

HBase @ Phonepe

Task 1:

- Create a Pseudo-distributed mode hbase cluster, with local storage.

Prerequisites:

Install and configure a single node hadoop on the machine.

Steps:

1. Install the hbase version that supports your hadoop version.
Extract the tar file and move it into the /usr/local/hbase.

```
hduser@hadoop-VirtualBox:~$ sudo wget http://dlcdn.apache.org/hbase/2.3.7/hbase-2.3.7-bin.tar.gz
--2022-03-11 18:20:37--  http://dlcdn.apache.org/hbase/2.3.7/hbase-2.3.7-bin.tar.gz
Resolving dlcdn.apache.org (dlcdn.apache.org)... 151.101.2.132, 2a04:4e42::644
Connecting to dlcdn.apache.org (dlcdn.apache.org)|151.101.2.132|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 272812222 (260M) [application/x-gzip]
Saving to: 'hbase-2.3.7-bin.tar.gz'

hbase-2.3.7-bin.tar.gz          100%[=====] 260.17M  7.32MB/s   in 41s

2022-03-11 18:21:19 (6.30 MB/s) - 'hbase-2.3.7-bin.tar.gz' saved [272812222/272812222]

hduser@hadoop-VirtualBox:~$ sudo tar xzf hbase-2.3.7-bin.tar.gz
hduser@hadoop-VirtualBox:~$ sudo mv hbase-2.3.7-bin.tar.gz /usr/local/hbase
```

2. Configure the hbase file.

Go to .bashrc file and add the following to that file.

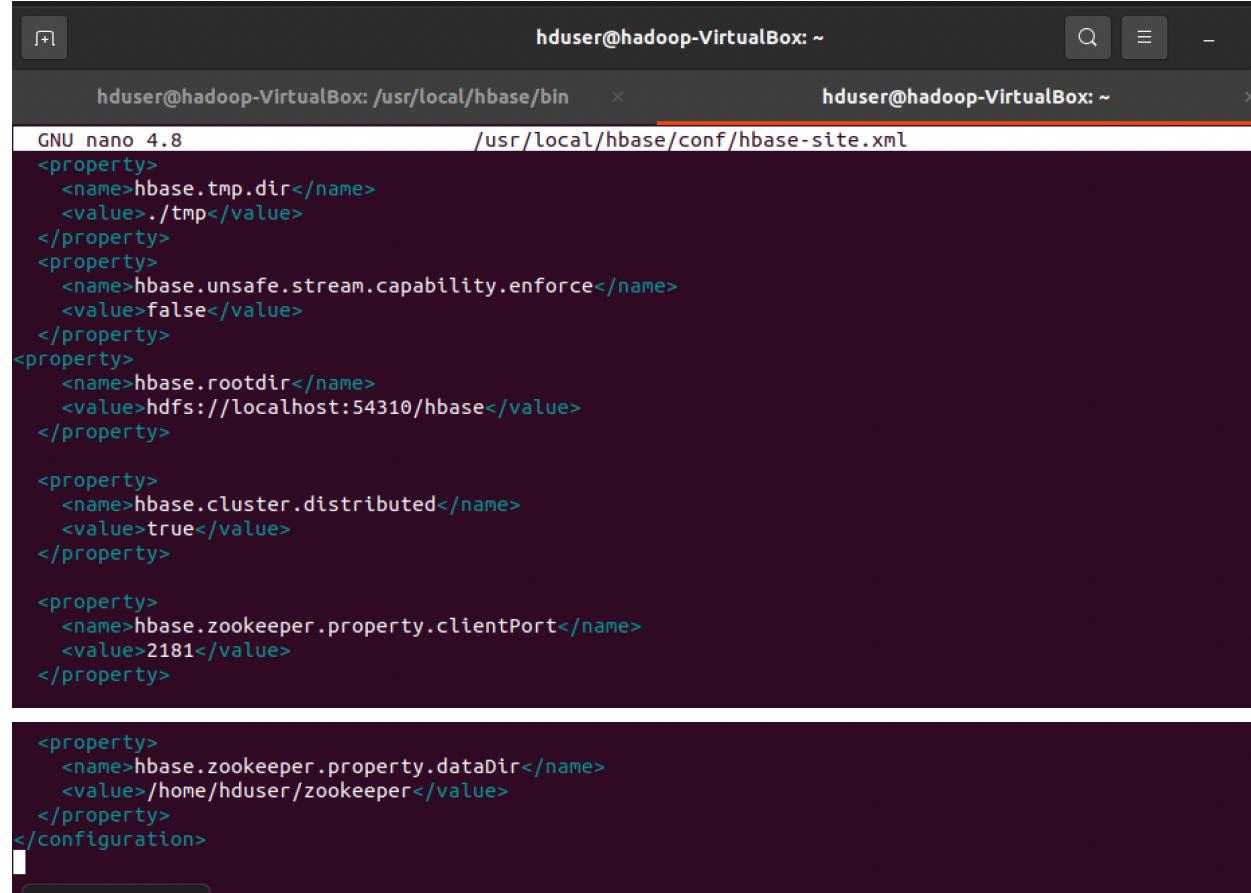
```
# HBASE VARIABLES
export HBASE_HOME=/usr/local/hbase
export PATH=$PATH:$HBASE_HOME/bin
# HBASE VARIABLES END
```

Add the following configuration to the /usr/local/hbase/conf/hbase-env.sh

```
hduser@hadoop-VirtualBox:~$ nano /usr/local/hbase/conf/hbase-env.sh
# so try to keep things idempotent unless you want to take an even deeper look
# into the startup scripts (bin/hbase, etc.)

# The java implementation to use. Java 1.8+ required.
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HBASE_REGIONSERVERS=/usr/local/hbase/conf/regionservers
export HBASE_MANAGES_ZK=true
# Extra Java CLASSPATH elements. Optional.
# export HBASE_CLASSPATH=
```

Add the following to the /usr/local/hbase/conf/hbase-site.xml



The screenshot shows a terminal window titled "hduser@hadoop-VirtualBox: ~". It has two tabs: "GNU nano 4.8" and "/usr/local/hbase/conf/hbase-site.xml". The current tab displays XML code for hbase-site.xml. The XML includes properties for hbase.tmp.dir, hbase.unsafe.stream.capability.enforce, hbase.rootdir, hbase.cluster.distributed, hbase.zookeeper.property.clientPort, hbase.zookeeper.property.dataDir, and a closing

```
GNU nano 4.8                               /usr/local/hbase/conf/hbase-site.xml
<property>
  <name>hbase.tmp.dir</name>
  <value>./tmp</value>
</property>
<property>
  <name>hbase.unsafe.stream.capability.enforce</name>
  <value>false</value>
</property>
<property>
  <name>hbase.rootdir</name>
  <value>hdfs://localhost:54310/hbase</value>
</property>

<property>
  <name>hbase.cluster.distributed</name>
  <value>true</value>
</property>

<property>
  <name>hbase.zookeeper.property.clientPort</name>
  <value>2181</value>
</property>

<property>
  <name>hbase.zookeeper.property.dataDir</name>
  <value>/home/hduser/zookeeper</value>
</property>
</configuration>
```

All the configuration are done and hbase is completely setup on the system.

3. Start the dfs and the hbase now and check what all are present in the system with jps.

```
hduser@hadoop-VirtualBox:/usr/local/hbase$ start-dfs.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denoted in a future release
22/03/11 21:24:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
Enter passphrase for key '/home/hduser/.ssh/id_rsa':
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-namenode-hadoop-VirtualBox.out
Enter passphrase for key '/home/hduser/.ssh/id_rsa':
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-hadoop-VirtualBox.out
localhost: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hduser-secondarynamenode-hadoop-VirtualBox.out

hduser@hadoop-VirtualBox:/usr/local/hbase$ jps
2272 NameNode
2804 Jps
2471 DataNode
2684 SecondaryNameNode
```

```
hduser@hadoop-VirtualBox:/usr/local/hbase$ start-hbase.sh
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/client-facing-thirdparty/slf4j-log4j12-1.7.30.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]
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/client-facing-thirdparty/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]

hduser@hadoop-VirtualBox:/usr/local/hbase$ jps
2272 NameNode
3186 HQuorumPeer
3302 HMaster
2471 DataNode
3704 Jps
3482 HRegionServer
2684 SecondaryNameNode
```

- Create a table.

Create a table in the hbase shell using the create ‘table-name’ ‘column-family-name’

```
hbase(main):001:0> create 'emp', 'details'
Created table emp
Took 2.0782 seconds
=> Hbase::Table - emp
```

- Insert data into the table.

Insert the data into the table with the help of put ‘table-name’ , ‘id’
‘column-family-region’,’column-family-region-value’

```
hbase(main):002:0> put 'emp', '1' , 'details:name', 'Janvi'
Took 0.3245 seconds
hbase(main):003:0> put 'emp', '1' , 'details:address', 'hisar,haryana'
Took 0.0052 seconds
hbase(main):004:0> put 'emp', '1' , 'details:age', '20'
Took 0.0055 seconds
hbase(main):005:0> scan 'emp'
ROW                                COLUMN+CELL
 1                                  column=details:address, timestamp=2022-03-11T22:15:22.441, value=hisar,haryana
 1                                  column=details:age, timestamp=2022-03-11T22:15:46.909, value=20
 1                                  column=details:name, timestamp=2022-03-11T22:14:59.683, value=Janvi
1 row(s)
Took 0.0513 seconds
```

- Split the table.

Splits the table cannot be one after the creation of the table.

```
hbase(main):006:0> create 'emp', 'details', SPLIT=> ['a','b','c']
ERROR: Table already exists: emp!
For usage try 'help "create"'
Took 0.7694 seconds
```

There is a process named pre-splitting. We have to split the table while creating the table only.

```
hbase(main):007:0> create 'student', 'details', SPLIT=> ['a','b','c']
Created table student
Took 2.1871 seconds
=> Hbase::Table - student
```

By this , we can list the number of tables listed .

```
hbase(main):008:0> list
TABLE
emp
student
2 row(s)
Took 0.0546 seconds
=> ["emp", "student"]
```