

Pre-Milestone 2 Monitoring

WQD7005 – DATA MINING

ZULKANAIN BIN HASAN

WQD180031

- Step 1: Verifying JAVA Installation. ...
- Step 2: Verifying Hadoop Installation. ...
- Step 4: Installing Hive. ...
- Step 5: Configuring Hive. ...
- Step 6: Downloading and Installing Apache Derby. ...
- Step 7: Configuring Metastore of Hive. ...
- Step 8: Verifying Hive Installation. ...

Step 1: Verifying JAVA Installation. ...

```
File Edit View Search Terminal Help
(base) zulkanh@elitebook:~$ java -version
openjdk version "11.0.4" 2019-07-16
OpenJDK Runtime Environment (build 11.0.4+11-post-Ubuntu-1ubuntu218.04.3)
OpenJDK 64-Bit Server VM (build 11.0.4+11-post-Ubuntu-1ubuntu218.04.3, mixed mode, sharing)
(base) zulkanh@elitebook:~$
```

File Edit View Search Terminal Help

(base) **zulkanh@elitebook**:~\$ sudo update-alternatives --config java

[sudo] password for zulkanh:

There are 2 choices for the alternative java (providing /usr/bin/java).

Selection	Path	Priority	Status
* 0	/usr/lib/jvm/java-11-openjdk-amd64/bin/java	1101	auto m
ode			
1	/usr/lib/jvm/java-11-openjdk-amd64/bin/java	1101	manual
mode			
2	/usr/lib/jvm/java-8-openjdk-amd64/jre/bin/java	1081	manual
mode			

Press <enter> to keep the current choice[*], or type selection number:

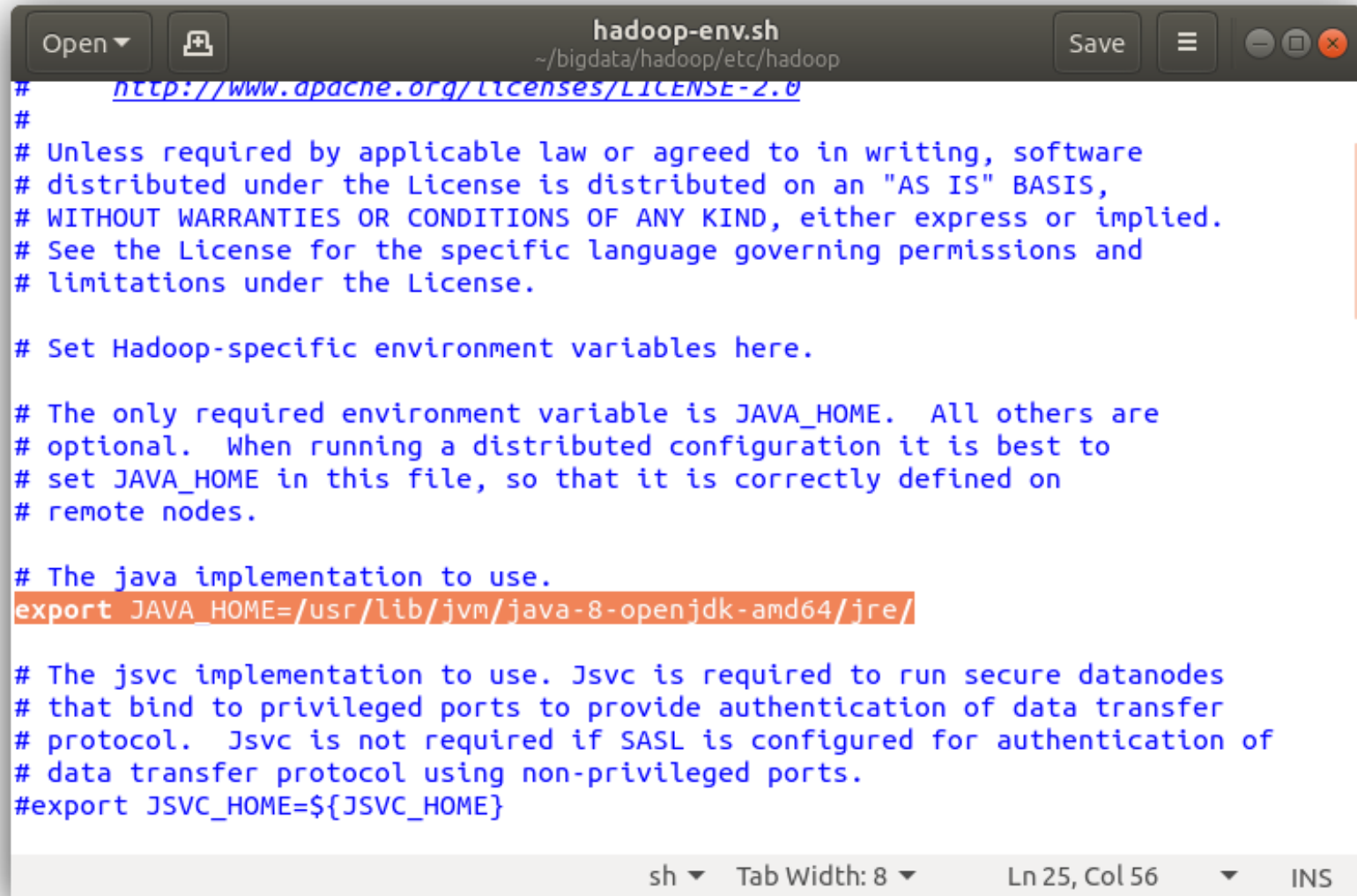
File Edit View Search Terminal Help

(base) **zulkanh@elitebook**:~\$ update-java-alternatives --list

java-1.11.0-openjdk-amd64	1111	/usr/lib/jvm/java-1.11.0-openjdk-amd64
---------------------------	------	--

java-1.8.0-openjdk-amd64	1081	/usr/lib/jvm/java-1.8.0-openjdk-amd64
--------------------------	------	---------------------------------------

(base) **zulkanh@elitebook**:~\$ █



```
hadoop-env.sh
~/bigdata/hadoop/etc/hadoop

# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use.
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre/

# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
# protocol. Jsvc is not required if SASL is configured for authentication of
# data transfer protocol using non-privileged ports.
#export JSVC_HOME=${JSVC_HOME}
```

sh ▾ Tab Width: 8 ▾ Ln 25, Col 56 ▾ INS

Step 2: Verifying Hadoop Installation. ...

```
File Edit View Search Terminal Help
(base) zulkanh@elitebook:~$ java -version
openjdk version "11.0.4" 2019-07-16
OpenJDK Runtime Environment (build 11.0.4+11-post-Ubuntu-1ubuntu218.04.3)
OpenJDK 64-Bit Server VM (build 11.0.4+11-post-Ubuntu-1ubuntu218.04.3, mixed mode, sharing)
(base) zulkanh@elitebook:~$ hadoop version
Hadoop 2.7.7
Subversion Unknown -r c1aad84bd27cd79c3d1a7dd58202a8c3ee1ed3ac
Compiled by stevel on 2018-07-18T22:47Z
Compiled with protoc 2.5.0
From source with checksum 792e15d20b12c74bd6f19a1fb886490
This command was run using /home/zulkanh/bigdata/hadoop/share/hadoop/common/hadoop-common-2.7.7.jar
(base) zulkanh@elitebook:~$
```



The image shows a text editor window titled ".bashrc" with a file icon and a "Save" button. The editor contains a bash script for setting up environment variables. The script includes an "else" block for updating the PATH, followed by "fi" and "unset __conda_setup". A comment "# <<< conda init <<<" is present. Below this, several sections are highlighted with orange backgrounds, each containing a comment and an "export" command to update the PATH. These sections are for "bigdata/hadoop", "bigdata/sqoop", "bigdata/hbase", "bigdata/hive", and "bigdata/pig". The "bigdata/pig" section also includes an "export JAVA_HOME" command. The status bar at the bottom indicates "sh", "Tab Width: 8", "Ln 138, Col 52", and "INS".

```
else
    \export PATH="/home/zulkanh/anaconda3/bin:$PATH"
fi
unset __conda_setup
# <<< conda init <<<

# bigdata/hadoop
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/bin
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/sbin

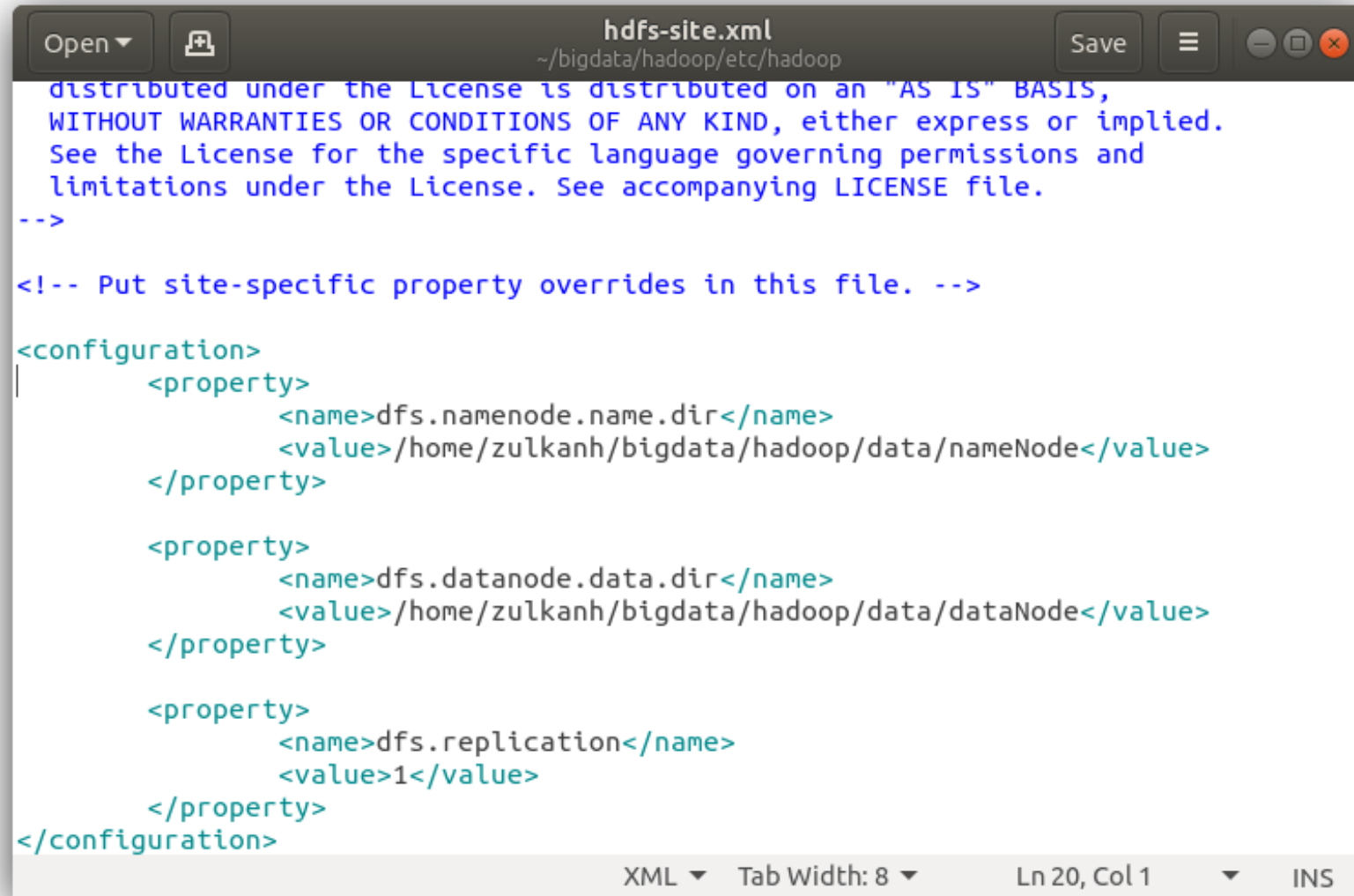
# bigdata/sqoop
export PATH=$PATH:/home/zulkanh/bigdata/sqoop/bin

# bigdata/hbase
export PATH=$PATH:/home/zulkanh/bigdata/hbase/bin

#bigdata/hive
export PATH=$PATH:/home/zulkanh/bigdata/hive/bin

#bigdata/pig
export PATH=$PATH:/home/zulkanh/bigdata/pig/bin
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre/
```

sh Tab Width: 8 Ln 138, Col 52 INS



```
hdfs-site.xml
~/bigdata/hadoop/etc/hadoop

distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

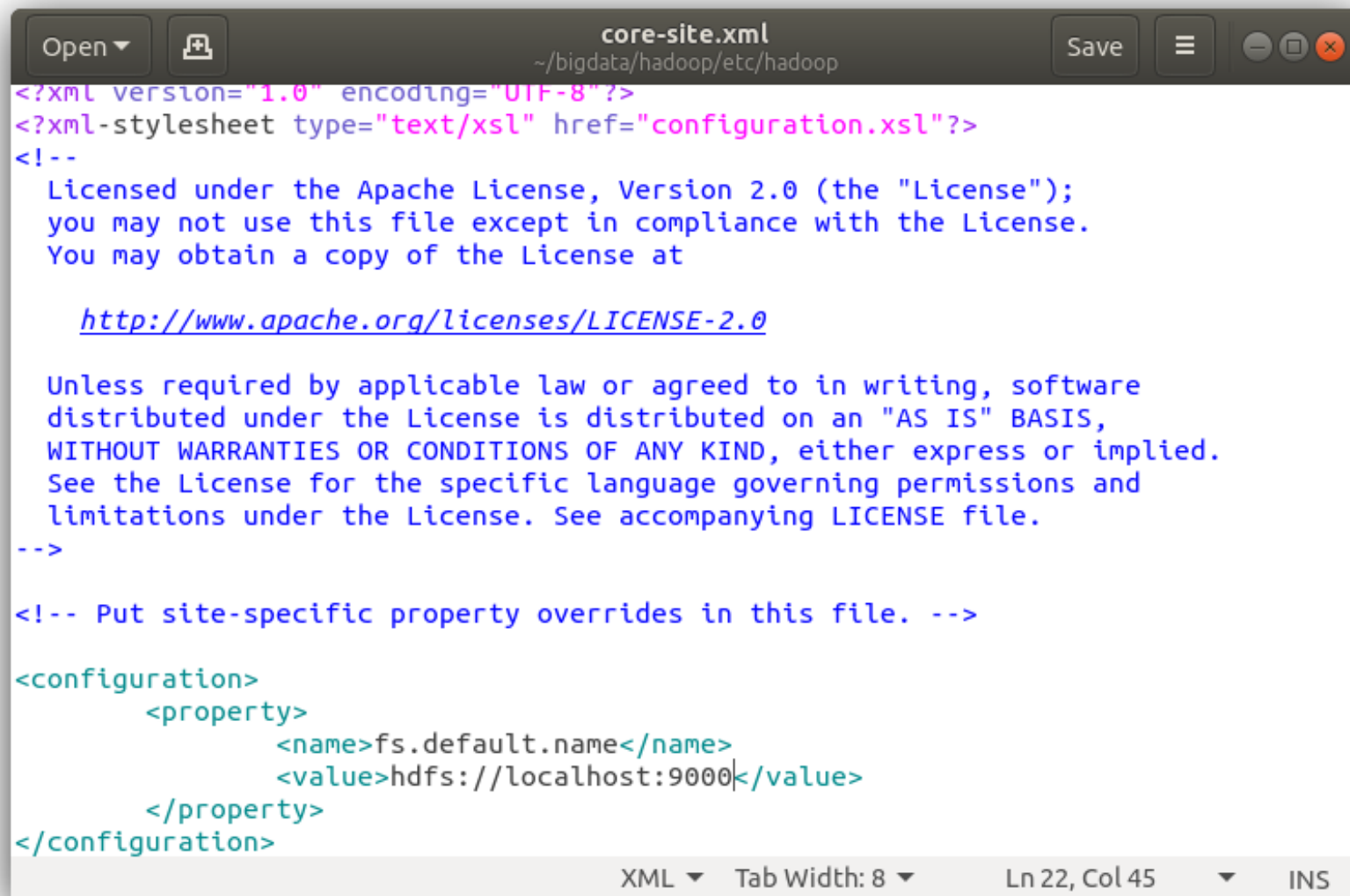
<!-- Put site-specific property overrides in this file. -->

<configuration>
|   <property>
        <name>dfs.namenode.name.dir</name>
        <value>/home/zulkanh/bigdata/hadoop/data/nameNode</value>
    </property>

    <property>
        <name>dfs.datanode.data.dir</name>
        <value>/home/zulkanh/bigdata/hadoop/data/dataNode</value>
    </property>

    <property>
        <name>dfs.replication</name>
        <value>1</value>
    </property>
</configuration>

XML ▾ Tab Width: 8 ▾ Ln 20, Col 1 ▾ INS
```

```
core-site.xml
~/bigdata/hadoop/etc/hadoop

Open Save
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

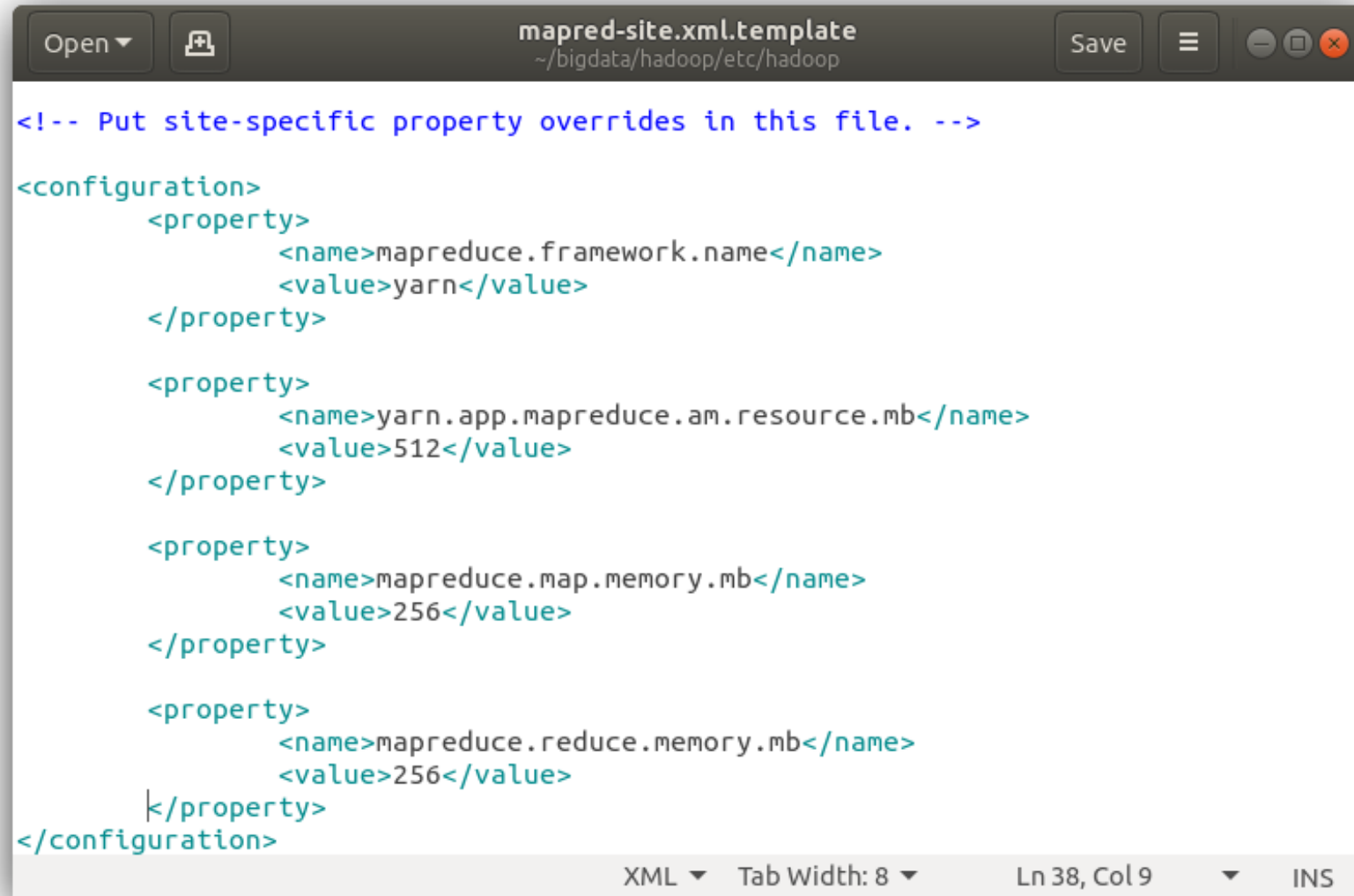
    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

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

XML Tab Width: 8 Ln 22, Col 45 INS



```
<!-- Put site-specific property overrides in this file. -->

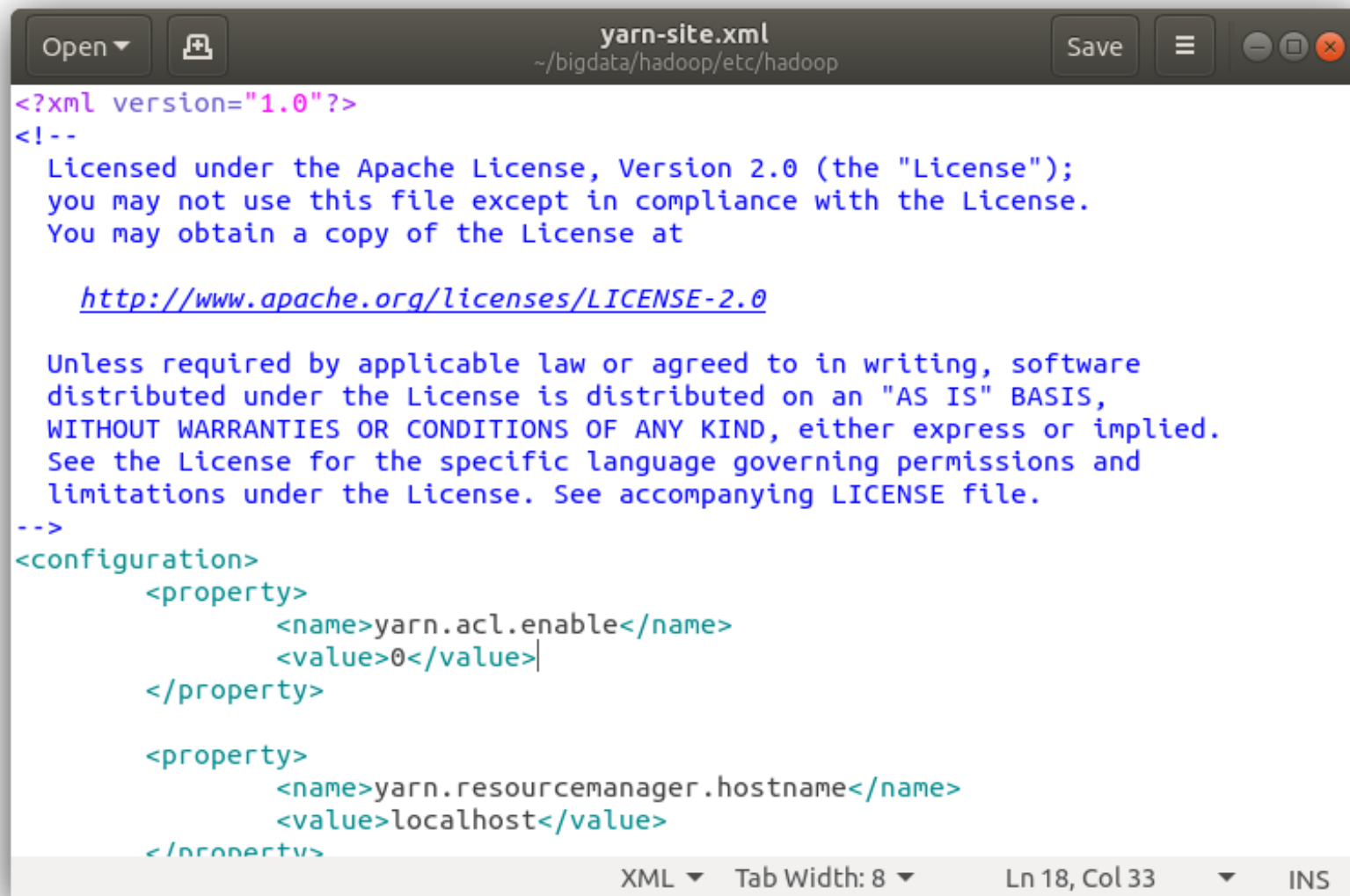
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>

  <property>
    <name>yarn.app.mapreduce.am.resource.mb</name>
    <value>512</value>
  </property>





  <property>
    <name>mapreduce.map.memory.mb</name>
    <value>256</value>
  </property>

  <property>
    <name>mapreduce.reduce.memory.mb</name>
    <value>256</value>
  </property>
</configuration>
```

XML ▾ Tab Width: 8 ▾ Ln 38, Col 9 ▾ INS



The image shows a code editor window with the title bar 'yarn-site.xml' and the file path '~/.bigdata/hadoop/etc/hadoop'. The editor contains XML code with a license notice and configuration properties. The license notice is enclosed in a comment block. The configuration section defines two properties: 'yarn.acl.enable' with a value of '0' and 'yarn.resourcemanager.hostname' with a value of 'localhost'. The status bar at the bottom indicates the file is in XML format, with a tab width of 8, and the cursor is at line 18, column 33 in insert mode.

```
Open ▾  yarn-site.xml ~/.bigdata/hadoop/etc/hadoop Save   
```

```
<?xml version="1.0"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<configuration>
  <property>
    <name>yarn.acl.enable</name>
    <value>0</value>
  </property>

  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>localhost</value>
  </property>
</configuration>
```

XML ▾ Tab Width: 8 ▾ Ln 18, Col 33 ▾ INS

File Edit View Search Terminal Help

```
(base) zulkanh@elitebook:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [localhost]
zulkanh@localhost's password:
localhost: starting namenode, logging to /home/zulkanh/bigdata/hadoop/logs/hadoop-zulkanh-namenode-elitebook.out
zulkanh@localhost's password:
localhost: starting datanode, logging to /home/zulkanh/bigdata/hadoop/logs/hadoop-zulkanh-datanode-elitebook.out
Starting secondary namenodes [0.0.0.0]
zulkanh@0.0.0.0's password:
0.0.0.0: starting secondarynamenode, logging to /home/zulkanh/bigdata/hadoop/logs/hadoop-zulkanh-secondarynamenode-elitebook.out
starting yarn daemons
starting resourcemanager, logging to /home/zulkanh/bigdata/hadoop/logs/yarn-zulkanh-resourcemanager-elitebook.out
zulkanh@localhost's password:
localhost: starting nodemanager, logging to /home/zulkanh/bigdata/hadoop/logs/yarn-zulkanh-nodemanager-elitebook.out
(base) zulkanh@elitebook:~$ jps
12273 SecondaryNameNode
11921 NameNode
12436 ResourceManager
12644 Jps
7576 RunJar
12601 NodeManager
12079 DataNode
(base) zulkanh@elitebook:~$
```

g

Bookmarks x Assignment x Hive-Instal x Apache Der x localhost x All Applicat x Namenode x SecondaryN x +

< > C | localhost:50070/dfshealth.html#tab-overview

Bookmarks Ba... BDM Google Sheets SAS@ Logon... Login | Tablea...

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Overview 'localhost:9000' (active)

Started:	Sun Oct 06 07:56:23 MYT 2019
Version:	2.7.7, rc1aad84bd27cd79c3d1a7dd58202a8c3ee1ed3ac
Compiled:	2018-07-18T22:47Z by stevel from branch-2.7.7
Cluster ID:	CID-5a298543-686a-4079-8bdb-8c7216807f9a
Block Pool ID:	BP-763754000-192.168.0.6-1555601472303

Summary

Security is off.

Safemode is off.

Hadoop Overview

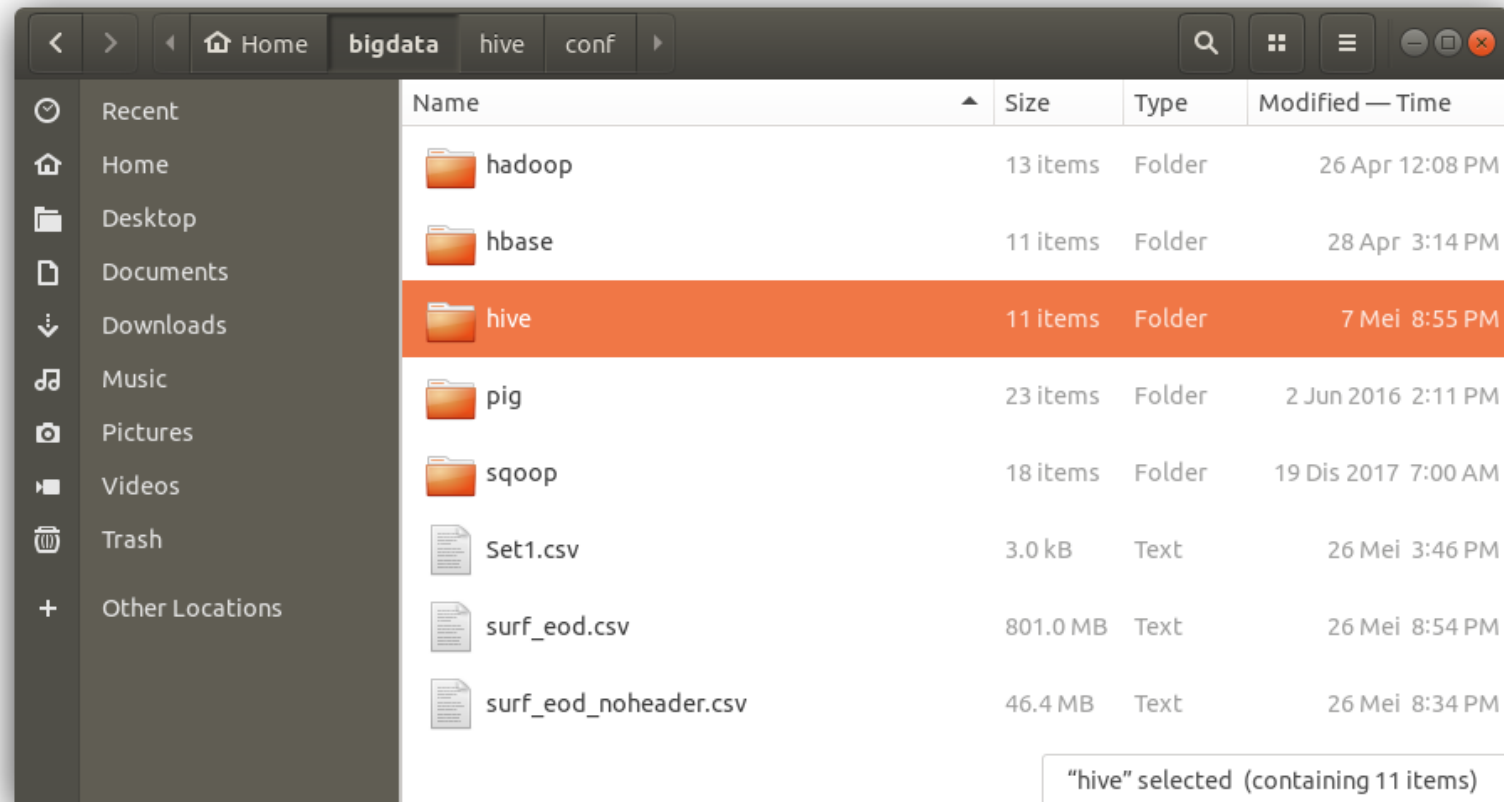
Overview

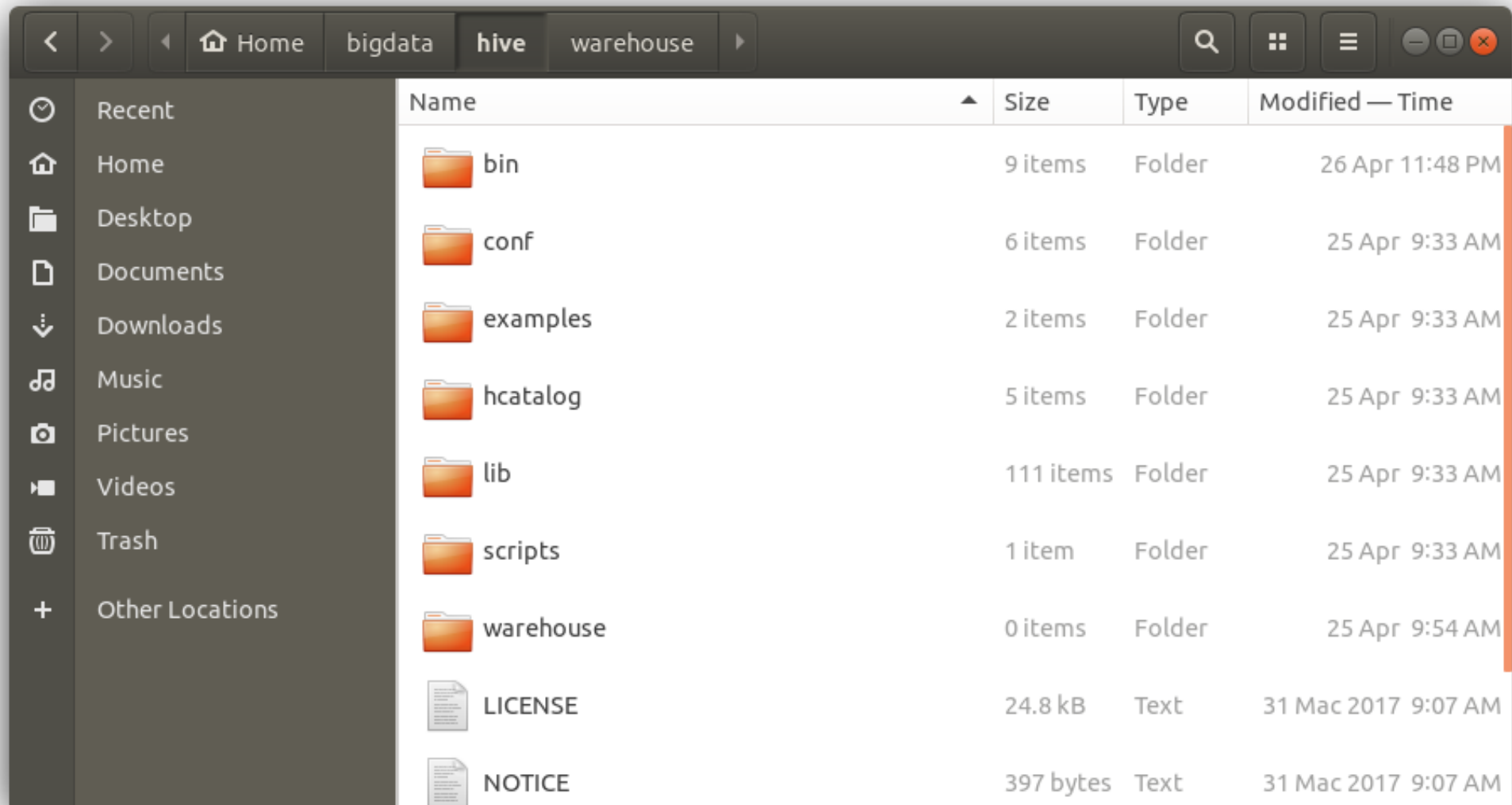
Version	2.7.7
Compiled	2018-07-18T22:47Z by stevel from branch-2.7.7
NameNode Address	localhost:9000
Started	10/6/2019, 7:55:31 AM
Last Checkpoint	Never
Checkpoint Period	3600 seconds
Checkpoint Transactions	1000000

Checkpoint Image URI



Step 4: Installing Hive. ...





Step 5: Configuring Hive. ...



```
hive-config.sh
~/bigdata/hive/bin

HIVE_CONF_DIR=$confdir
;;
--auxpath)
shift
HIVE_AUX_JARS_PATH=$1
shift
;;
*)
break;
;;
esac
done

# Allow alternate conf dir location.
HIVE_CONF_DIR="${HIVE_CONF_DIR:-$HIVE_HOME/conf}"

export HIVE_CONF_DIR=$HIVE_CONF_DIR
export HIVE_AUX_JARS_PATH=$HIVE_AUX_JARS_PATH

# Default to use 256MB
export HADOOP_HEAPSIZE=${HADOOP_HEAPSIZE:-256}

export HADOOP_HOME=/home/zulkanh/bigdata/hadoop
```

sh Tab Width: 8 Ln 72, Col 48 INS



```
else
    \export PATH="/home/zulkanh/anaconda3/bin:$PATH"
fi
unset __conda_setup
# <<< conda init <<<

# bigdata/hadoop
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/bin
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/sbin

# bigdata/sqoop
export PATH=$PATH:/home/zulkanh/bigdata/sqoop/bin

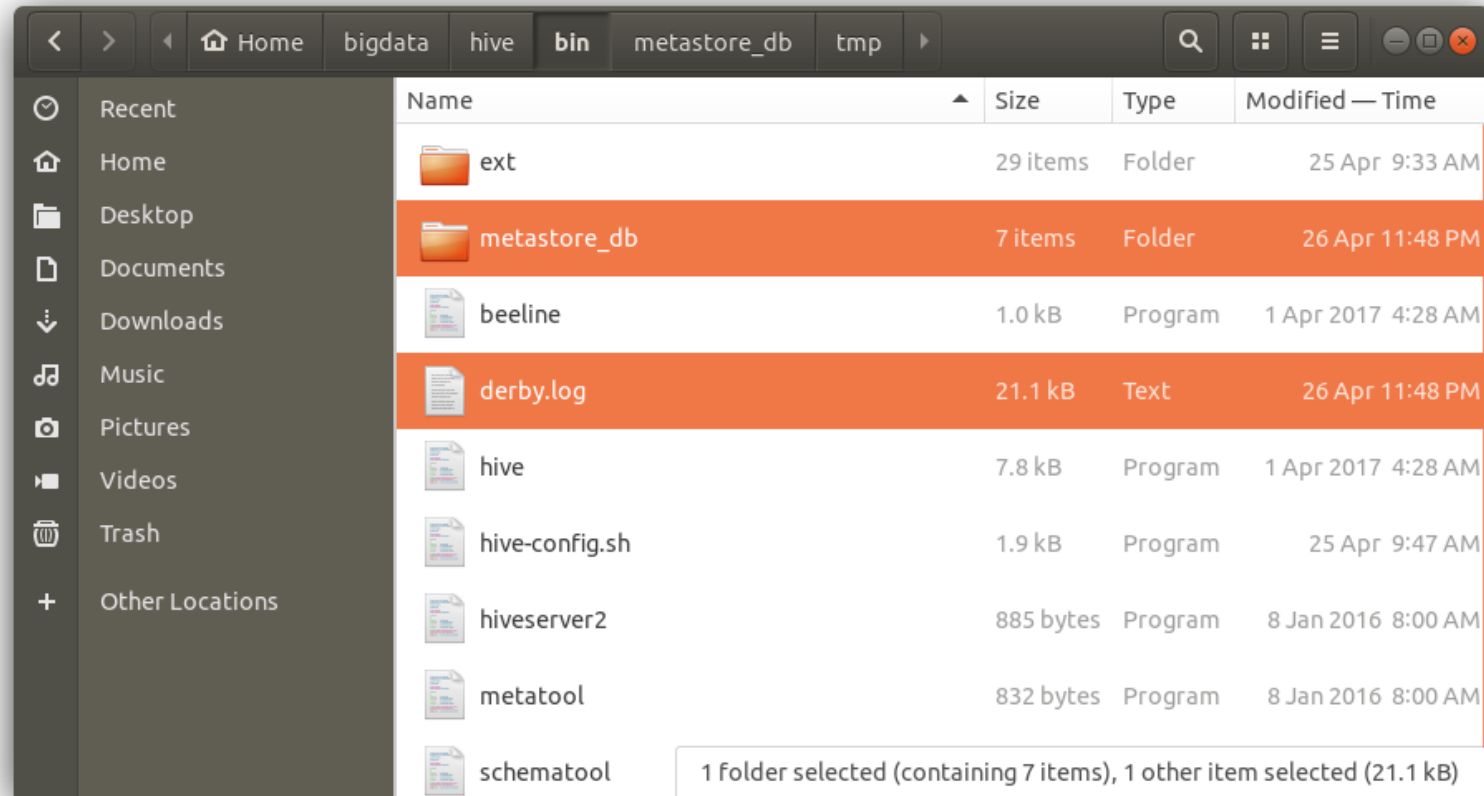
# bigdata/hbase
export PATH=$PATH:/home/zulkanh/bigdata/hbase/bin

#bigdata/hive
export PATH=$PATH:/home/zulkanh/bigdata/hive/bin

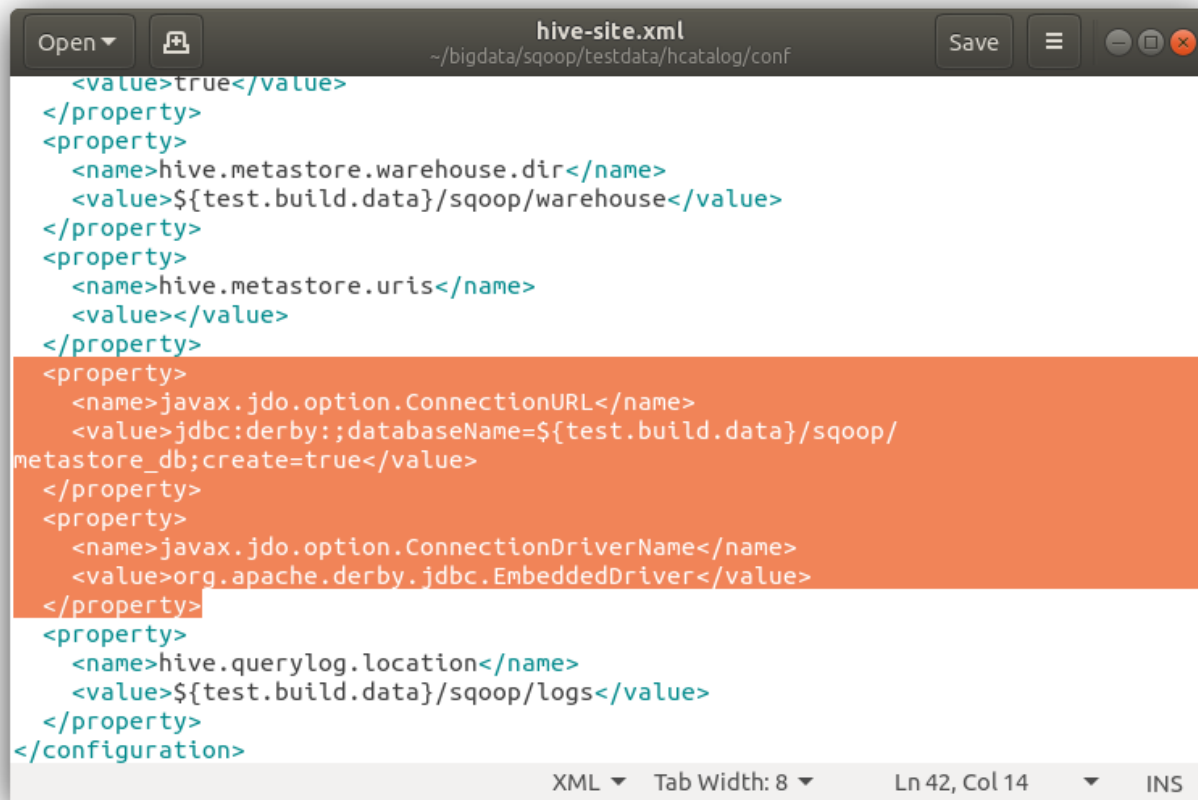
#bigdata/pig
export PATH=$PATH:/home/zulkanh/bigdata/pig/bin
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre/
```

sh ▾ Tab Width: 8 ▾ Ln 147, Col 49 ▾ INS

Step 6: Downloading and Installing Apache Derby. ...



Step 7: Configuring Metastore of Hive. ...



The screenshot shows a code editor window titled "hive-site.xml" with the file path "~/bigdata/sqoop/testdata/hcatalog/conf". The editor contains XML configuration for the Hive metastore. A section of the code is highlighted in orange, showing the JDBC connection details for the metastore database.

```
<value>true</value>
</property>
<property>
  <name>hive.metastore.warehouse.dir</name>
  <value>${test.build.data}/sqoop/warehouse</value>
</property>
<property>
  <name>hive.metastore.uris</name>
  <value></value>
</property>
<property>
  <name>javax.jdo.option.ConnectionURL</name>
  <value>jdbc:derby:::databaseName=${test.build.data}/sqoop/
metastore_db;create=true</value>
</property>
<property>
  <name>javax.jdo.option.ConnectionDriverName</name>
  <value>org.apache.derby.jdbc.EmbeddedDriver</value>
</property>
<property>
  <name>hive.querylog.location</name>
  <value>${test.build.data}/sqoop/logs</value>
</property>
</configuration>
```

The status bar at the bottom indicates "XML", "Tab Width: 8", "Ln 42, Col 14", and "INS".

Step 8: Verifying Hive Installation. ...

```
File Edit View Search Terminal Help
(base) zulkanh@elitebook:~$ hive;

Logging initialized using configuration in jar:file:/home/zulkanh/bigdata/hive/lib/hive-common-1.2.2.jar!/hive-log4j.properties
hive> create database wqd7005;
OK
Time taken: 1.095 seconds
hive> show databases;
OK
default
wqd7005
wqd7007
Time taken: 0.315 seconds, Fetched: 3 row(s)
hive> 
```