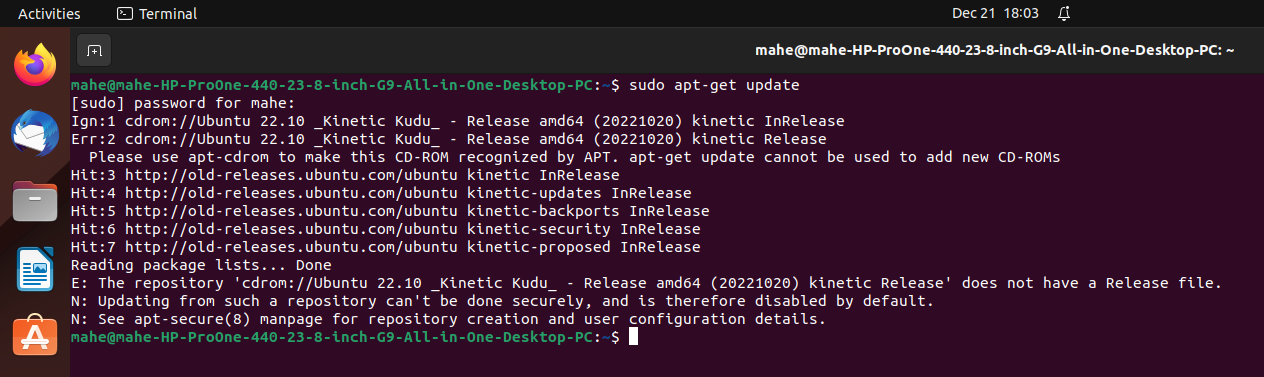
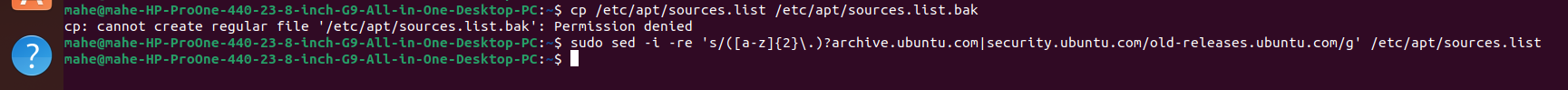
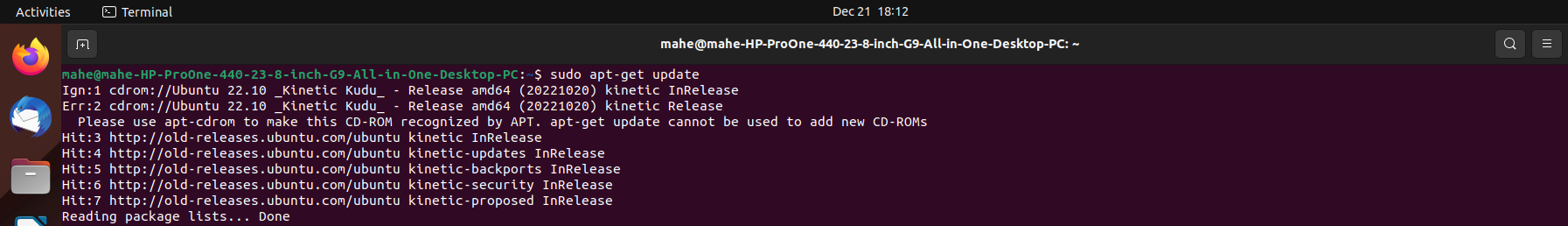
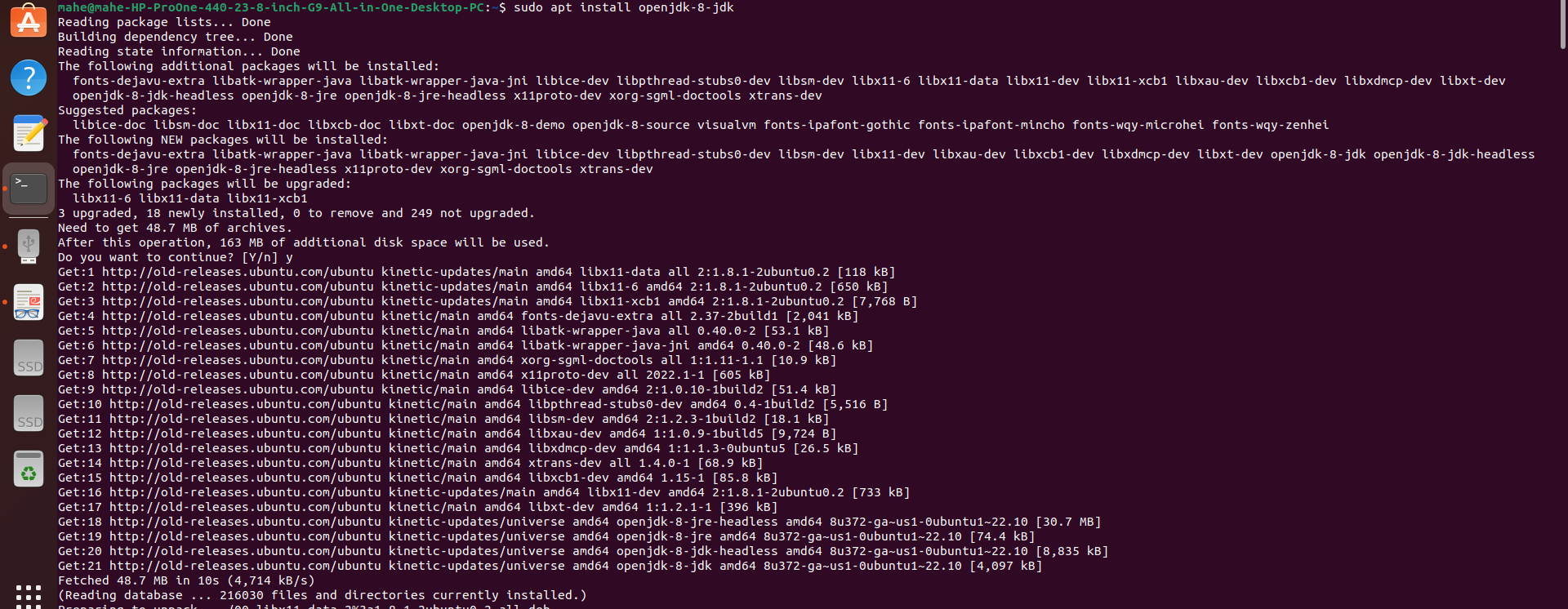
1. sudo apt-get update

2. cp /etc/apt/sources.list /etc/apt/sources.list.bak

3. sudo sed -i -re 's/([a-z]{2}\.)?archive.ubuntu.com|security.ubuntu.com/old-releases.ubuntu.com/g' /etc/apt/sources.list

4. sudo apt-get update

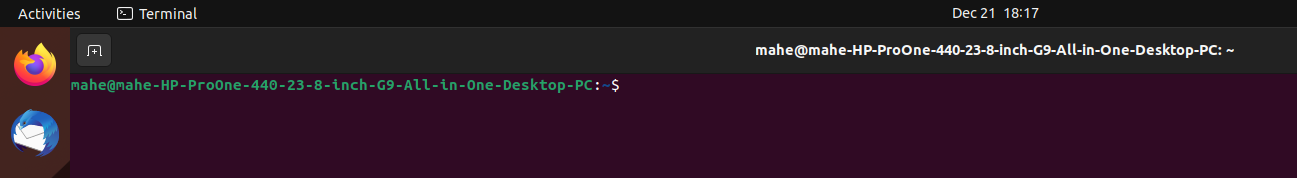
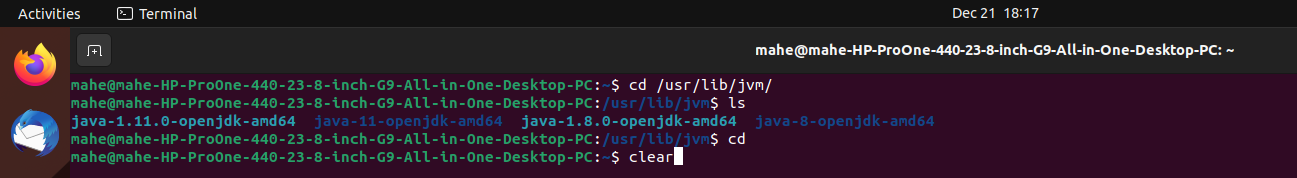
5. sudo apt install openjdk-8-jdk

6. cd /usr/lib/jvm/

7. ls

8. cd

9. clear

10. sudo nano .bashrc

11. export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

export PATH=$PATH:/usr/lib/jvm/java-8-openjdk-amd64/bin

export HADOOP\_HOME=~/hadoop-3.3.6/

export PATH=$PATH:$HADOOP\_HOME/bin

export PATH=$PATH:$HADOOP\_HOME/sbin

export HADOOP\_MAPRED\_HOME=$HADOOP\_HOME

export YARN\_HOME=$HADOOP\_HOME

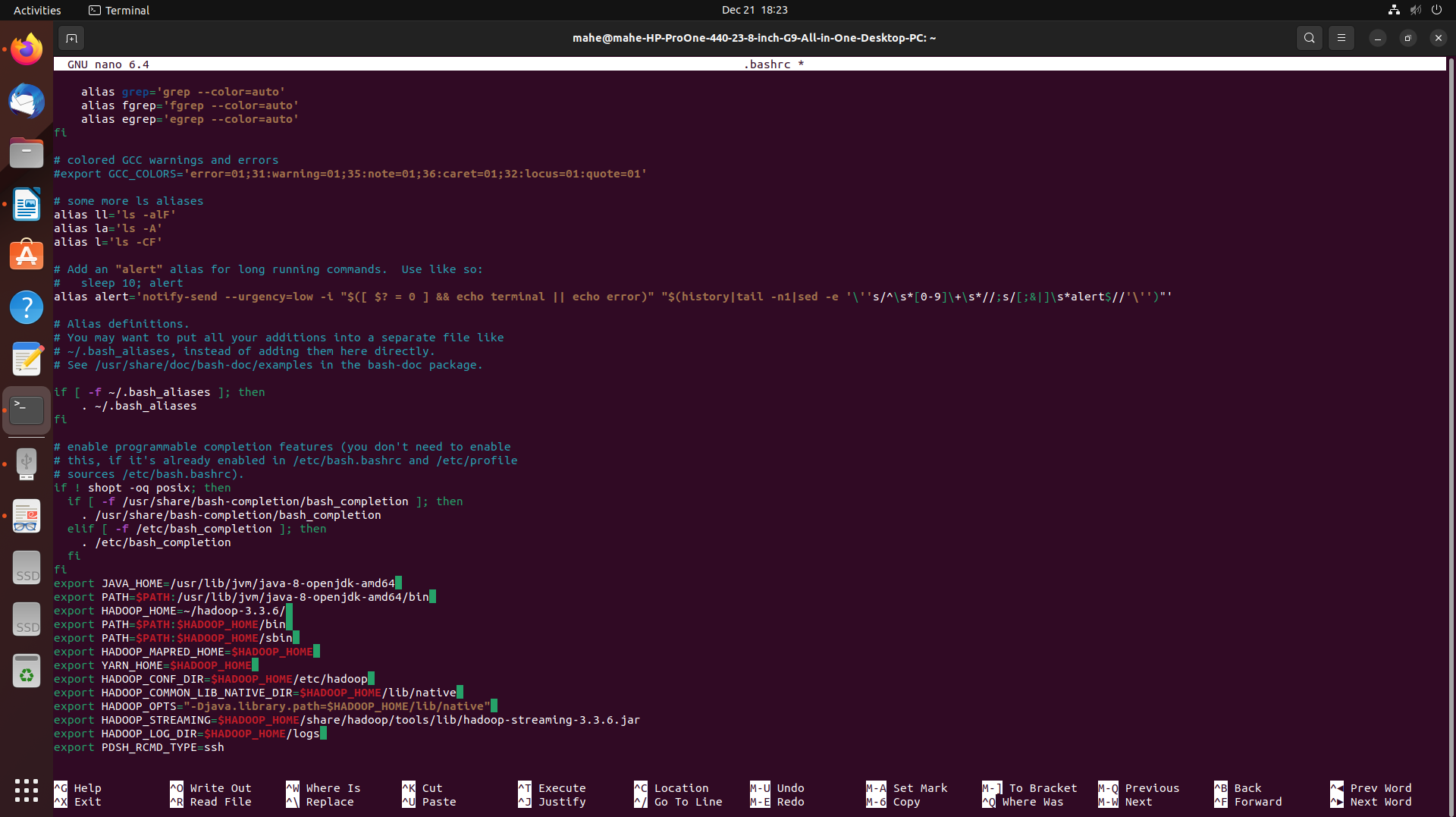
export HADOOP\_CONF\_DIR=$HADOOP\_HOME/etc/hadoop

export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_HOME/lib/native

export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_HOME/lib/native"

export HADOOP\_STREAMING=$HADOOP\_HOME/share/hadoop/tools/lib/hadoop-streaming-3.3.6.jar

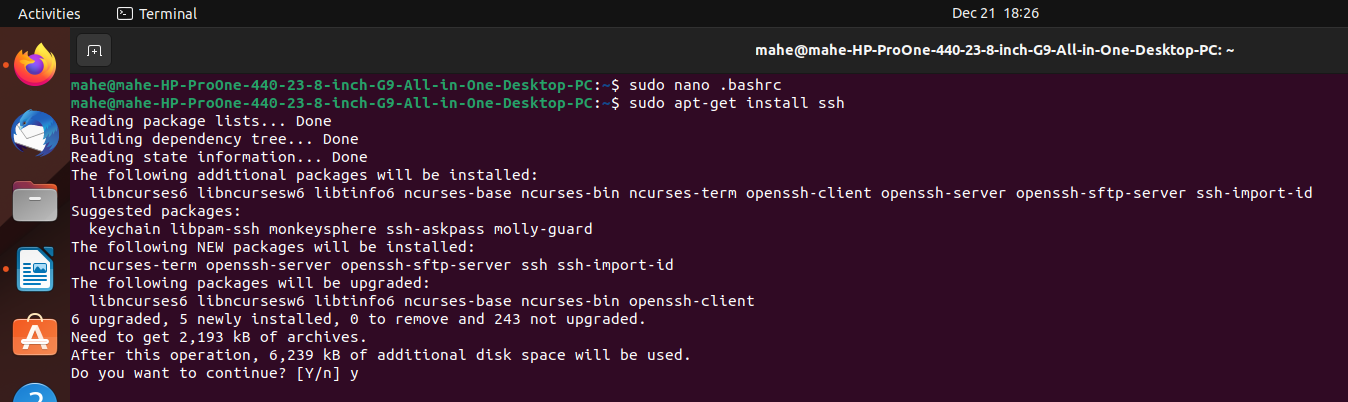
export HADOOP\_LOG\_DIR=$HADOOP\_HOME/logs

export PDSH\_RCMD\_TYPE=ssh

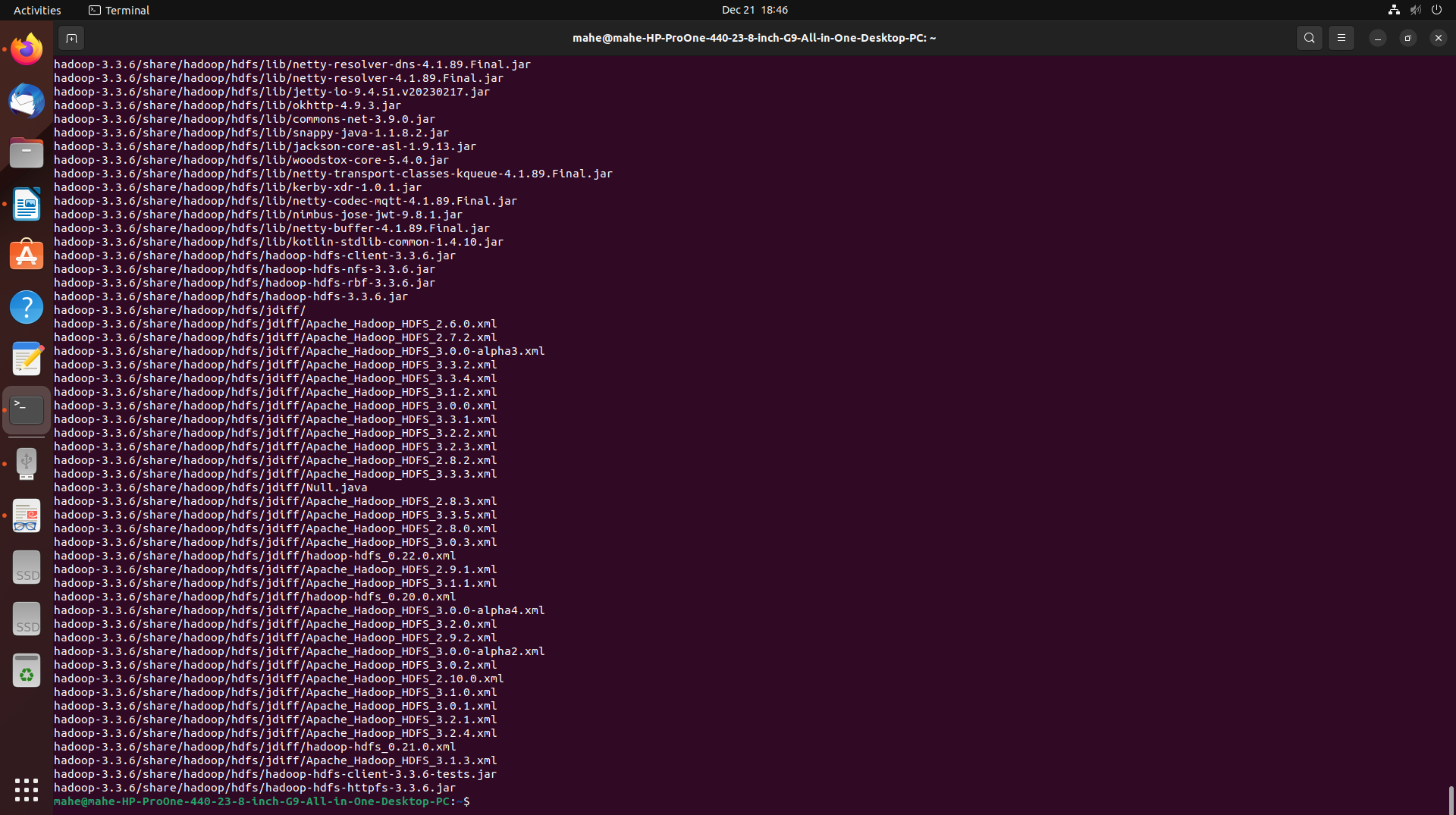
12. ctrl+o

13. ctrl+x

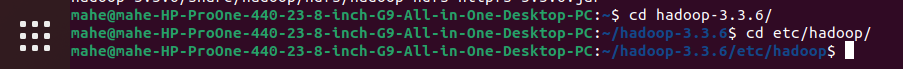
14. sudo apt-get install ssh

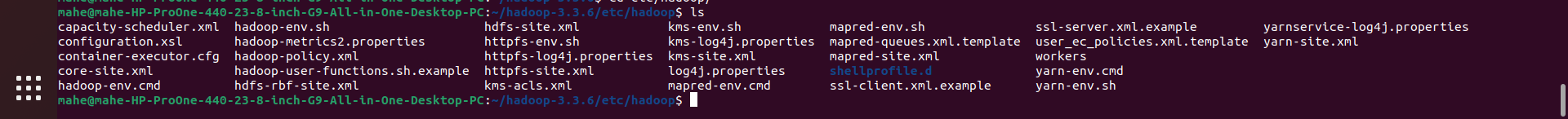
15. download <https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz>

16. tar -zxvf ~/Downloads/hadoop-3.3.6.tar.gz

17. cd hadoop-3.3.6/

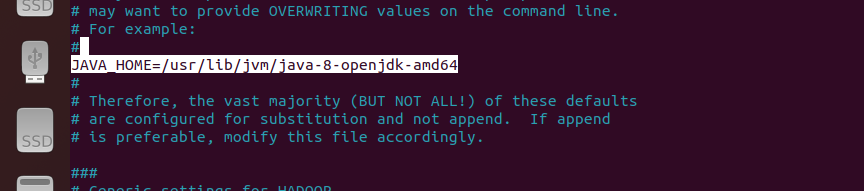
18. cd etc/hadoop/

19. ls



20. sudo nano hadoop-env.sh

21. JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

22. sudo nano core-site.xml

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value> </property>

<property>

<name>hadoop.proxyuser.dataflair.groups</name> <value>\*</value>

</property>

<property>

<name>hadoop.proxyuser.dataflair.hosts</name> <value>\*</value>

</property>

<property>

<name>hadoop.proxyuser.server.hosts</name> <value>\*</value>

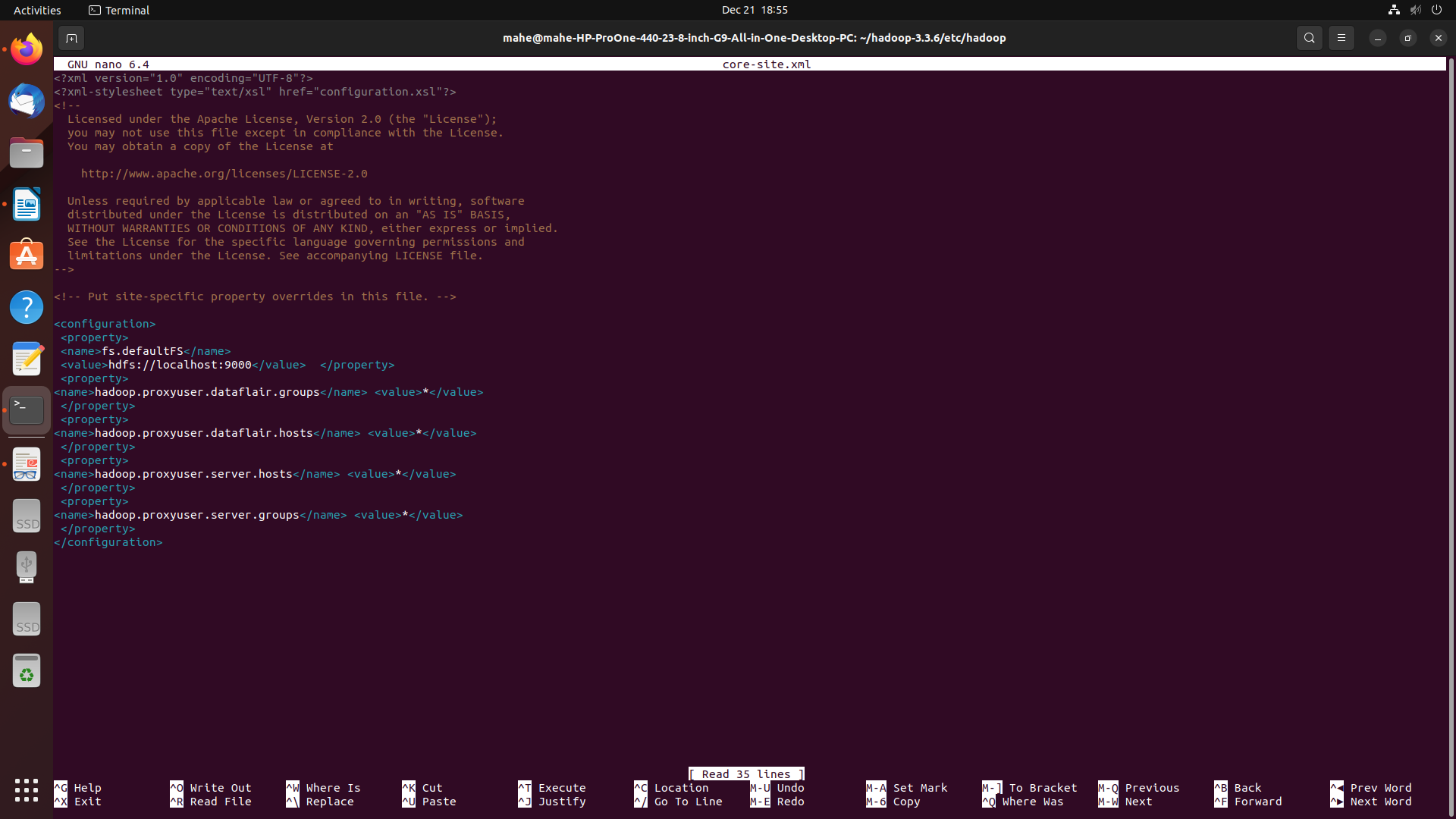
</property>

<property>

<name>hadoop.proxyuser.server.groups</name> <value>\*</value>

</property>

</configuration>

23. sudo nano hdfs-site.xml

<configuration>

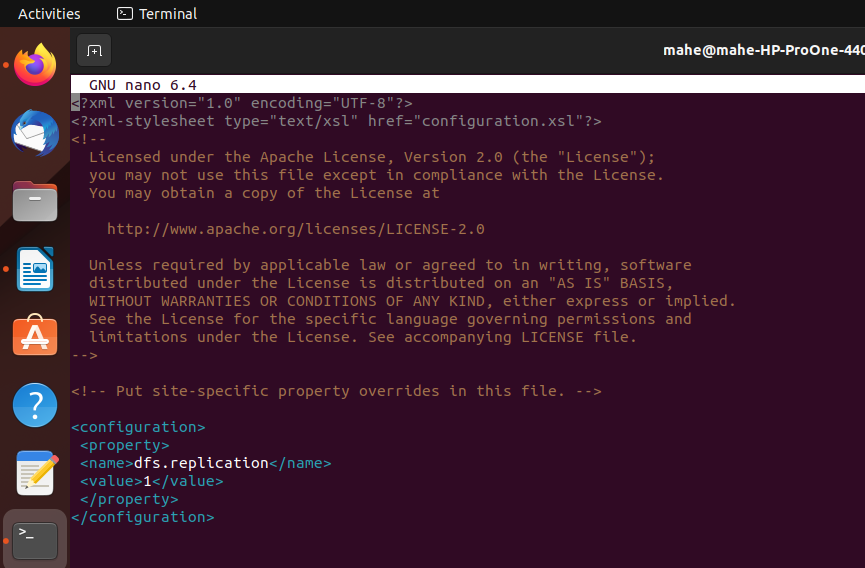
<property>

<name>dfs.replication</name>

<value>1</value>

</property>

</configuration>

24. sudo nano mapred-site.xml

<configuration>

<property>

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

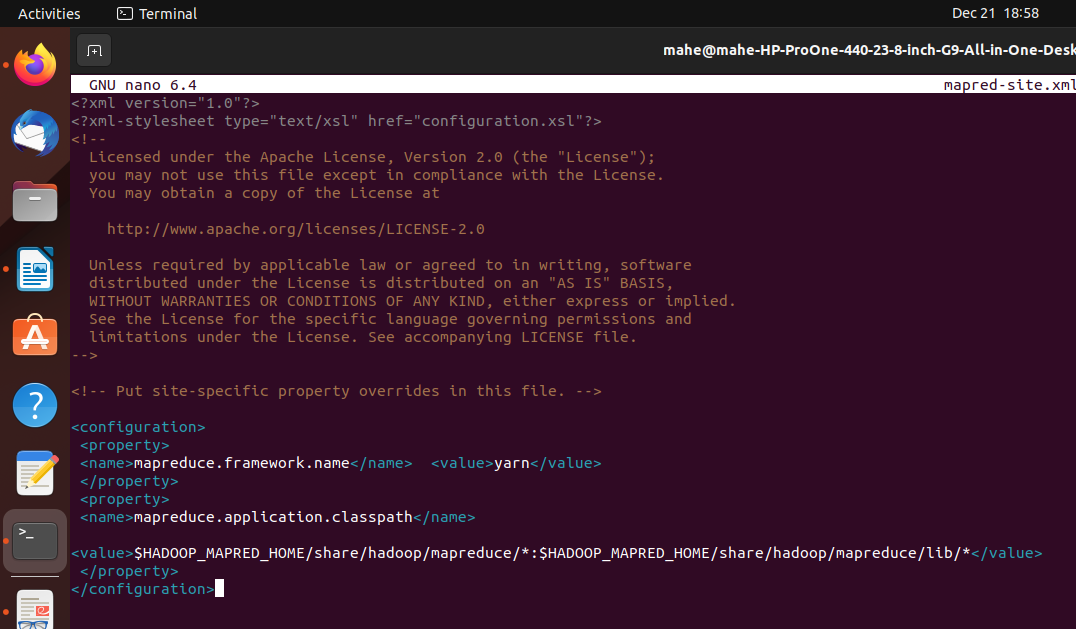
</property>

<property>

<name>mapreduce.application.classpath</name>

<value>$HADOOP\_MAPRED\_HOME/share/hadoop/mapreduce/\*:$HADOOP\_MAPRED\_HOME/share/hadoop/mapreduce/lib/\*</value>

</property>

</configuration>

25. sudo nano yarn-site.xml

<configuration>

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

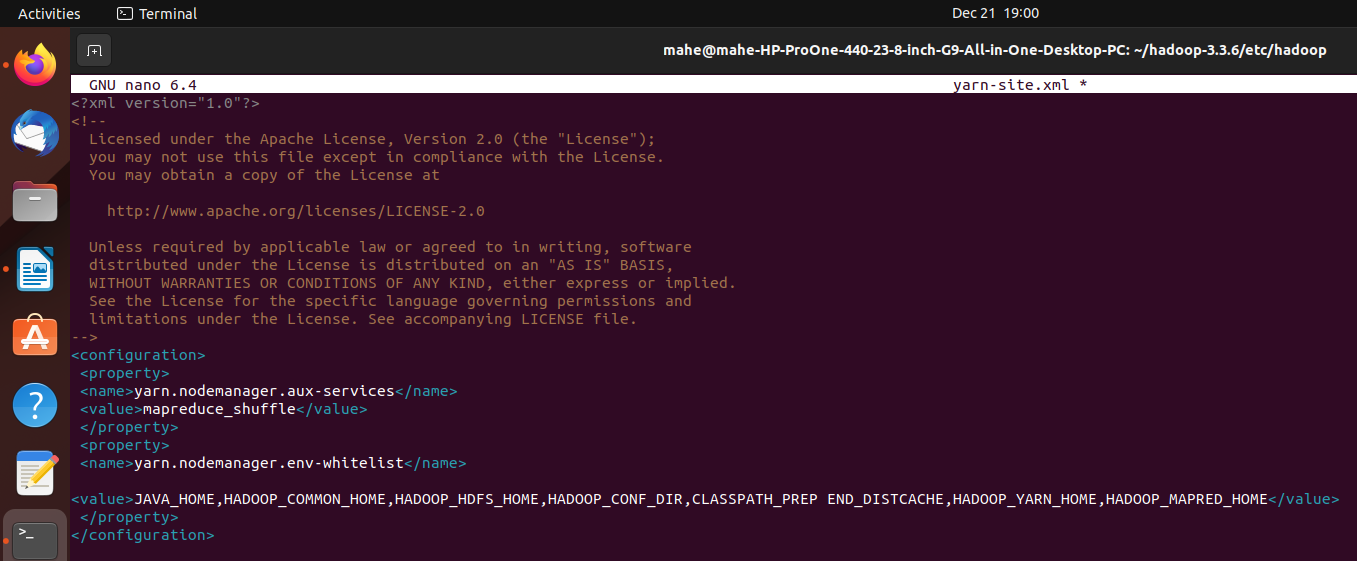
<property>

<name>yarn.nodemanager.env-whitelist</name>

<value>JAVA\_HOME,HADOOP\_COMMON\_HOME,HADOOP\_HDFS\_HOME,HADOOP\_CONF\_DIR,CLASSPATH\_PREP END\_DISTCACHE,HADOOP\_YARN\_HOME,HADOOP\_MAPRED\_HOME</value>

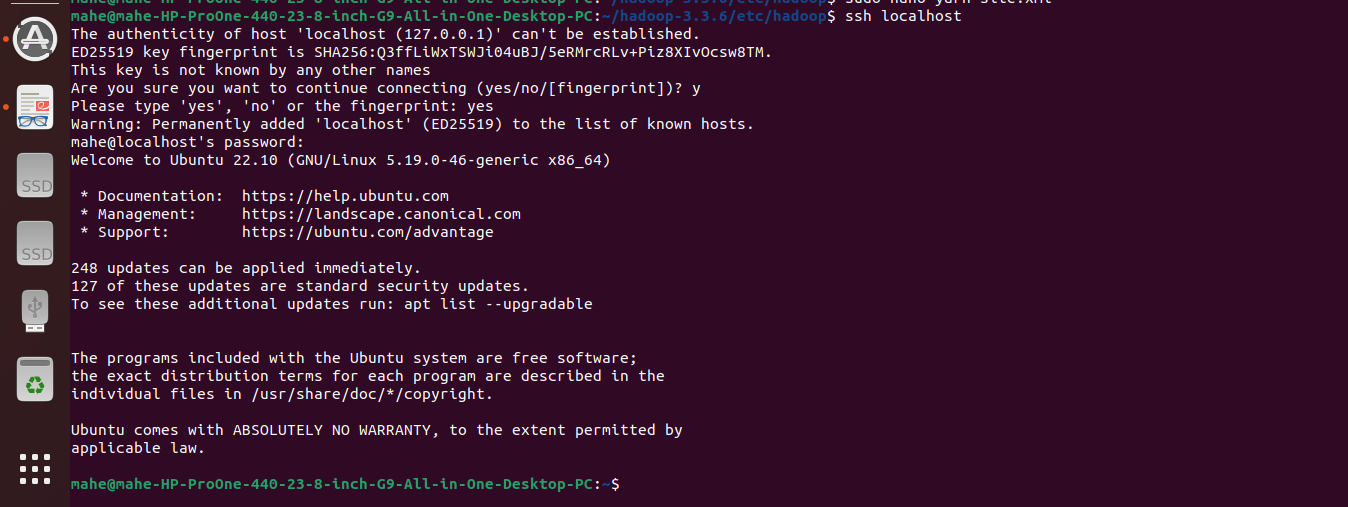
</property>

</configuration>

26. ssh localhost

pw:[Mahe@123](mailto:Mahe@123)

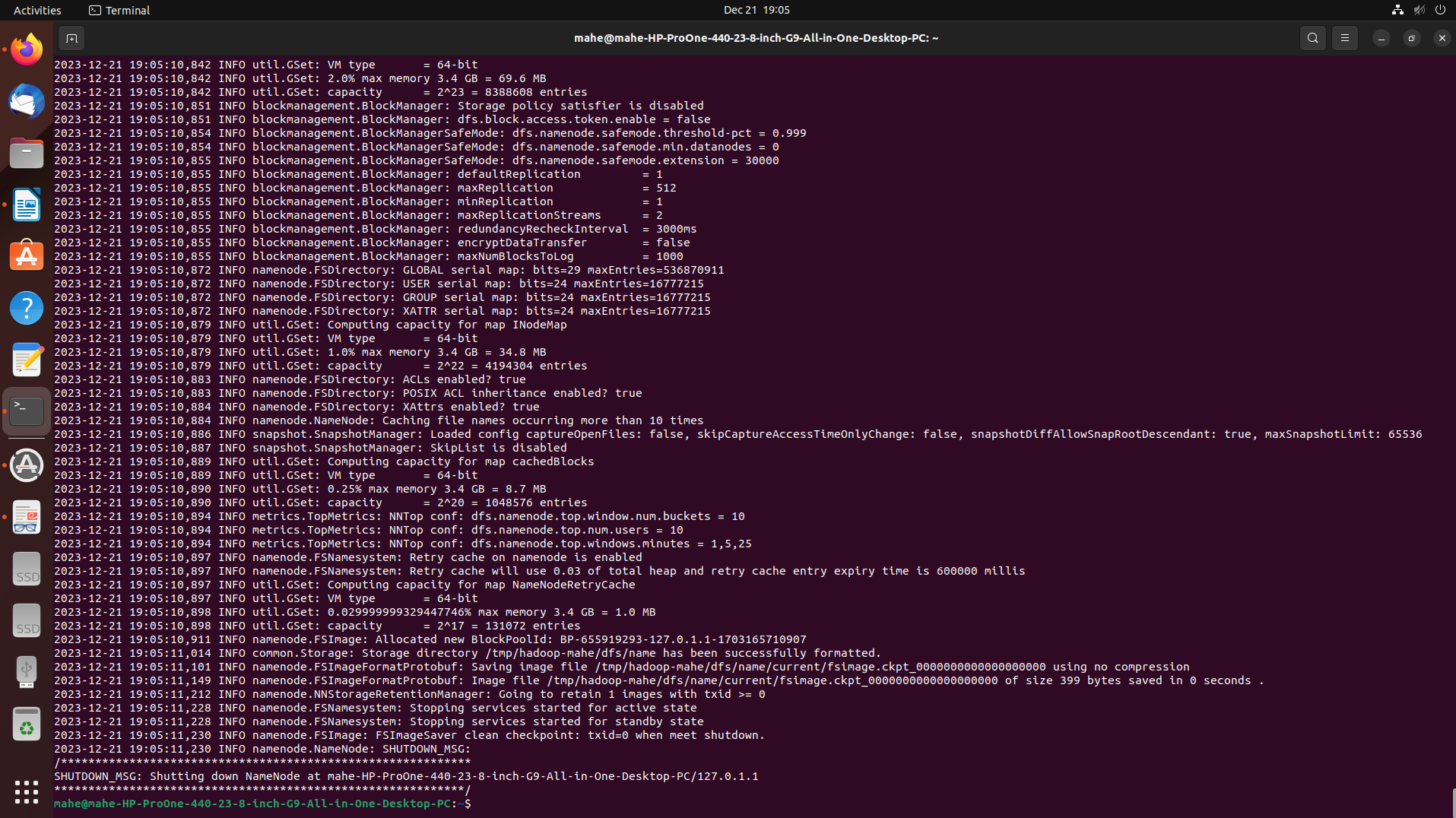
type yes

27. ssh-keygen -t rsa -P '' -f ~/.ssh/id\_rsa

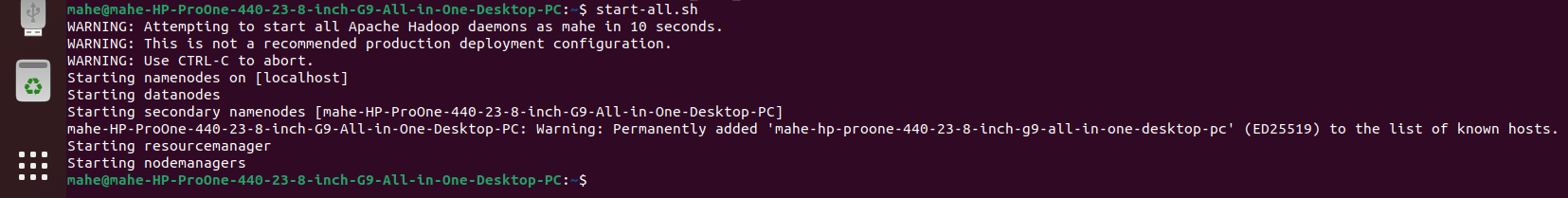
cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys

28. chmod 0600 ~/.ssh/authorized\_keys

29. hadoop-3.3.6/bin/hdfs namenode -format

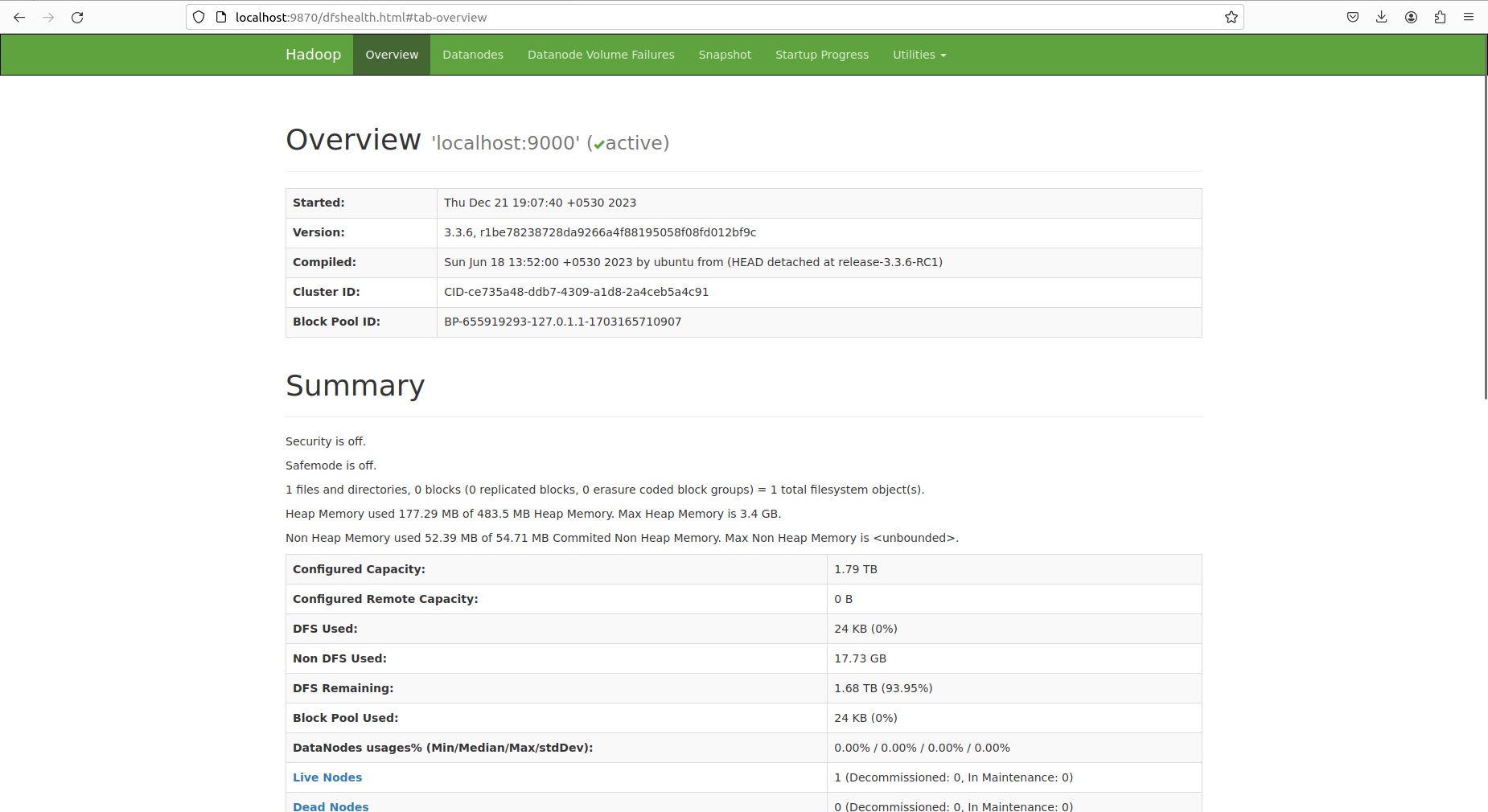
30. export PDSH\_RCMD\_TYPE=ssh

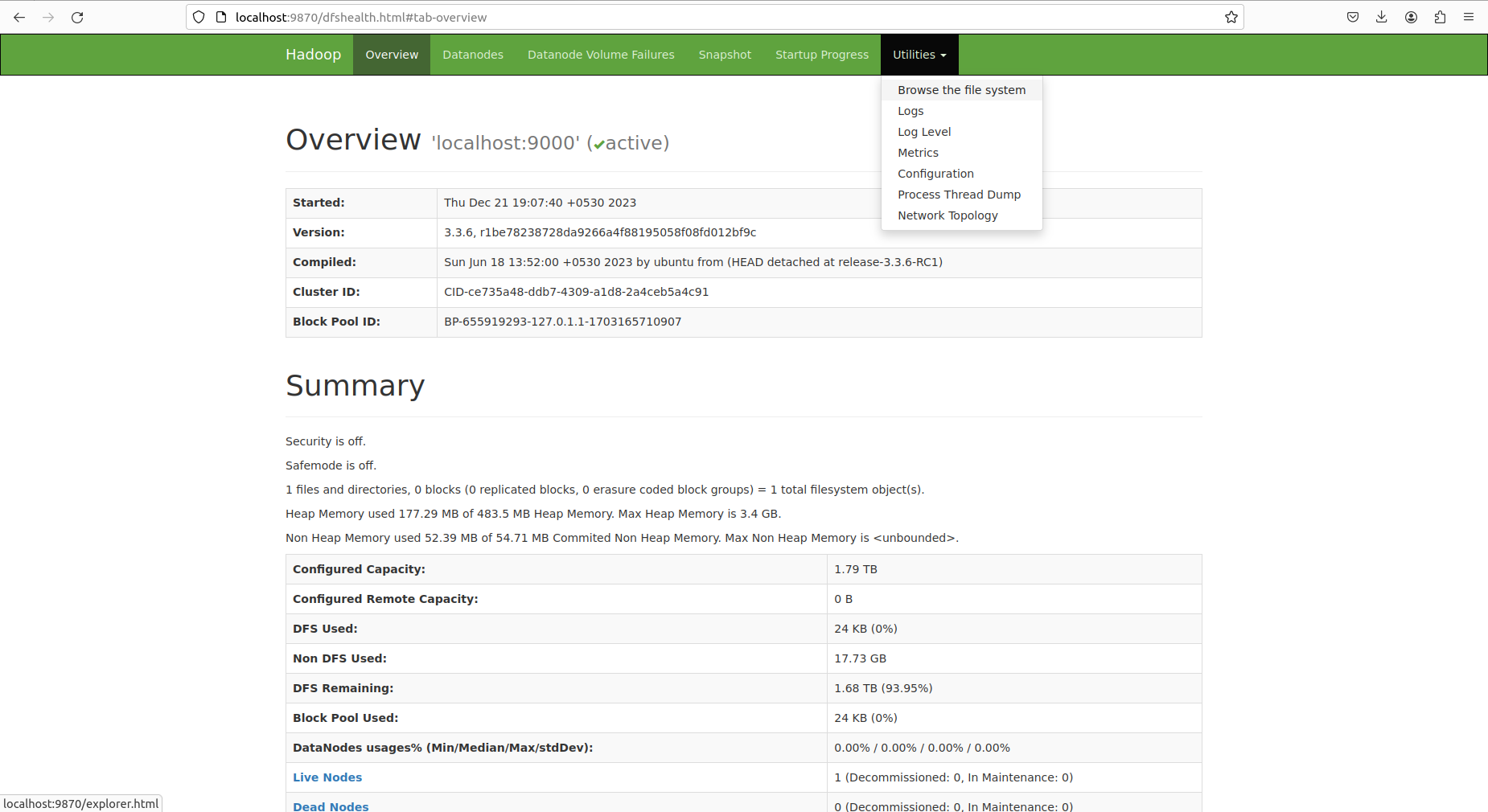
31. start-all.sh

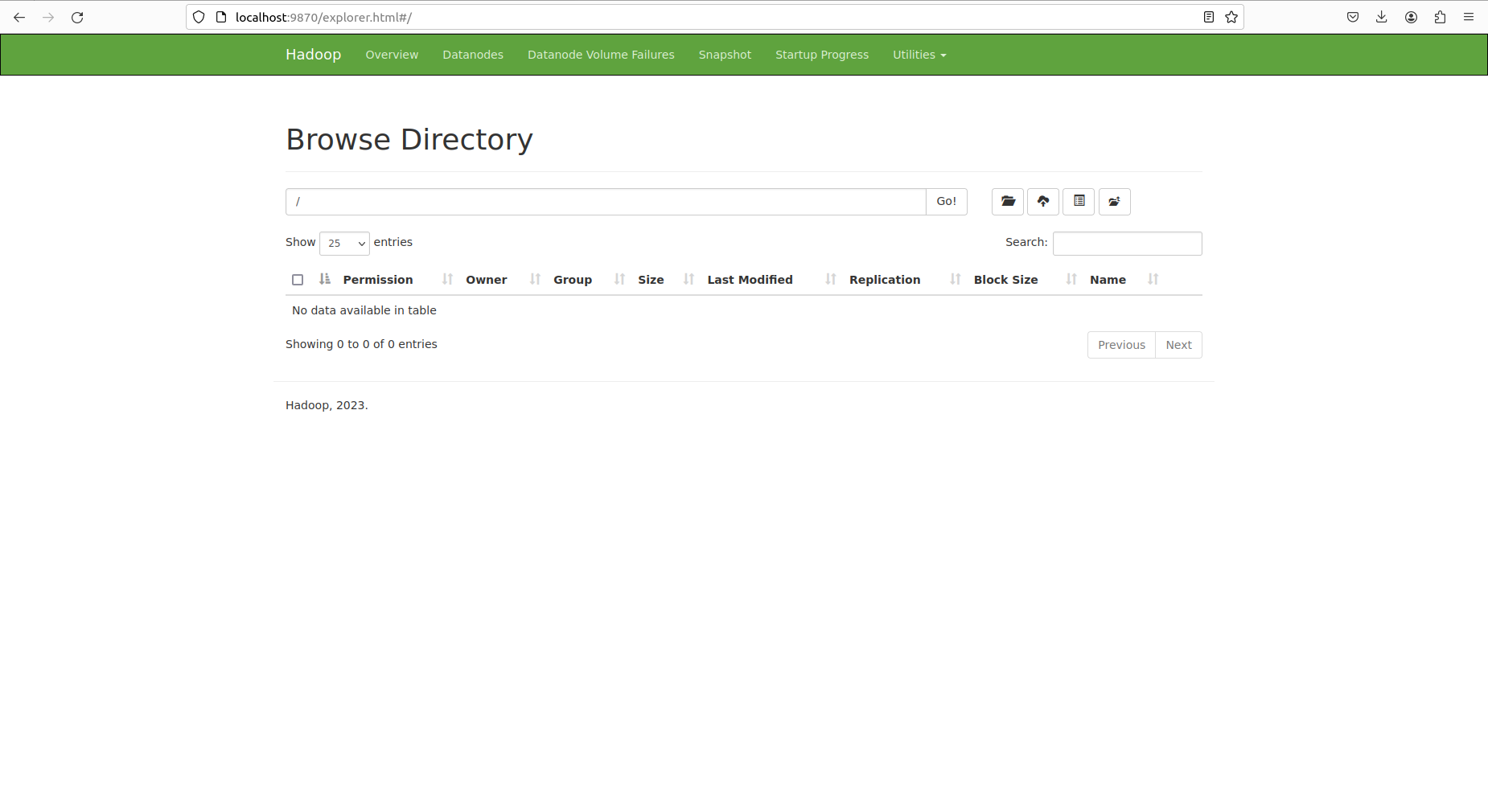


localhost:9870

copy above shown localhost and paste in any browser

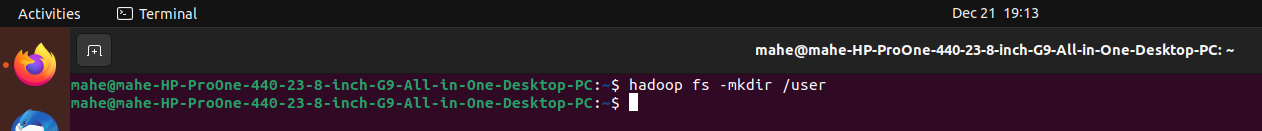


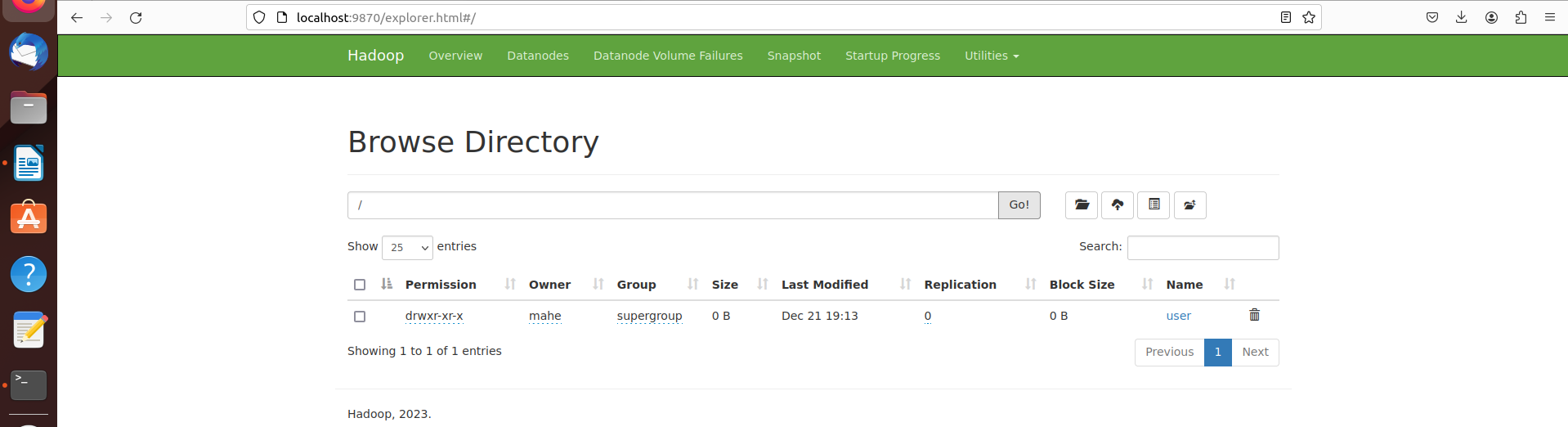




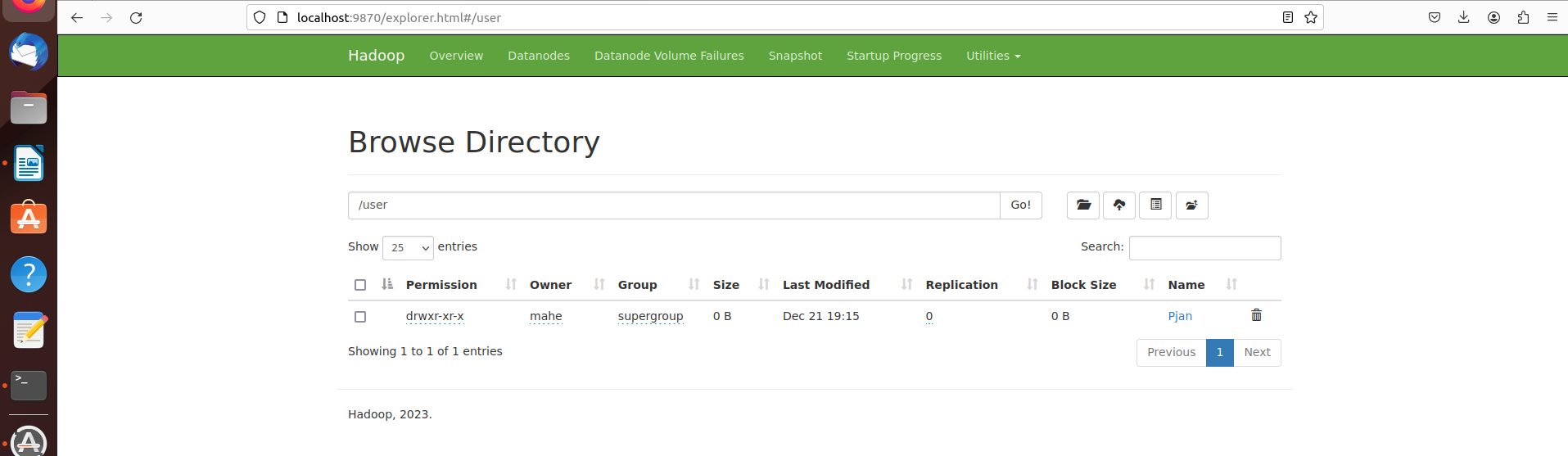
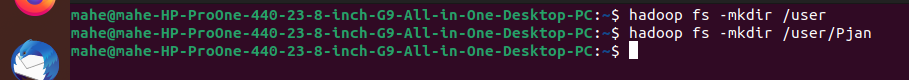
32. clear

33. hadoop fs -mkdir /user

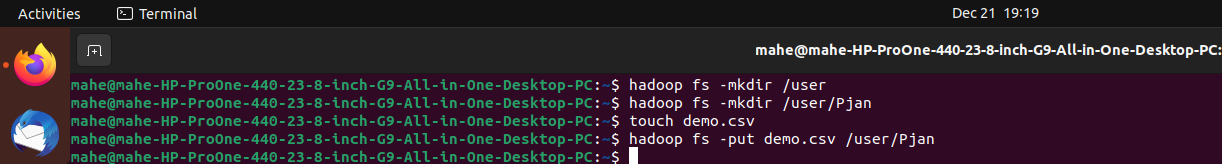
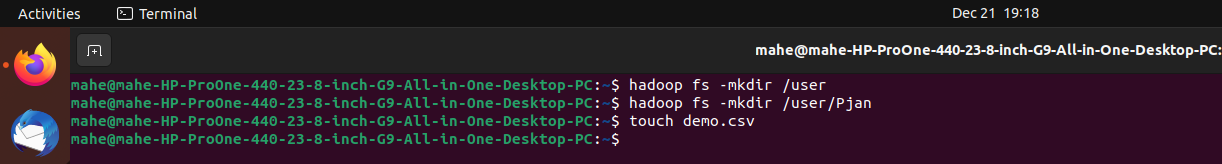


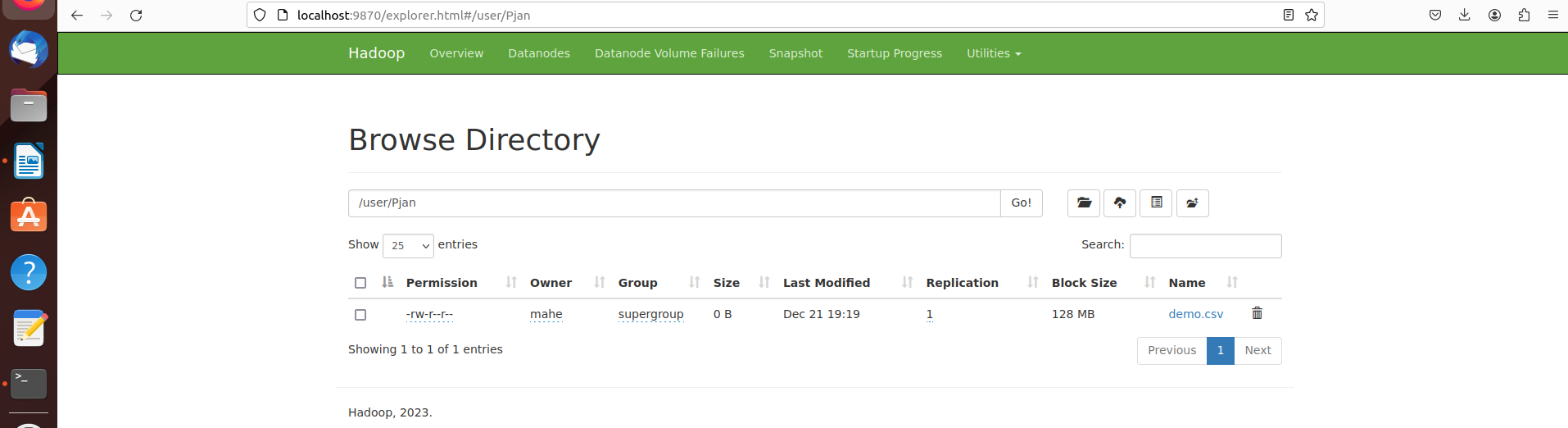


34. hadoop fs -mkdir /user/Pjan



35. touch demo.csv

36. hadoop fs -put demo.csv /user/Pjan



32. to stop use : stop-all.sh