**Installing Hadoop in Ubuntu**

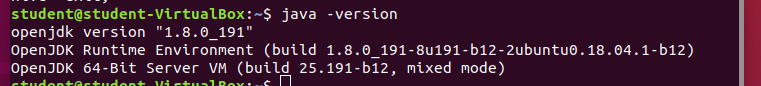
1. To update latest distributions of software:

* *sudo apt-get update*

2. Hadoop requires Java 8 to be installed. To install java version 8:

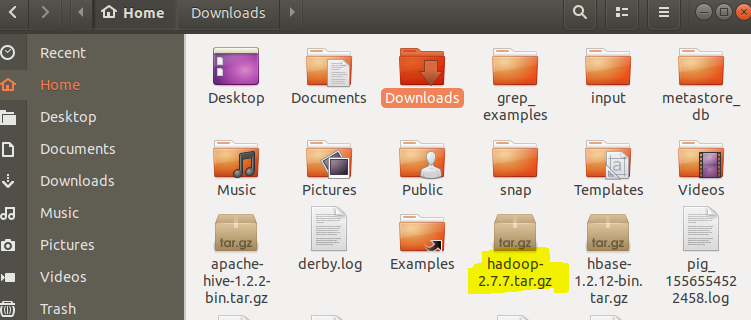
* *sudo apt-get install openssh-server*
* *sudo apt-get install openjdk-8-jdk*

3. Verifying JAVA 8 Installation:

* *java -version*

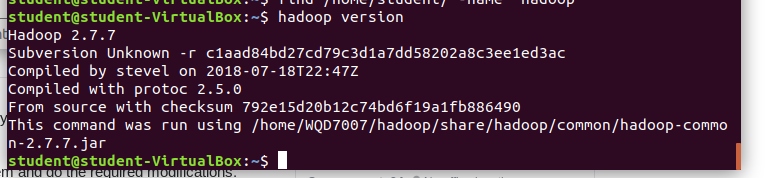
4. Download and extract Hadoop:

* *wget* [*https://www-eu.apache.org/dist/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz*](https://www-eu.apache.org/dist/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz)



5. Unzip the Hadoop file:

* *tar -xzf hadoop-2.7.7.tar.gz*

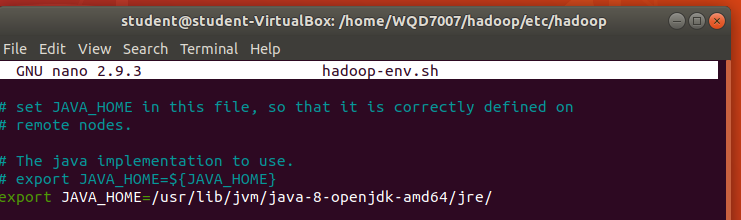


6. Move the file to your own directory:

* *sudo mkdir /home/{yourname}*
* *sudo mv hadoop-2.7.7 /home/{yourname}/hadoop/*

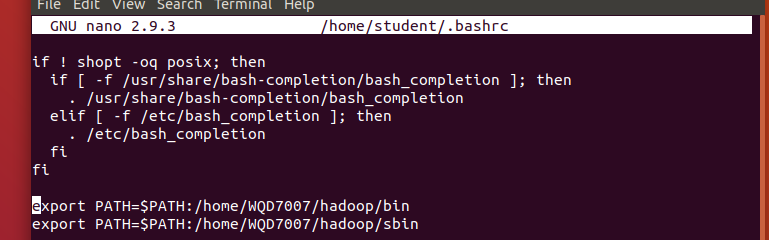
7. Set JAVA\_HOME in /home/{yourname}/hadoop/etc/hadoop/hadoop-env.sh:

* *nano /home/{yourname}/hadoop/etc/hadoop/hadoop-env.sh*

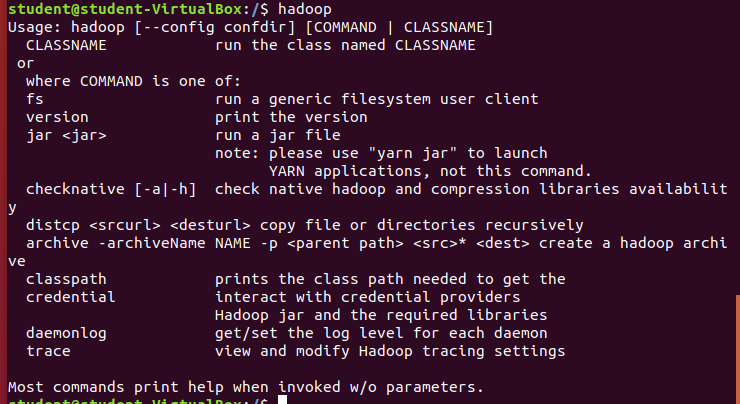


8. Export path for Hadoop in ~/.bashr to allow easy Hadoop function access:

* *export PATH=$PATH:/home/{yourname}/hadoop/bin*
* *export PATH=$PATH:/home/{yourname}/hadoop/sbin*

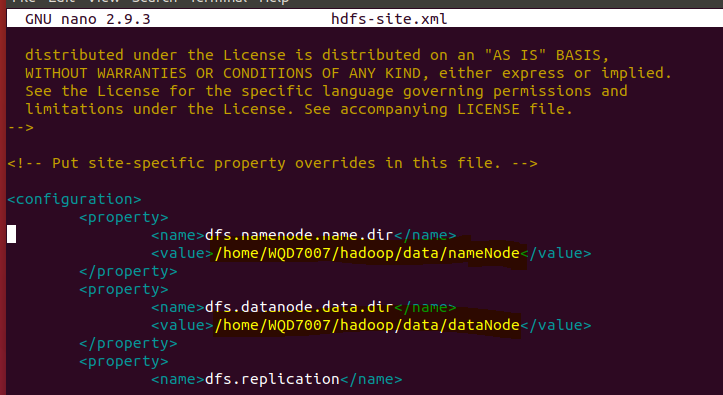


9. Run Hadoop:

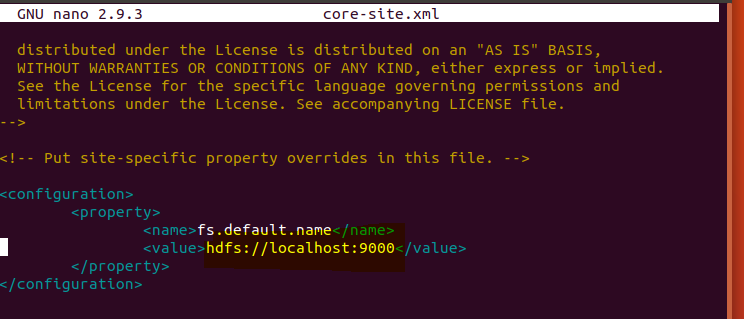


10. Update hdfs-site.xml in /home/{yourname}/hadoop/etc/hadoop folder using nano. This file contains the configuration properties that Hadoop uses when starting up. Save and close this file.

* nano /home/{yourname}/hadoop/etc/Hadoop/hdfs-site.xml



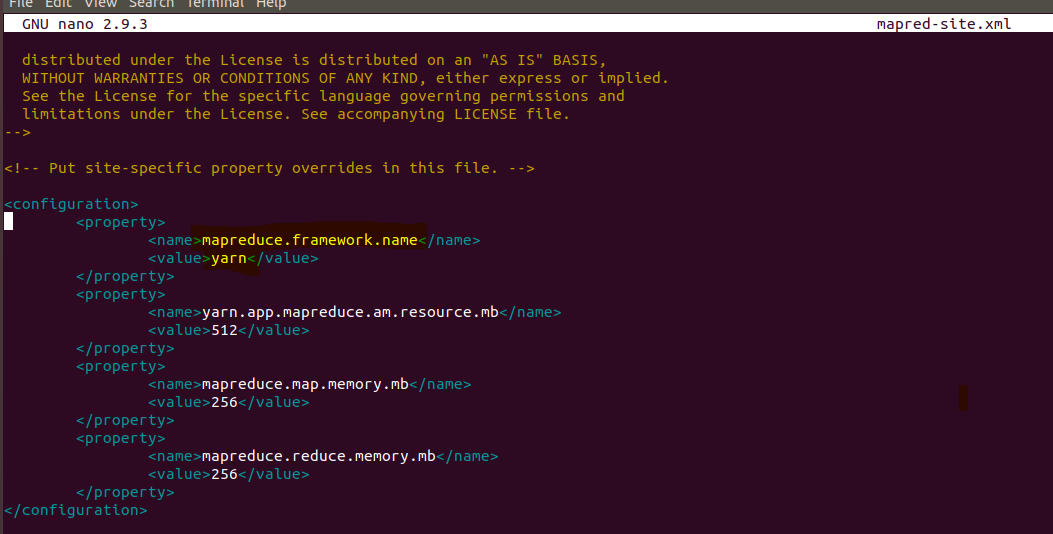
11. Update core-site.xml in /home/{yourname}/hadoop/etc/hadoop using nano. You can change ‘localhost’ to your PC’s IP address. Save and close this file.



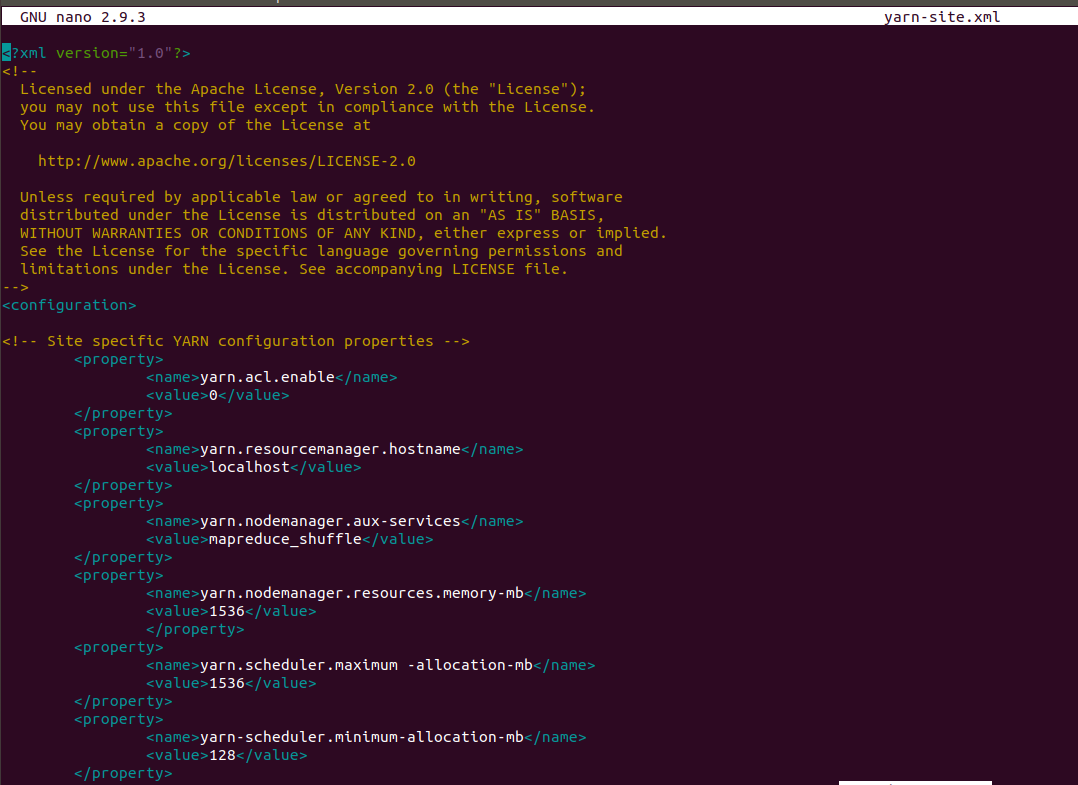
12. Rename/copy the mapred-site.xml.template in /home/{yourname}/hadoop/etc/hadoop to

mapred-site.xml. This file is used to specify which framework is being used for MapReduce. Then, update mapred-site.xml. Save and close the file.

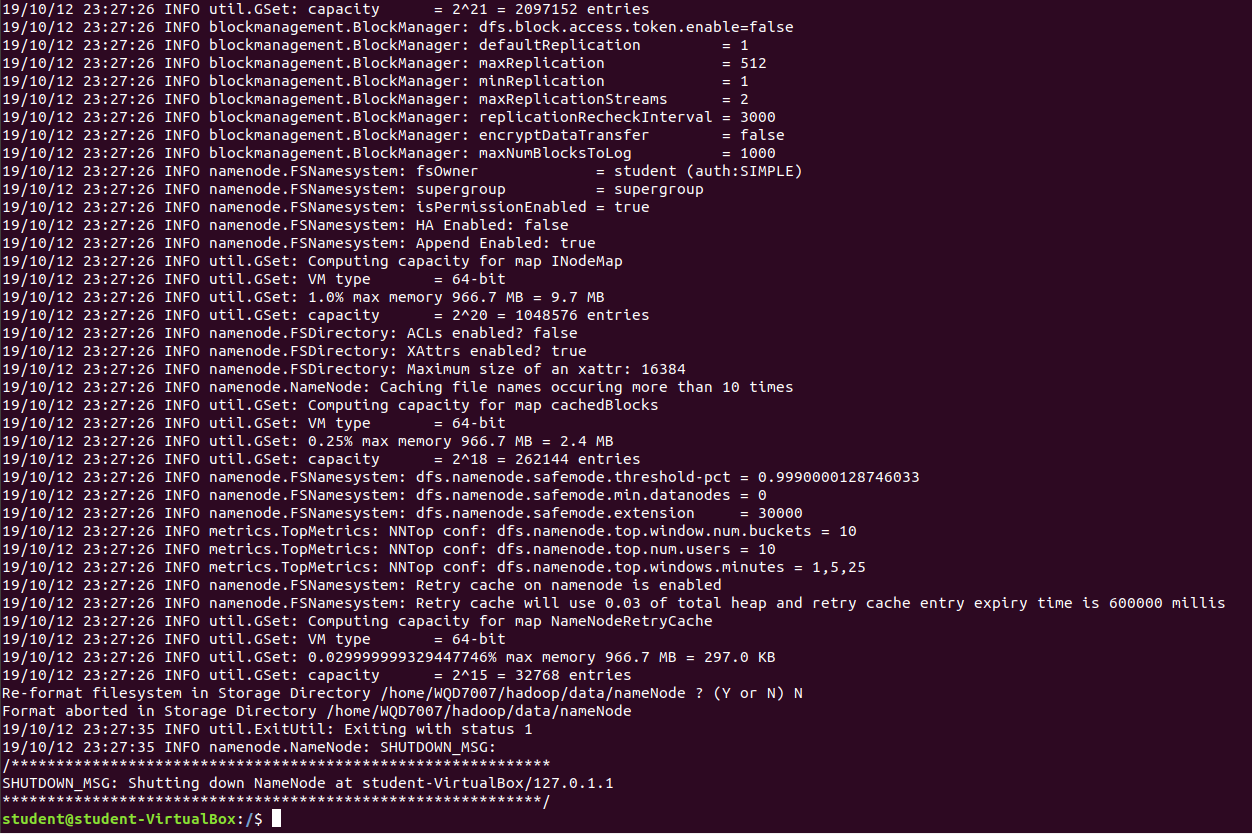
* *cp /home/{yourname}/hadoop/etc/hadoop/mapred-site.xml.template /home/{yourname}/hadoop/etc/hadoop/mapred-site.xml*
* *nano /home/{yourname}/hadoop/etc/hadoop/mapred-site.xml*



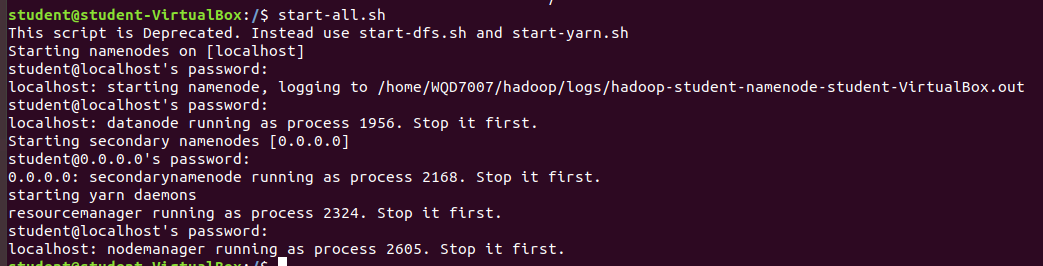
13. Update yarn-site.xml in /home/{yourname}/hadoop/etc/hadoop folder using nano. This file contains the configuration properties that Mapreduce uses when starting up. Save and close this file.



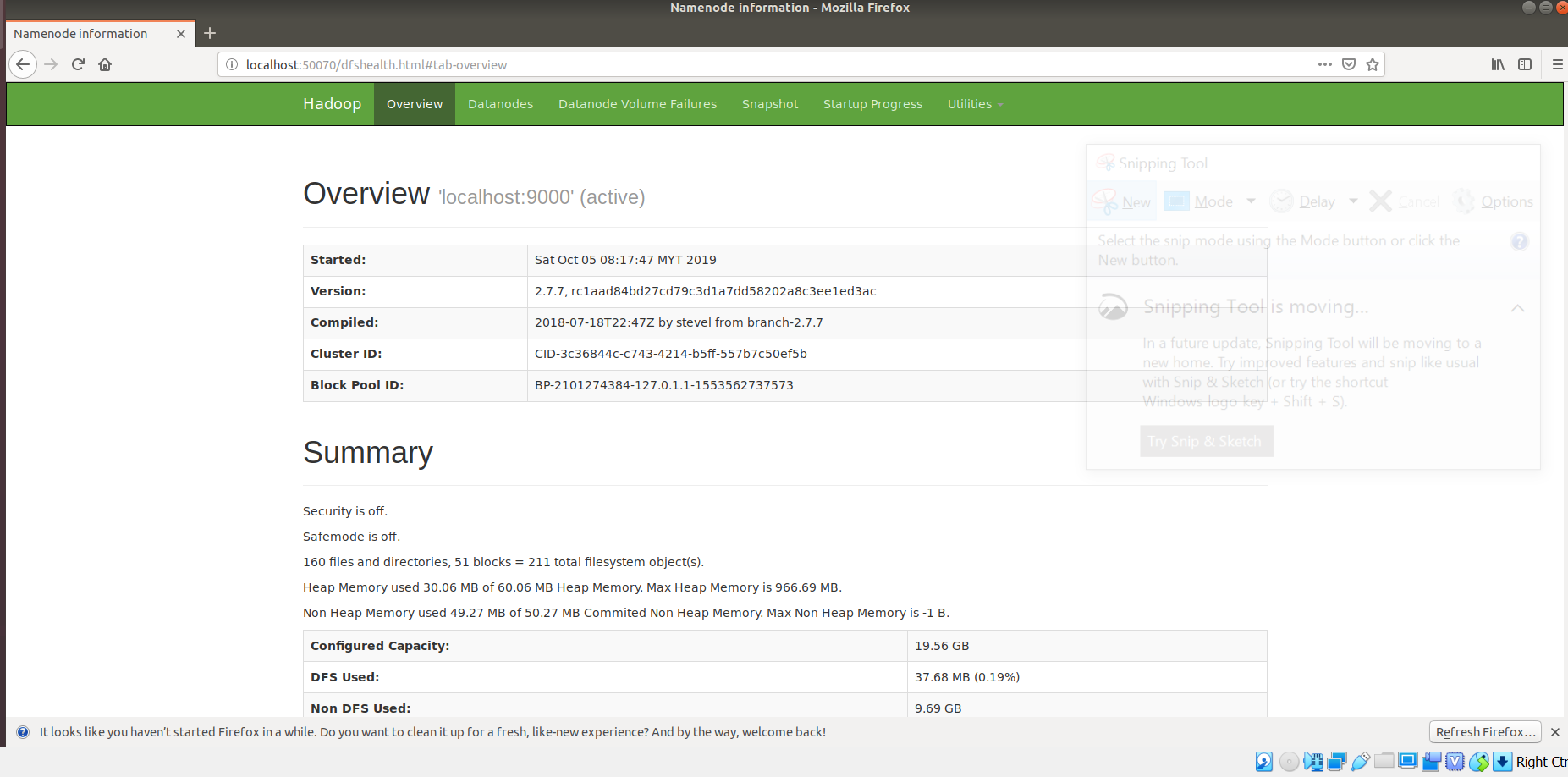
14. Run *hdfs namenode -format.*



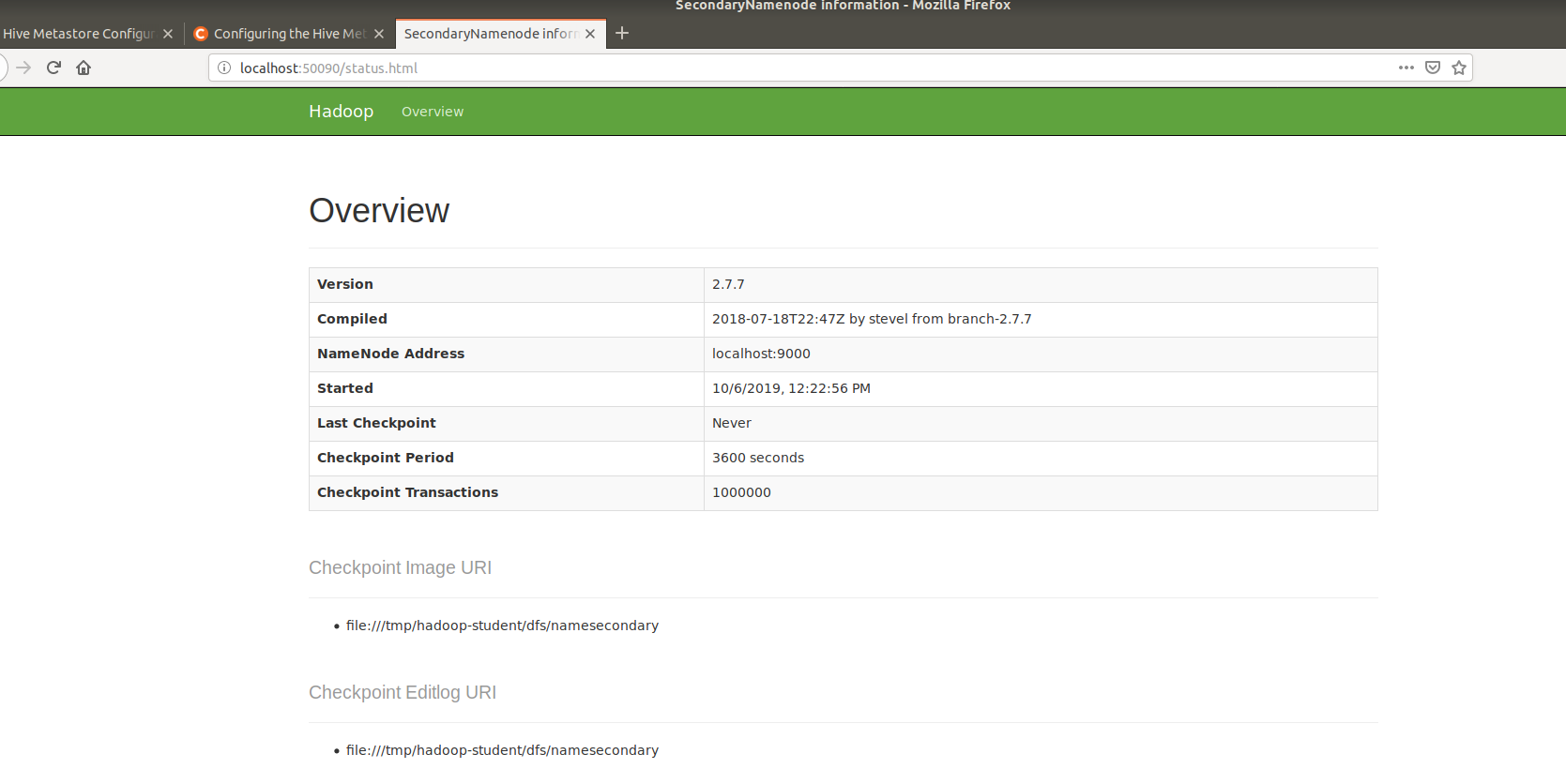
15.Run *start-all.sh* (or run *start-dfs.sh* and *start-yarn.sh* separately)



16. Browselocalhost:50070 in your browser for namenode:



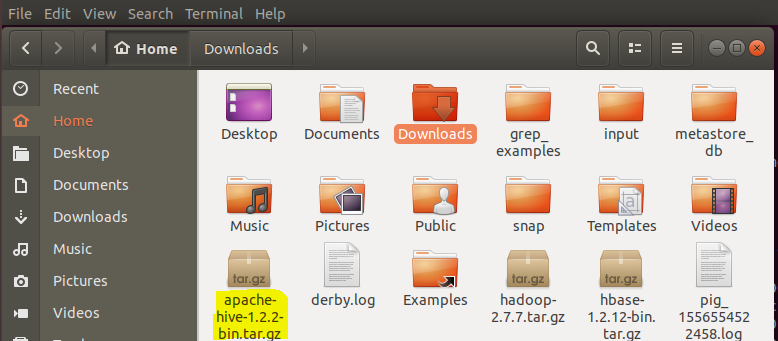
17. Browse localhost:50090 in your browser for secondary namenode:



**Installing Hive in Ubuntu**

1. Download and install hive using:

* *wget* [*https://www.apache.org/dist/hive/hive-1.2.2/apache-hive-1.2.2-bin.tar.gz*](https://www.apache.org/dist/hive/hive-1.2.2/apache-hive-1.2.2-bin.tar.gz)



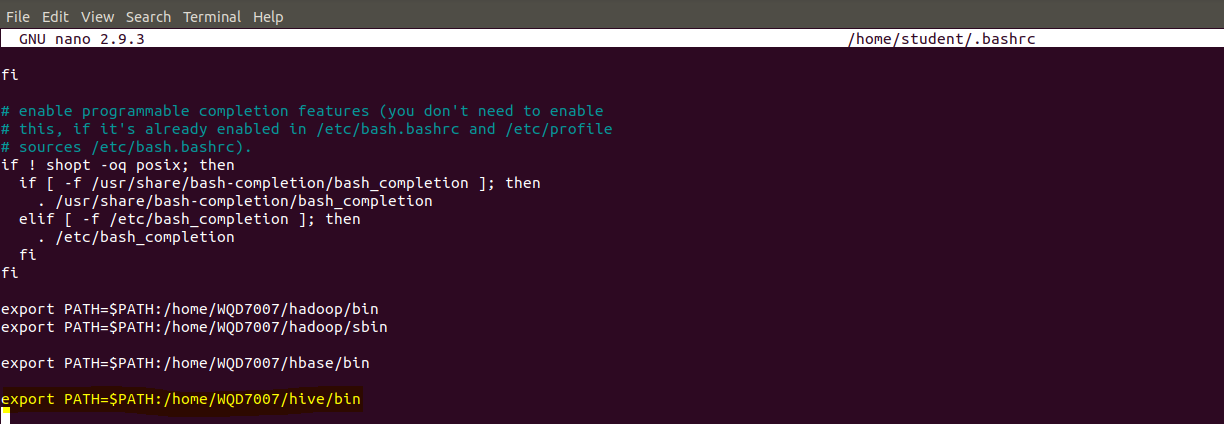
2. Unzip the hive folder:

* *tar -xzf apache-hive-1.2.2-bin.tar.gz*

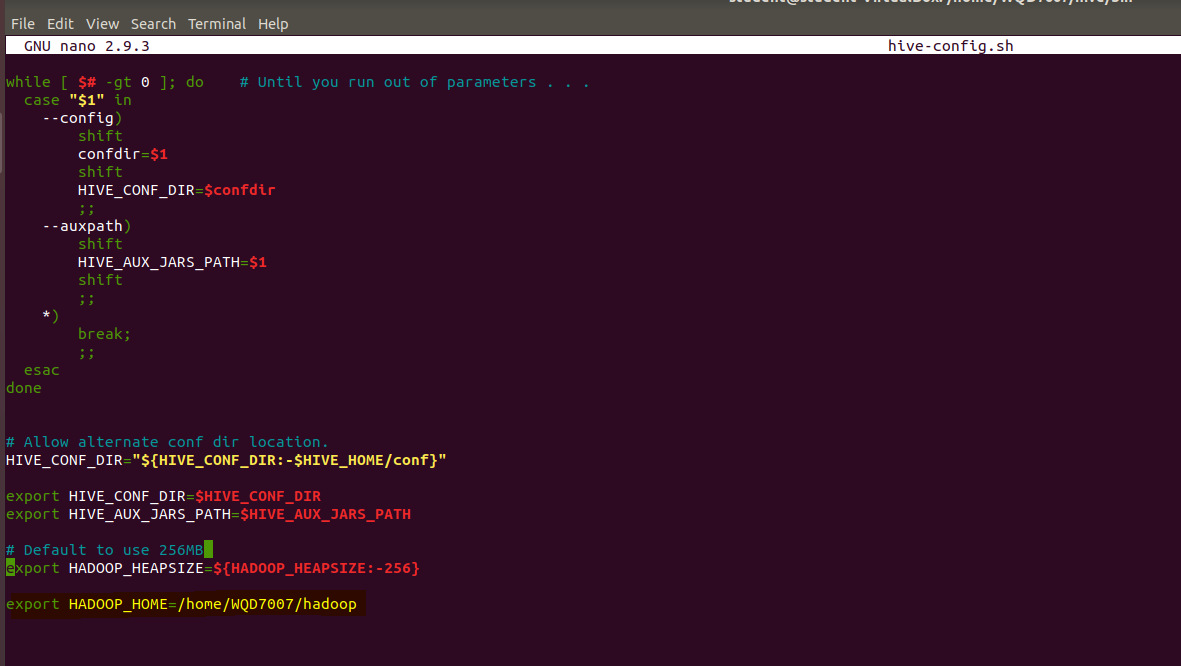
3. Move the hive folder to your own directory:

* mv apache-hive-1.2.2-bin /home/{yourname}/hive/

4. In *nano ~/.bashrc*, set export PATH=$PATH:/home/{yourname}/hive/bin



5. In hive bin folder, *nano hive-config.sh* and add export HADOOP\_HOME=/home/wlhoo/hadoop at the end of the file to connect hive with Hadoop.



6. Run Hive and hive is successfully installed.

