

EX.No:4 Install Hadoop single mode cluster and run simple applications

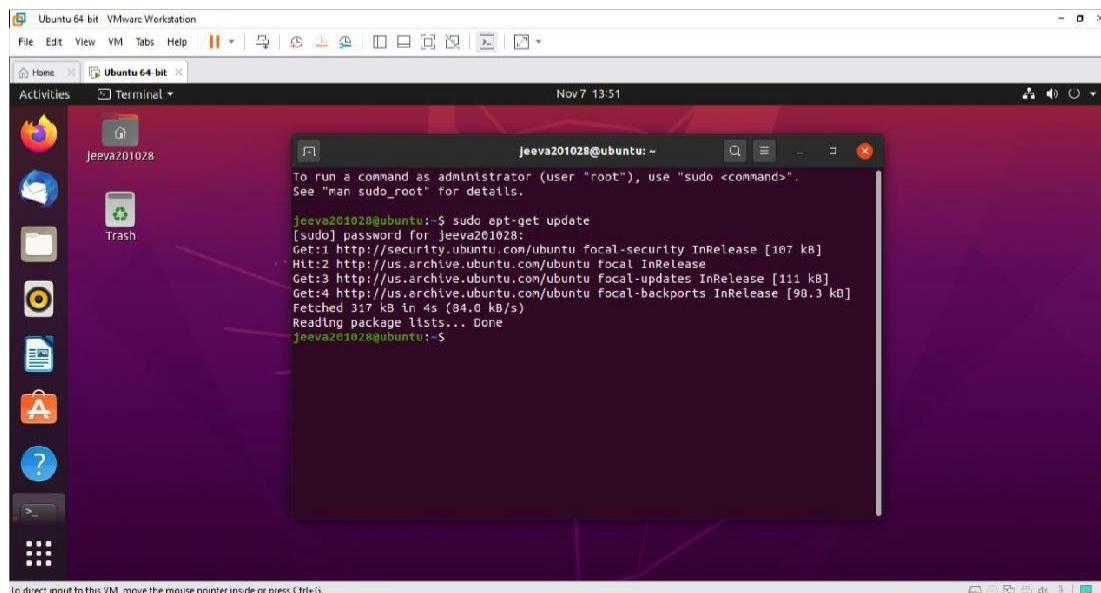
Date:

Aim:

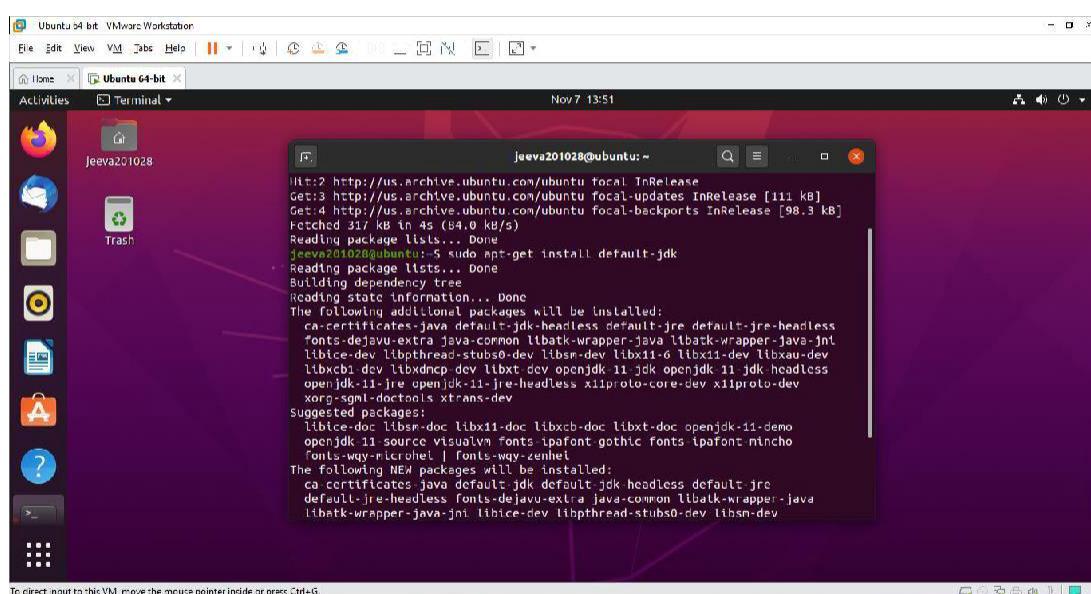
To install hadoop single mode cluster and run a simple word count application.

Procedure:

Update the packages



Installing java



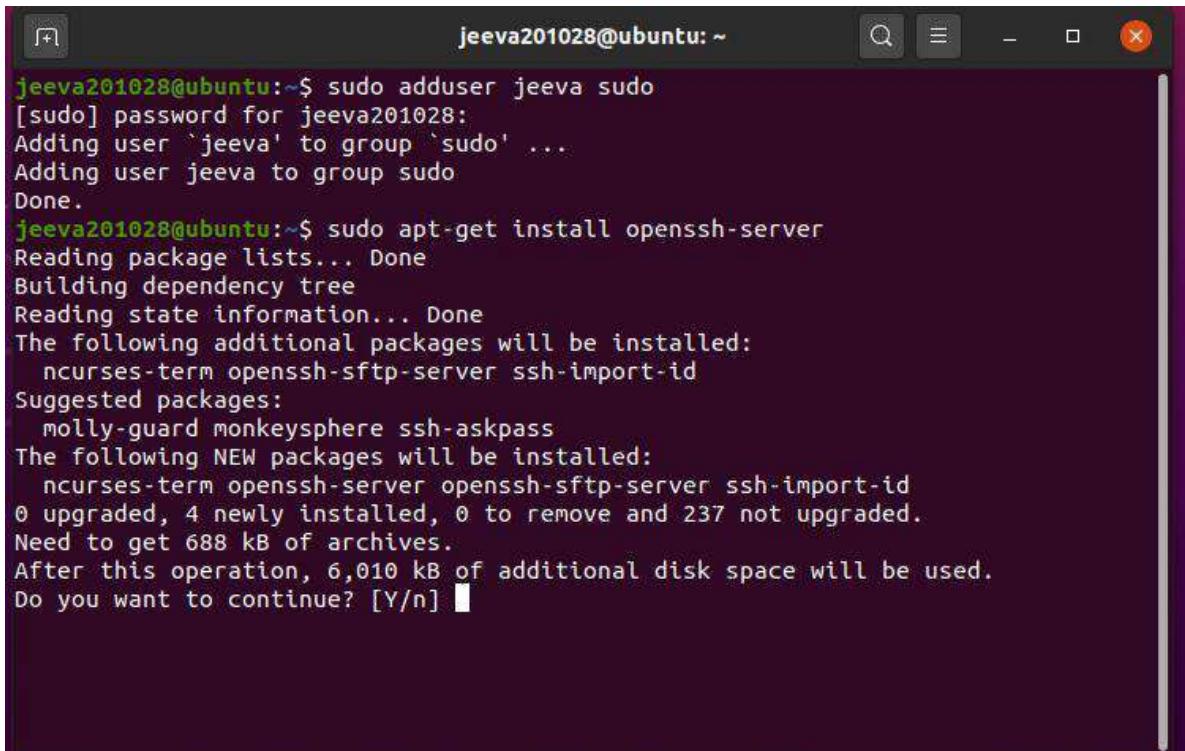
Assigning a dedicated user on hadoop to perform operations

```
jeeva201028@ubuntu:~$ sudo addgroup hadoop
[sudo] password for jeeva201028:
Adding group `hadoop' (GID 1001) ...
Done.
jeeva201028@ubuntu:~$ sudo adduser --ingroup hadoop jeeva
Adding user `jeeva' ...
Adding new user `jeeva' (1001) with group `hadoop' ...
Creating home directory `/home/jeeva' ...
Copying files from `/etc/skel' ...
New password:
Retype new password:
passwd: password updated successfully
Changing the user information for jeeva
Enter the new value, or press ENTER for the default
      Full Name []: Jeeva
      Room Number []: 1
      Work Phone []: 98765
      Home Phone []: 43210
      Other []:
Is the information correct? [Y/n] Y
jeeva201028@ubuntu:~$
```

Adding user to sudo list

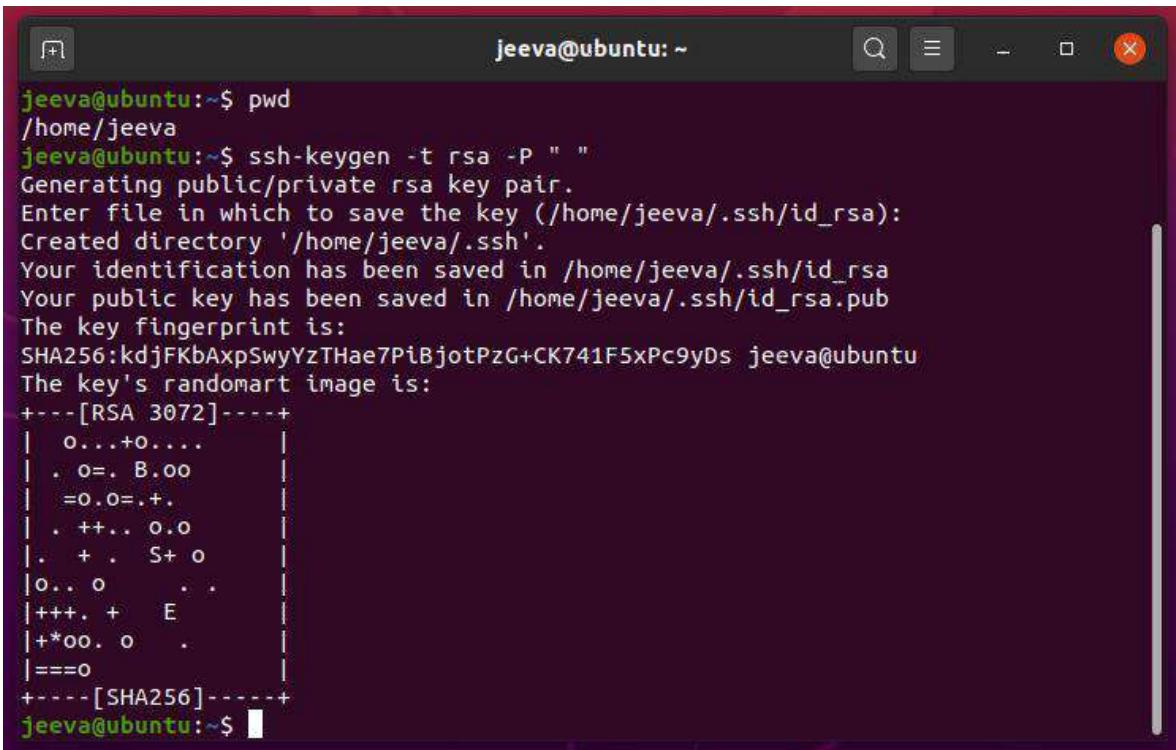
```
jeeva201028@ubuntu:~$ sudo adduser jeeva sudo
[sudo] password for jeeva201028:
Adding user `jeeva' to group `sudo' ...
Adding user jeeva to group sudo
Done.
jeeva201028@ubuntu:~$
```

Next install a package ssh(secured shell login)



```
jeeva201028@ubuntu:~$ sudo adduser jeeva sudo
[sudo] password for jeeva201028:
Adding user `jeeva' to group `sudo' ...
Adding user jeeva to group sudo
Done.
jeeva201028@ubuntu:~$ sudo apt-get install openssh-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ncurses-term openssh-sftp-server ssh-import-id
Suggested packages:
  molly-guard monkeysphere ssh-askpass
The following NEW packages will be installed:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
0 upgraded, 4 newly installed, 0 to remove and 237 not upgraded.
Need to get 688 kB of archives.
After this operation, 6,010 kB of additional disk space will be used.
Do you want to continue? [Y/n] 
```

Next step is key generation and add a key to the file



```
jeeva@ubuntu:~$ pwd
/home/jeeva
jeeva@ubuntu:~$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/jeeva/.ssh/id_rsa):
Created directory '/home/jeeva/.ssh'.
Your identification has been saved in /home/jeeva/.ssh/id_rsa
Your public key has been saved in /home/jeeva/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:kdjFKbAxpSwyYzTHae7PiBjotPzG+CK741F5xPc9yDs jeeva@ubuntu
The key's randomart image is:
+---[RSA 3072]----+
| o...+o.... |
| . o=. B.oo |
| =o.o=+. |
| . +... o.o |
| . + . S+ o |
| o.. o . . |
| ++++ + E |
| +*oo. o . |
| ===o |
+---[SHA256]-----+
jeeva@ubuntu:~$ 
```

To check whether ssh is installed properly by logging in and after that exit from it

```
jeeva201028@ubuntu:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:3w08Zvw2Q8dtjtAaSLLWahjenZ7zywgGfYjr0Im8eYM.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
jeeva201028@localhost's password:
Welcome to Ubuntu 20.04.1 LTS (GNU/Linux 5.4.0-52-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

237 updates can be installed immediately.
100 of these updates are security updates.
To see these additional updates run: apt list --upgradable

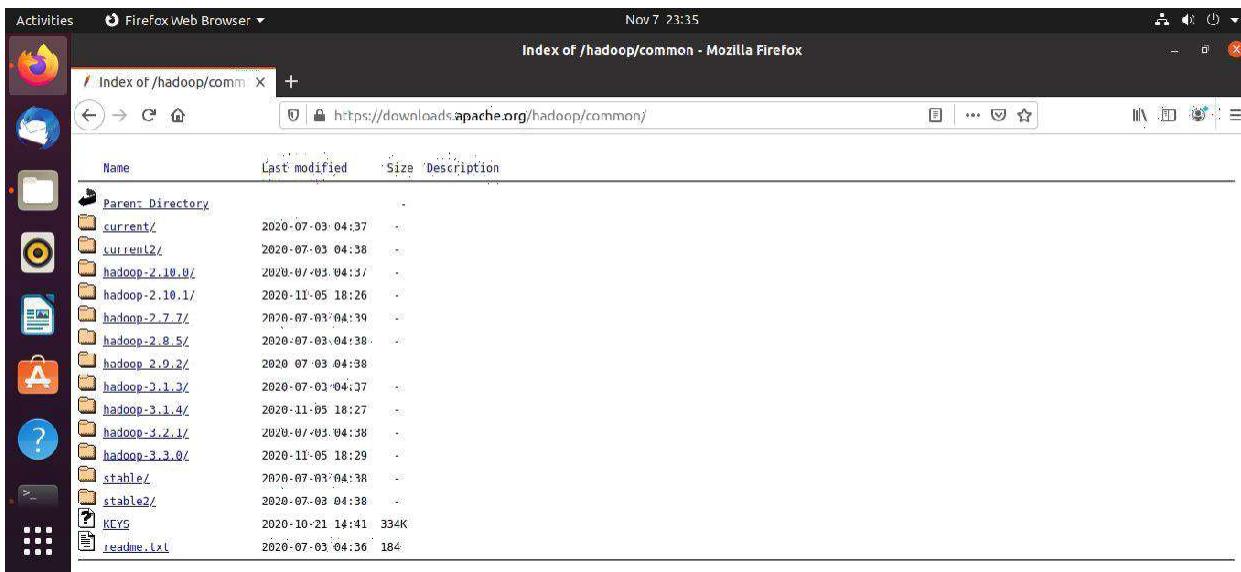
Your Hardware Enablement Stack (HWE) is supported until April 2025.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*copyright.

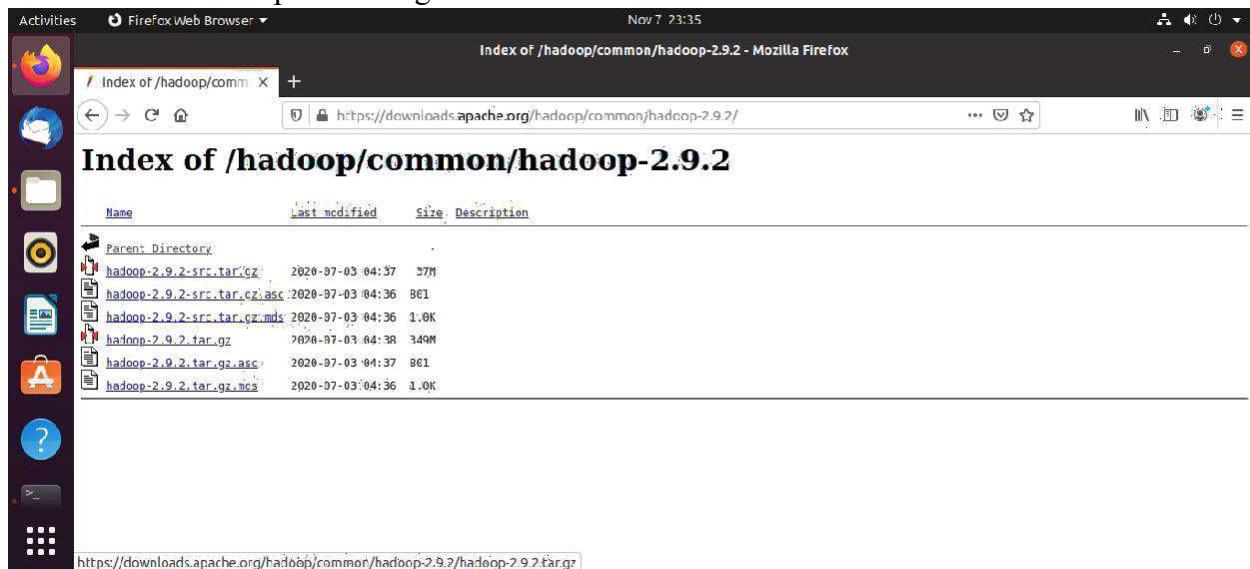
Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

jeeva201028@ubuntu:~$ exit
logout
Connection to localhost closed.
```

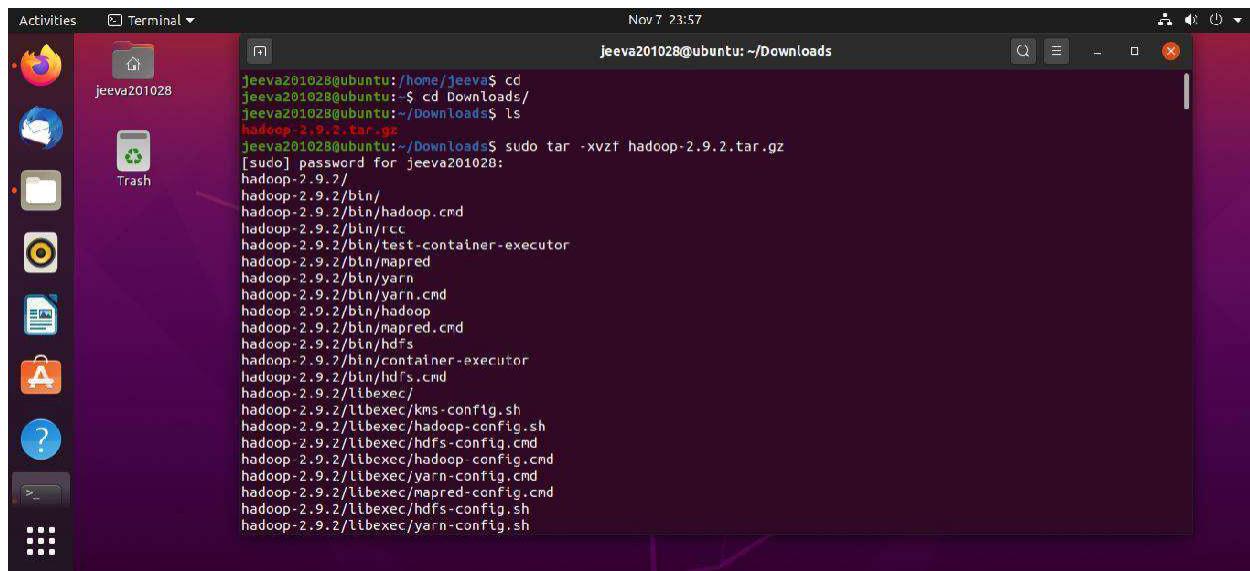
Download Hadoop using this link : <https://downloads.apache.org/hadoop/common/>



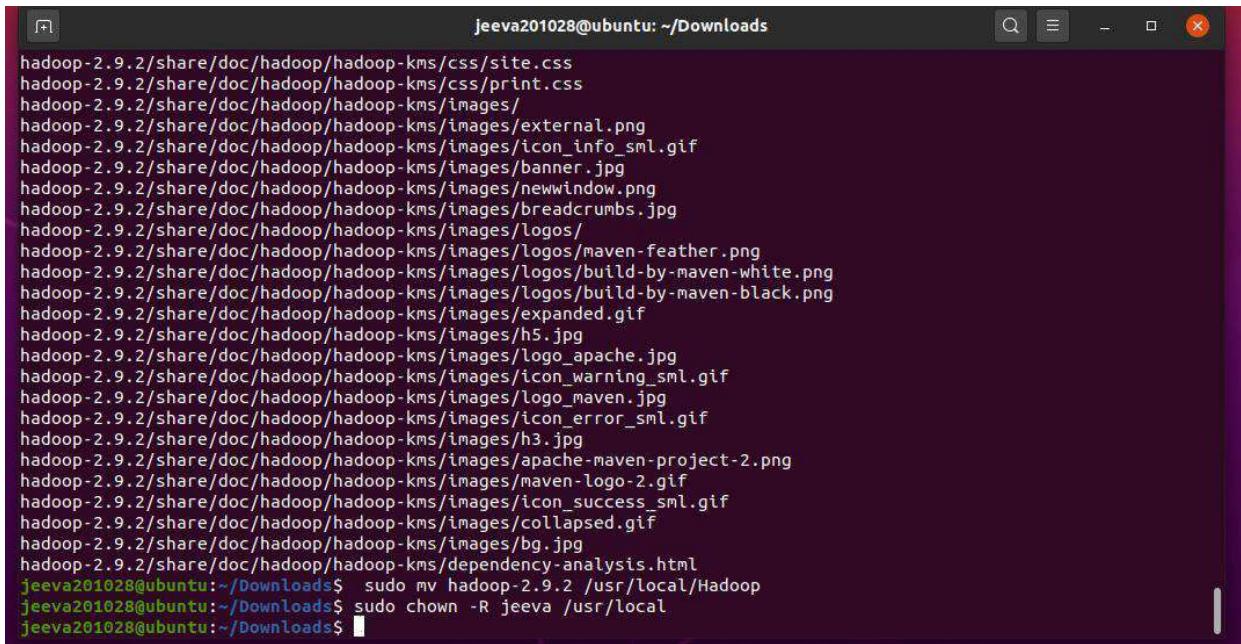
Download the hadoop-2.9.2.tar.gz file



Now extract the tar file

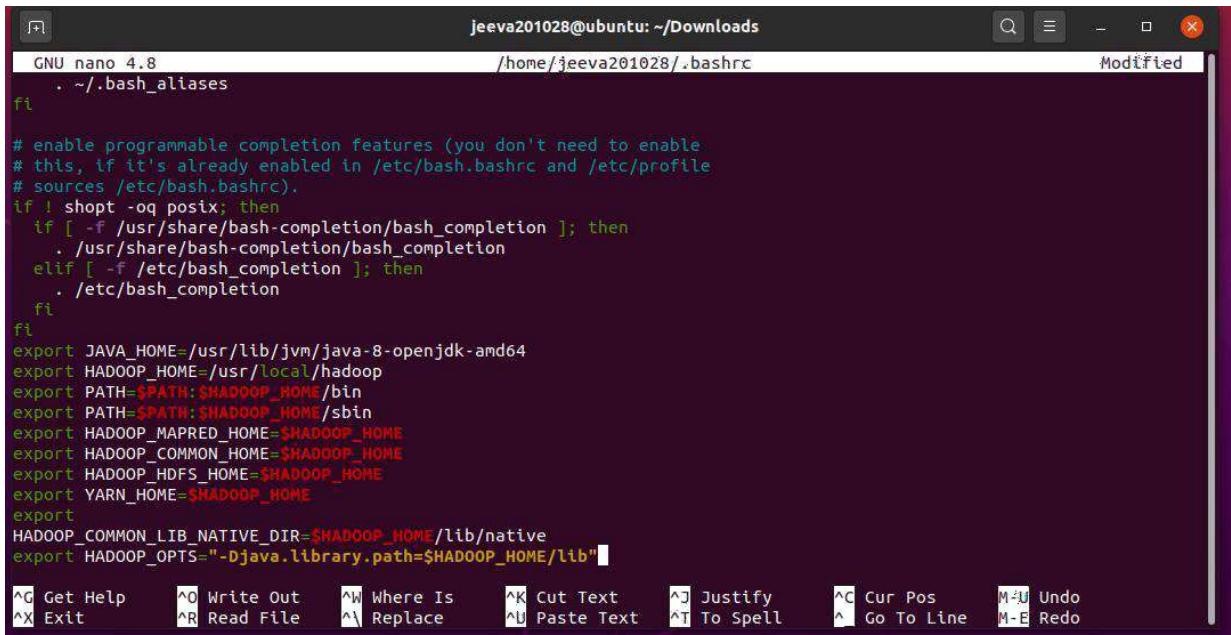


Move the file to local and Change the ownership of hadoop folder



```
jeeva201028@ubuntu: ~/Downloads
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/css/site.css
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/css/print.css
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/external.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/icon_info_sml.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/banner.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/newwindow.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/breadcrumbs.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/logos/
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/logos/maven-feather.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/logos/build-by-maven-white.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/logos/build-by-maven-black.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/expanded.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/h5.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/Logo_apache.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/icon_warning_sml.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/logo_maven.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/icon_error_sml.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/h3.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/apache-maven-project-2.png
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/maven-logo-2.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/icon_success_sml.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/collapsed.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/bg.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/dependency-analysis.html
jeeva201028@ubuntu:~/Downloads$ sudo mv hadoop-2.9.2 /usr/local/Hadoop
jeeva201028@ubuntu:~/Downloads$ sudo chown -R jeeva /usr/local
jeeva201028@ubuntu:~/Downloads$
```

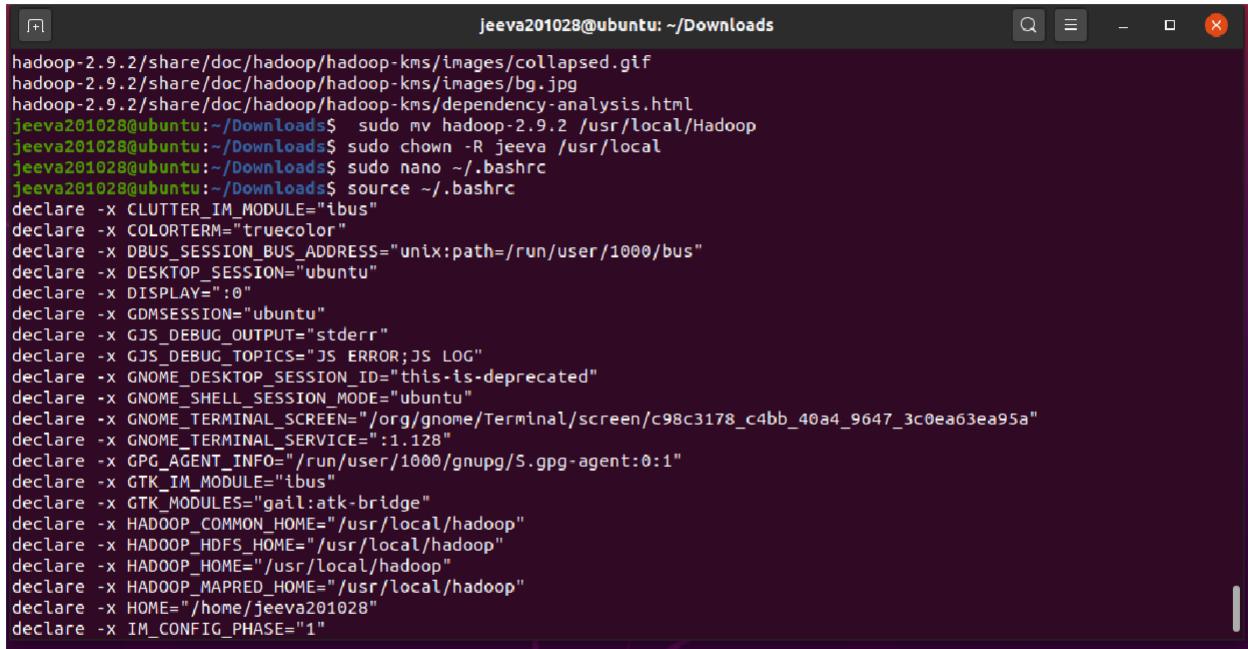
In .bashrc file add the following and make it at source



```
GNU nano 4.8
jeeva201028@ubuntu: ~/Downloads
/home/jeeva201028/.bashrc
Modified

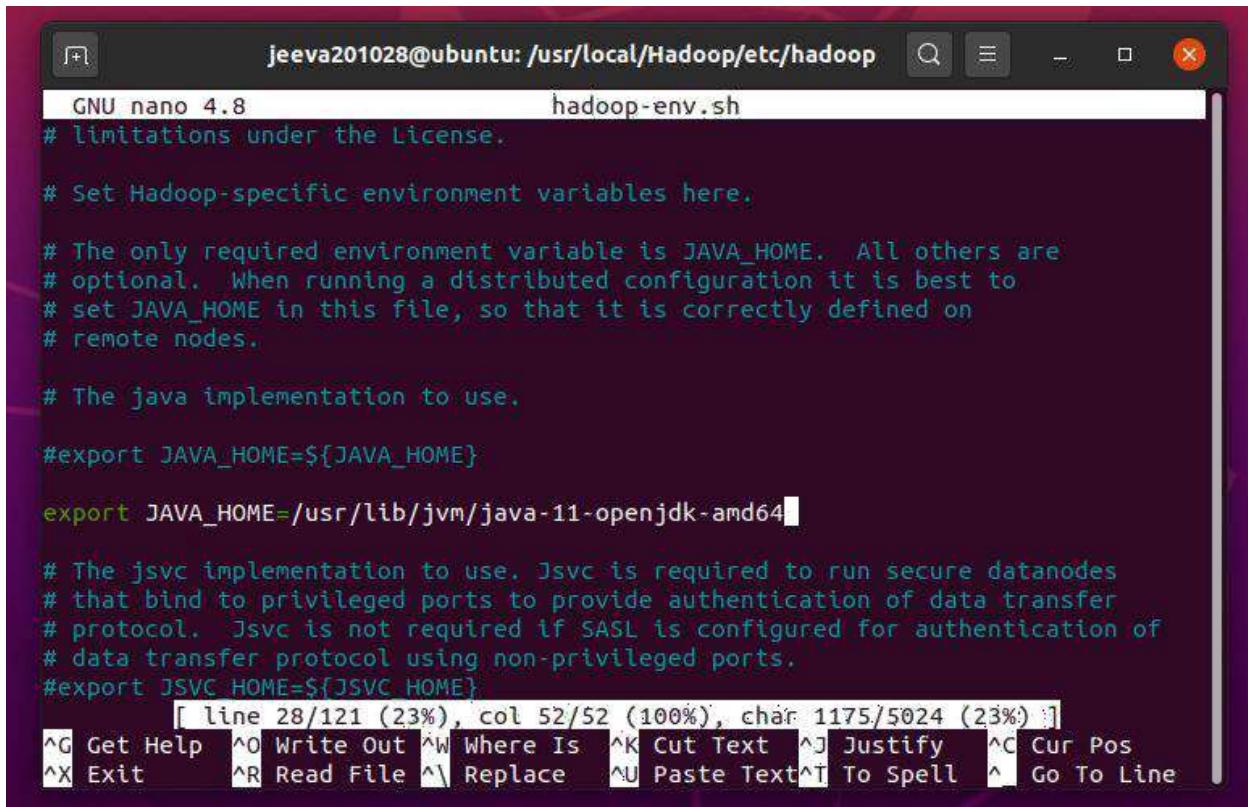
.
fi

# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
  if [ -f /usr/share/bash-completion/bash_completion ]; then
    . /usr/share/bash-completion/bash_completion
  elif [ -f /etc/bash_completion ]; then
    . /etc/bash_completion
  fi
fi
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"
```



```
jeeva201028@ubuntu: ~/Downloads
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/collapsed.gif
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/images/bg.jpg
hadoop-2.9.2/share/doc/hadoop/hadoop-kms/dependency-analysis.html
jeeva201028@ubuntu:~/Downloads$ sudo mv hadoop-2.9.2 /usr/local/Hadoop
jeeva201028@ubuntu:~/Downloads$ sudo chown -R jeeva /usr/local
jeeva201028@ubuntu:~/Downloads$ sudo nano ~/.bashrc
jeeva201028@ubuntu:~/Downloads$ source ~/.bashrc
declare -x CLUTTER_IM_MODULE="ibus"
declare -x COLORTERM="truecolor"
declare -x DBUS_SESSION_BUS_ADDRESS="unix:path=/run/user/1000/bus"
declare -x DESKTOP_SESSION="ubuntu"
declare -x DISPLAY=:0"
declare -x GDMSESSION="ubuntu"
declare -x GJS_DEBUG_OUTPUT="stderr"
declare -x GJS_DEBUG_TOPICS="JS ERROR;JS LOG"
declare -x GNOME_DESKTOP_SESSION_ID="this-is-deprecated"
declare -x GNOME_SHELL_SESSION_MODE="ubuntu"
declare -x GNOME_TERMINAL_SCREEN="/org/gnome/Terminal/screen/c98c3178_c4bb_40a4_9647_3c0ea63ea95a"
declare -x GNOME_TERMINAL_SERVICE=:1.128"
declare -x GPG_AGENT_INFO="/run/user/1000/gnupg/S.gpg-agent:0:1"
declare -x GTK_IM_MODULE="ibus"
declare -x GTK_MODULES="gail:atk-bridge"
declare -x HADOOP_COMMON_HOME="/usr/local/hadoop"
declare -x HADOOP_HDFS_HOME="/usr/local/hadoop"
declare -x HADOOP_HOME="/usr/local/hadoop"
declare -x HADOOP_MAPRED_HOME="/usr/local/hadoop"
declare -x HOME="/home/jeeva201028"
declare -x IM_CONFIG_PHASE="1"
```

Edit the hadoop-env.sh file as following



```
GNU nano 4.8          hadoop-env.sh
# limitations under the License.

# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

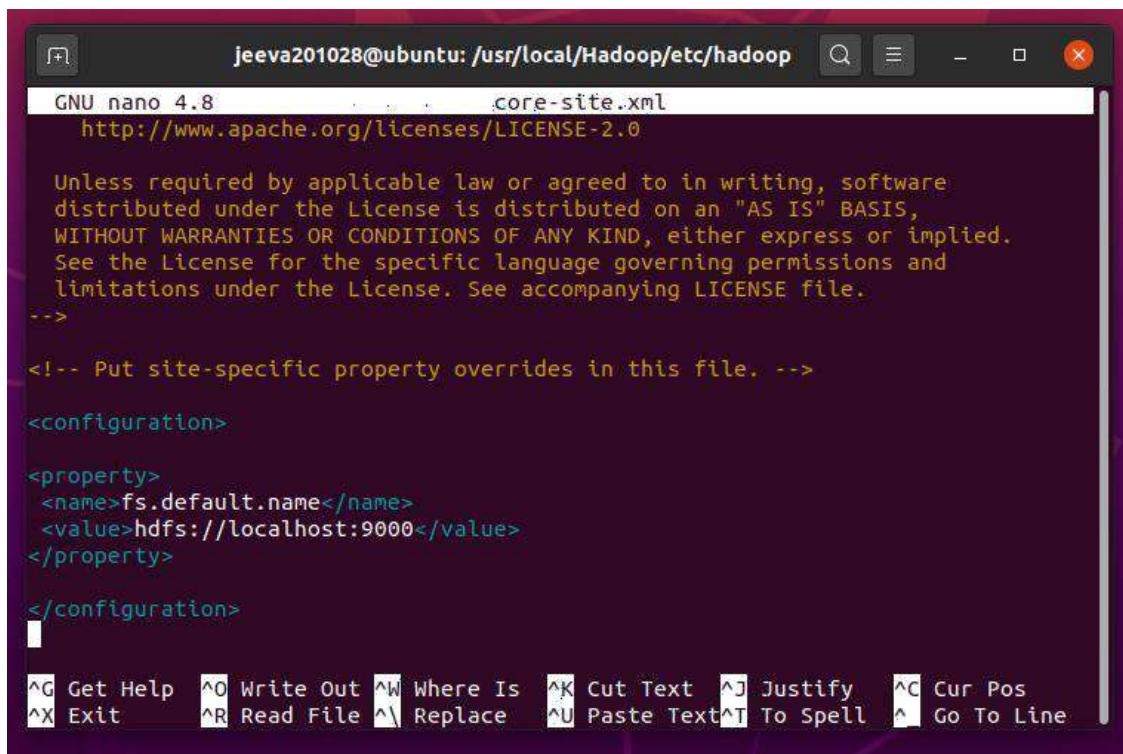
# The java implementation to use.

#export JAVA_HOME=${JAVA_HOME}

export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64

# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
# protocol. Jsvc is not required if SASL is configured for authentication of
# data transfer protocol using non-privileged ports.
#export JSVC_HOME=${JSVC_HOME}
[ line 28/121 (23%), col 52/52 (100%), char 1175/5024 (23%) ]
^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit      ^R Read File ^\ Replace   ^U Paste Text^T To Spell ^_ Go To Line
```

Edit the core-site.xml



The screenshot shows a terminal window titled "jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop". The file being edited is "core-site.xml". The terminal displays the XML configuration for HDFS, including the default file system name and its value.

```
GNU nano 4.8          core-site.xml
http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.

-->

<!-- Put site-specific property overrides in this file. -->

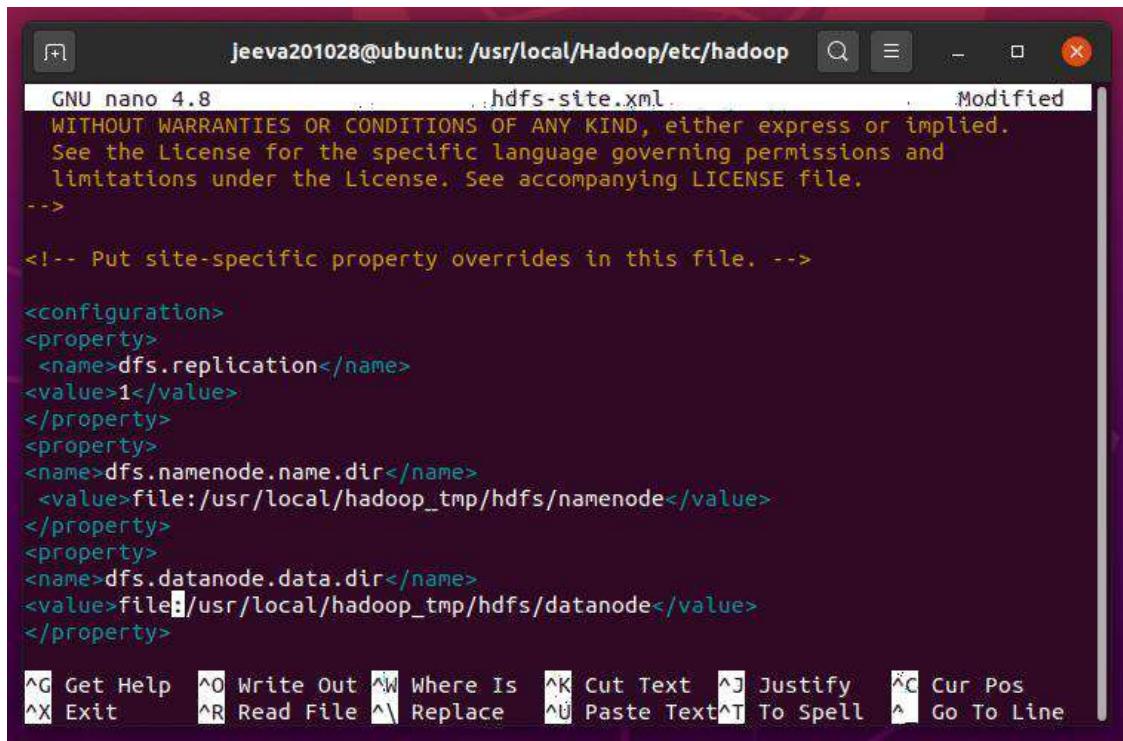
<configuration>

<property>
  <name>fs.default.name</name>
  <value>hdfs://localhost:9000</value>
</property>

</configuration>

^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit      ^R Read File ^\ Replace   ^U Paste Text^T To Spell ^_ Go To Line
```

Edit the hdfs-site.xml



The screenshot shows a terminal window titled "jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop". The file being edited is "hdfs-site.xml". The terminal displays the configuration for DFS, specifically setting the replication factor to 1 and defining the namenode and datanode data directories.

```
GNU nano 4.8          hdfs-site.xml
Modified
http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.

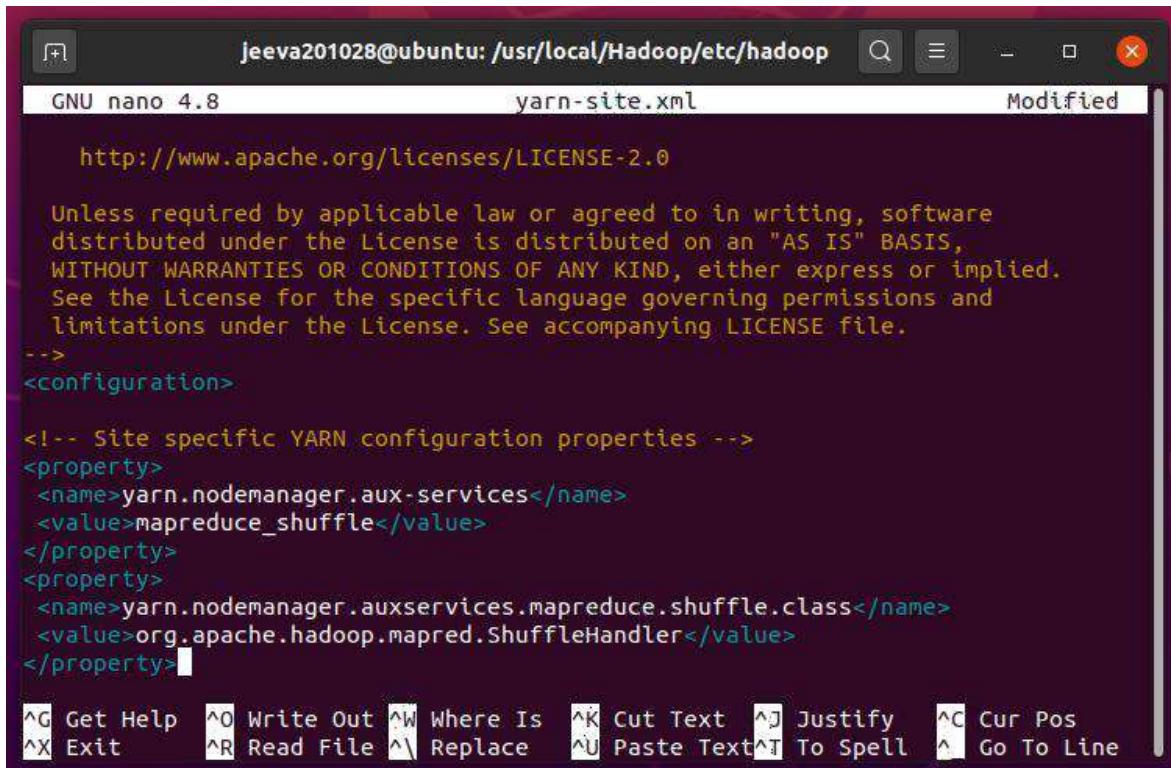
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>
<property>
  <name>dfs.namenode.name.dir</name>
  <value>file:/usr/local/hadoop_tmp/hdfs/namenode</value>
</property>
<property>
  <name>dfs.datanode.data.dir</name>
  <value>file:/usr/local/hadoop_tmp/hdfs/datanode</value>
</property>

^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit      ^R Read File ^\ Replace   ^U Paste Text^T To Spell ^_ Go To Line
```

Edit the yarn-site.xml



The screenshot shows a terminal window titled "jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop". The file being edited is "yarn-site.xml", which is marked as "Modified". The content of the file includes the Apache License header and configuration properties for YARN. The nano editor interface is visible at the bottom.

```
jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop$ nano yarn-site.xml
yarn-site.xml Modified

http://www.apache.org/licenses/LICENSE-2.0

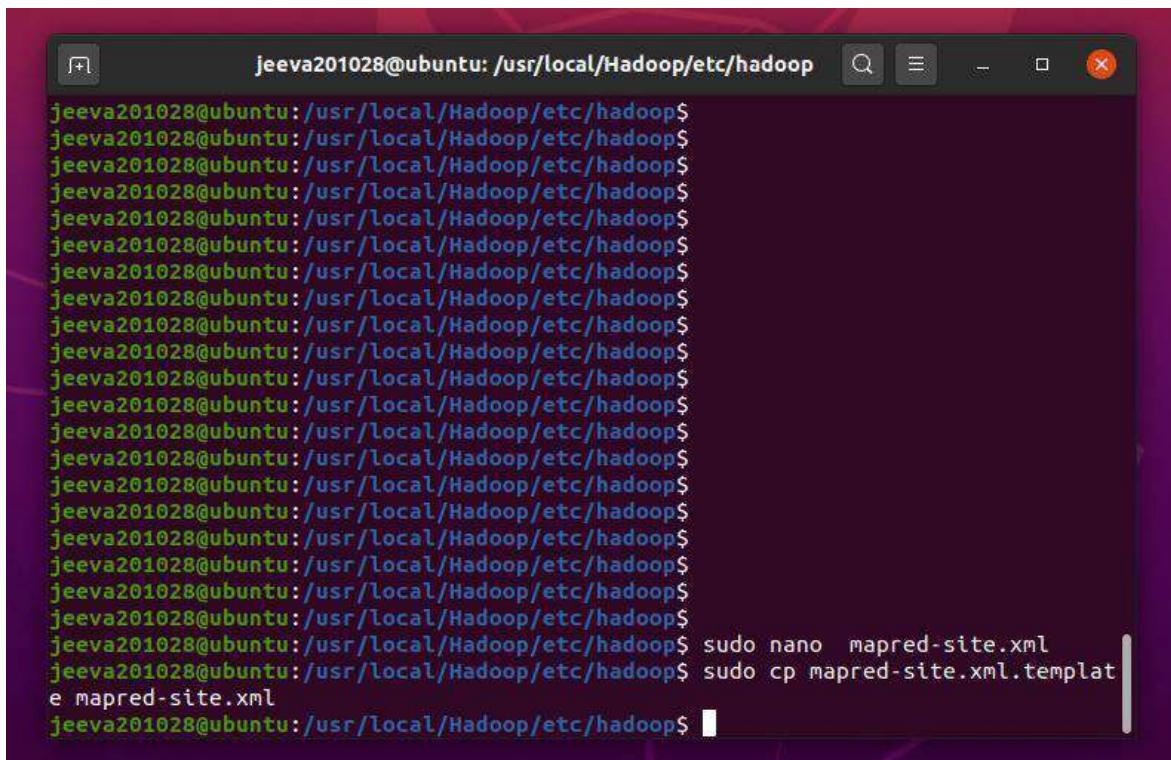
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.

-->
<configuration>

<!-- Site specific YARN configuration properties -->
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>

^G Get Help  ^O Write Out  ^W Where Is  ^K Cut Text  ^J Justify  ^C Cur Pos
^X Exit      ^R Read File  ^\ Replace   ^U Paste Text^T To Spell  ^L Go To Line
```

Edit the mapred-site.xml by remaning the mapred-site.xml.template



The screenshot shows a terminal window titled "jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop". The user has run several commands to navigate to the directory and then used the "sudo nano mapred-site.xml" command to edit the file. The terminal prompt is visible at the bottom.

```
jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop$ 
jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop$ sudo nano mapred-site.xml
jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop$ sudo cp mapred-site.xml.template
e mapred-site.xml
jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop$ 
```

The screenshot shows a terminal window titled "jeeva201028@ubuntu: /usr/local/Hadoop/etc/hadoop". The file being edited is "mapred-site.xml". The content of the file is the Apache License 2.0, followed by XML configuration overrides. The XML includes setting the mapreduce.framework.name to yarn.

```
GNU nano 4.8          mapred-site.xml          Modified
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.

-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
</property>
```

Create the following directory

The screenshot shows a terminal window titled "jeeva201028@ubuntu: ~/Desktop". The user runs several commands to create a directory structure. They first navigate to the Desktop directory, then use sudo to edit mapred-site.xml, copy its template, and edit mapred-site.xml again. Finally, they create a tmp directory and its subdirectories for HDFS and Hadoop.

```
jeeva201028@ubuntu:/usr/local/Hadoop/etc/hadoop$ 
jeeva201028@ubuntu:/usr/local/Hadoop/etc/hadoop$ sudo nano mapred-site.xml
jeeva201028@ubuntu:/usr/local/Hadoop/etc/hadoop$ sudo cp mapred-site.xml.template mapred-site.xml
jeeva201028@ubuntu:/usr/local/Hadoop/etc/hadoop$ sudo nano mapred-site.xml
jeeva201028@ubuntu:/usr/local/Hadoop/etc/hadoop$ cd
jeeva201028@ubuntu:~$ cd Desktop/
jeeva201028@ubuntu:~/Desktop$ sudo mkdir -p /usr/local/hadoop_tmp
jeeva201028@ubuntu:~/Desktop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/namenode
jeeva201028@ubuntu:~/Desktop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/datanode
jeeva201028@ubuntu:~/Desktop$ sudo chown -R jeeva /usr/local/hadoop_tmp
jeeva201028@ubuntu:~/Desktop$ 
```

Format hdfs namenode

```
jeeva201028@ubuntu:~$ hdfs
Usage: hdfs [-config confdir] [-loglevel loglevel] COMMAND
  where COMMAND is one of:
    dfs          run a filesystem command on the file systems supported in Hadoop.
    classpath    prints the classpath
    namenode -format  format the DFS filesystem
    secondarynamenode  run the DFS secondary namenode
    namenode     run the DFS namenode
    journalnode  run the DFS journalnode
    zkfc         run the ZK Failover Controller daemon
    datanode     run a DFS datanode
    debug        run a Debug Admin to execute HDFS debug commands
    dfsadmin    run a DFS admin client
    dfrouter    run the DFS router
    dfrouteradmin  manage Router-based Federation
    haadmin     run a DFS HA admin client
    fsck        run a DFS filesystem checking utility
    balancer    run a cluster balancing utility
    jmxget      get JMX exported values from NameNode or DataNode.
    mover       run a utility to move block replicas across
                storage types
    oiv         apply the offline fsimage viewer to an fsimage
    oiv_legacy  apply the offline fsimage viewer to an legacy fsimage
    oev         apply the offline edits viewer to an edits file
    fetchdt    fetch a delegation token from the NameNode
    getconf     get config values from configuration
    groups     get the groups which users belong to
    snapshotdiff  diff two snapshots of a directory or diff the
                current directory contents with a snapshot
    lssnapshottableDir  list all snapshottable dirs owned by the current user
                        Use -help to see options
    portmap     run a portmap service
    nfs3        run an NFS version 3 gateway
    cacheadmin  configure the HDFS cache
    crypto      configure HDFS encryption zones
    storagepolicies  list/get/set block storage policies
    version     print the version

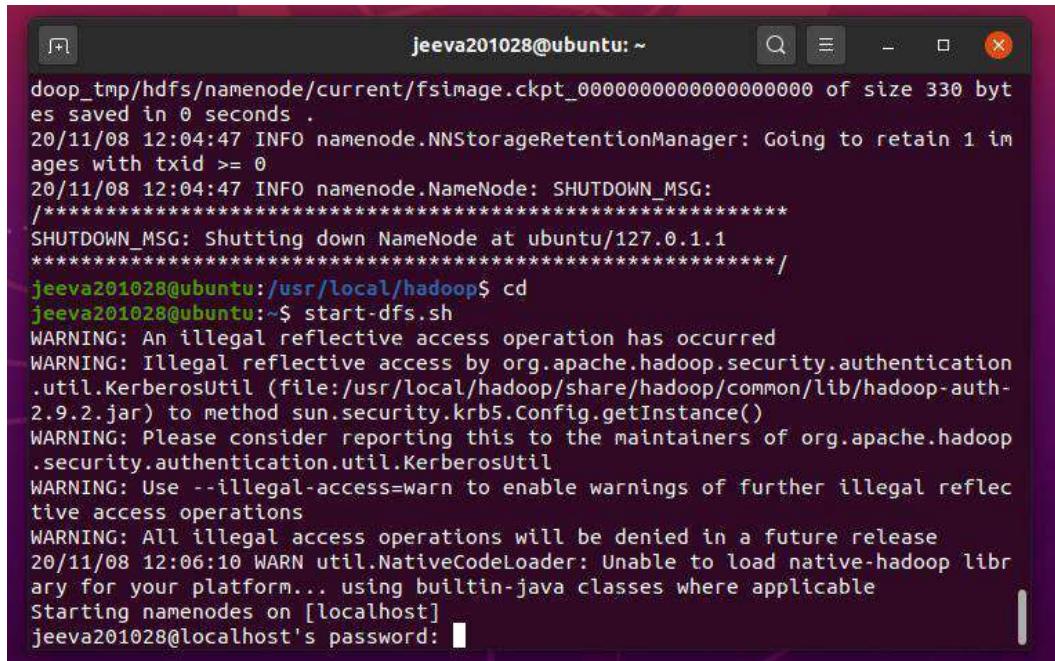
Most commands print help when invoked w/o parameters.
jeeva201028@ubuntu:~$ hdfs namenode -format

jeeva201028@ubuntu:~$ 
20/11/08 04:19:35 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total
heap and retry cache entry expiry time is 600000 millis
20/11/08 04:19:35 INFO util.GSet: Computing capacity for map NameNodeRetryCache
20/11/08 04:19:35 INFO util.GSet: VM type      = 64-bit
20/11/08 04:19:35 INFO util.GSet: 0.02999999329447746% max memory 1000 MB = 307
.2 KB
20/11/08 04:19:35 INFO util.GSet: capacity      = 2^15 = 32768 entries
20/11/08 04:19:35 INFO namenode.FSImage: Allocated new BlockPoolId: BP-725375456
-127.0.1.1-1604837975223
20/11/08 04:19:35 INFO common.Storage: Storage directory /usr/local/hadoop_tmp/h
dfs/namenode has been successfully formatted.
20/11/08 04:19:35 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/lo
cal/hadoop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 using no c
ompression
20/11/08 04:19:35 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/had
oop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 329 bytes
saved in 0 seconds.
20/11/08 04:19:35 INFO namenode.NNStorageRetentionManager: Going to retain 1 ima
ges with txid >= 0
20/11/08 04:19:35 INFO namenode.NameNode: SHUTDOWN_MSG:
*****
SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
*****
```

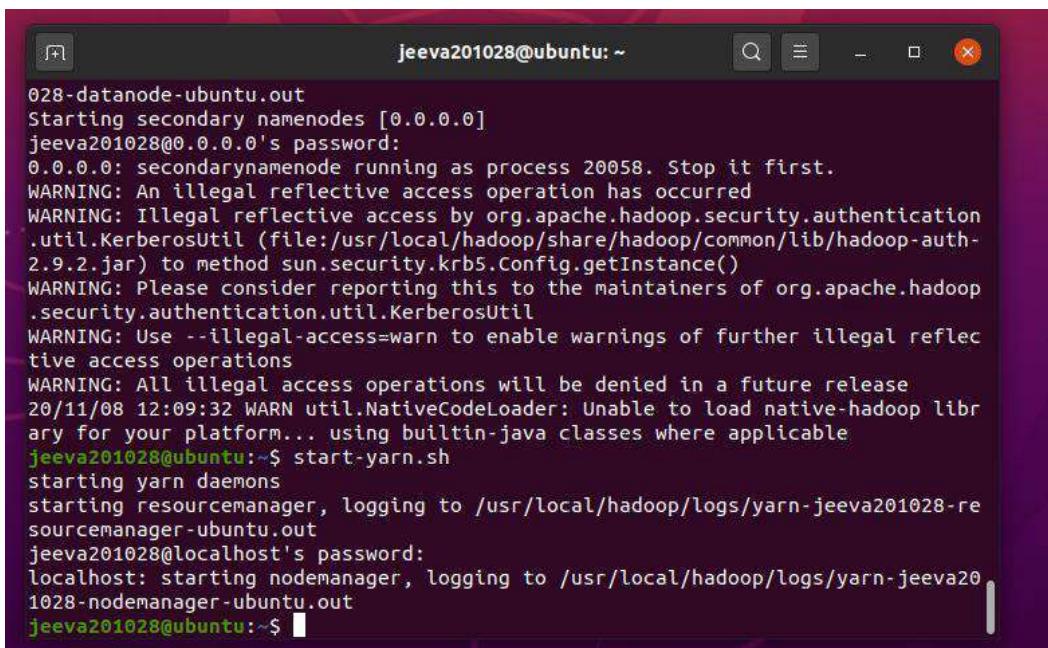
```
$ start-dfs.sh
```

```
$ start-yarn.sh
```

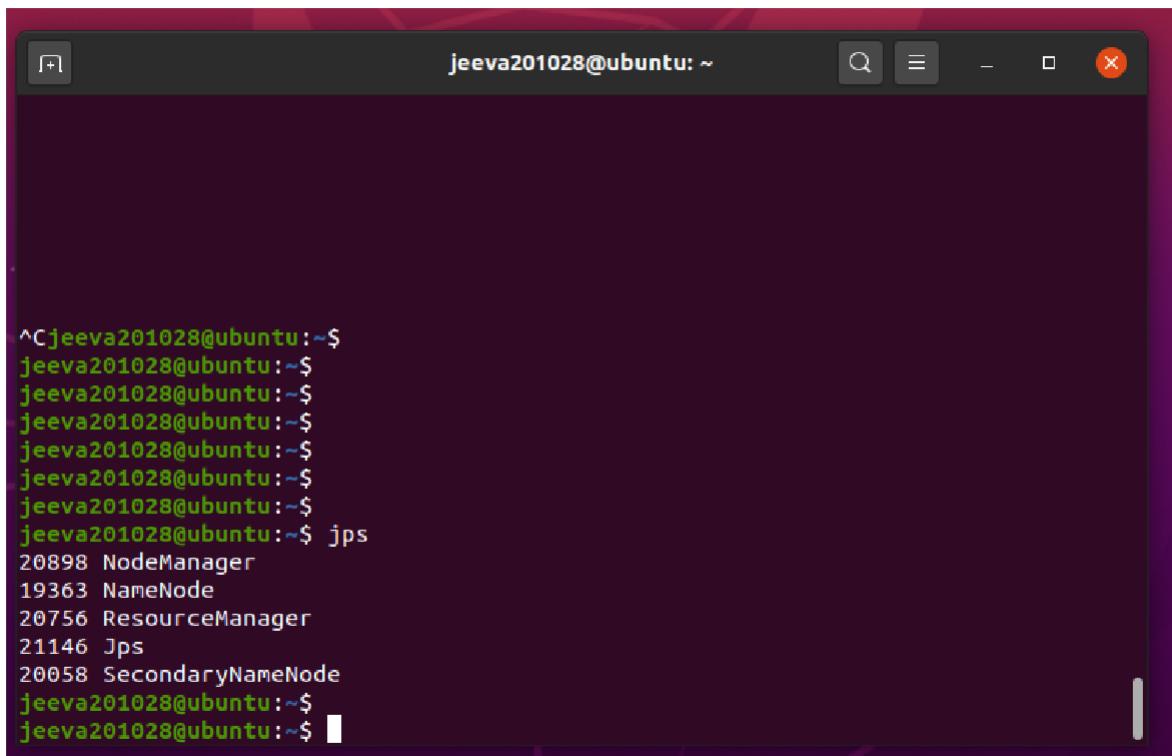
To check whether hadoop is correctly installed or not: \$ jps



```
jeeva201028@ubuntu: ~
doop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 330 bytes saved in 0 seconds .
20/11/08 12:04:47 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
20/11/08 12:04:47 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
*****
jeeva201028@ubuntu:/usr/local/hadoop$ cd
jeeva201028@ubuntu:~$ start-dfs.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:06:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
jeeva201028@localhost's password: ■
```



```
028-datanode-ubuntu.out
Starting secondary namenodes [0.0.0.0]
jeeva201028@0.0.0.0's password:
0.0.0.0: secondarynamenode running as process 20058. Stop it first.
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:09:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
jeeva201028@ubuntu:~$ start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-jeeva201028-re
sourcemanager-ubuntu.out
jeeva201028@localhost's password:
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-jeeva20
1028-nodemanager-ubuntu.out
jeeva201028@ubuntu:~$ ■
```

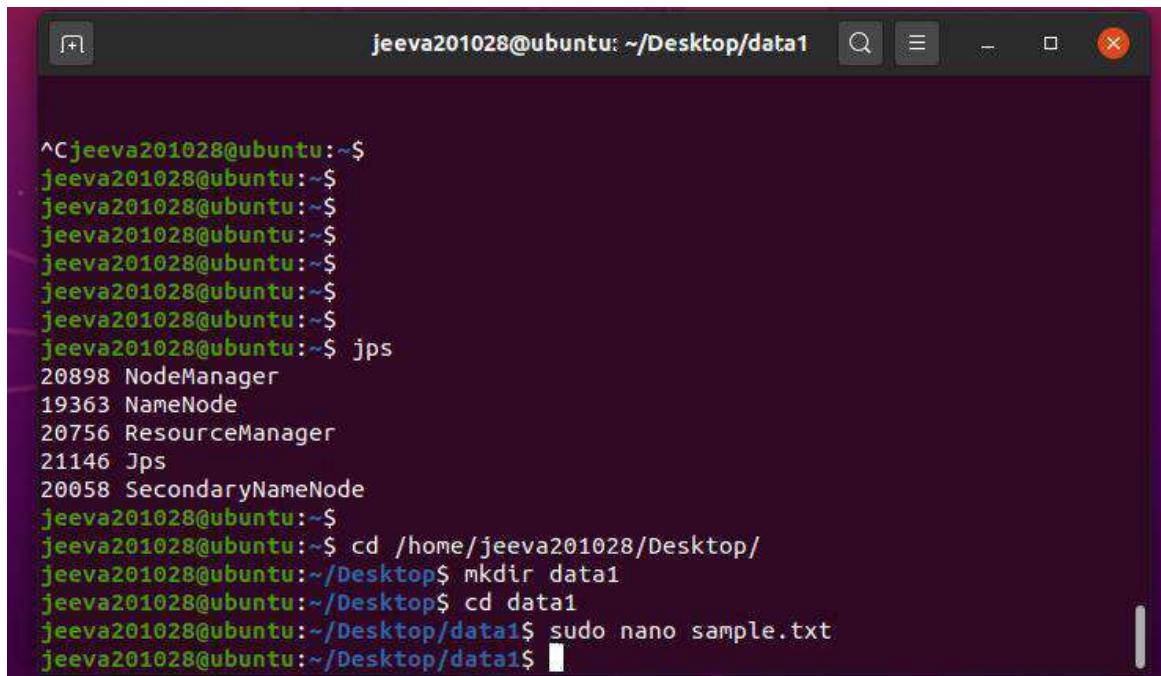


A screenshot of a terminal window titled "jeeva201028@ubuntu: ~". The window shows the command "jps" being run, which lists several Java processes running on the system. The processes listed are NodeManager (20898), NameNode (19363), ResourceManager (20756), Jps (21146), and SecondaryNameNode (20058).

```
^Cjeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$ jps  
20898 NodeManager  
19363 NameNode  
20756 ResourceManager  
21146 Jps  
20058 SecondaryNameNode  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$
```

To execute word count program:

Create a directory and add a text file in it named “sample.txt”



A screenshot of a terminal window titled "jeeva201028@ubuntu: ~/Desktop/data1". The user runs "jps" to check processes, then navigates to their desktop directory with "cd /home/jeeva201028/Desktop/". They create a new directory "data1" with "mkdir data1", change into it with "cd data1", and finally open a new text file "sample.txt" for editing with "sudo nano sample.txt".

```
^Cjeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~$ jps  
20898 NodeManager  
19363 NameNode  
20756 ResourceManager  
21146 Jps  
20058 SecondaryNameNode  
jeeva201028@ubuntu:~$  
jeeva201028@ubuntu:~/Desktop$ mkdir data1  
jeeva201028@ubuntu:~/Desktop$ cd data1  
jeeva201028@ubuntu:~/Desktop/data1$ sudo nano sample.txt  
jeeva201028@ubuntu:~/Desktop/data1$
```

Input: ‘sample.txt’ file

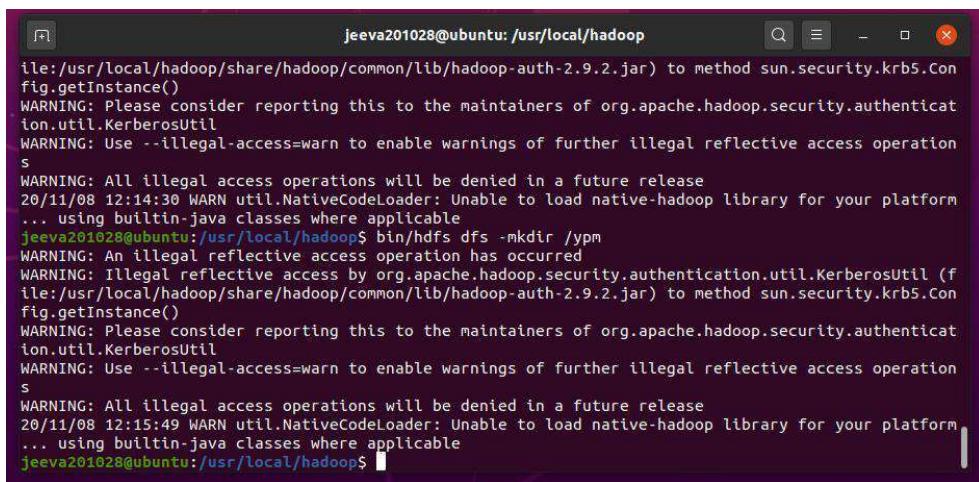
```
hello
how are you
communication
components
computational
computer
computing
coordinate
compiler
distributed file system
hadoop
single node cluster
word count
```

Get Help Write Out Where Is Cut Text Justify Cur Pos Undo
Exit Read File Replace Paste Text To Spell Go To Line Redo
Mark Text To Bracket
Copy Text Where Was

Now run bin/hdfs dfs -mkdir /user

```
jeeva201028@ubuntu: /usr/local/hadoop
20058 SecondaryNameNode
jeeva201028@ubuntu:~$ cd /home/jeeva201028/Desktop/
jeeva201028@ubuntu:~/Desktop$ mkdir data1
jeeva201028@ubuntu:~/Desktop$ cd data1
jeeva201028@ubuntu:~/Desktop/data1$ sudo nano sample.txt
jeeva201028@ubuntu:~/Desktop/data1$ cd
jeeva201028@ubuntu:~$ cd /usr/local/hadoop
jeeva201028@ubuntu:/usr/local/hadoop$ bin/hdfs dfs -mkdir /user
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:14:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

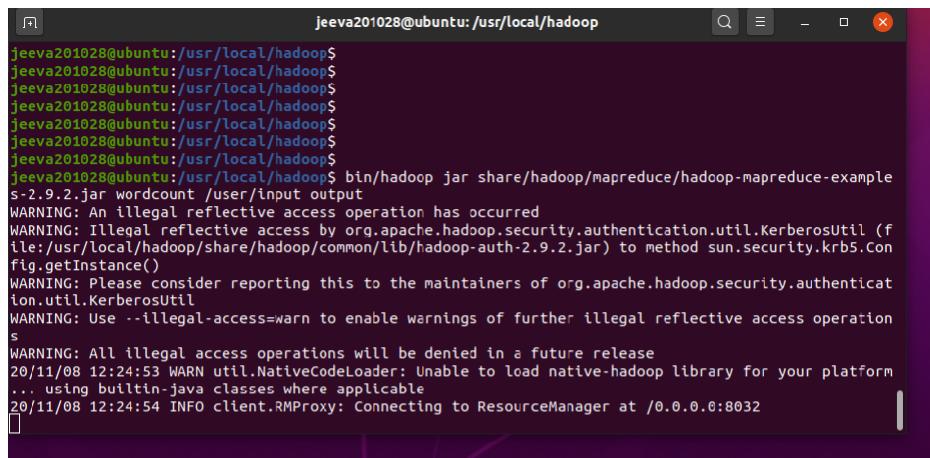
Then run bin/hdfs dfs -mkdir /ypm



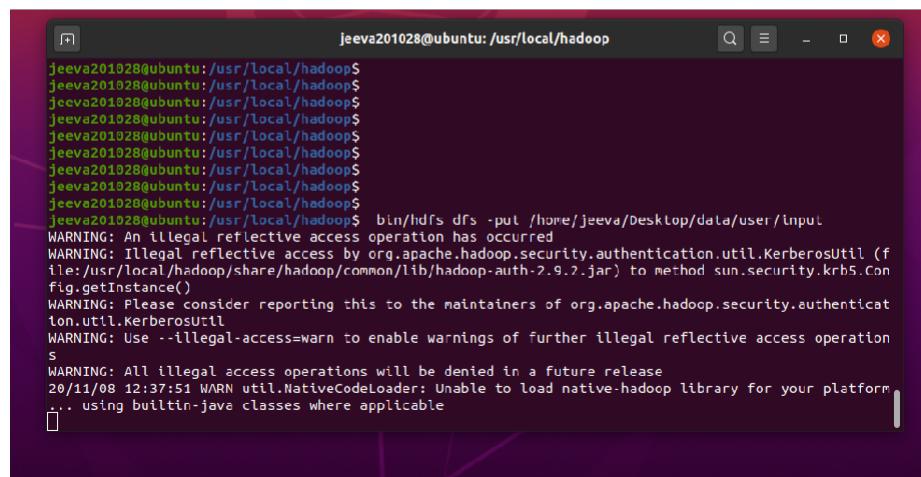
```
jeeva201028@ubuntu:/usr/local/hadoop
file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:14:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
... using builtin-java classes where applicable
jeeva201028@ubuntu:/usr/local/hadoop$ bin/hdfs dfs -mkdir /ypm
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:15:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
... using builtin-java classes where applicable
jeeva201028@ubuntu:/usr/local/hadoop$
```

Run the program //word count prgm will be in the jar file by default which we are using now.

```
$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.9.2.jar
wordcount /user/input output
```



```
jeeva201028@ubuntu:/usr/local/hadoop
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.9.2.jar wordcount /user/input output
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:24:53 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
... using builtin-java classes where applicable
20/11/08 12:24:54 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0:8032
```



```
jeeva201028@ubuntu:/usr/local/hadoop
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ bin/hdfs dfs -put /home/jeeva/Desktop/data/user/input
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/08 12:37:51 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
... using builtin-java classes where applicable
```

```
jeeva201028@ubuntu: ~
Total megabyte-milliseconds taken by all reduce tasks=2942976
Map-Reduce Framework
  Map input records=16
  Map output records=23
  Map output bytes=274
  Map output materialized bytes=285
  Input split bytes=108
  Combine input records=23
  Combine output records=20
  Reduce input groups=20
  Reduce shuffle bytes=285
  Reduce input records=20
  Reduce output records=20
  Spilled Records=40
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=62
  CPU time spent (ms)=1510
  Physical memory (bytes) snapshot=455553624
  Virtual memory (bytes) snapshot=4149194752
  Total committed heap usage (bytes)=273678336
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=186
File Output Format Counters
  Bytes Written=199
```

Now run the following command to see the output

```
$ bin/hdfs dfs -cat output/*
```

```
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ 
jeeva201028@ubuntu:/usr/local/hadoop$ bin/hdfs dfs -cat output/*
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use -illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/07 17:25:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
are      1
cluster  1
communication  1
compiler    1
components   1
computational  1
computer    2
computing   1
coordinate   1
count      1
distributed  1
file       1
hadoop     3
hello      1
how       1
node       1
single     1
system     1
word       1
you       1
```

Using gui: In browser open port 50700, <http://localhost:50070> (Hadoop must be running)

The screenshot shows the Hadoop Overview page. At the top, there's a navigation bar with links for Apps, Gmail, YouTube, and Maps. Below it is a green header bar with tabs for Hadoop (which is selected), Overview, Datanodes, DataNode Volume Failures, Snapshot, Startup Progress, and Utilities. The main content area has a title "Overview 'localhost:9000' (active)". It displays various cluster metrics: Started: Sat Nov 07 15:45:51 +0530 2020; Version: 2.9.2, r02bfebea31ca687bc2f8471dc641b66ed2c6704; Compiled: Tue Nov 13 18:12:00 +0530 2018 by ajisaka from branch-2.9.2; Cluster ID: CID-41245562-d807-4C0b-89Bd-e279108927c1; Block Pool ID: BP-424539900-127.0.1.1-1604743859@7. The page also includes a "Summary" section with security and safemode status, and a detailed DFS usage table.

Overview 'localhost:9000' (active)

Started:	Sat Nov 07 15:45:51 +0530 2020
Version:	2.9.2, r02bfebea31ca687bc2f8471dc641b66ed2c6704
Compiled:	Tue Nov 13 18:12:00 +0530 2018 by ajisaka from branch-2.9.2
Cluster ID:	CID-41245562-d807-4C0b-89Bd-e279108927c1
Block Pool ID:	BP-424539900-127.0.1.1-1604743859@7

Summary

Security is off.
Safemode is off.
20 files and directories, 5 blocks = 25 total filesystem object(s).
Heap Memory used 73.32 MB of 111 MB Heap Memory. Max Heap Memory is 1000 MB
Non Heap Memory used 50.57 MB of 54.81 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	97.63 GB
DFS Used:	296 kB (0%)
Non DFS Used:	15.43 GB
DFS Remaining:	77.2 GB (79.07%)

The screenshot shows the Datanode Information page. At the top, there's a navigation bar with links for Apps, Gmail, YouTube, and Maps. Below it is a green header bar with tabs for Hadoop (selected), Overview, Datanodes (selected), DataNode Volume Failures, Snapshot, Startup Progress, and Utilities. A dropdown menu under Utilities shows "Browse the file system" and "Logs". The main content area has a title "Datanode Information". It includes a "Datanode usage histogram" which is mostly empty except for one bar at the bottom. Below it is a table titled "In operation" showing one datanode entry: Node: 104959010 (127.0.0.1:59010), Http Address: http://127.0.0.1:59075, Last contact: 0s, Last Block Report: 1141, Capacity: 97.03 GB, Blocks: 5, Block pended: 0, Version: 2.9.2. The footer shows the URL localhost:5070/explorer.html.

Datanode Information

✓ In service ⚡ Down ⚡ Decommissioned ⚡ Decommissioned & dead ⚡ In Maintenance & dead

Datanode usage histogram



In operation

Show	25	entries	Search				
Node	Http Address	Last contact	Last Block Report	Capacity	Blocks	Block pended	Version
✓ 104959010 (127.0.0.1:59010)	http://127.0.0.1:59075	0s	1141	97.03 GB	5	0 kB (0%)	2.9.2

Showing 1 to 1 of 1 entries

PerPage: 1 Red

localhost:5070/explorer.html

localhost:50070/explorer.html/

Apps Gmail YouTube Maps

Hadoop Overview Datanodes DataNode Volume Failures Snapshot Status Progress Utilities

Browse Directory

/

Show 25 entries

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr--r--	Jensv201028	supergroup	0 B	Nov 07 17:11	0	0 B	tmp	
drwxr-xr-x	Jensv201028	supergroup	0 B	Nov 07 17:10	0	0 B	user	
drwxr-xr-x	Jensv201028	supergroup	0 B	Nov 07 17:00	0	0 B	yes	

Showing 1 to 3 of 3 entries

Previous 1 Next

Hadoop, 2018.

localhost:50070/explorer.html#/user/input

Apps Gmail YouTube Maps

Hadoop Overview Datanodes DataNode Volume Failures Snapshot Status Progress Utilities

Browse Directory

/user/input

Show 25 entries

	Permission	Owner
drwxr-xr-x	Jensv201028	lesha

Showing 1 to 1 of 1 entries

Hadoop, 2018.

File information - sample.txt

Download Read the file (first 32K) Tail the file (last 32K)

Block information Block 0

Block ID: 107374L829
Block Pool ID: BP-43432390-127.0.1-1604743859075
Generation Stamp: 1801
Size: 198
Availability:
• Success

File contents:

```
Hello  
Hello are you  
complete  
communication  
components  
distributed  
complete  
complete  
complete
```

Close

localhost:50070/explorer.html#/user/output

Apps Gmail YouTube Maps

Hadoop Overview Datanodes DataNode Volume Failures Snapshot Status Progress Utilities

Browse Directory

/user/output

Show 25 entries

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	Jensv201028	supergroup	0 B	Nov 07 17:10	0	0 B	output	

Showing 1 to 1 of 1 entries

Previous 1 Next

Hadoop, 2018.

Name	Size	Last Modified	Replication	Block Size
_SUCCESS	0 B	Nov 07 17:15	1	128 MB
part+00000	199 B	Nov 07 17:15	1	128 MB

After completing stop hadoop \$ stop-all.sh

```
jeeva201028@ubuntu:/usr/local/hadoop$ stop-all.sh
This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.2.jar) to method sun.security.krb5.Config.getINSTANCE()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
20/11/09 00:21:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Stopping namenodes on [localhost]
jeeva201028@localhost's password: 
```

Result:

Thus the installation of hadoop single mode cluster and execution of word count program is done and the output is obtained successfully.