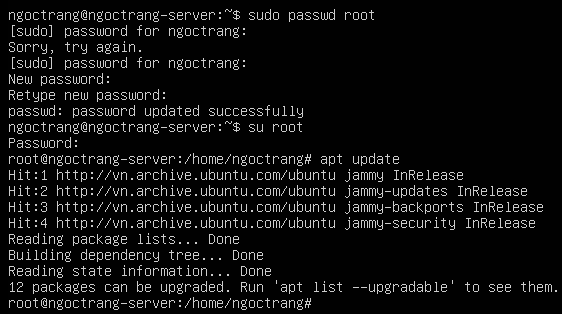
**SV: Trần Thị Ngoc Trang**

MSSV: 21133109

**Cài đặt Hadoop Ecosystem**

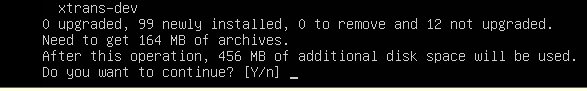
# **Cài đặt OpenJDK**

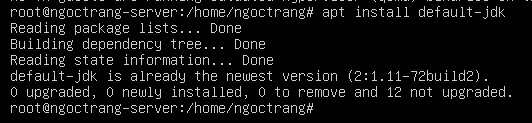
# apt update



# reboot

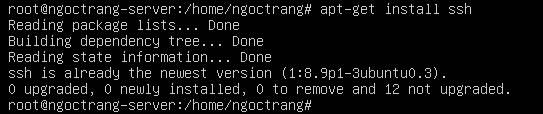
# **apt install default-jdk**



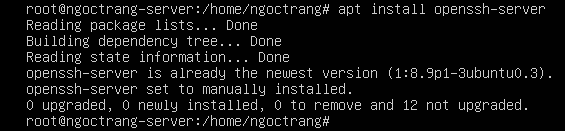


# **Cài đặt SSH**

# apt-get install ssh



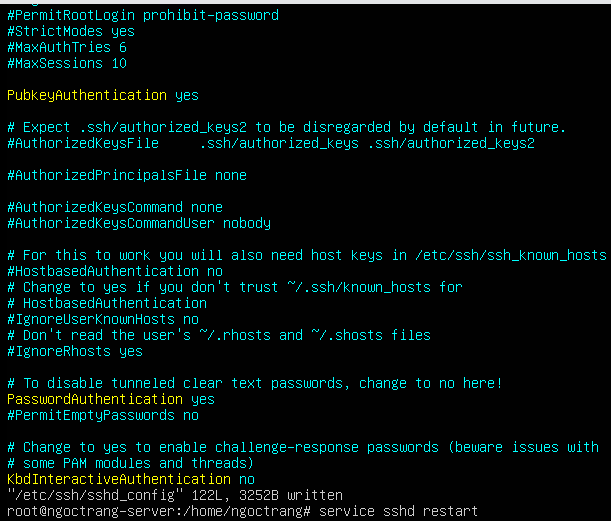
# apt install openssh-server



# reboot

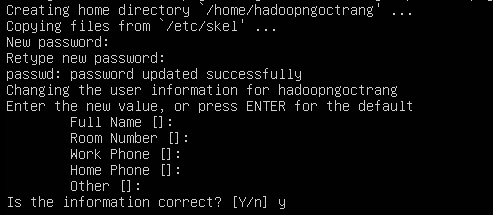
# vim /etc/ssh/sshd\_config

# service sshd restart



# **Tạo user hadoop**

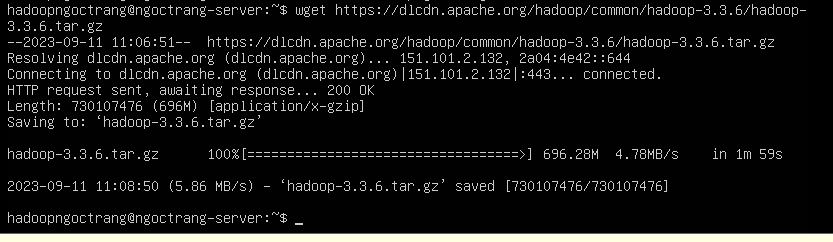
# adduser hadoopngoctrang



# **Cài đặt Hadoop 3.3.6**

# su hadoopngoctrang

wget <https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz>



# tar -xzf hadoop-3.3.6.tar.gz

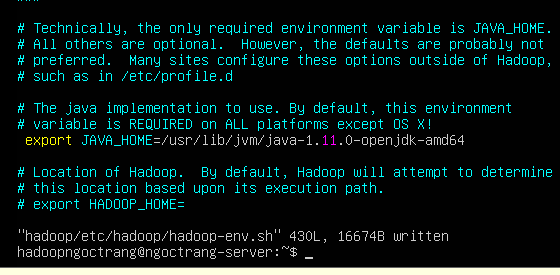
# mv hadoop-3.3.6 hadoop



# **Thiết lập JAVA\_HOME**

# vim ~/hadoop/etc/hadoop/hadoop-env.sh

# export JAVA\_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64



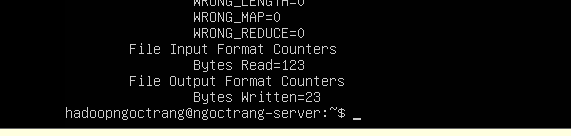
# **Standalone Operation**

# mkdir input

# cp hadoop/etc/hadoop/\*.xml input

# hadoop/bin/hadoop jar hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar grep input output 'dfs[a-z.]+'

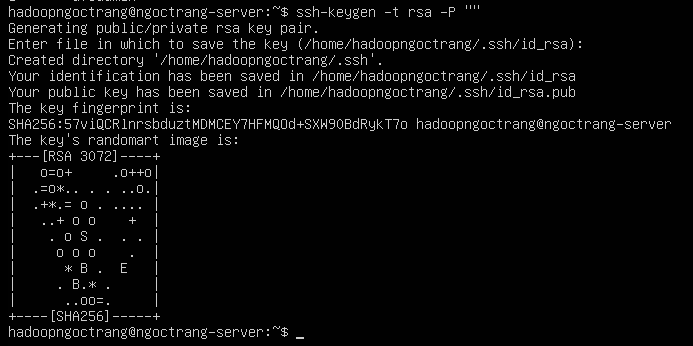
# cat output/\*





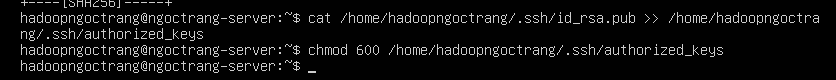
# **Pseudo-Distributed Operation**

# ssh-keygen -t rsa -P ""



# cat /home/hadoopngoctrang/.ssh/id\_rsa.pub >> /home/hadoopngoctrang/.ssh/authorized\_keys

# chmod 600 /home/hadoopngoctrang/.ssh/authorized\_keys



## **Cấu hình core-site.xml**

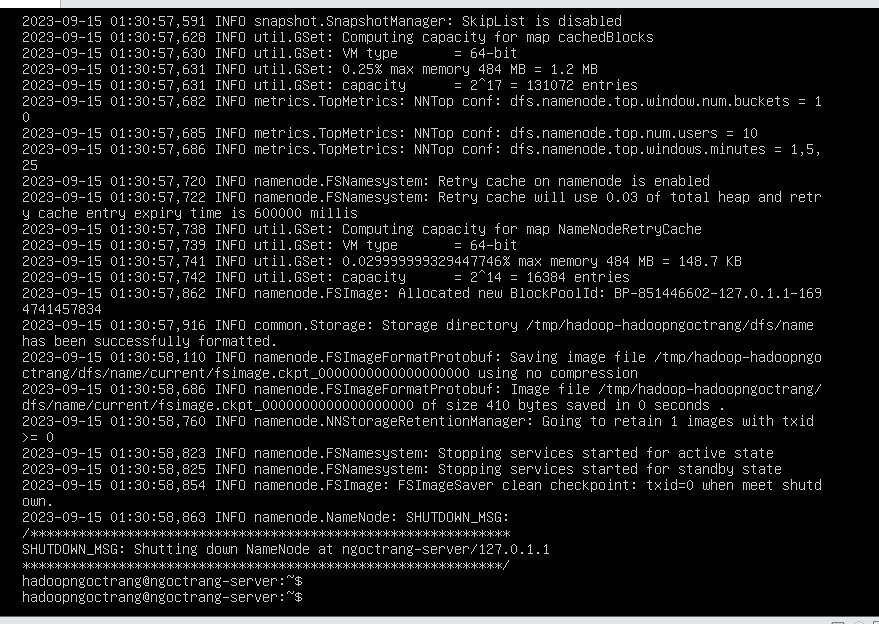


## **Câu**



## 

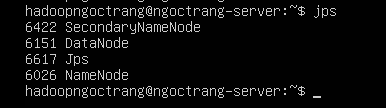
$ hadoop/bin/hdfs namenode -format



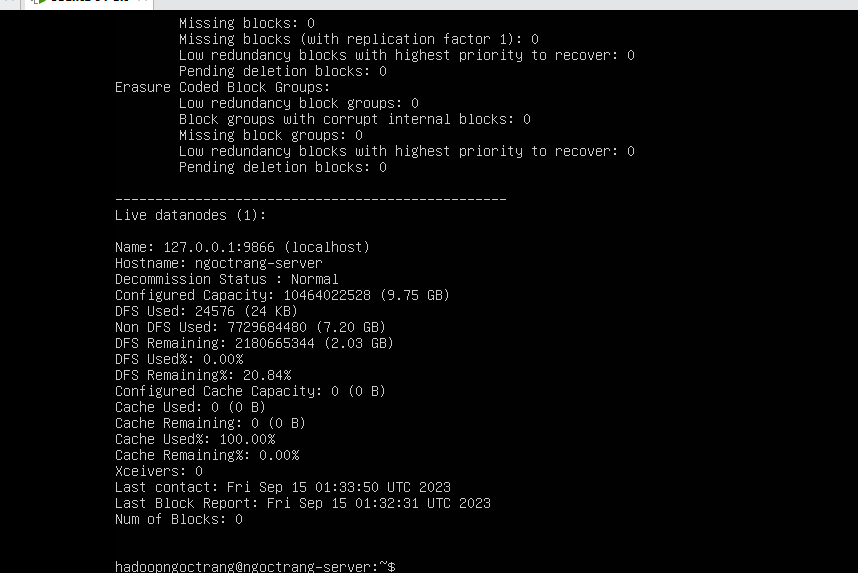
$ hadoop/sbin/start-dfs.sh

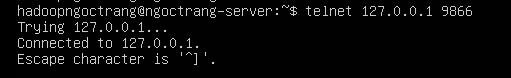
****

$ jps



$ ~/hadoop/bin/hdfs dfsadmin -report





$ ~/hadoop/bin/hdfs dfs -mkdir /user

$ ~/hadoop/bin/hdfs dfs -mkdir /user/hadoopngoctrang

$ ~/hadoop/bin/hdfs dfs -mkdir /user/hadoopngoctrang/input



$ ~/hadoop/bin/hdfs dfs -put hadoop/etc/hadoop/\*.xml /user/hadoopngoctrang/input



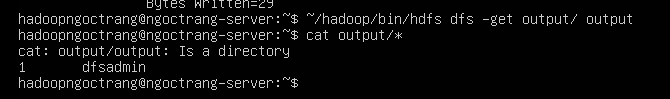
$ ~/hadoop/bin/hadoop jar hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar grep input **output** ‘dfs[a-z.]+’



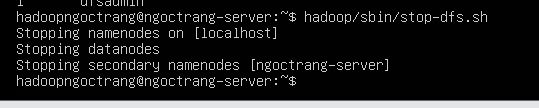
vim

$ ~/hadoop/bin/hdfs dfs -get output/ **output**

$ cat output/\*



$ hadoop/sbin/stop-dfs.sh

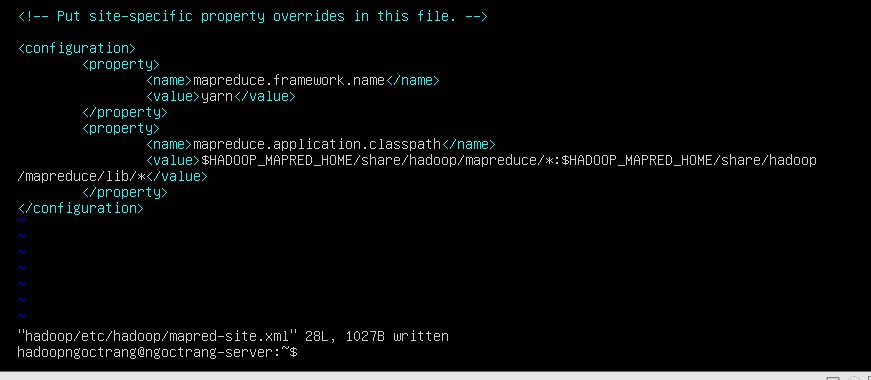


## **B**

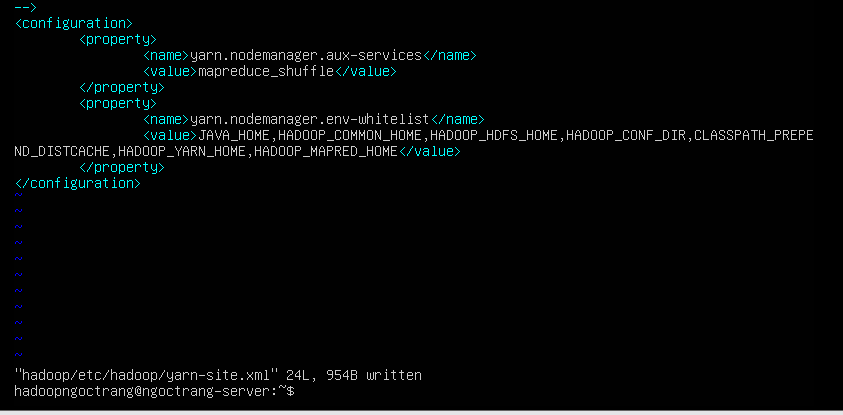
## **vim ~/.bashrc**



$ vim hadoop/etc/hadoop/mapred-site.xml

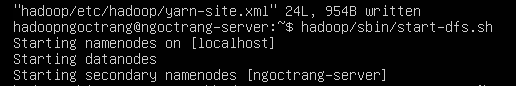


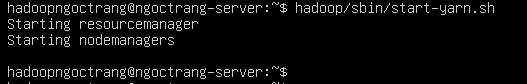
## **Cấu hình file yarn-site.xml**



$ hadoop/sbin/start-dfs.sh

$ hadoop/sbin/start-yarn.sh





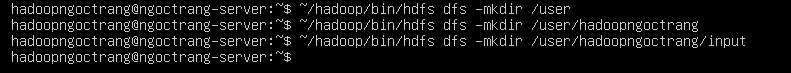
$ rm -rf output

$ ~/hadoop/bin/hdfs dfs -rm -r /user

$ ~/hadoop/bin/hdfs dfs -mkdir /user

$ ~/hadoop/bin/hdfs dfs -mkdir /user/hadoopngoctrang

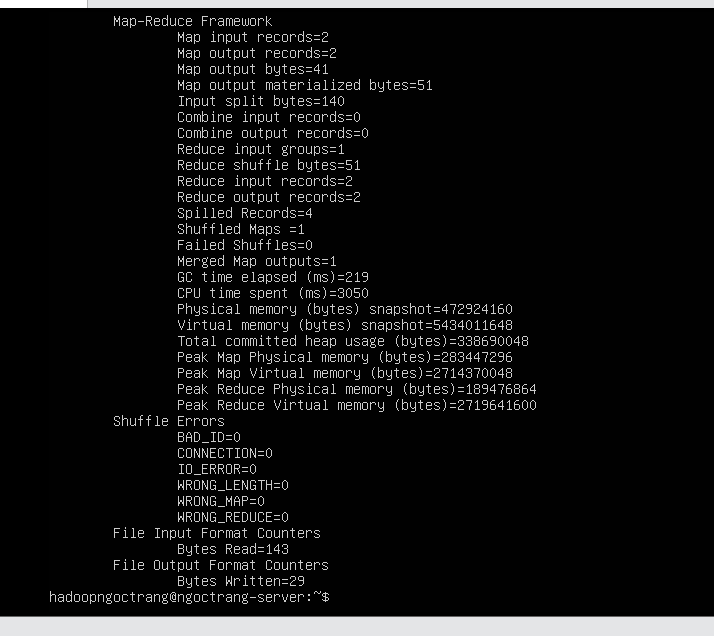
$ ~/hadoop/bin/hdfs dfs -mkdir /user/hadoopngoctrang/input



$ ~/hadoop/bin/hdfs dfs -put hadoop/etc/hadoop/\*.xml /user/hadoopngoctrang/input

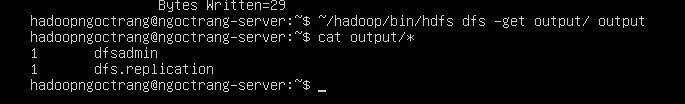


$ ~/hadoop/bin/hadoop jar hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar grep input **output** ‘dfs[a-z.]+’



$ ~/hadoop/bin/hdfs dfs -get output/ **output**

$ cat output/\*



$ hadoop/sbin/stop-dfs.sh

$ hadoop/sbin/stop-yarn.sh

