Hadoop Installation

→ To check whether java is present or not, enter the command > which java

op: /usr/bin/java

→ For the installation of hadoop download the hadoop latest version (hadoop 2.9.2) from the google in the .tar.gz format (archive format)

→ From the downloads, decompress the file from .tar.gz format to .tar

>gunzip hadoop-2.9.2.tar.gz

op: hadoop-2.9.2.tar

→ Now decompress the tar file.

> tar xvf hadoop-2.9.2.tar

op: hadoop-2.9.2

→ To ~~install~~ check hadoop, go to hadoop-2.9.2 → bin → hadoop

→ To check the updates present or not

sudo apt-get install updates

op: Enter password: @Prec@123

→ sudo apt-get install ssh

→ sudo apt-get install rsync

→ Now go to the hadoop file

→ hadoop → etc → hadoop → hadoop-env.sh

check whether the java home directory is present as /usr or not.

→ In the etc → configuration.xsl write the following code about <html>

```
<configuration>
<property>
<name> fs.defaultFS </name>
<value> hdfs://localhost:9000 </value>
</property>
</configuration>
```

→ In hdfs-site.xml

```
<configuration>
<property>
<name> dfs.replication </name>
<value> 1 </value>
</property>
<property>
<name> fs.permission </name>
<value> false </value>
</property>
</configuration>
```

→ In the terminal we need to check shellscript

> ssh localhost > echo ssh

ssh

→ cd /home/eclab220/hadoop-2.9.2/hadoop2

→ > ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
(rsa → space)

→ > cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
(. → dot)

→ > chmod 0600 ~/.ssh/authorized_keys

→ > bin/hdfs namenode -format

↳ error occurs if we are not in hadoop file directory

→ > sbin/start-dfs.sh → enter password

gprec@123 3 times

→ Go to browser http://localhost:50070/

↳ To check whether hadoop is installed or onot

↓

or

type > jps in terminal

Op: JPS

secondary name node

Data node

name node.

→ jps

op: ---- Jps

→ echo hadoop

op: hadoop

→ echo ssh

op: ssh

→ echo sshd

op: sshd

→ ~~sbin/stop-dfs.sh~~

cd hadoop-2.9.2/.

→ sbin/stop-dfs.sh

→ bin/hadoop namenode -format

→ sbin/start-dfs.sh (password: gpreece123)

→ jps

Commands

→ hadoop fs -ls

op: No such file or directory

→ bin/hadoop fs -mkdir /demo (→ space)

→ Consider some file xyz (eg: NOTICE.txt)

bin/hadoop fs -put /home/cclab220/hadoop-2.9.2/ (→ space)
NOTICE.txt /demo (→ space)

→ bin/hadoop fs -ls /demo → shows permissions

→ bin/hadoop fs -cat /demo

To Run MapReduce Job, write the below code:

(bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.9.2.jar grep /demo/NOTICE.txt /myoutput 'This')

bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.9.2.jar grep /demo/NOTICE.txt /output 'This'