BigData Concepts

01 January 2024 09:43

Hbase Configuration

```
hadoop@mainserver1:/usr/local/hadoop/etc/hadoop$ cd
hadoop@mainserver1:~$ cat /etc/hosts
127.0.0.1 localhost
127.0.1.1 mainserver1

# The following lines are desirable for IPv6 capable hosts
::1     ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters

192.168.56.100  mainserver1
hadoop@mainserver1:~$
```

$ su
# **nano** /etc/bash.bashrc

```
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HADOOP_HOME=/usr/local/hadoop
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export HADOOP_YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
export CLASSPATH=`hadoop classpath`
export PDSH_RCMD_TYPE=ssh
#HBASE CONFIGS
export HBASE_HOME="/usr/local/hbase"
export PATH="$HBASE_HOME/bin:$PATH"
```

$ nano /usr/local/hadoop/etc/hadoop/core-site.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<configuration>
     <property>
         <name>fs.default.name</name>
         <value>hdfs://localhost:9000</value>
     </property>
</configuration>
<save and exit>
```

$nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
<property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>

  <property>
    <name>dfs.name.dir</name>
    <value>file:///home/hadoop/hadoopinfra/hdfs/namenode </value>
  </property>
```

```
  <property>
    <name>dfs.data.dir</name>
    <value>file:///home/hadoop/hadoopinfra/hdfs/datanode </value>
  </property>
</configuration>
```

<save and exit>

$ nano /usr/local/hadoop/etc/hadoop/mapred-site.xml

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

<save and exit>

$ nano /usr/local/hadoop/etc/hadoop/yarn-site.xml

```
<?xml version="1.0"?>

      <property>
          <name>yarn.nodemanager.aux-services</name>
          <value>mapreduce_shuffle</value>
       </property>
</configuration>
```

$ **nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh**

```
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HADOOP_OPTS=" -Djava.net.preferIPv4Stack=true"
```

<save and exit>

$ nano slaves

```
localhost
```

<save and exit>

$ nano /usr/local/hbase/conf/regionservers

```
localhost
```

<save and exit>

$ nano /usr/local/hbase/conf/hbase-site.xml

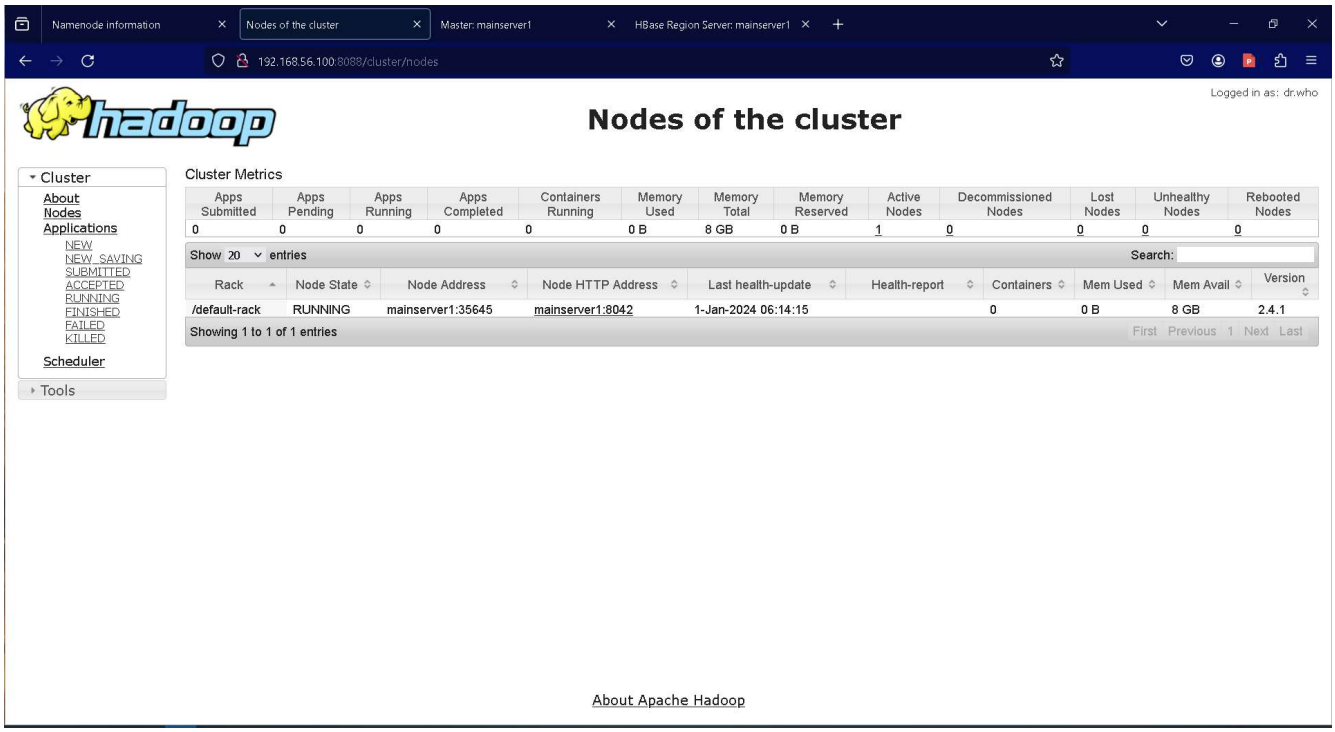
```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<configuration>
 <property>
   <name>hbase.rootdir</name>
   <value>hdfs://localhost:9000/hbase</value>
 </property>
 <property>
    <name>hbase.zookeeper.property.dataDir</name>
    <value>hdfs://localhost:9000/zookeeper</value>
  </property>
  <property>
    <name>hbase.cluster.distributed</name>
    <value>true</value>
  </property>
</configuration>
```

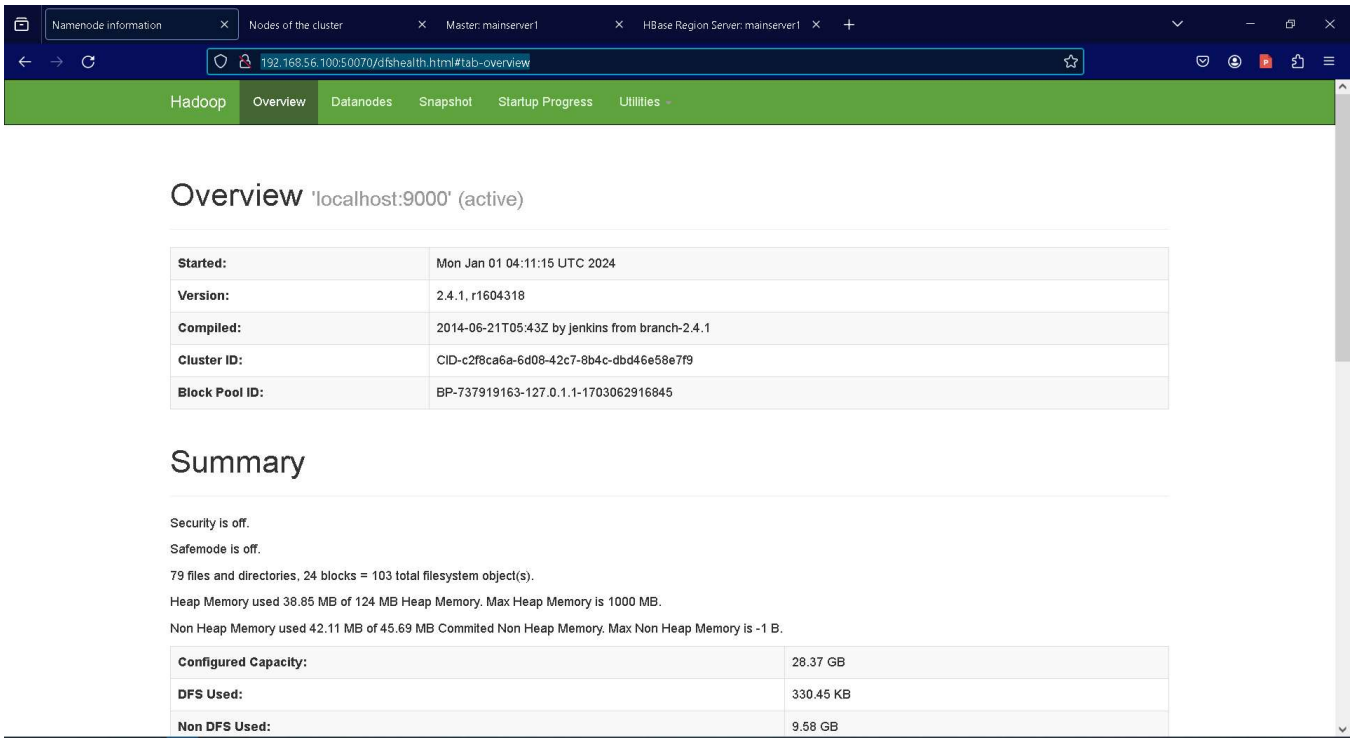
<save and exit>

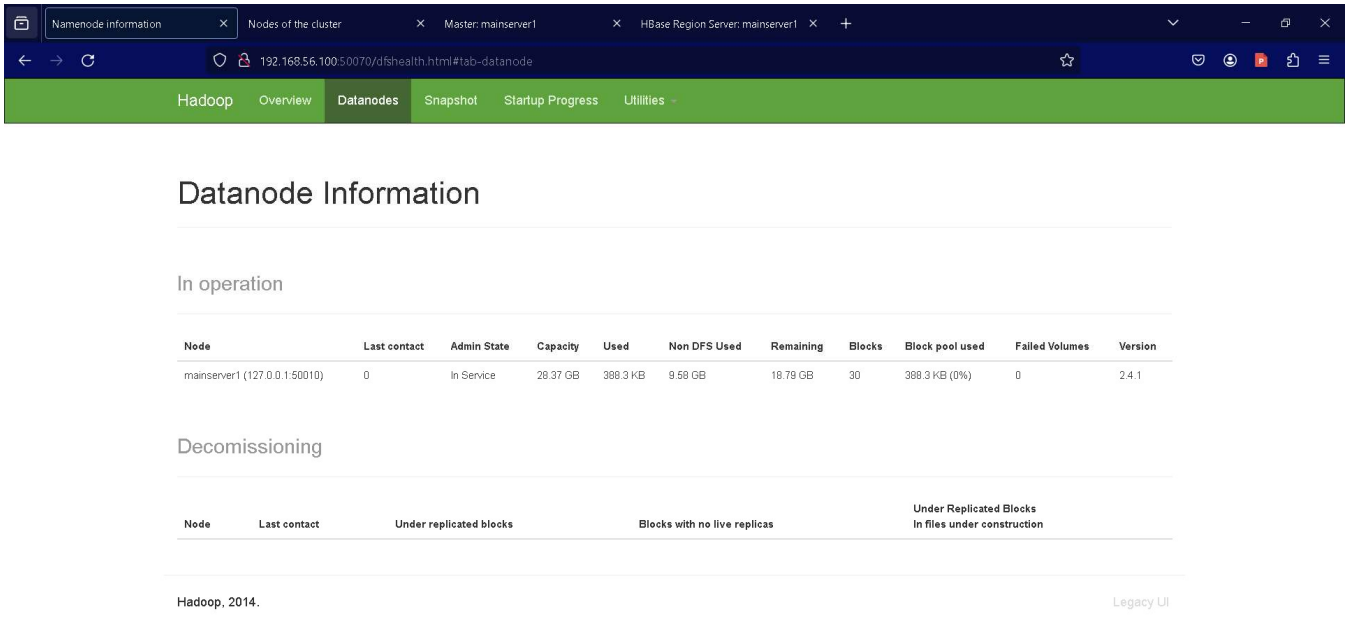
$ cd /home/hadoop/hadoopinfra
$ mkdir hbase zookeeper
$ chown -R hadoop.hadoop hbase
$ chown -R hadoop.hadoop zookeeper
$ start-dfs.sh
$ start-yarn.sh
$ start-hbase.sh
$ jps (output should be as shown below)

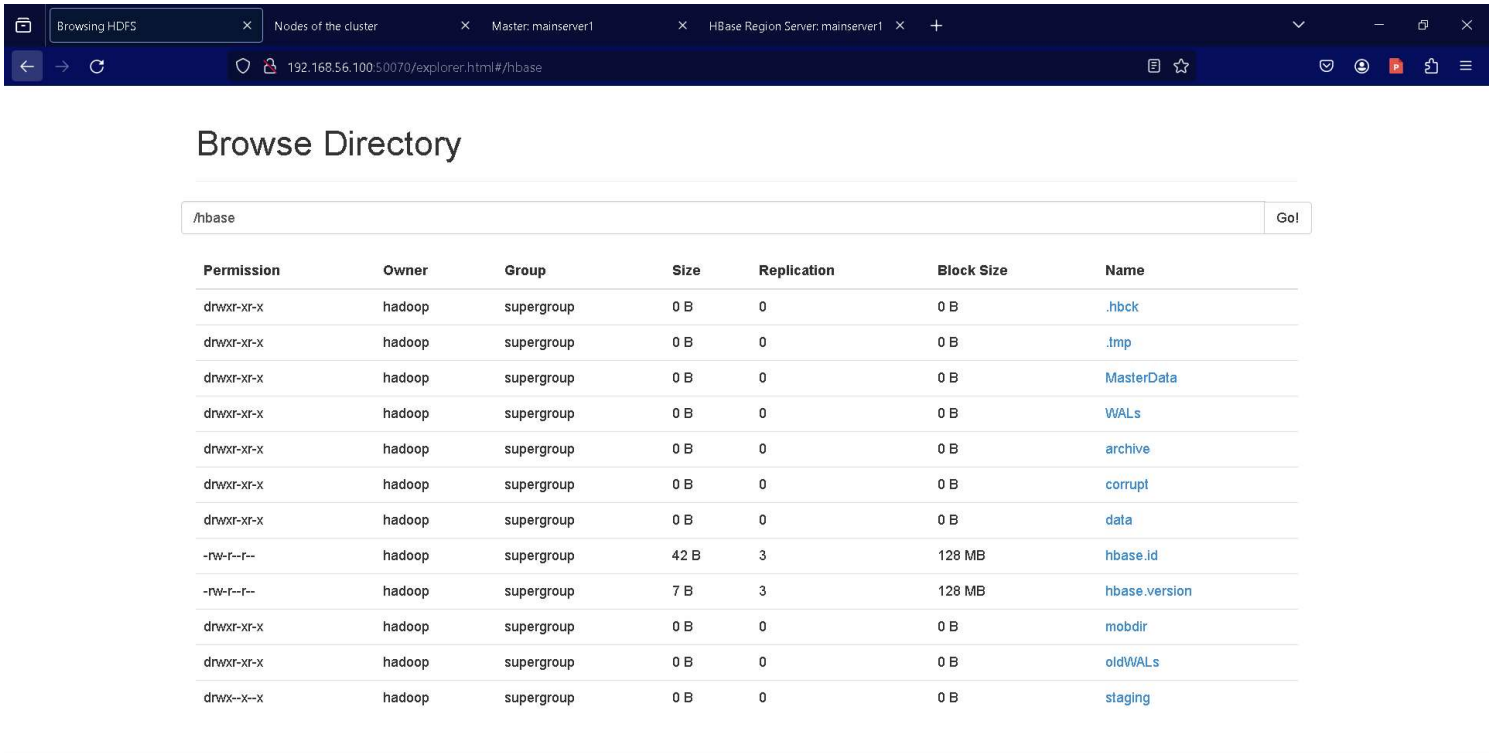
```
hadoop@mainserver1:/usr/local/hbase/conf$ jps
1794 SecondaryNameNode
2131 NodeManager
6117 Jps
1527 DataNode
1959 ResourceManager
1368 NameNode
2859 HMaster
2766 HQuorumPeer
2991 HRegionServer
hadoop@mainserver1:/usr/local/hbase/conf$
```

Browser related details for port 8088, 50070, 16010 and 16030 should be as shown below




Namenode information
Nodes of the cluster
Master: mainserver1
HBase Region Server: mainserver1
192.168.56.100:50070/dfshealth.html#tab-overview
Hadoop
Overview
Datanodes
Snapshot
Startup Progress
Utilities
Overview 'localhost:9000' (active)
Started: Mon Jan 01 04:11:15 UTC 2024
Version: 2.4.1, r1604318
Compiled: 2014-06-21T05:43Z by jenkins from branch-2.4.1
Cluster ID: CID-c2f8ca6a-6d08-42c7-8b4c-dbd46e58e7f9
Block Pool ID: BP-737919163-127.0.1.1-1703062916845
Summary
Security is off.
Safemode is off.
79 files and directories, 24 blocks = 103 total filesystem object(s).
Heap Memory used 38.85 MB of 124 MB Heap Memory. Max Heap Memory is 1000 MB.
Non Heap Memory used 42.11 MB of 45.69 MB Commited Non Heap Memory. Max Non Heap Memory is -1 B.
Configured Capacity: 28.37 GB
DFS Used: 330.45 KB
Non DFS Used: 9.58 GB


Namenode information
Nodes of the cluster
Master: mainserver1
HBase Region Server: mainserver1
192.168.56.100:50070/dfshealth.html#tab-datanode
Hadoop
Overview
Datanodes
Snapshot
Startup Progress
Utilities
Datanode Information
In operation
Node
Last contact
Admin State
Capacity
Used
Non DFS Used
Remaining
Blocks
Block pool used
Failed Volumes
Version
mainserver1 (127.0.0.1:50010)
0
In Service
28.37 GB
388.3 KB
9.58 GB
18.79 GB
30
388.3 KB (0%)
0
2.4.1
Decomissioning
Node
Last contact
Under replicated blocks
Blocks with no live replicas
Under Replicated Blocks
In files under construction
Hadoop, 2014.
Legacy UI



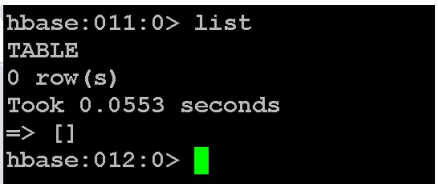
hadoop@mainserver1:~$ ls -l /home/hadoop/hadoopinfra/
total 12
drwxrwxrwx 12 hadoop hadoop 4096 Dec 31 12:51 hbase
drwxrwxrwx 4 hadoop hadoop 4096 Jan 1 04:45 hdfs
drwxrwxrwx 3 hadoop hadoop 4096 Dec 31 11:57 zookeeper

$ hbase shell

```
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
For Reference, please visit: http://hbase.apache.org/2.0/book.html#shell
Version 2.4.17, r7fd096f39b4284da9a71da3ce67c48d259ffa79a, Fri Mar 31 18:10:45 UTC 2023
Took 0.0113 seconds
hbase:001:0>
```

HBase Command:-

1. Create a table
2. scan
3. List a table
4. Enable a table
5. Disable a table
6. Describe table
7. Alter a table
8. Drow a table
9. Exists
10. Showdown
11. Create data
12. Update data
13. Read data
14. Delete date
15. Count and truncate
16. security
17. Client API



1. **Start HBase Shell:**

   hbase shell

2. **Create a Table:**

   create 'mytable', 'col1', 'col2'

   This command creates a table named 'mytable' with two column families: 'col1' and 'col2'.

3. **List Tables:**

   list

4. **Describe a Table:**

   describe 'mytable'

5. **Put Data into a Table:**

   put 'mytable', 'row1', 'col1:col1', 'value1'

   This command inserts a value ('value1') into column 'col1' of column family 'col1' for row 'row1' in 'mytable'.

   put 'mytable', 'row2', 'col1:col1', 'value2'

   put 'mytable', 'row3', 'col1:col1', 'value2'

6. **Get Data from a Table:**

   get 'mytable', 'row1'

   Retrieves the entire row with key 'row1' from 'mytable'.

7. **Scan a Table:**

   scan 'mytable'

   Scans and displays all rows in 'mytable'.

8. **Delete Data from a Table:**

   delete 'mytable', 'row1', 'col1:col1'

   Deletes the value in 'col1' of 'col1' for row 'row1' in 'mytable'.

9. **Disable and Drop a Table:**

   disable 'mytable' drop 'mytable'

   Disables and drops the 'mytable' table.

10. **Count Number of Rows in a Table:**

    count 'mytable'

    Returns the count of rows in 'mytable'.

11. **List Column Families of a Table:**

    list 'mytable'

    Lists all column families of 'mytable'.