Pre-Milestone 2 Monitoring

WQD7005 – DATA MINING

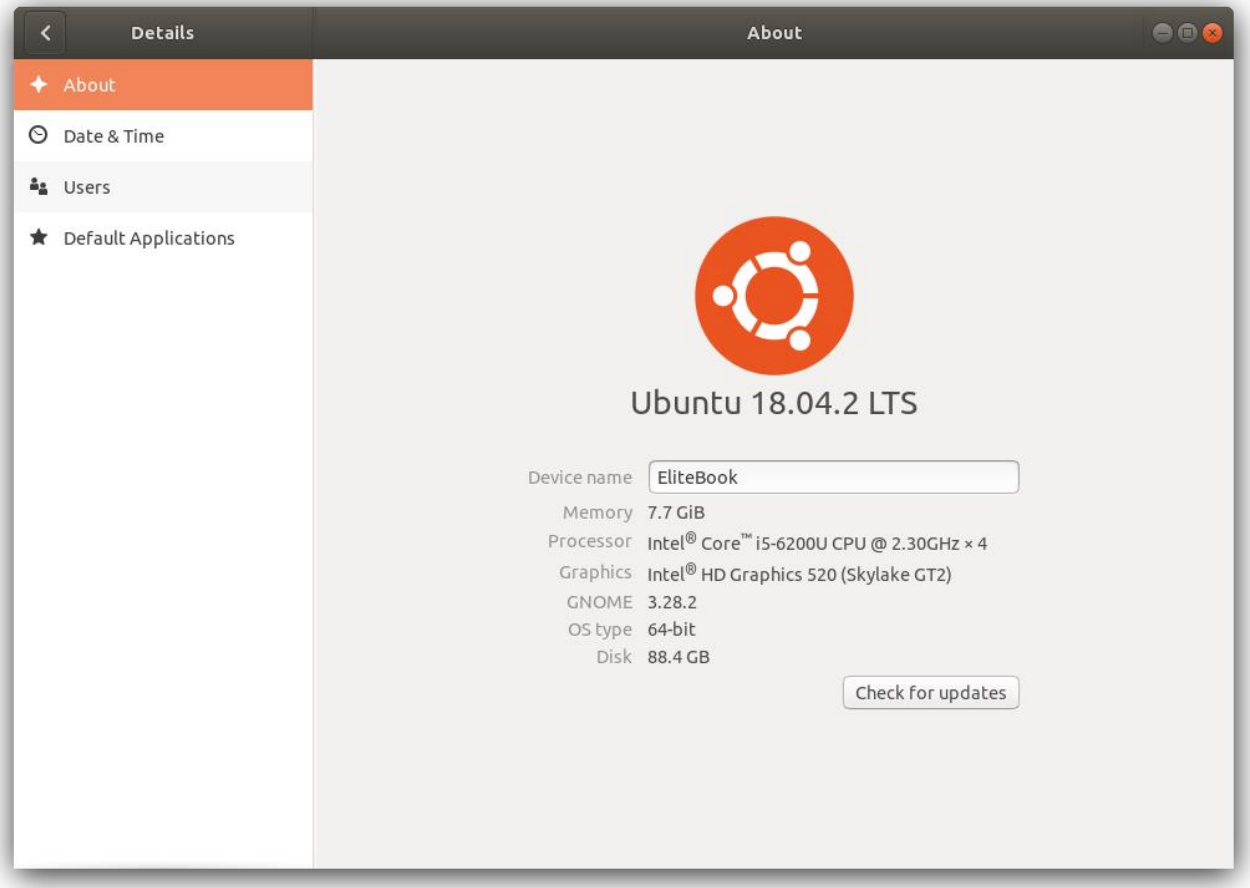
ZULKANAIN BIN HASAN

WQD180031

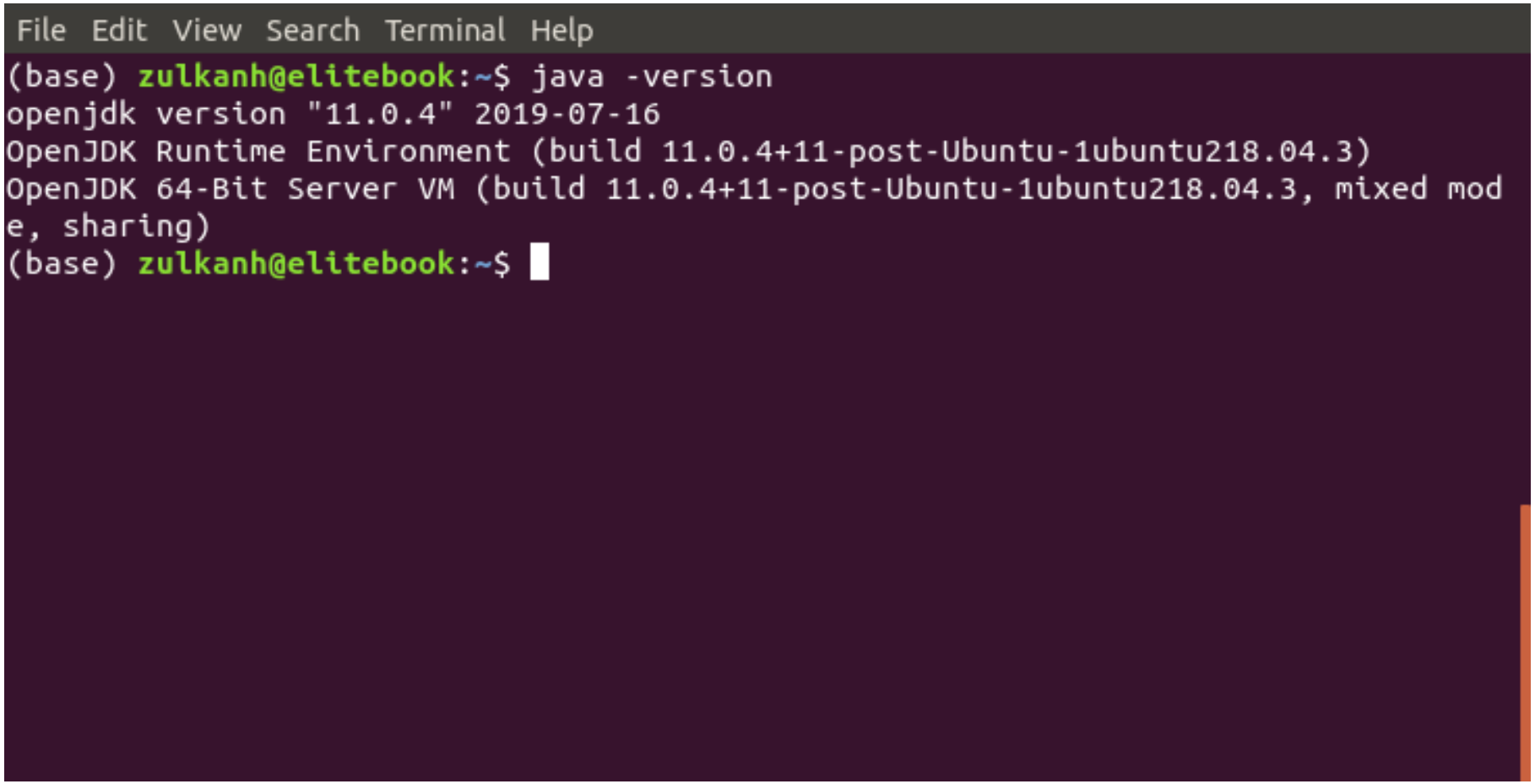
- Step 1: Verifying JAVA Installation. ...
- Step 2: Verifying Hadoop Installation. ...
- Step 4: Installing Hive. ...
- Step 5: Configuring Hive. ...
- Step 6: Downloading and Installing Apache Derby. ...
- Step 7: Configuring Metastore of Hive. ...
- Step 8: Verifying Hive Installation. ...

Computer Specification. ...

- I am running Hadoop using Ubuntu OS.
- In order to install Hadoop, I set my Laptop to dual boot OS (Win 10 and Ubuntu 18.04.2 LTS).

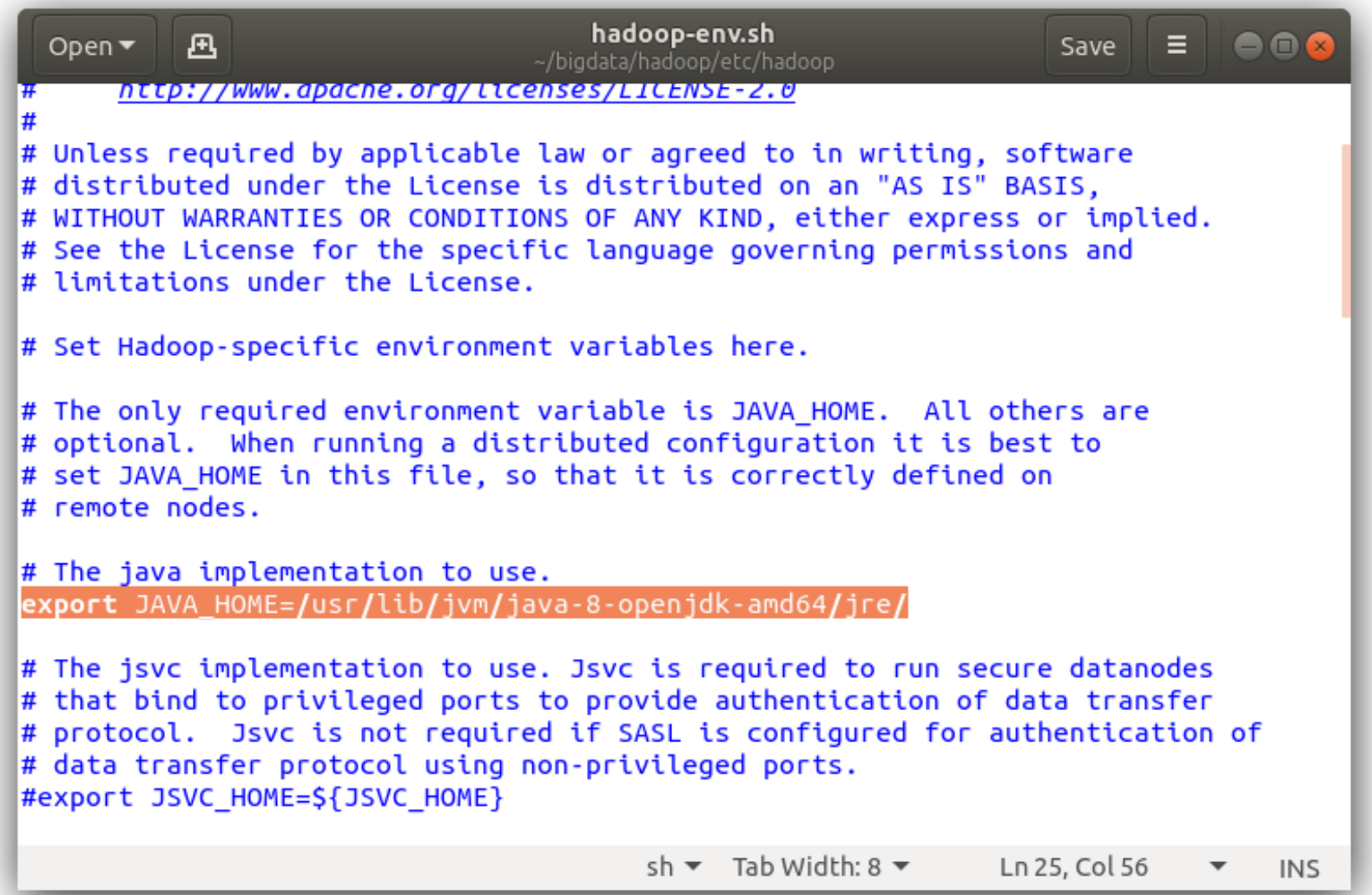


Step 1: Verifying JAVA Installation. ...

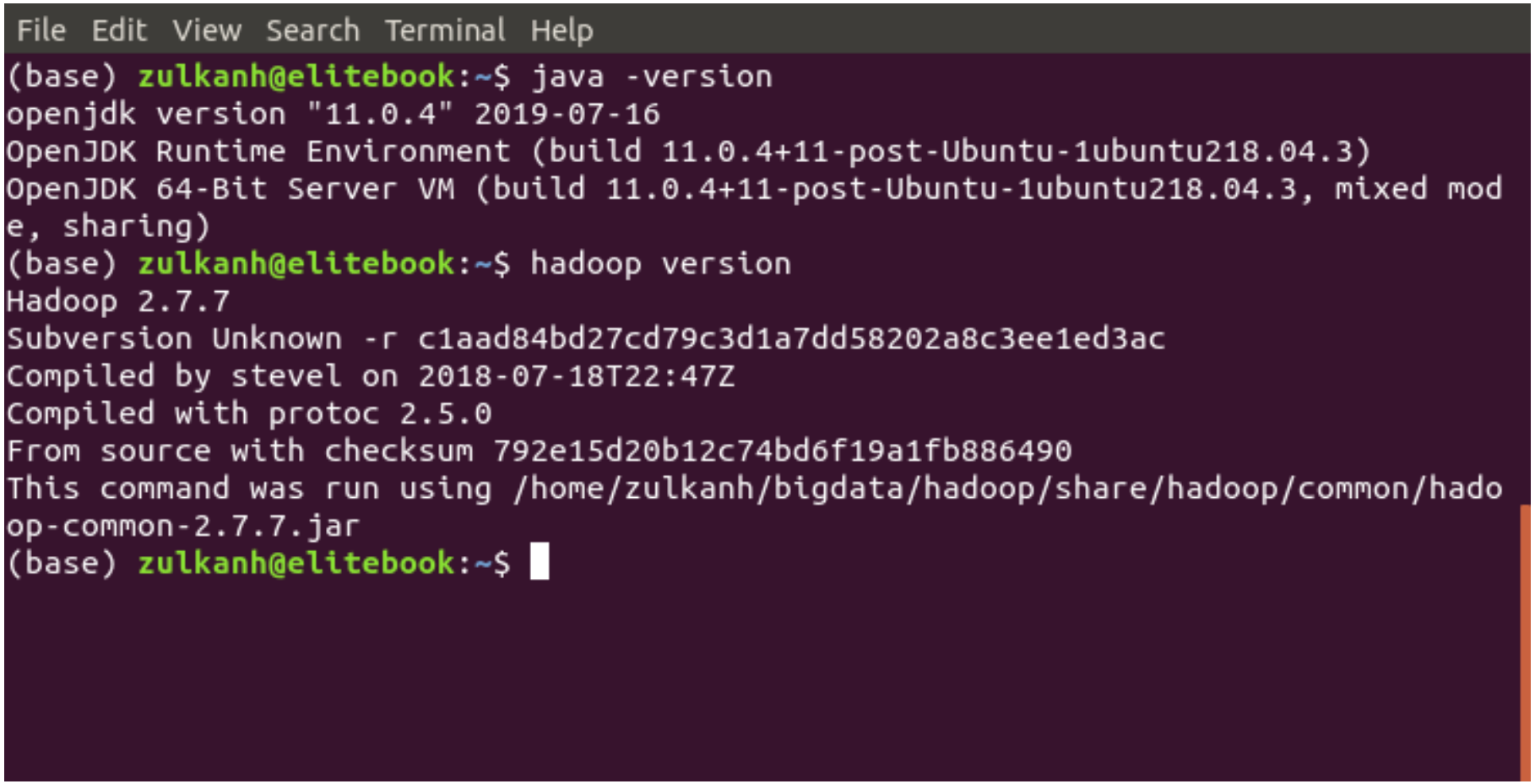


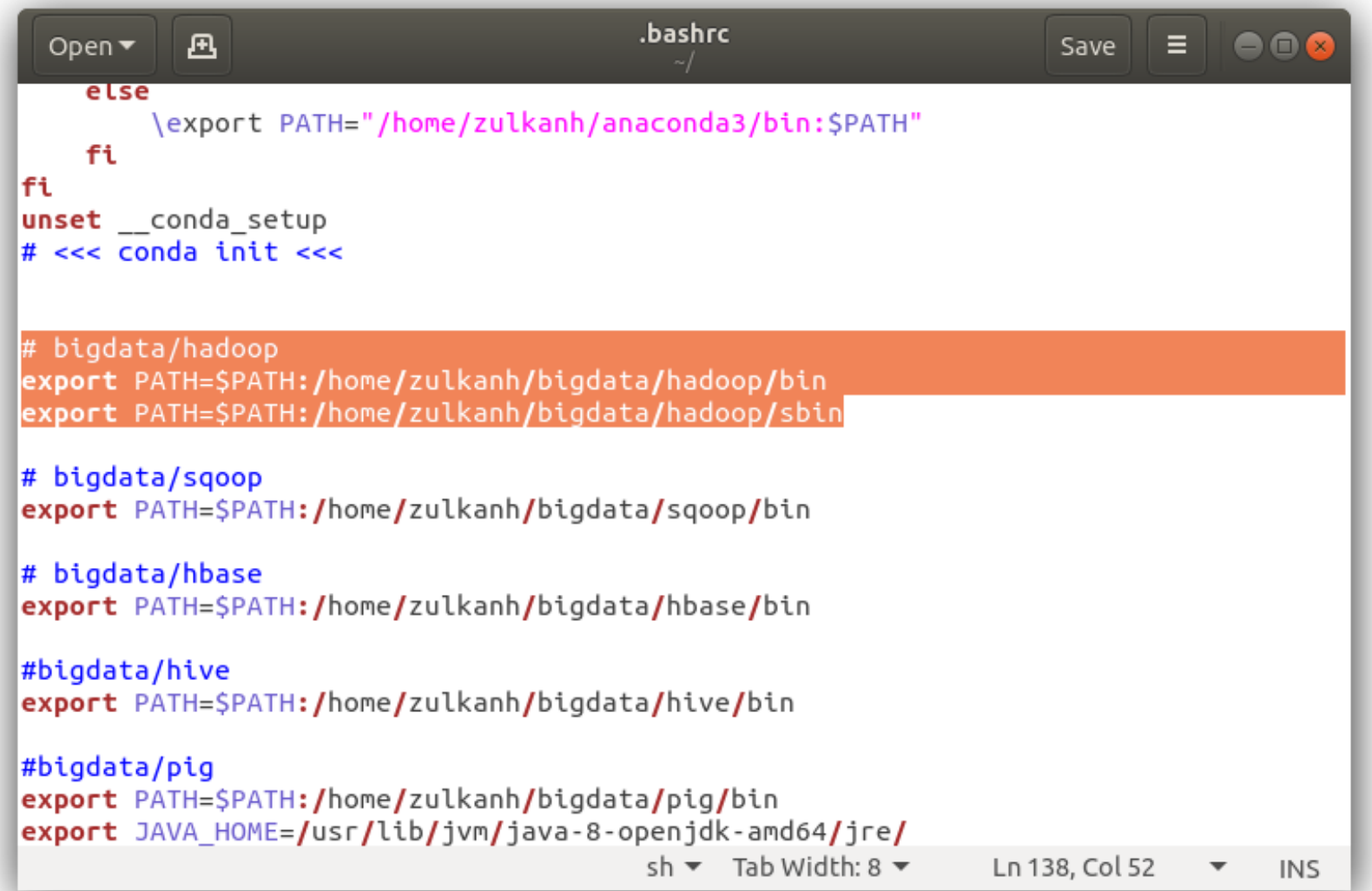
```
File  Edit  View  Search  Terminal  Help
(base) zulkanh@elitebook:~$ sudo update-alternatives --config java
[sudo] password for zulkanh:
There are 2 choices for the alternative java (providing /usr/bin/java).

  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:~$
```


Open
hadoop-env.sh
~/bigdata/hadoop/etc/hadoop
Save
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. ...

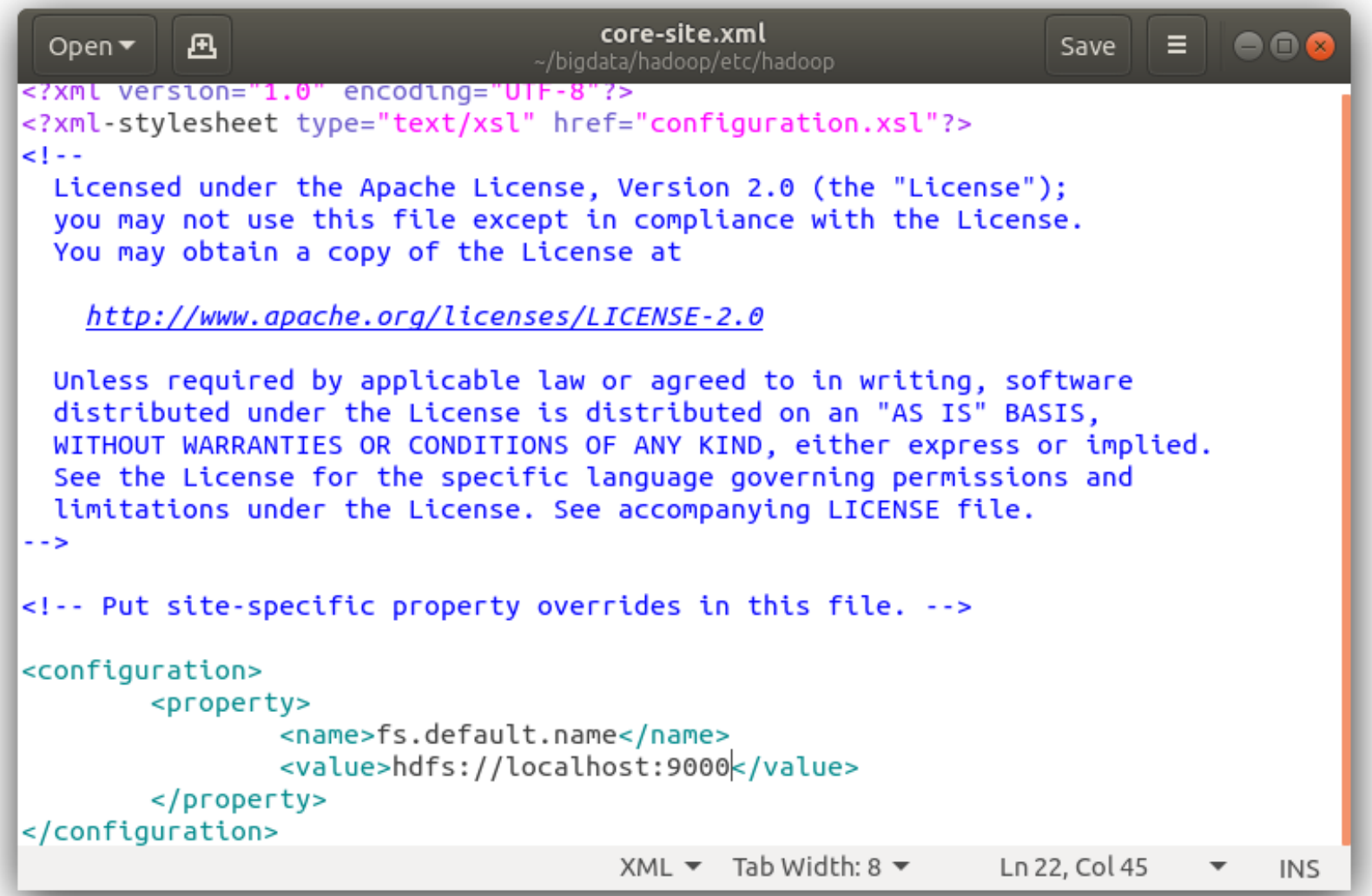


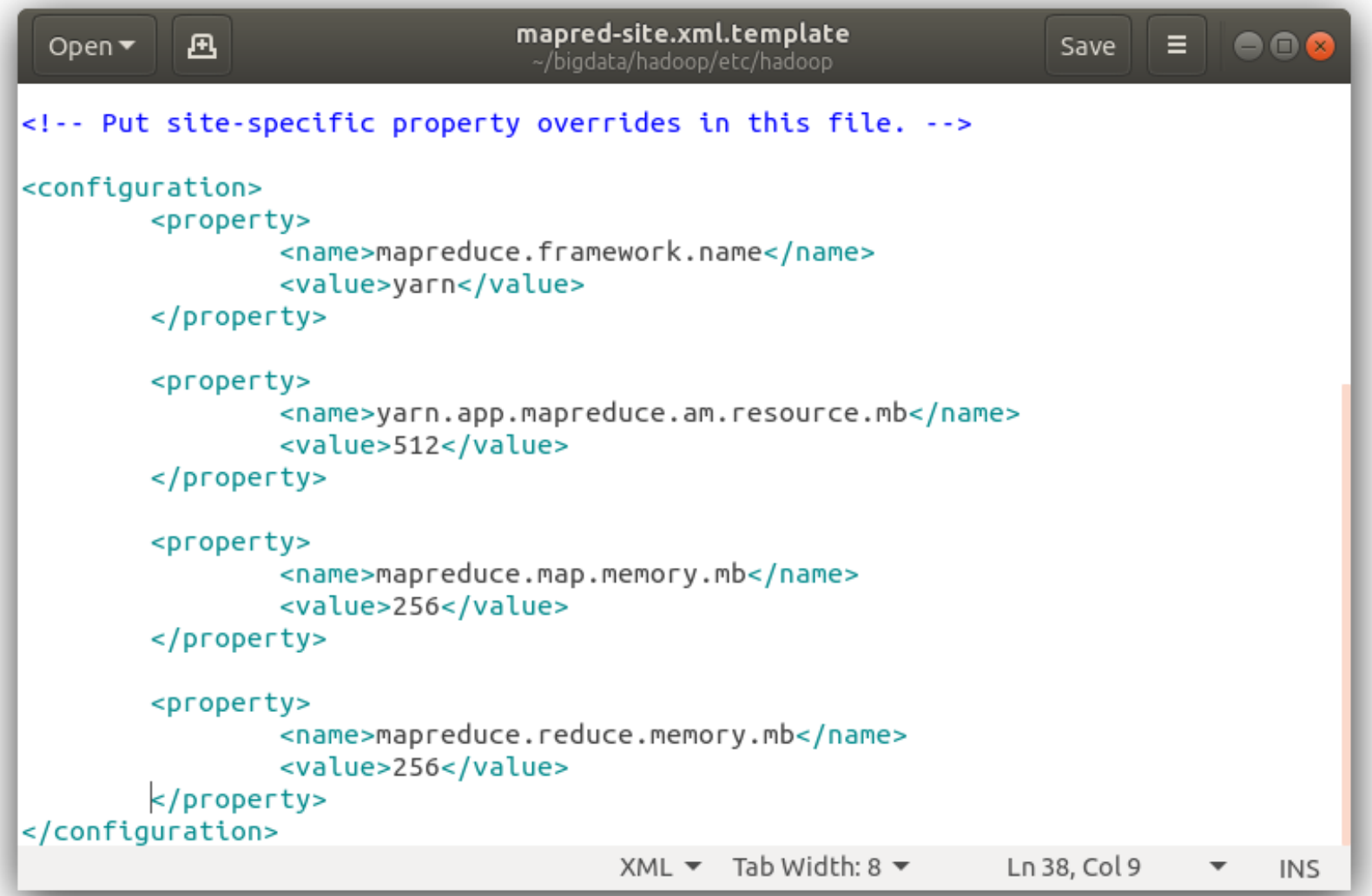

Open
.bashrc
~/
Save
else
\export PATH="/home/zulkanh/anaconda3/bin:$PATH"
fi
fi
unset __conda_setup
<<< conda init <<<
bigdata/hadoop
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/bin
export PATH=$PATH:/home/zulkanh/bigdata/hadoop/sbin
bigdata/sqoop
export PATH=$PATH:/home/zulkanh/bigdata/sqoop/bin
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

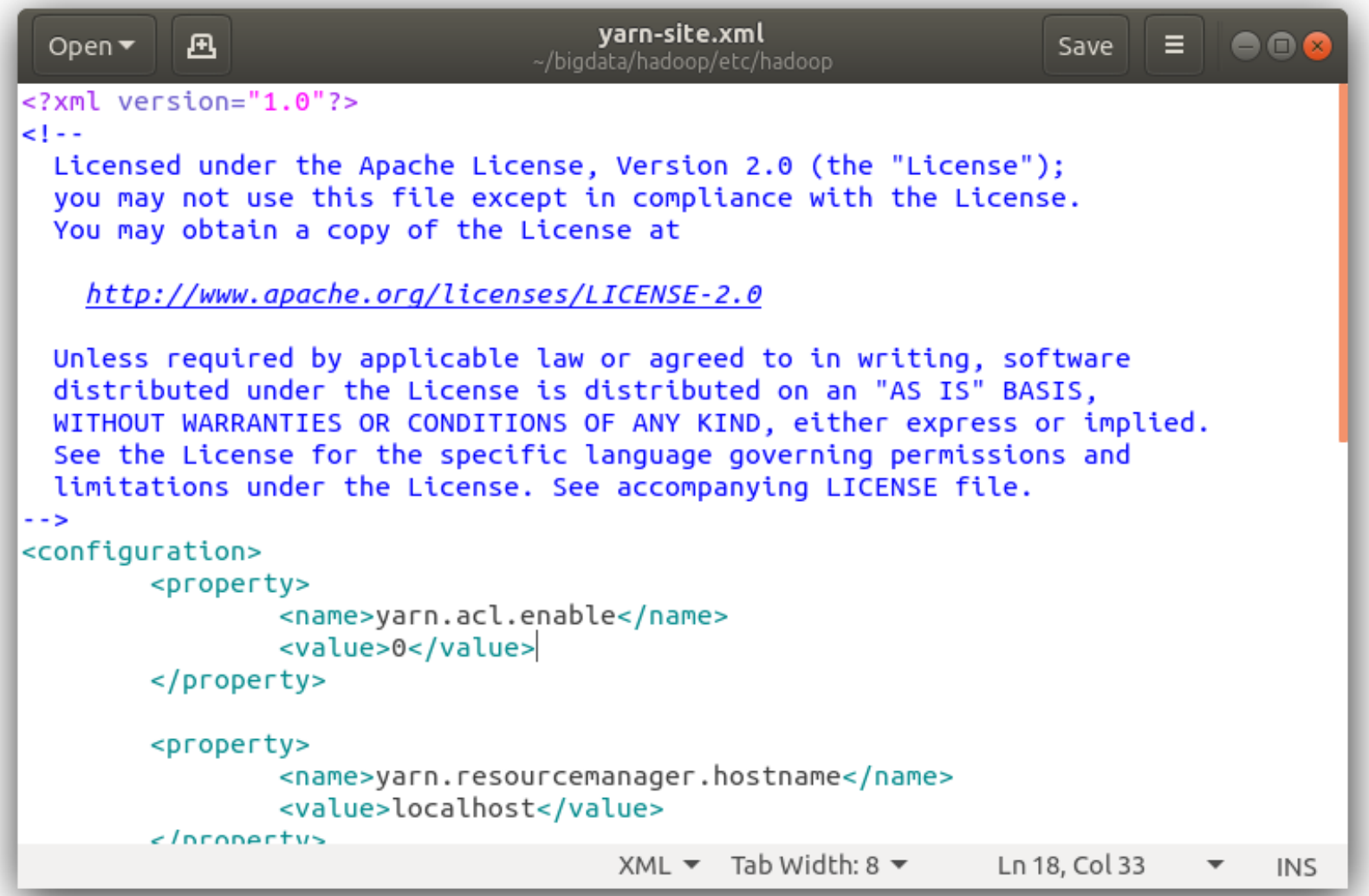
Open | hdfs-site.xml | ~/bigdata/hadoop/etc/hadoop | Save

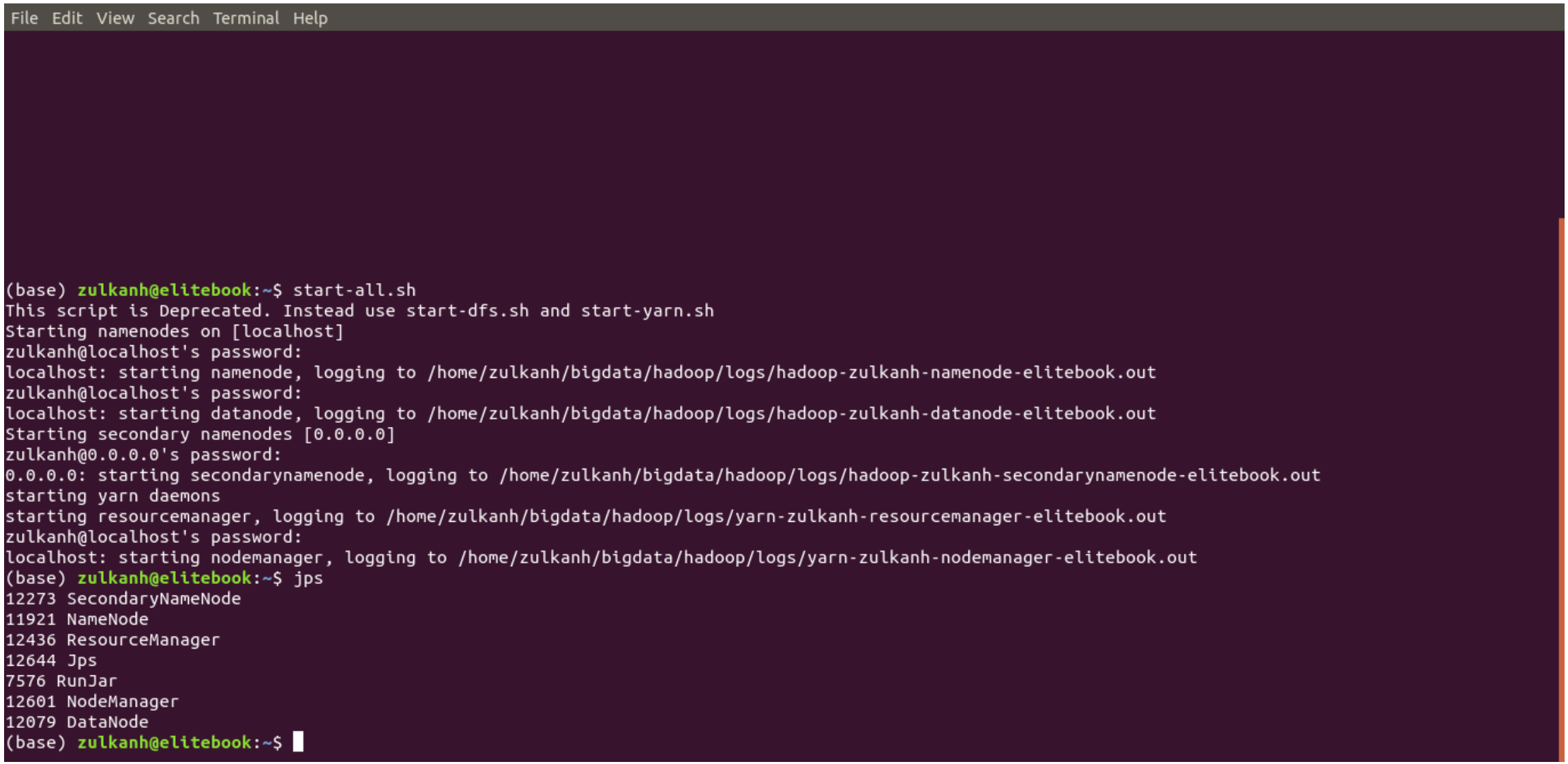
```
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- 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>
```

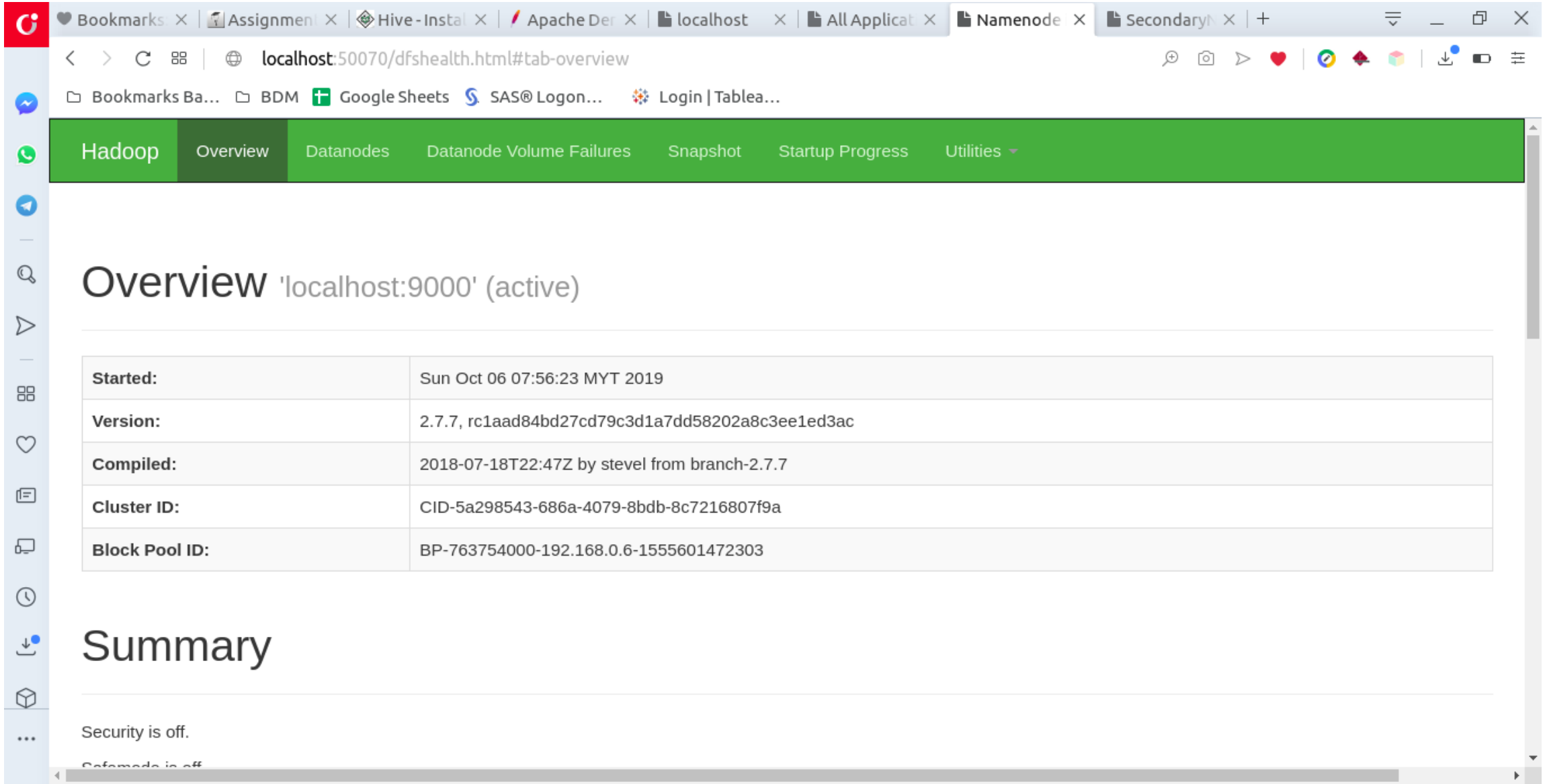
XML | Tab Width: 8 | Ln 20, Col 1 | INS

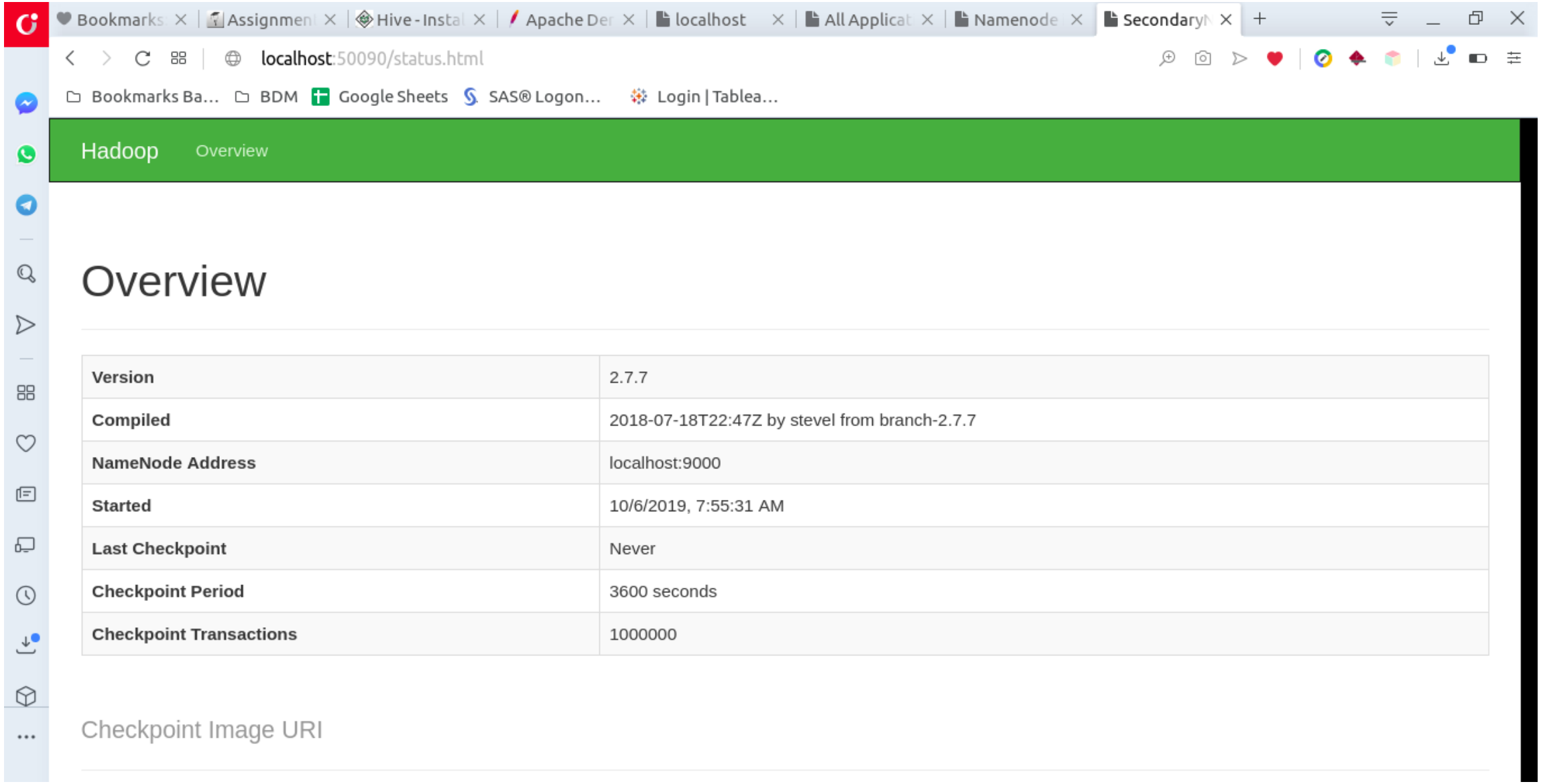

Open
core-site.xml
~/bigdata/hadoop/etc/hadoop
Save
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a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

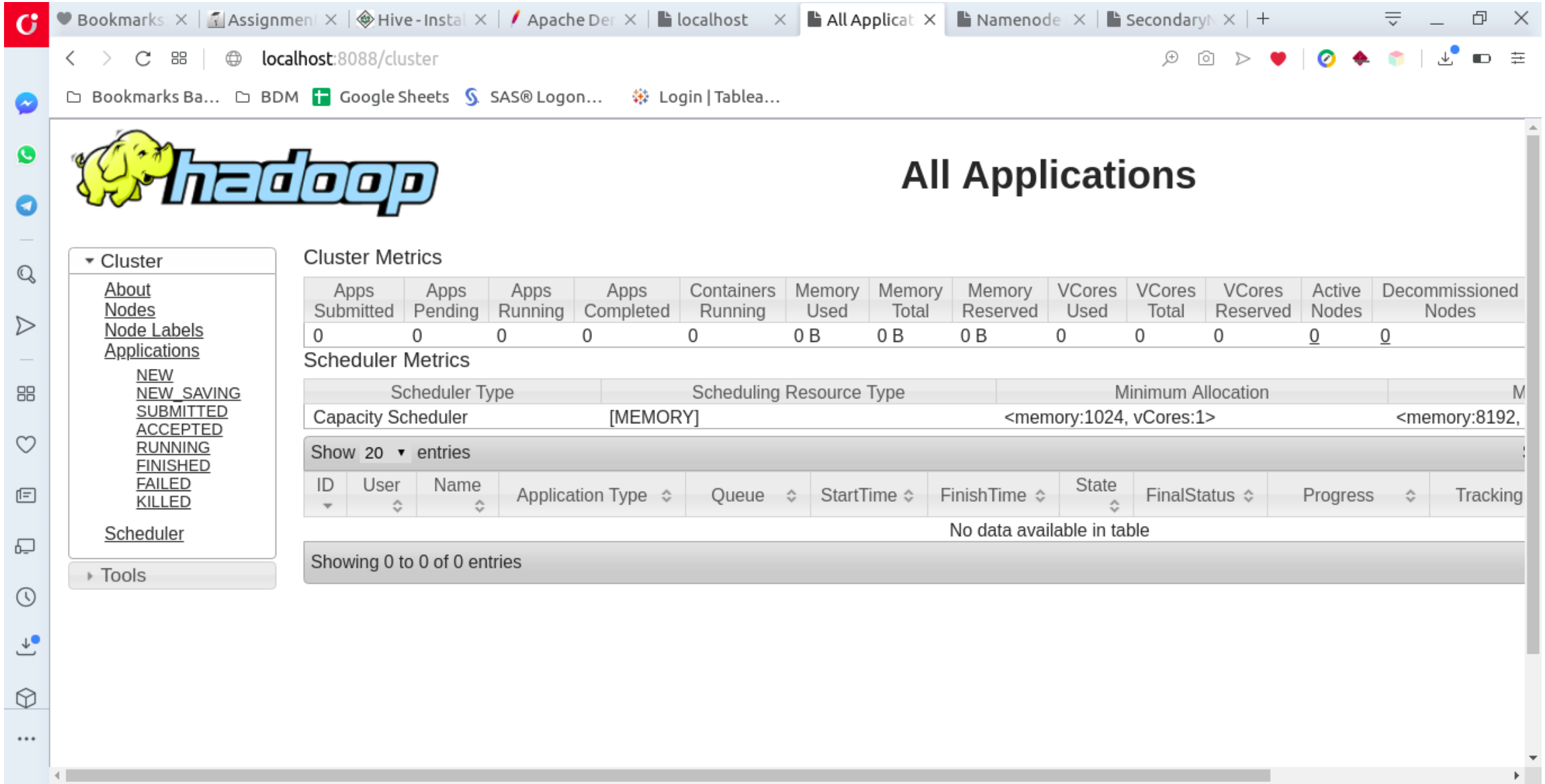

Open
mapred-site.xml.template
~/bigdata/hadoop/etc/hadoop
Save
<!-- Put site-specific property overrides in this file. -->
<configuration>
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
<property>
<name>yarn.app.mapreduce.am.resource.mb</name>
<value>512</value>
</property>
<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


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

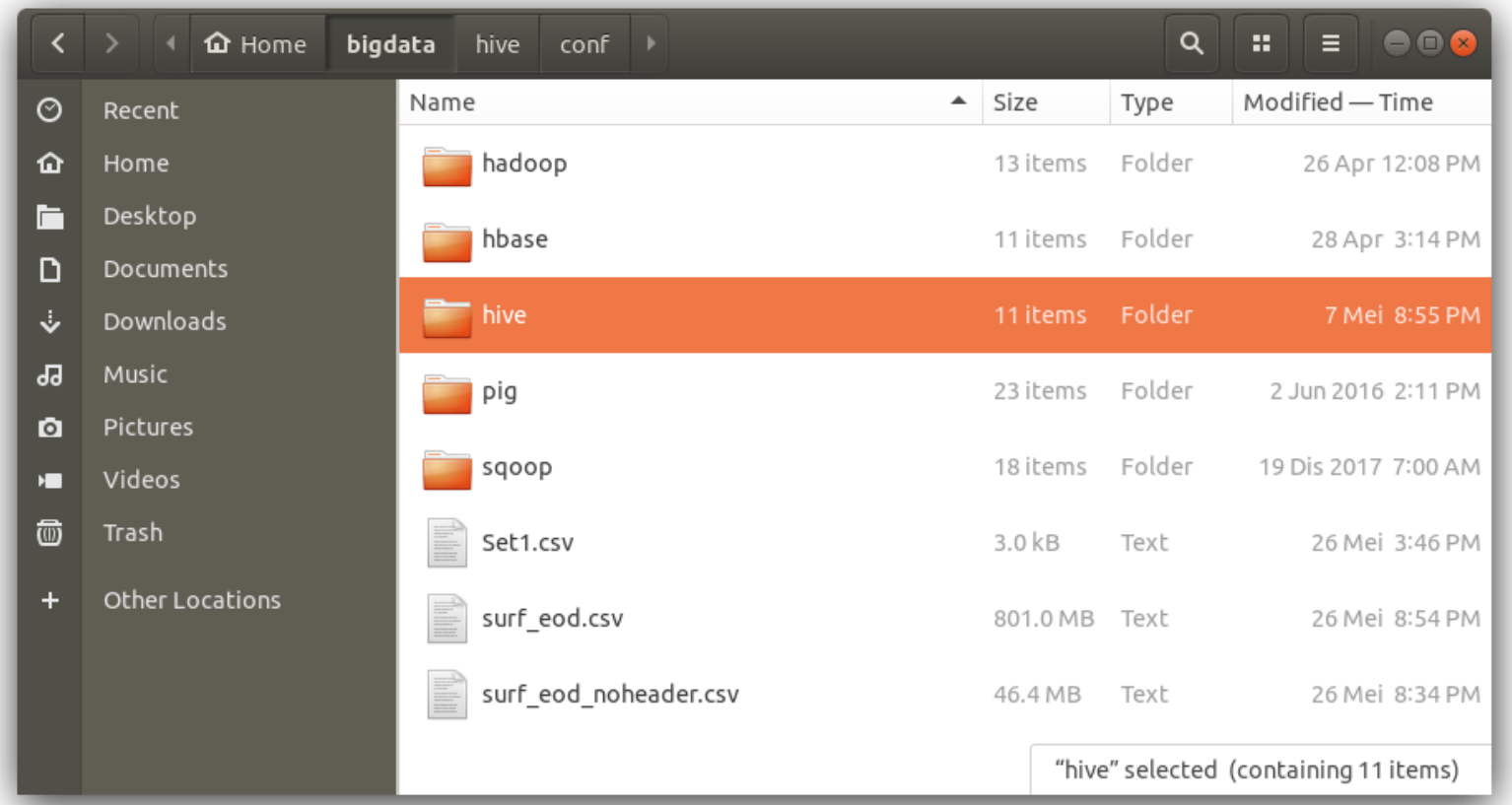

File Edit View Search Terminal Help
(base) zulkanh@elitebook:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [localhost]
zulkanh@localhost's password:
localhost: starting namenode, logging to /home/zulkanh/bigdata/hadoop/logs/hadoop-zulkanh-namenode-elitebook.out
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:~$

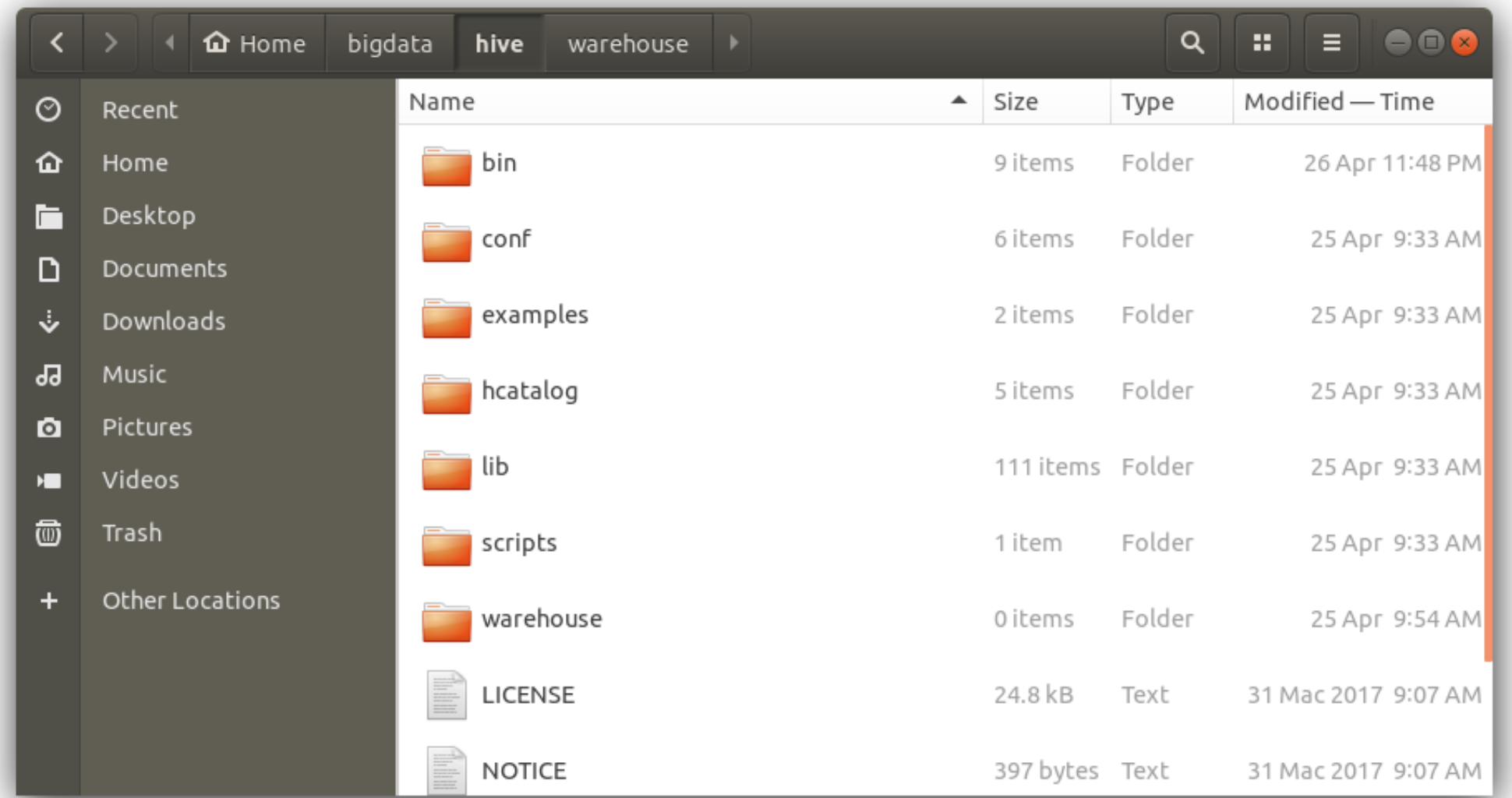

Bookmarks
Assignmen
Hive - Instal
Apache Der
localhost
All Applicat
Namenode
Secondary
localhost:50070/dfshealth.html#tab-overview
Bookmarks Ba...
BDM
Google Sheets
SAS® Logon...
Login | Tablea...
Hadoop
Overview
Datanodes
Datanode Volume Failures
Snapshot
Startup Progress
Utilities
Overview 'localhost:9000' (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.


Bookmarks
Assignmen
Hive - Instal
Apache Der
localhost
All Applicat
Namenode
SecondaryN
localhost:50090/status.html
Bookmarks Ba...
BDM
Google Sheets
SAS® Logon...
Login | Tablea...
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

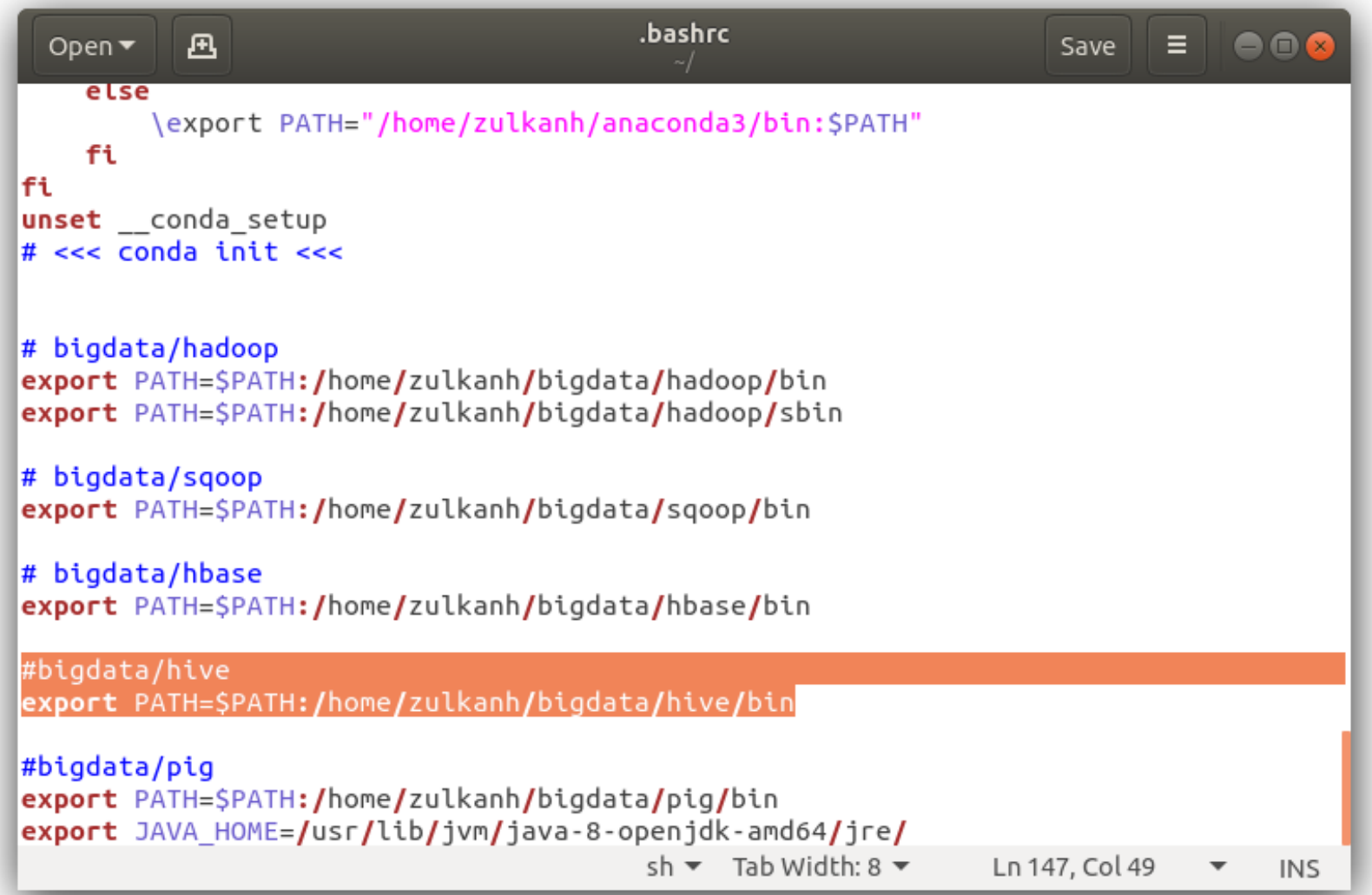

Bookmarks
Assignmen
Hive - Instal
Apache Der
localhost
All Applicat
Namenode
Secondary
localhost:8088/cluster
Bookmarks Ba...
BDM
Google Sheets
SAS® Logon...
Login | Tablea...
hadoop
All Applications
Cluster
About
Nodes
Node Labels
Applications
NEW
NEW_SAVING
SUBMITTED
ACCEPTED
RUNNING
FINISHED
FAILED
KILLED
Scheduler
Tools
Cluster Metrics
Apps Submitted
Apps Pending
Apps Running
Apps Completed
Containers Running
Memory Used
Memory Total
Memory Reserved
VCores Used
VCores Total
VCores Reserved
Active Nodes
Decommissioned Nodes
0 0 0 0 0 0 B 0 B 0 B 0 0 0 0 0
Scheduler Metrics
Scheduler Type
Scheduling Resource Type
Minimum Allocation
Capacity Scheduler
[MEMORY]
<memory:1024, vCores:1>
<memory:8192,
Show 20 entries
ID
User
Name
Application Type
Queue
StartTime
FinishTime
State
FinalStatus
Progress
Tracking
No data available in table
Showing 0 to 0 of 0 entries

Step 4: Installing Hive. ...




Home
bigdata
hive
warehouse
Recent
Home
Desktop
Documents
Downloads
Music
Pictures
Videos
Trash
Other Locations
Name
Size
Type
Modified — Time
bin
9 items
Folder
26 Apr 11:48 PM
conf
6 items
Folder
25 Apr 9:33 AM
examples
2 items
Folder
25 Apr 9:33 AM
hcatalog
5 items
Folder
25 Apr 9:33 AM
lib
111 items
Folder
25 Apr 9:33 AM
scripts
1 item
Folder
25 Apr 9:33 AM
warehouse
0 items
Folder
25 Apr 9:54 AM
LICENSE
24.8 kB
Text
31 Mac 2017 9:07 AM
NOTICE
397 bytes
Text
31 Mac 2017 9:07 AM

Step 5: Configuring Hive. ...


Open
.bashrc
~/
Save
    else
        \export PATH="/home/zulkanh/anaconda3/bin:$PATH"
    fi
fi
unset __conda_setup
<<< conda init <<<

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

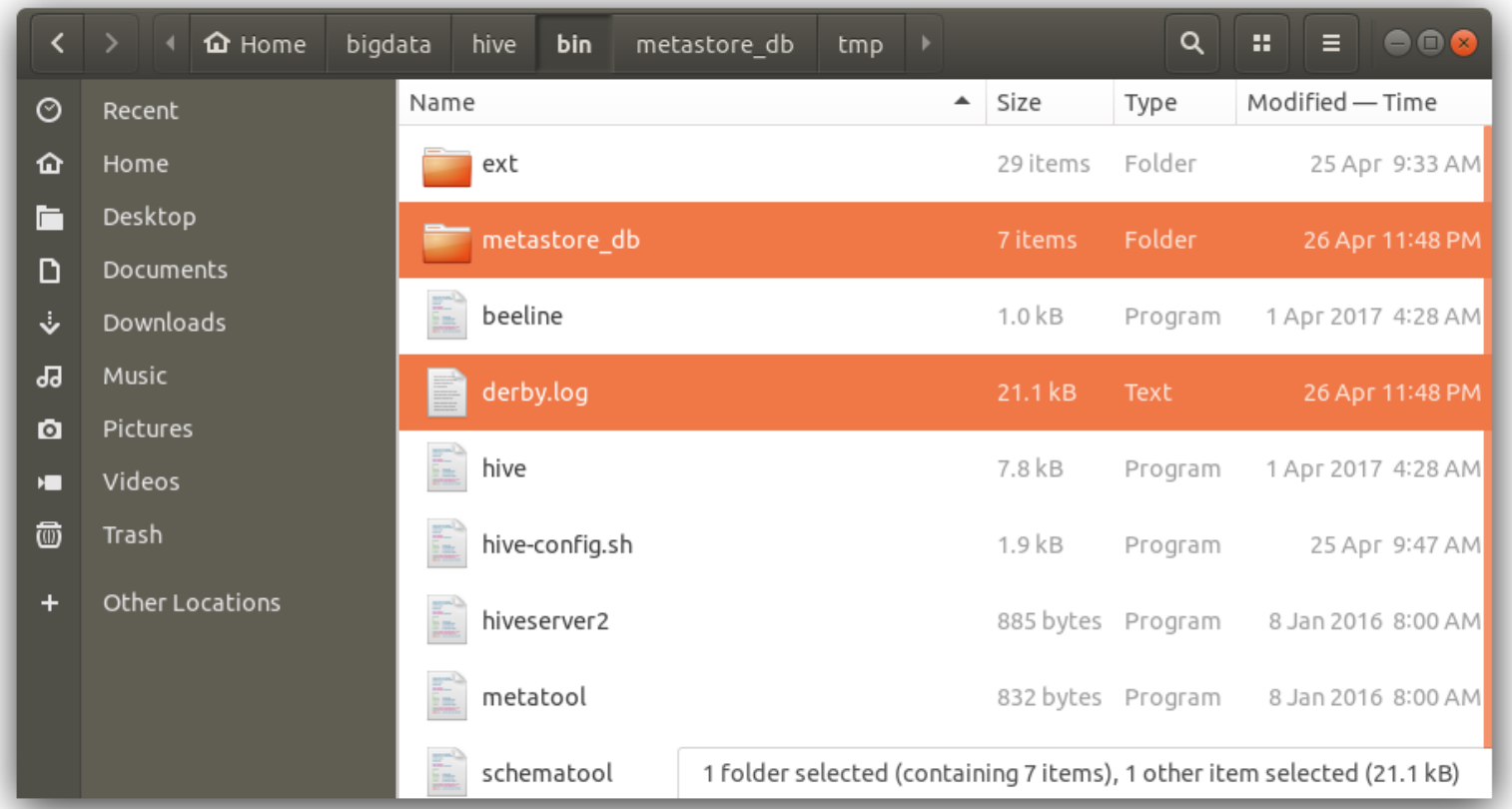
bigdata/sqoop
export PATH=$PATH:/home/zulkanh/bigdata/sqoop/bin

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

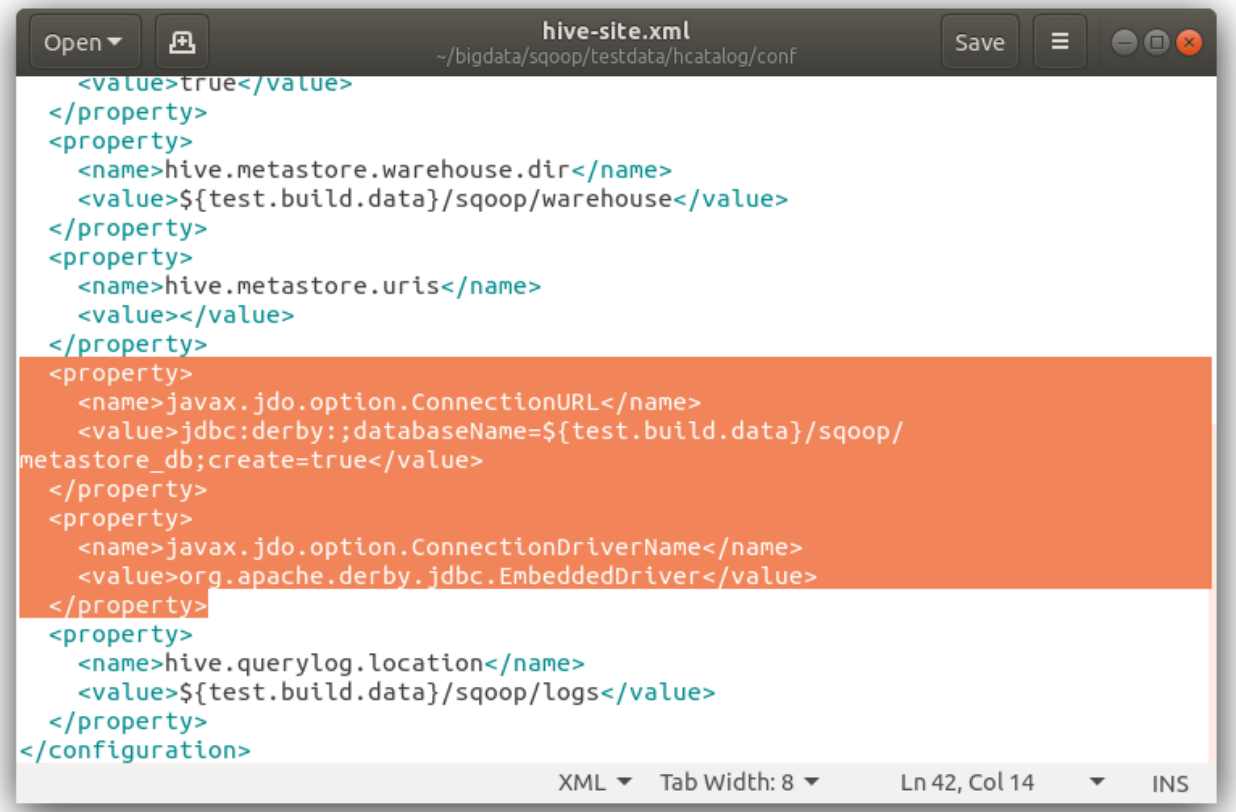
#bigdata/hive
export PATH=$PATH:/home/zulkanh/bigdata/hive/bin

#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. ...



Step 8: Verifying Hive Installation. ...