Output links:

Key steps:

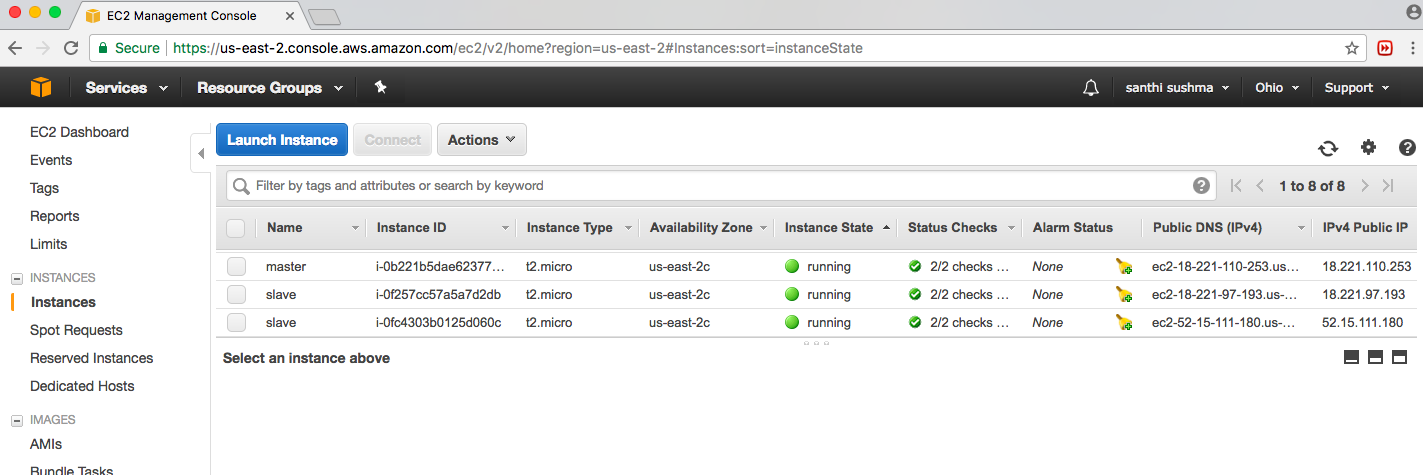
1) Installation of Hadoop and JAVA in 3 EC2 instances - one being name node(master) and two data nodes(slaves)

2) Test java code on local Hadoop setup and export jar file along with libraries.

3) Run the jar file from instance.

Step 1: Installation of Hadoop in EC2:

1) Launch three EC2 instances with one namenode(master) and two datanodes(slaves)



Nodes used in setup

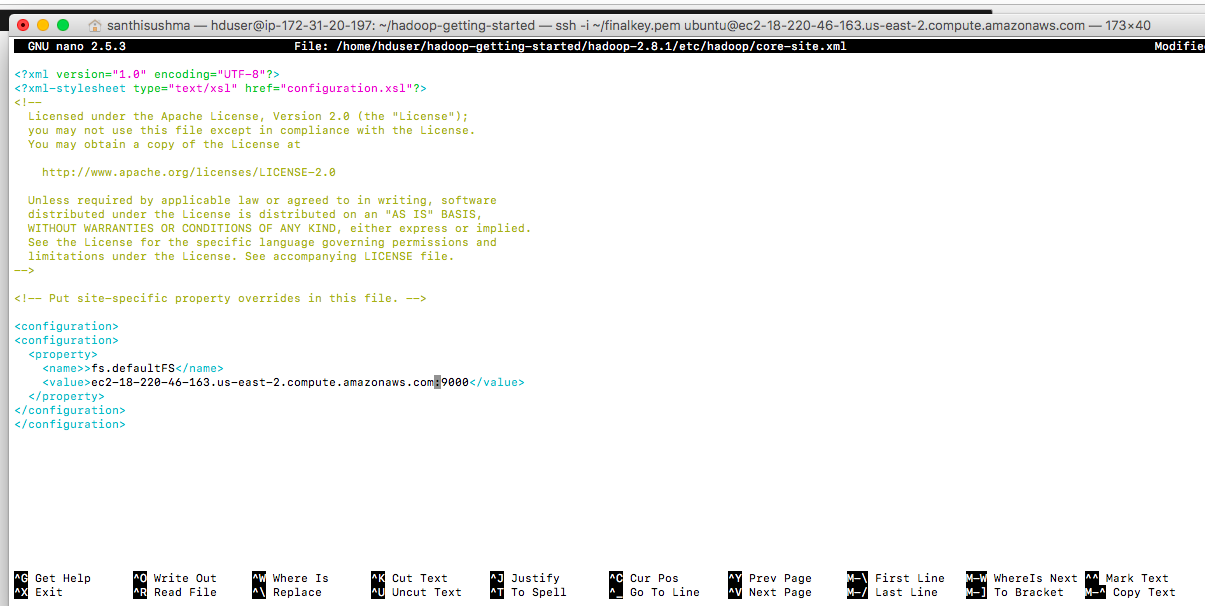
2) Login to nodes from your terminal to set up java and Hadoop

3) Enable password less login from name to data nodes

4) Configure core-site, hfs and yarn-site xml

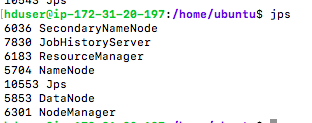
5) Update path and few parameters

6) Once set up is done, format name node and start services with command “$configpath/start-all.sh”

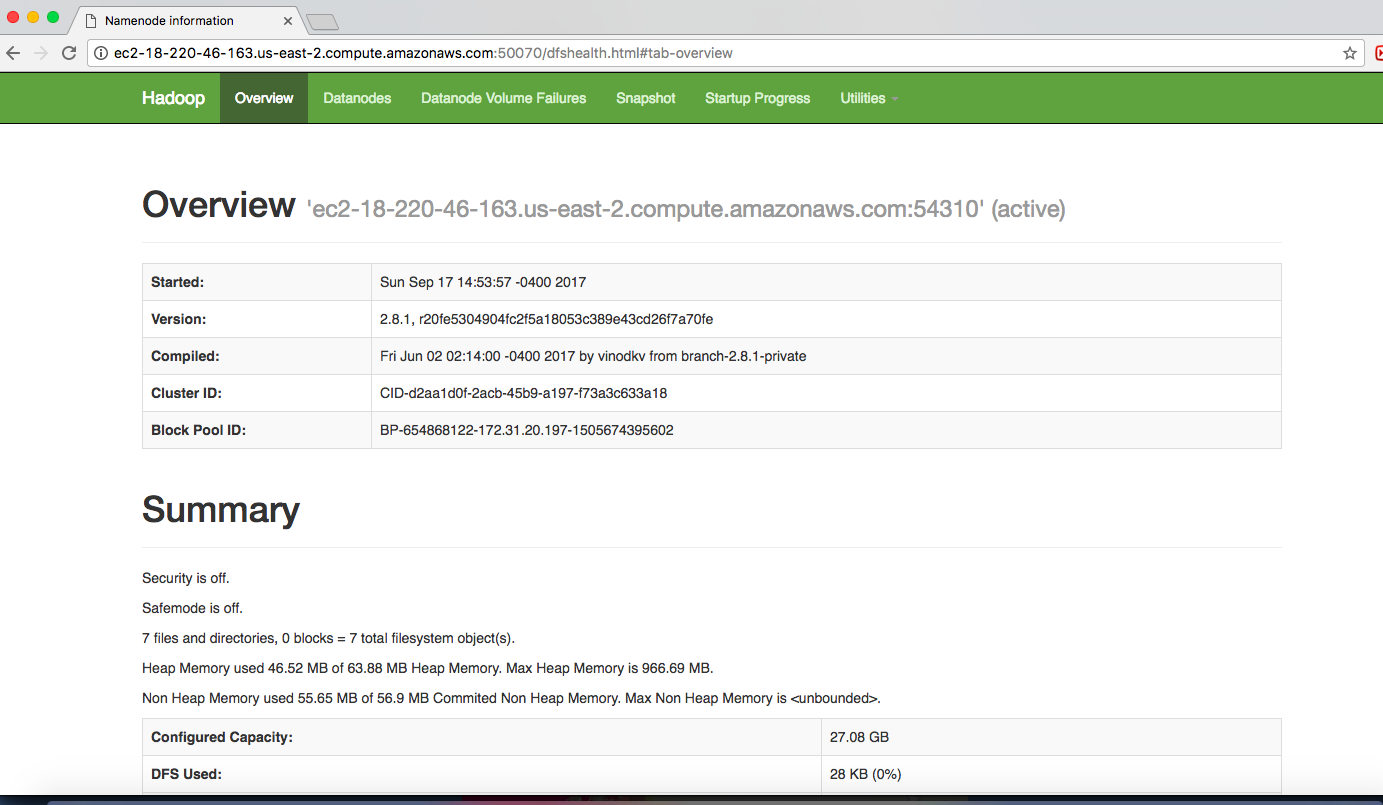


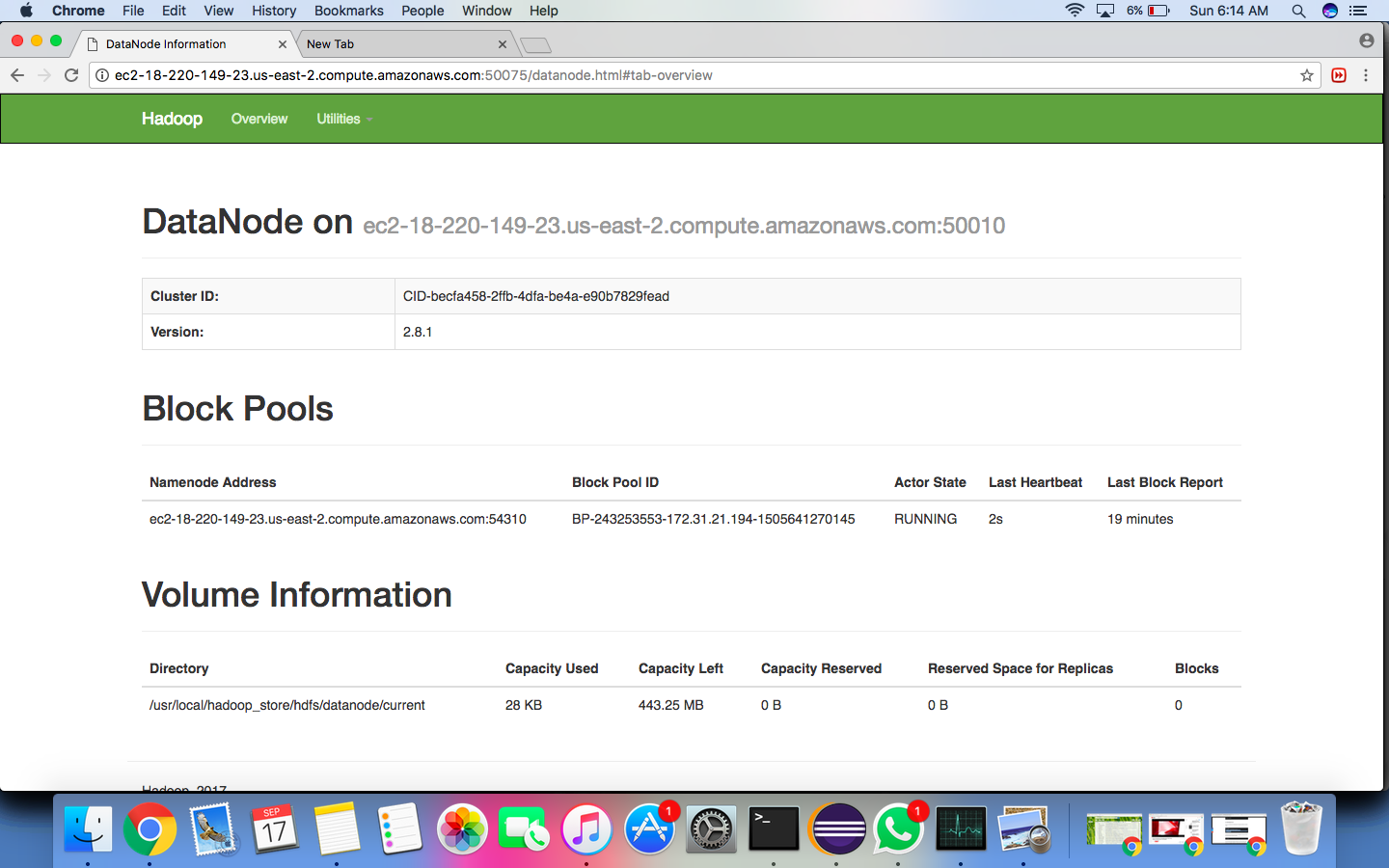
Sample core-site.xml

7) Use jps command to see whether all nodes are running in name node and data nodes

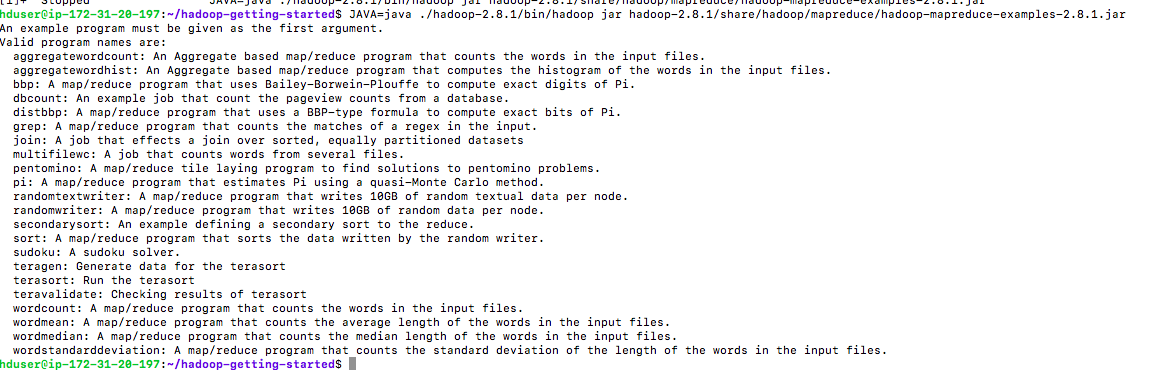


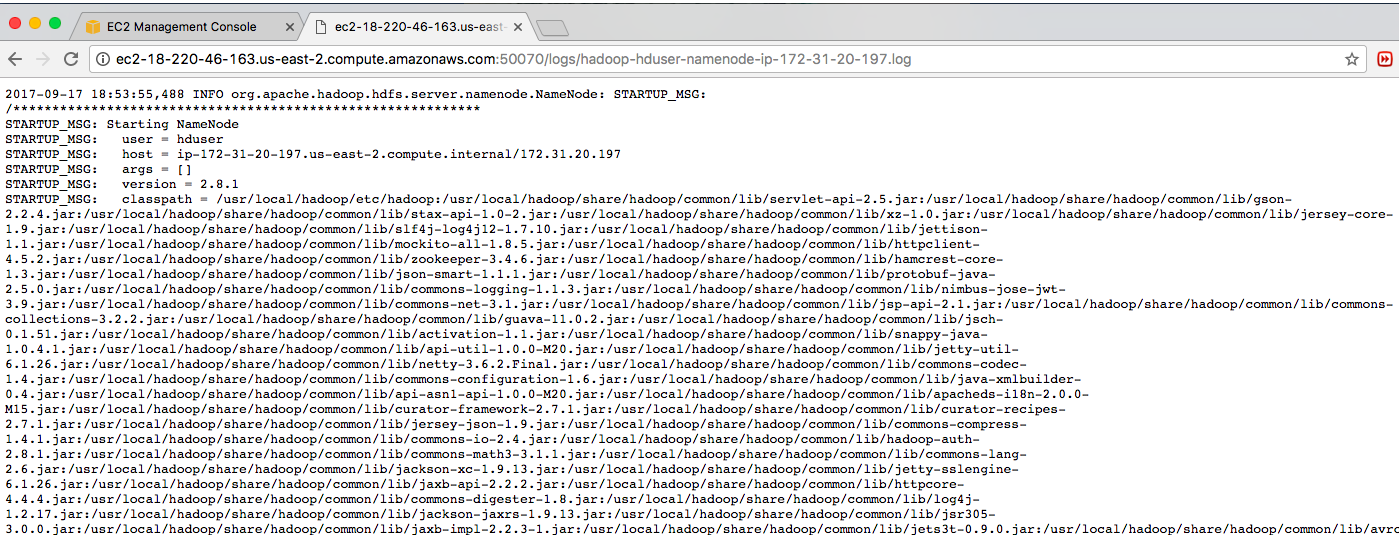
8) Verify if all data nodes are installed through <http://publicdns:50070>

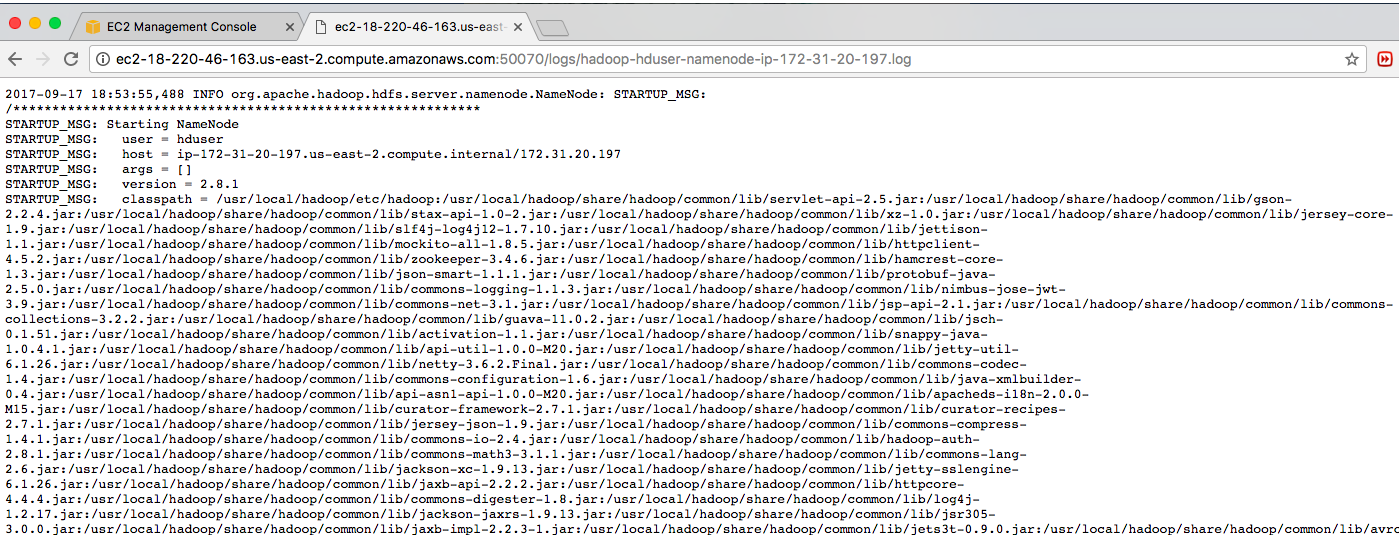




9) Once everything is running test if required predefined jar file exists in the library







10) When jar file exists import jar file and run the code for result.