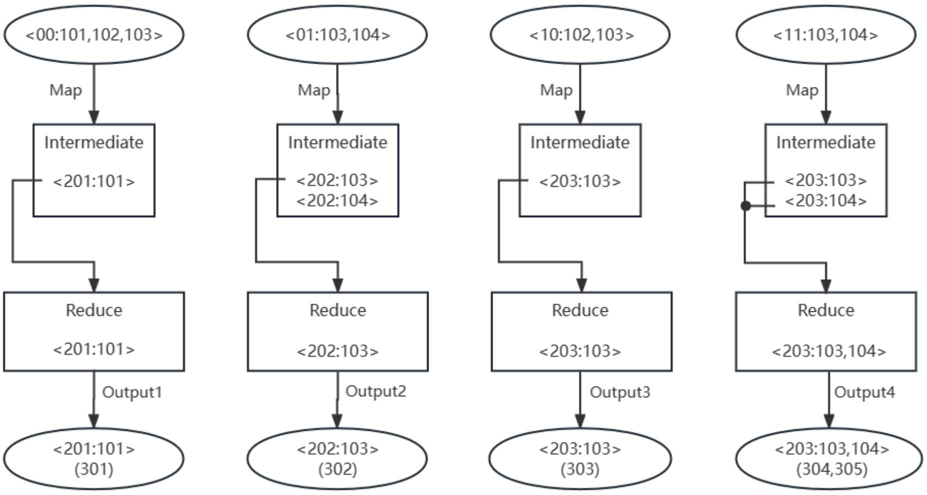
编译环境变量后，查询 Java 版本信息，验证配置是否正确：



```
[root@masterNode1 ~]# source /etc/profile
[root@masterNode1 ~]# java -version
java version "1.8.0_162"
Java(TM) SE Runtime Environment (build 1.8.0_162-b12)
Java HotSpot(TM) 64-Bit Server VM (build 25.162-b12, mixed mode)
```

第四次作业





上图，主题层多边形从 101 编码到 104 被提交到主机，然后在将所有资源随机分为四项任务后接收多边形 ID 和它们的数据，每个 map 任务产生一个由网格的中间键值对组成的 ID 和一个多边形的 ID，它指示多边形与网格相交。随后，中间键值具有相同网格 ID 的对被输入到相同的 reduce 任务中。这些多边形与网格执行交叉测试，那些与网格不相交的多边形被从键值对列表中删除。类似地，裁剪层的网格索引多边形被生成。网格索引的建立结果如下图。这个 map 任务的输出是剪辑层多边形的 ID 和有可能与前者相交的主题层的 ID 的中间键值对。在 reduce 阶段，每个任务都有具有相同剪辑层多边形 ID 的中间键值对。reduce 任务的输出是裁剪层多边形与主题多边形相交得到的多边形。