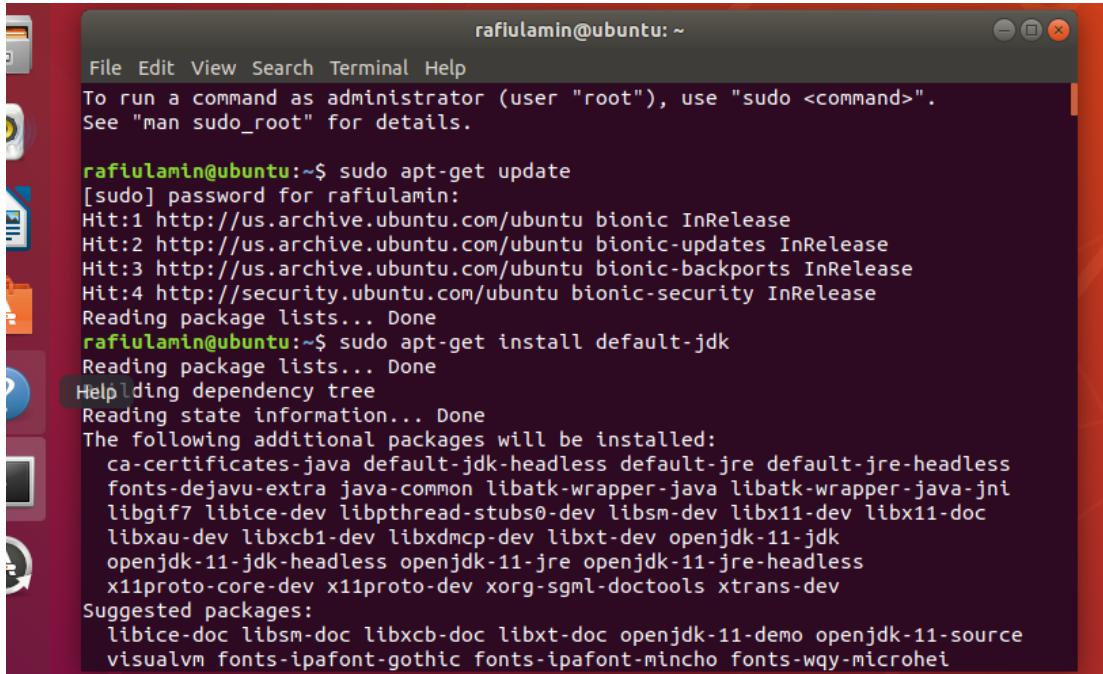


Project 1

Name: Md. Rafiul Amin

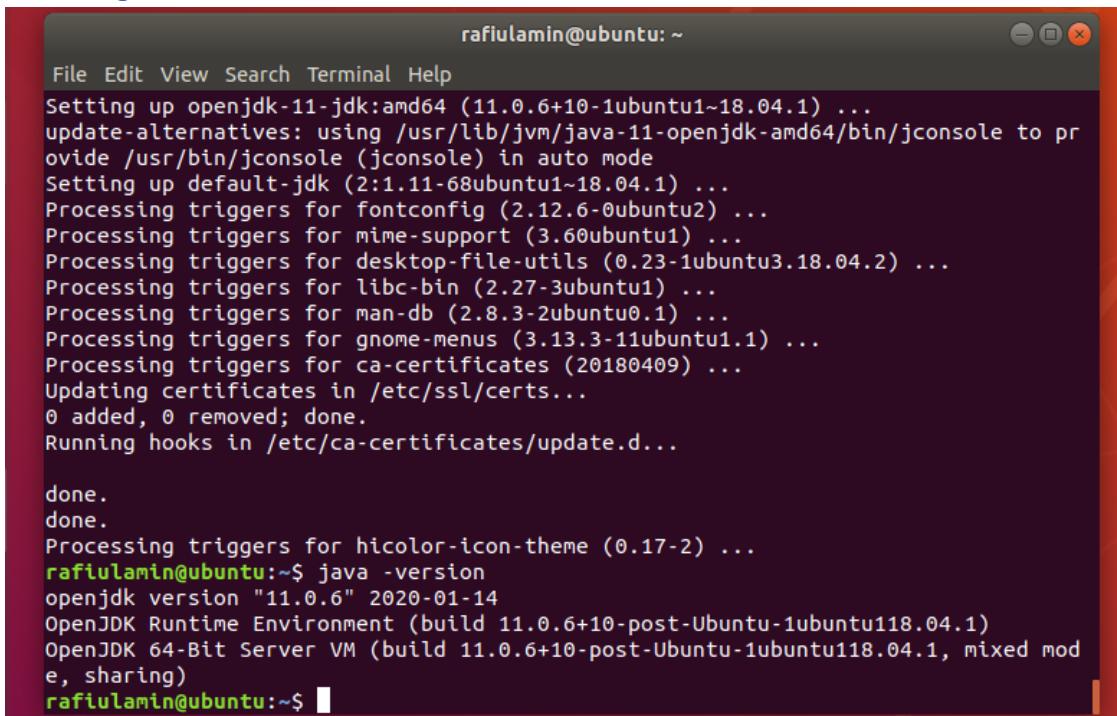
Student ID: 1637374

Installing Java:



```
rafiulamin@ubuntu:~$ sudo apt-get update
[sudo] password for rafiulamin:
Hit:1 http://us.archive.ubuntu.com/ubuntu bionic InRelease
Hit:2 http://us.archive.ubuntu.com/ubuntu bionic-updates InRelease
Hit:3 http://us.archive.ubuntu.com/ubuntu bionic-backports InRelease
Hit:4 http://security.ubuntu.com/ubuntu bionic-security InRelease
Reading package lists... Done
rafiulamin@ubuntu:~$ sudo apt-get install default-jdk
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ca-certificates-java default-jdk-headless default-jre default-jre-headless
  fonts-dejavu-extra java-common libatk-wrapper-java libatk-wrapper-java-jni
  libgif7 libice-dev libpthread-stubs0-dev libsm-dev libx11-dev libx11-doc
  libxau-dev libxcb1-dev libxdmcp-dev libxt-dev openjdk-11-jdk
  openjdk-11-jdk-headless openjdk-11-jre openjdk-11-jre-headless
  x11proto-core-dev x11proto-dev xorg-sgml-doctools xtrans-dev
Suggested packages:
  libice-doc libsm-doc libxcb-doc libxt-doc openjdk-11-demo openjdk-11-source
  visualvm fonts-ipafont-gothic fonts-ipafont-mincho fonts-wqy-microhei
```

Checking Java version:



```
rafiulamin@ubuntu:~$ Setting up openjdk-11-jdk:amd64 (11.0.6+10-1ubuntu1~18.04.1) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Setting up default-jdk (2:1.11-68ubuntu1~18.04.1) ...
Processing triggers for fontconfig (2.12.6-0ubuntu2) ...
Processing triggers for mime-support (3.60ubuntu1) ...
Processing triggers for desktop-file-utils (0.23-1ubuntu3.18.04.2) ...
Processing triggers for libc-bin (2.27-3ubuntu1) ...
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
Processing triggers for gnome-menus (3.13.3-11ubuntu1.1) ...
Processing triggers for ca-certificates (20180409) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...

done.
done.
Processing triggers for hicolor-icon-theme (0.17-2) ...
rafiulamin@ubuntu:~$ java -version
openjdk version "11.0.6" 2020-01-14
OpenJDK Runtime Environment (build 11.0.6+10-post-Ubuntu-1ubuntu118.04.1)
OpenJDK 64-Bit Server VM (build 11.0.6+10-post-Ubuntu-1ubuntu118.04.1, mixed mode, sharing)
```

Downloading Hadoop:

The screenshot shows a file download interface with the following details:

File Information:

- Name: hadoop-2.8.0.tar.gz
- Completed — 410 MB

Downloaded Files:

Name	Last Modified	Size	Description
Parent Directory			
CHANGES.md	2017-03-25 00:33	527K	
CHANGES.md.asc	2017-03-25 00:33	801	
CHANGES.md.md5	2017-03-25 00:33	136	
CHANGES.md.mds	2017-03-25 00:33	927	
RELEASENOTES.md	2017-03-25 00:33	56K	
RELEASENOTES.md.asc	2017-03-25 00:33	801	
RELEASENOTES.md.md5	2017-03-25 00:33	146	
RELEASENOTES.md.mds	2017-03-25 00:33	962	
hadoop-2.8.0-rat.txt	2017-03-25 00:33	1.2M	
hadoop-2.8.0-rat.txt.asc	2017-03-25 00:33	801	
hadoop-2.8.0-rat.txt.md5	2017-03-25 00:33	156	
hadoop-2.8.0-rat.txt.mds	2017-03-25 00:33	1.0K	
hadoop-2.8.0-site.tar.gz	2017-03-25 00:33	120M	
hadoop-2.8.0-site.tar.gz.asc	2017-03-25 00:33	801	
hadoop-2.8.0-site.tar.gz.md5	2017-03-25 00:33	164	
hadoop-2.8.0-site.tar.gz.mds	2017-03-25 00:33	1.0K	
hadoop-2.8.0-src.tar.gz	2017-03-25 00:33	33M	
hadoop-2.8.0-src.tar.gz.asc	2017-03-25 00:33	801	
hadoop-2.8.0-src.tar.gz.md5	2017-03-25 00:33	162	
hadoop-2.8.0-src.tar.gz.mds	2017-03-25 00:33	1.0K	
hadoop-2.8.0.tar.gz	2017-03-26 21:17	410M	
hadoop-2.8.0.tar.gz.asc	2017-03-25 00:33	801	

Extract Hadoop files:

```
rafiulamin@ubuntu: ~/Downloads
File Edit View Search Terminal Help
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
Processing triggers for gnome-menus (3.13.3-11ubuntu1.1) ...
Processing triggers for ca-certificates (20180409) ...
Updating certificates in /etc/ssl/certs...
 0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
done.
Processing triggers for hicolor-icon-theme (0.17-2) ...
rafiulamin@ubuntu:$ java -version
openjdk version "11.0.6" 2020-01-14
OpenJDK Runtime Environment (build 11.0.6+10-post-Ubuntu-1ubuntu118.04.1)
OpenJDK 64-Bit Server VM (build 11.0.6+10-post-Ubuntu-1ubuntu118.04.1, mixed mode, sharing)
rafiulamin@ubuntu:$ cd downloads
bash: cd: downloads: No such file or directory
rafiulamin@ubuntu:$ ls
Desktop  Downloads  Music  Public  Videos
Documents examples.desktop  Pictures  Templates
rafiulamin@ubuntu:$ cd Downloads
rafiulamin@ubuntu:~/Downloads$ ls
hadoop-2.8.0.tar.gz
rafiulamin@ubuntu:~/Downloads$
```

```
ice-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/test/
hadoop-2.8.0/share/hadoop/yarn/test/hadoop-yarn-server-tests-2.8.0-tests.jar
rafiulamin@ubuntu:~/Downloads$ tar -xzvf hadoop-2.8.0.tar.gz
```

Move Hadoop files to the new directory (/usr/local/hadoop):

```
Index of /dist/hadoop/common
rafiulamin@ubuntu: ~/Downloads
File Edit View Search Terminal Help
ice-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-sources.j
ar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-test-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-test-s
ources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-
sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-
test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-sources.ja
r
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-test-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-test-s
ources.jar
hadoop-2.8.0/share/hadoop/yarn/test/
hadoop-2.8.0/share/hadoop/yarn/test/hadoop-yarn-server-tests-2.8.0-tests.jar
rafiulamin@ubuntu:~/Downloads$ sudo mv hadoop-2.8.0
```

```
rafiulamin@ubuntu: ~/Downloads
File Edit View Search Terminal Help
ice-2.8.0-test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-sources.j
ar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-common-2.8.0-test-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-nodemanager-2.8.0-test-s
ources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-
sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-resourcemanager-2.8.0-
test-sources.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-sources.ja
r
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-test-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-test-s
ources.jar
hadoop-2.8.0/share/hadoop/yarn/test/
hadoop-2.8.0/share/hadoop/yarn/test/hadoop-yarn-server-tests-2.8.0-tests.jar
rafiulamin@ubuntu:~/Downloads$ sudo mv hadoop-2.8.0 /usr/local/hadoop
```

Edit File “bashrc.sh” to set paths:

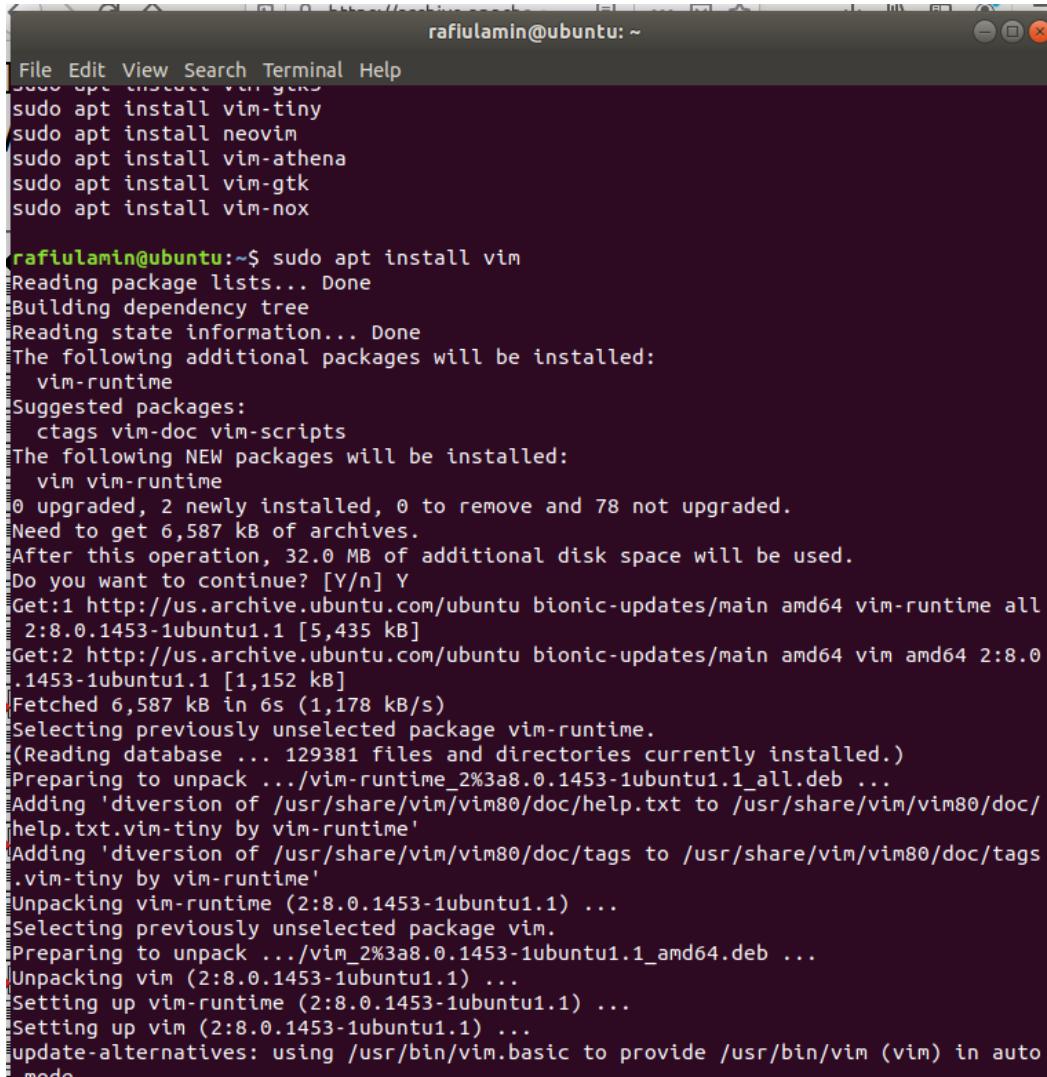
```
rafiulamin@ubuntu: ~
File Edit View Search Terminal Help
r
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-tests-2.8.0-test-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-sourc
es.jar
hadoop-2.8.0/share/hadoop/yarn/sources/hadoop-yarn-server-web-proxy-2.8.0-test-s
ources.jar
hadoop-2.8.0/share/hadoop/yarn/test/
hadoop-2.8.0/share/hadoop/yarn/test/hadoop-yarn-server-tests-2.8.0-tests.jar
rafiulamin@ubuntu:~/Downloads$ sudo mv hadoop-2.8.0 /usr/local/hadoop
rafiulamin@ubuntu:~/Downloads$ cd
rafiulamin@ubuntu:~$ vim ~/.bashrc

Command 'vim' not found, but can be installed with:

sudo apt install vim
sudo apt install vim-gtk3
sudo apt install vim-tiny
sudo apt install neovim
sudo apt install vim-athena
sudo apt install vim-gtk
sudo apt install vim-nox

rafiulamin@ubuntu:~$
```

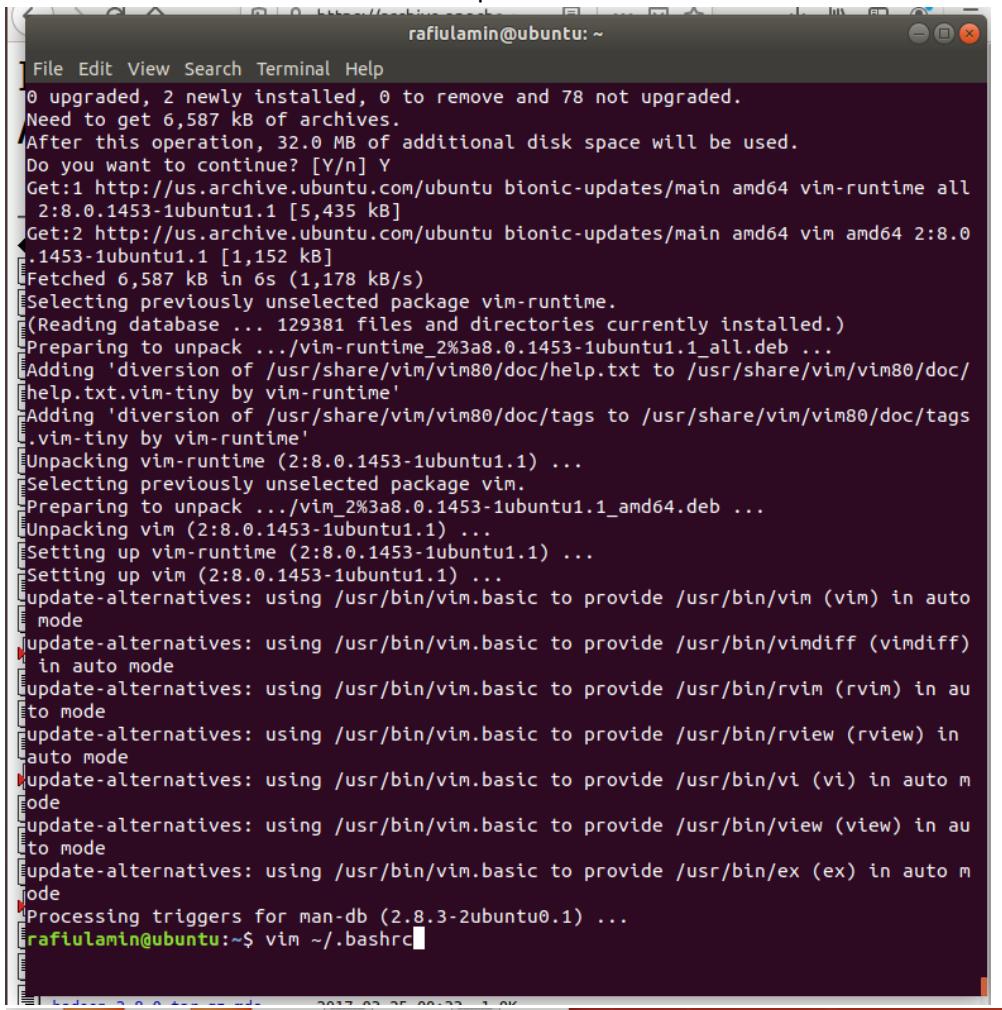
Install vim in order to open and edit the files.



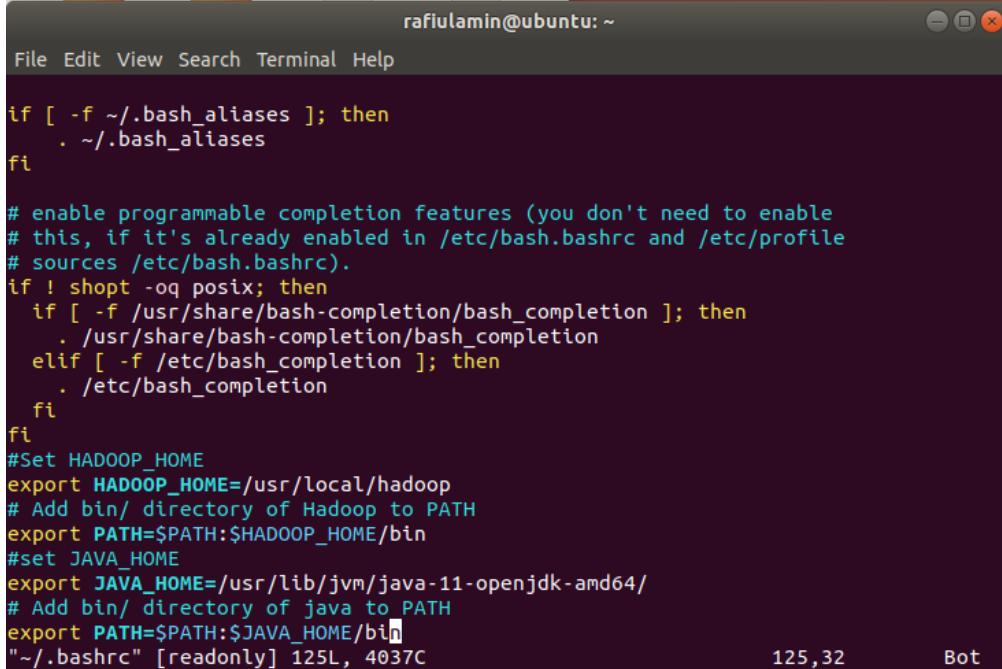
The screenshot shows a terminal window titled "rafiulamin@ubuntu: ~". The user has run the command "sudo apt install vim" and is viewing the output of the package manager. The output shows the selection of vim-runtime and vim packages, their download from the archive, unpacking, and setting up. The terminal interface includes a menu bar at the top with options like File, Edit, View, Search, Terminal, and Help. The window has standard Linux-style window controls (minimize, maximize, close) in the top right corner.

```
rafiulamin@ubuntu:~$ sudo apt install vim
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  vim-runtime
Suggested packages:
  ctags vim-doc vim-scripts
The following NEW packages will be installed:
  vim vim-runtime
0 upgraded, 2 newly installed, 0 to remove and 78 not upgraded.
Need to get 6,587 kB of archives.
After this operation, 32.0 MB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 vim-runtime all 2:8.0.1453-1ubuntu1.1 [5,435 kB]
Get:2 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 vim amd64 2:8.0.1453-1ubuntu1.1 [1,152 kB]
Fetched 6,587 kB in 6s (1,178 kB/s)
Selecting previously unselected package vim-runtime.
(Reading database ... 129381 files and directories currently installed.)
Preparing to unpack .../vim-runtime_2%3a8.0.1453-1ubuntu1.1_all.deb ...
Adding 'diversion of /usr/share/vim/vim80/doc/help.txt to /usr/share/vim/vim80/doc/help.txt.vim-tiny by vim-runtime'
Adding 'diversion of /usr/share/vim/vim80/doc/tags to /usr/share/vim/vim80/doc/tags.vim-tiny by vim-runtime'
Unpacking vim-runtime (2:8.0.1453-1ubuntu1.1) ...
Selecting previously unselected package vim.
Preparing to unpack .../vim_2%3a8.0.1453-1ubuntu1.1_amd64.deb ...
Unpacking vim (2:8.0.1453-1ubuntu1.1) ...
Setting up vim-runtime (2:8.0.1453-1ubuntu1.1) ...
Setting up vim (2:8.0.1453-1ubuntu1.1) ...
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/vim (vim) in auto mode
```

Now we can edit “bashrc.sh” file to set paths.



```
rafiulamin@ubuntu: ~
File Edit View Search Terminal Help
0 upgraded, 2 newly installed, 0 to remove and 78 not upgraded.
Need to get 6,587 kB of archives.
After this operation, 32.0 MB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 vim-runtime all
2:8.0.1453-1ubuntu1.1 [5,435 kB]
Get:2 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 vim amd64 2:8.0
.1453-1ubuntu1.1 [1,152 kB]
Fetched 6,587 kB in 6s (1,178 kB/s)
Selecting previously unselected package vim-runtime.
(Reading database ... 129381 files and directories currently installed.)
Preparing to unpack .../vim-runtime_2%3a8.0.1453-1ubuntu1.1_all.deb ...
Adding 'diversion of /usr/share/vim/vim80/doc/help.txt to /usr/share/vim/vim80/doc/
help.txt.vim-tiny by vim-runtime'
Adding 'diversion of /usr/share/vim/vim80/doc/tags to /usr/share/vim/vim80/doc/tags
.vim-tiny by vim-runtime'
Unpacking vim-runtime (2:8.0.1453-1ubuntu1.1) ...
Selecting previously unselected package vim.
Preparing to unpack .../vim_2%3a8.0.1453-1ubuntu1.1_amd64.deb ...
Unpacking vim (2:8.0.1453-1ubuntu1.1) ...
Setting up vim-runtime (2:8.0.1453-1ubuntu1.1) ...
Setting up vim (2:8.0.1453-1ubuntu1.1) ...
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/vim (vim) in auto
 mode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/vimdiff (vimdiff)
 in auto mode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/rvim (rvim) in au
to mode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/rview (rview) in
auto mode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/vi (vi) in auto m
ode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/view (view) in au
to mode
update-alternatives: using /usr/bin/vim.basic to provide /usr/bin/ex (ex) in auto m
ode
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
rafiulamin@ubuntu:~$ vim ~/.bashrc
```



```
rafiulamin@ubuntu: ~
File Edit View Search Terminal Help

if [ -f ~/.bash_aliases ]; then
    . ~/.bash_aliases
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
#Set HADOOP_HOME
export HADOOP_HOME=/usr/local/hadoop
# Add bin/ directory of Hadoop to PATH
export PATH=$PATH:$HADOOP_HOME/bin
#set JAVA_HOME
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64/
# Add bin/ directory of java to PATH
export PATH=$PATH:$JAVA_HOME/bin
"~/.bashrc" [readonly] 125L, 4037C          125,32      Bot
```

Edit files for configuration:

Change directory to “/usr/local/Hadoop/etc/hadoop”.

```
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop
File Edit View Search Terminal Help
Thunderbird Mail buntu:~$ source ~/.bashrc
rafiulamin@ubuntu:~$ cd /usr/local/hadoop
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin include libexec NOTICE.txt sbin
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu:/usr/local/hadoop$ cd etc
rafiulamin@ubuntu:/usr/local/hadoop/etc$ ls
hadoop
rafiulamin@ubuntu:/usr/local/hadoop/etc$ cd hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$
```

The terminal window shows the user navigating to the Hadoop configuration directory. The command history at the bottom of the window includes the source of the bashrc file and the navigation commands.

Configuring “core-site.xml” file.

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop
File Edit View Search Terminal Help
rafiulamin@ubuntu:~$ source ~/.bashrc
rafiulamin@ubuntu:~$ cd /usr/local/hadoop
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin include libexec NOTICE.txt sbin
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu:/usr/local/hadoop$ cd etc
rafiulamin@ubuntu:/usr/local/hadoop/etc$ ls
hadoop
rafiulamin@ubuntu:/usr/local/hadoop/etc$ cd hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim core-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$
```

The terminal window shows the user opening the core-site.xml file in vim editor. A tooltip for "Software Updater" is visible on the left side of the screen.

Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop

```
File Edit View Search Terminal Help
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
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>fs.defaultFS</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>
-- INSERT --
hadoop-2.8.0-rat.txt.mds 2017-03-25 00:33 156
```

Make “mapred-site.xml” file. We copy the template and then edit it.

Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu:~\$ source ~/.bashrc
rafiulamin@ubuntu:~\$ cd /usr/local/hadoop
rafiulamin@ubuntu:/usr/local/hadoop\$ ls
bin include libexec NOTICE.txt sbin
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu:/usr/local/hadoop\$ cd etc
rafiulamin@ubuntu:/usr/local/hadoop/etc\$ ls
hadoop
rafiulamin@ubuntu:/usr/local/hadoop/etc\$ cd hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop\$ vim core-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop\$ vim core-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop\$ sudo cp \$HADOOP_HOME/etc/hadoop/
mapred-site.xml.template \$HADOOP_HOME/etc/hadoop/mapred-site.xml

Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox

rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop

File Edit View Search Terminal Help

```
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu:/usr/local/hadoop$ cd etc
rafiulamin@ubuntu:/usr/local/hadoop/etc$ ls
hadoop
rafiulamin@ubuntu:/usr/local/hadoop/etc$ cd hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim core-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim core-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ sudo cp $HADOOP_HOME/etc/hadoop/
mapred-site.xml.template $HADOOP_HOME/etc/hadoop/mapred-site.xml
[sudo] password for rafiulamin:
Isorry, try again.
[sudo] password for rafiulamin:
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml      httpfs-env.sh      mapred-env.sh
configuration.xsl           httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg       httpfs-signature.secret  mapred-site.xml
core-site.xml                httpfs-site.xml      mapred-site.xml.template
hadoop-env.cmd               kms-acls.xml        slaves
hadoop-env.sh                kms-env.sh          ssl-client.xml.example
hadoop-metrics2.properties   kms-log4j.properties  ssl-server.xml.example
hadoop-metrics.properties    kms-site.xml        yarn-env.cmd
hadoop-policy.xml            log4j.properties    yarn-env.sh
hdfs-site.xml                mapred-env.cmd     yarn-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$
```

hadoop-2.8.0-rat.txt.md5 2017-03-25 00:33 156

Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox

rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop

File Edit View Search Terminal Help

```
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
 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>
</configuration>
```

-- INSERT -- 2017-03-25 00:33 156 24,17 Bot

hadoop-2.8.0-rat.txt.md5 2017-03-25 00:33 156

Create “data” Directory and Corresponding Sub-directory:

```
Ind rafiulamin@ubuntu: /usr/local/hadoop
File Edit View Search Terminal Help
Sorry, try again.
[sudo] password for rafiulamin:
[rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml      httpfs-env.sh          mapred-env.sh
configuration.xsl           httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg       httpfs-signature.secret mapred-site.xml
core-site.xml                httpfs-site.xml        mapred-site.xml.template
hadoop-env.cmd               kms-acls.xml         slaves
hadoop-env.sh                kms-env.sh          ssl-client.xml.example
hadoop-metrics2.properties   kms-log4j.properties  ssl-server.xml.example
hadoop-metrics.properties    kms-site.xml         yarn-env.cmd
hadoop-policy.xml            log4j.properties     yarn-env.sh
hdfs-site.xml                mapred-env.cmd      yarn-site.xml
rafiulamin@ubuntu:/usr/local/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu:/usr/local/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu:/usr/local/hadoop$ cd /usr/local/hadoop
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin  include  libexec  NOTICE.txt  sbin
etc  lib      LICENSE.txt  README.txt  share
rafiulamin@ubuntu:/usr/local/hadoop$ mkdir data
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin  etc      lib      LICENSE.txt  README.txt  share
data  include  libexec  NOTICE.txt  sbin
rafiulamin@ubuntu:/usr/local/hadoop$
```

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
Ind rafiulamin@ubuntu: /usr/local/hadoop/data
File Edit View Search Terminal Help
core-site.xml          httpfs-site.xml          mapred-site.xml.template
hadoop-env.cmd          kms-acls.xml          slaves
hadoop-env.sh           kms-env.sh          ssl-client.xml.example
hadoop-metrics2.properties  kms-log4j.properties  ssl-server.xml.example
hadoop-metrics.properties  kms-site.xml         yarn-env.cmd
hadoop-policy.xml        log4j.properties     yarn-env.sh
hdfs-site.xml           mapred-env.cmd      yarn-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ cd /usr/local/hadoop
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin  include  libexec  NOTICE.txt  sbin
etc  lib      LICENSE.txt  README.txt  share
rafiulamin@ubuntu:/usr/local/hadoop$ mkdir data
rafiulamin@ubuntu:/usr/local/hadoop$ ls
bin  etc      lib      LICENSE.txt  README.txt  share
data  include  libexec  NOTICE.txt  sbin
rafiulamin@ubuntu:/usr/local/hadoop$ cd data
rafiulamin@ubuntu:/usr/local/hadoop/data$ ls
rafiulamin@ubuntu:/usr/local/hadoop/data$ mkdir namenode
rafiulamin@ubuntu:/usr/local/hadoop/data$ mkdir datanode
rafiulamin@ubuntu:/usr/local/hadoop/data$ ls
datanode  namenode
rafiulamin@ubuntu:/usr/local/hadoop/data$
```

We configure “hdfs-site.xml” to set the path for datanode and namenode:

rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop

```
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ cd /usr/local/hadoop
rafiulamin@ubuntu: /usr/local/hadoop$ ls
bin include libexec NOTICE.txt sbin
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu: /usr/local/hadoop$ mkdir data
rafiulamin@ubuntu: /usr/local/hadoop$ ls
bin etc lib LICENSE.txt README.txt share
data include libexec NOTICE.txt sbin
rafiulamin@ubuntu: /usr/local/hadoop$ cd data
rafiulamin@ubuntu: /usr/local/hadoop/data$ ls
rafiulamin@ubuntu: /usr/local/hadoop/data$ mkdir namenode
rafiulamin@ubuntu: /usr/local/hadoop/data$ mkdir datanode
rafiulamin@ubuntu: /usr/local/hadoop/data$ ls
datanode namenode
rafiulamin@ubuntu: /usr/local/hadoop/data$ cd /usr/local/hadoop/etc/hadoop/
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim hdfs-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$
```

rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop

```
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. -->

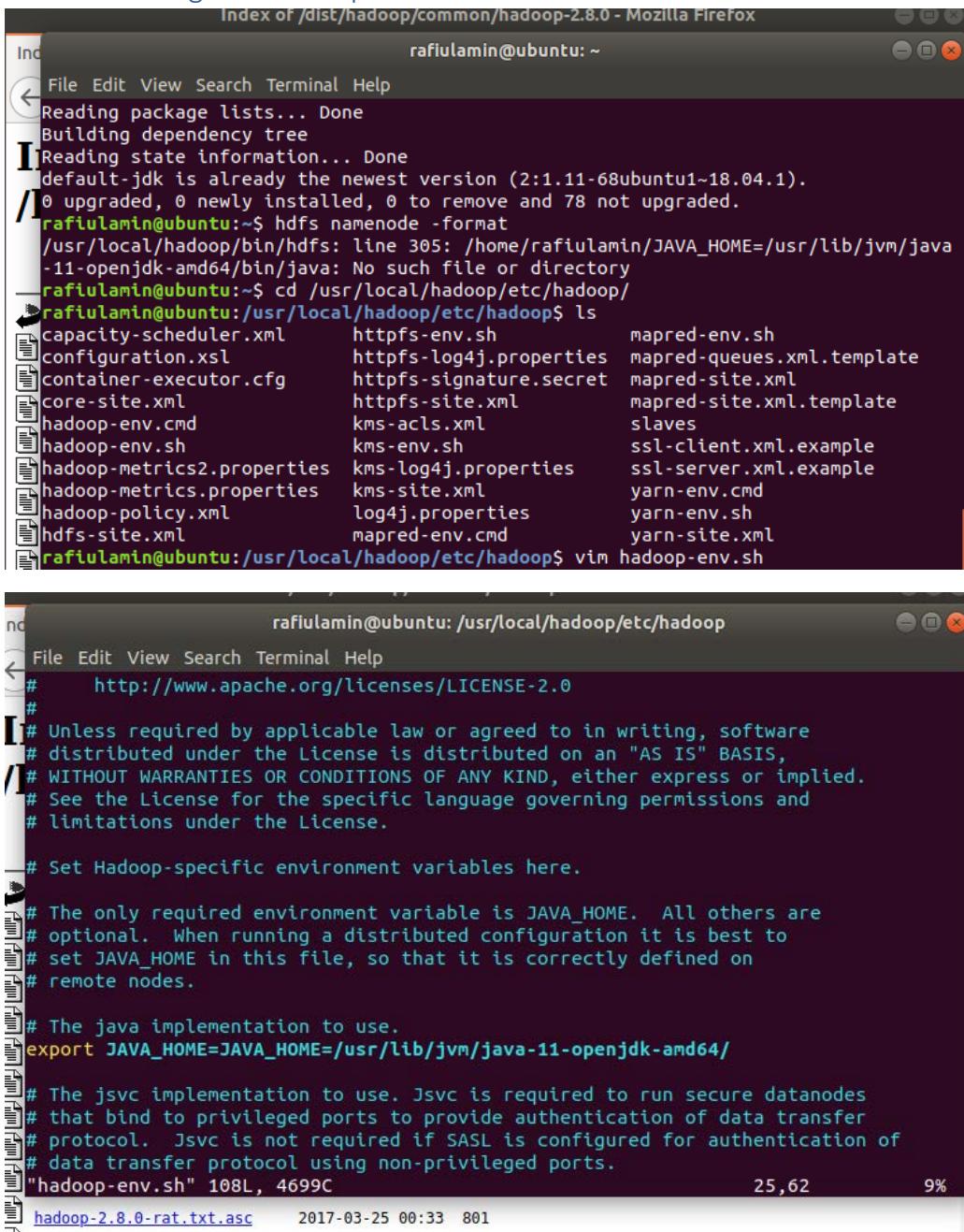
<configuration>
    <property>
        <name>dfs.replication</name>
        <value>1</value>
    </property>
    <property>
        <name>dfs.namenode.name.dir</name>
        <value>/usr/local/hadoop/data/namenode</value>
    </property>
    <property>
        <name>dfs.datanode.data.dir</name>
        <value>/usr/local/hadoop/data/datanode</value>
    </property>
</configuration>
-- INSERT --
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ 2017-03-25 00:33 301 31,2-9 Bot
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ hadoop-2.8.0-rat.txt.md5 2017-03-25 00:33 156
```

Configure “yarn-site.xml”:

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop
File Edit View Search Terminal Help
hadoop-policy.xml      log4j.properties      yarn-env.sh
hdfs-site.xml          mapred-env.cmd      yarn-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim mapred-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ cd /usr/local/hadoop
rafiulamin@ubuntu: /usr/local/hadoop$ ls
bin include libexec NOTICE.txt sbin
etc lib LICENSE.txt README.txt share
rafiulamin@ubuntu: /usr/local/hadoop$ mkdir data
rafiulamin@ubuntu: /usr/local/hadoop$ ls
bin etc lib LICENSE.txt README.txt share
data include libexec NOTICE.txt sbin
rafiulamin@ubuntu: /usr/local/hadoop$ cd data
rafiulamin@ubuntu: /usr/local/hadoop/data$ ls
rafiulamin@ubuntu: /usr/local/hadoop/data$ mkdir namenode
rafiulamin@ubuntu: /usr/local/hadoop/data$ mkdir datanode
rafiulamin@ubuntu: /usr/local/hadoop/data$ ls
datanode namenode
rafiulamin@ubuntu: /usr/local/hadoop/data$ cd /usr/local/hadoop/etc/hadoop/
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim hdfs-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim yarn-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$ vim yarn-site.xml
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop$
```

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/etc/hadoop
File Edit View Search Terminal Help
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.
-->
<configuration>
  <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>
</configuration>
~
```

Then we configure “hadoop-env.sh” and cha.



The screenshot shows a terminal window titled "Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox". The terminal session starts with package management output:

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
default-jdk is already the newest version (2:1.11-68ubuntu1~18.04.1).
0 upgraded, 0 newly installed, 0 to remove and 78 not upgraded.

rafiulamin@ubuntu:~$ hdfs namenode -format
/usr/local/hadoop/bin/hdfs: line 305: /home(rafiulamin)/JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64/bin/java: No such file or directory

rafiulamin@ubuntu:~$ cd /usr/local/hadoop/etc/hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ ls
```

Then, the user runs `vim hadoop-env.sh` to edit the configuration file:

```
capacity-scheduler.xml      httpfs-env.sh      mapred-env.sh
configuration.xsl           httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg       httpfs-signature.secret   mapred-site.xml
core-site.xml                httpfs-site.xml        mapred-site.xml.template
hadoop-env.cmd               kms-acls.xml         slaves
hadoop-env.sh                kms-env.sh          ssl-client.xml.example
hadoop-metrics2.properties  kms-log4j.properties  ssl-server.xml.example
hadoop-metrics.properties   kms-site.xml         yarn-env.cmd
hadoop-policy.xml            log4j.properties    yarn-env.sh
hdfs-site.xml                mapred-env.cmd      yarn-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim hadoop-env.sh
```

The terminal then shows the contents of `hadoop-env.sh`:

```
# 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.

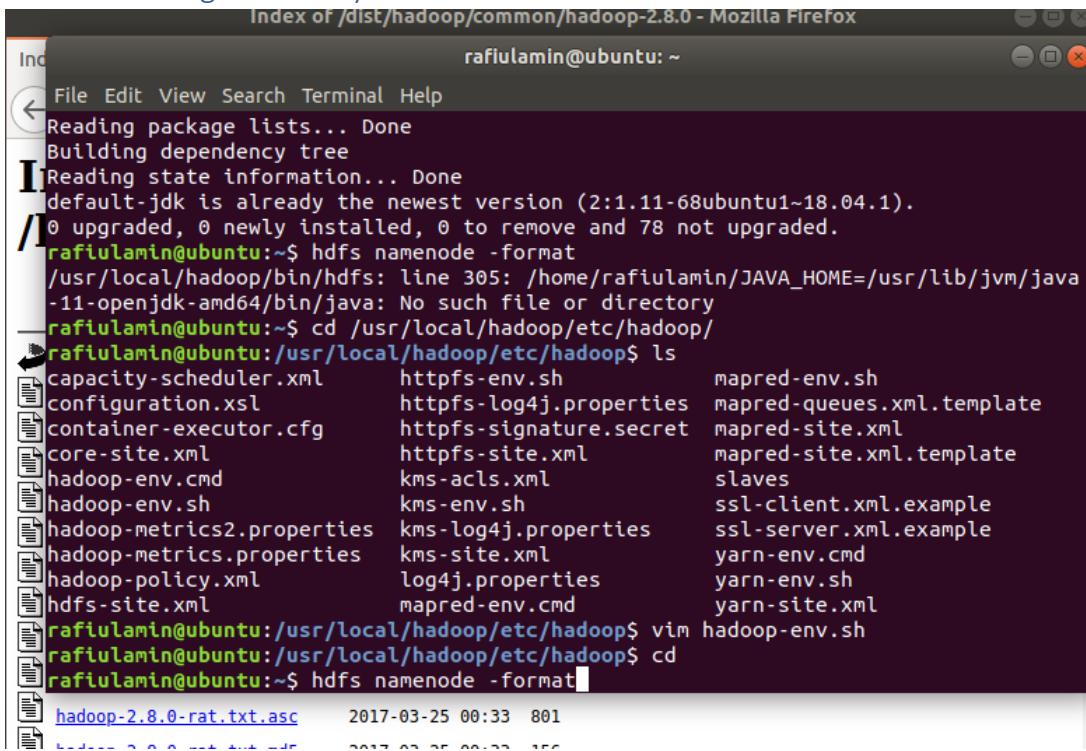
# Set Hadoop-specific environment variables here.

# 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=/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.
".hadoop-env.sh" 108L, 4699C                                25,62      9%
hadoop-2.8.0-rat.txt.asc 2017-03-25 00:33 801
```

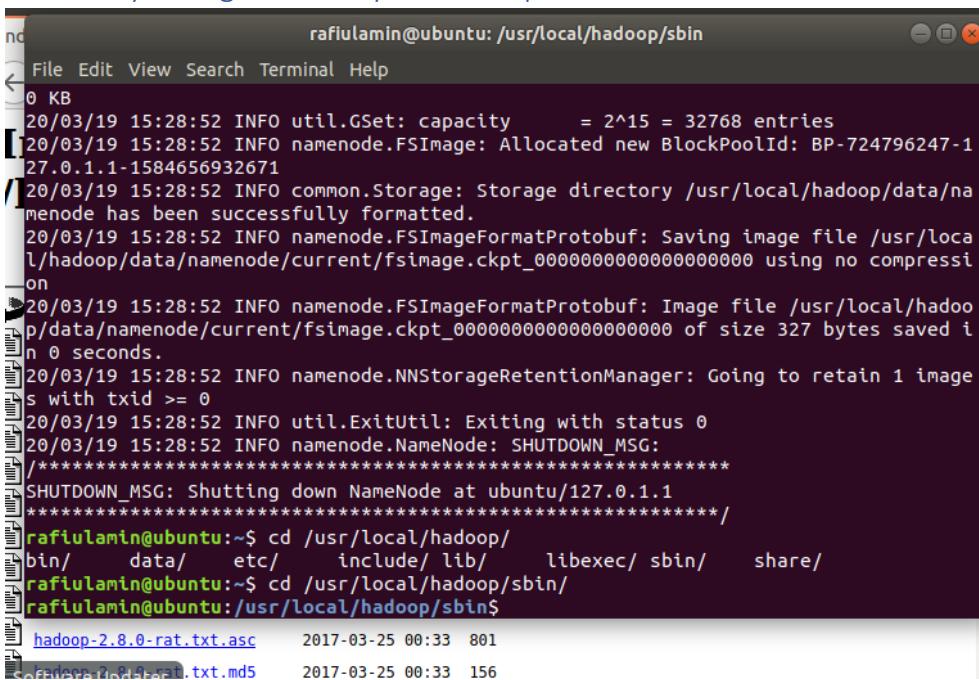
Then we change directory and format namenode:



```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: ~

File Edit View Search Terminal Help
Reading package lists... Done
Building dependency tree
Reading state information... Done
default-jdk is already the newest version (2:1.11-68ubuntu1~18.04.1).
0 upgraded, 0 newly installed, 0 to remove and 78 not upgraded.
rafiulamin@ubuntu:~$ hdfs namenode -format
/usr/local/hadoop/bin/hdfs: line 305: /home/rafiulamin/JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64/bin/java: No such file or directory
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/etc/hadoop/
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml      httpfs-env.sh          mapred-env.sh
configuration.xsl           httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg       httpfs-signature.secret mapred-site.xml
core-site.xml                httpfs-site.xml        mapred-site.xml.template
hadoop-env.cmd              kms-acls.xml          slaves
hadoop-env.sh               kms-env.sh           ssl-client.xml.example
hadoop-metrics2.properties  kms-log4j.properties  ssl-server.xml.example
hadoop-metrics.properties   kms-site.xml         yarn-env.cmd
hadoop-policy.xml            log4j.properties     yarn-env.sh
hdfs-site.xml                mapred-env.cmd      yarn-site.xml
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ vim hadoop-env.sh
rafiulamin@ubuntu:/usr/local/hadoop/etc/hadoop$ cd
rafiulamin@ubuntu:~$ hdfs namenode -format
rafiulamin@ubuntu:~$ hadoop-2.8.0-rat.txt.asc 2017-03-25 00:33 801
rafiulamin@ubuntu:~$ hadoop-2.8.0-rat.txt.mds 2017-03-25 00:33 156
```

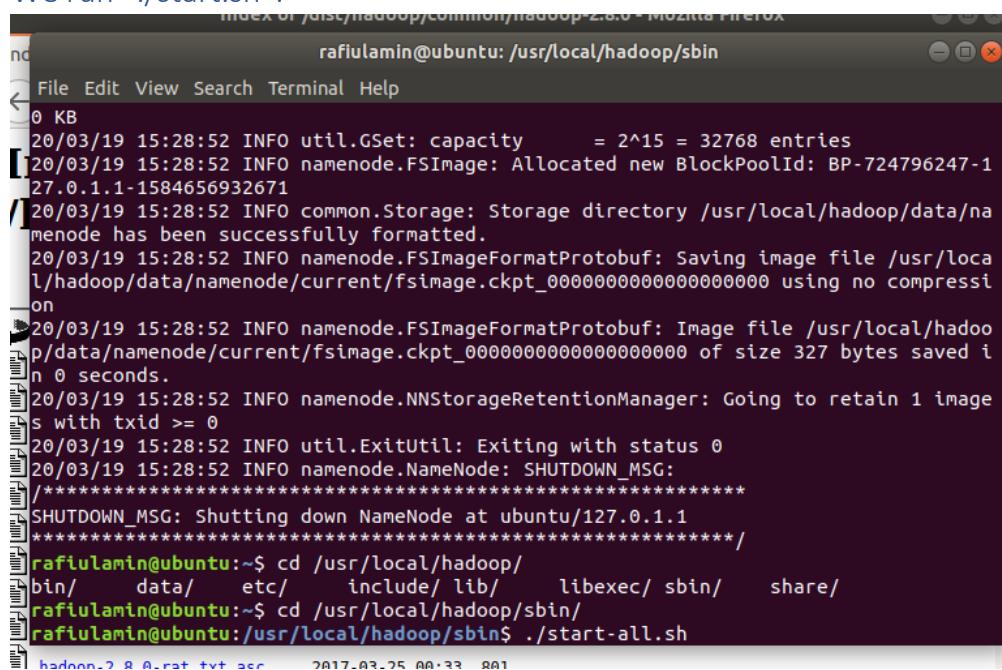
We finally change directory to Hadoop “sbin” in order to start all the clusters:



```
rafiulamin@ubuntu: /usr/local/hadoop/sbin

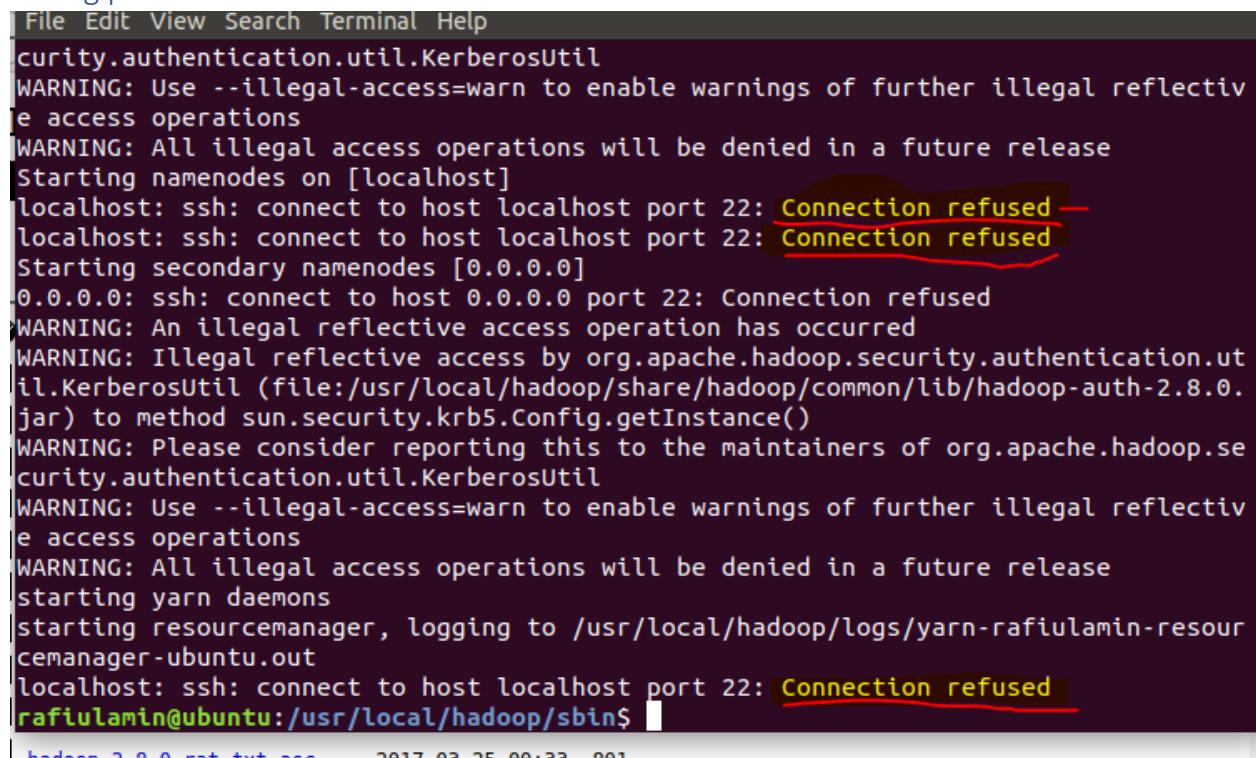
File Edit View Search Terminal Help
0 KB
20/03/19 15:28:52 INFO util.GSet: capacity      = 2^15 = 32768 entries
20/03/19 15:28:52 INFO namenode.FSImage: Allocated new BlockPoolId: BP-724796247-127.0.1.1-1584656932671
20/03/19 15:28:52 INFO common.Storage: Storage directory /usr/local/hadoop/data/namenode has been successfully formatted.
20/03/19 15:28:52 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
20/03/19 15:28:52 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_00000000000000000000 of size 327 bytes saved in 0 seconds.
20/03/19 15:28:52 INFO namenode.NNStorageRetentionManager: Going to retain 1 image(s) with txid >= 0
20/03/19 15:28:52 INFO util.ExitUtil: Exiting with status 0
20/03/19 15:28:52 INFO namenode.NameNode: SHUTDOWN_MSG:
*****SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1*****
*****SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1*****
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/
bin/    data/   etc/   include/ lib/   libexec/ sbin/   share/
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/sbin/
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ 
rafiulamin@ubuntu:~$ hadoop-2.8.0-rat.txt.asc 2017-03-25 00:33 801
rafiulamin@ubuntu:~$ hadoop-2.8.0-rat.txt.mds 2017-03-25 00:33 156
```

We run "./start.sh":



```
index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/sbin
File Edit View Search Terminal Help
0 KB
20/03/19 15:28:52 INFO util.GSet: capacity      = 2^15 = 32768 entries
20/03/19 15:28:52 INFO namenode.FSImage: Allocated new BlockPoolId: BP-724796247-1
27.0.1.1-1584656932671
20/03/19 15:28:52 INFO common.Storage: Storage directory /usr/local/hadoop/data/namenode has been successfully formatted.
20/03/19 15:28:52 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_000000000000000000000000 using no compression
20/03/19 15:28:52 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_000000000000000000000000 of size 327 bytes saved in 0 seconds.
20/03/19 15:28:52 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
20/03/19 15:28:52 INFO util.ExitUtil: Exiting with status 0
20/03/19 15:28:52 INFO namenode.NameNode: SHUTDOWN_MSG:
*****SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
*****
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/
bin/    data/    etc/    include/    lib/    libexec/    sbin/    share/
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/sbin/
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ ./start-all.sh
hadoop-2.8.0-start.txt.out 2017-03-25 00:33 801
```

Facing problem with connection refused:



```
File Edit View Search Terminal Help
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
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [0.0.0.0]
0.0.0.0: ssh: connect to host 0.0.0.0 port 22: Connection refused
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-2.8.0.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
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-rafiulamin-resourcemanager-ubuntu.out
localhost: ssh: connect to host localhost port 22: Connection refused
rafiulamin@ubuntu:/usr/local/hadoop/sbin$
```

Therefore we re-install open-ssh:

The image shows two terminal windows side-by-side. Both windows have a title bar 'rafiulamin@ubuntu: /usr/local/hadoop/sbin' and a status bar at the bottom showing the date and time.

Terminal 1 (Top):

```
rafiulamin@ubuntu: /usr/local/hadoop/sbin
File Edit View Search Terminal Help
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
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [0.0.0.0]
0.0.0.0: ssh: connect to host 0.0.0.0 port 22: Connection refused
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-2.8.0.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
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-rafiulamin-resourcemanager-ubuntu.out
localhost: ssh: connect to host localhost port 22: Connection refused
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ sudo apt-get remove openssh-client openssh-server
```

Terminal 2 (Bottom):

```
rafiulamin@ubuntu: /usr/local/hadoop/sbin
File Edit View Search Terminal Help
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
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-rafiulamin-resourcemanager-ubuntu.out
localhost: ssh: connect to host localhost port 22: Connection refused
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ sudo apt-get remove openssh-client openssh-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
Package 'openssh-server' is not installed, so not removed
The following package was automatically installed and is no longer required:
  libssl1.0.0
Use 'sudo apt autoremove' to remove it.
The following packages will be REMOVED:
  gnome-software-plugin-snap openssh-client snapd
0 upgraded, 0 newly installed, 3 to remove and 78 not upgraded.
After this operation, 69.5 MB disk space will be freed.
Do you want to continue? [Y/n] y
(Reading database ... 131147 files and directories currently installed.)
Removing gnome-software-plugin-snap (3.28.1-0ubuntu4.18.04.14) ...
Removing snapd (2.42.1+18.04) ...
Removing openssh-client (1:7.6p1-4ubuntu0.3) ...
Processing triggers for desktop-file-utils (0.23-1ubuntu3.18.04.2) ...
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
Processing triggers for gnome-menus (3.13.3-11ubuntu1.1) ...
Processing triggers for mime-support (3.60ubuntu1) ...
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ sudo apt-get install openssh-client openssh-server
```

We do the “namenode” formatting again and the run “start-all.sh” from Hadoop “sbin”:

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: ~

File Edit View Search Terminal Help
Unpacking ncurses-term (6.1-1ubuntu1.18.04) ...
Selecting previously unselected package openssh-sftp-server.
Preparing to unpack .../openssh-sftp-server_1%3a7.6p1-4ubuntu0.3_amd64.deb ...
Unpacking openssh-sftp-server (1:7.6p1-4ubuntu0.3) ...
Selecting previously unselected package openssh-server.
Preparing to unpack .../openssh-server_1%3a7.6p1-4ubuntu0.3_amd64.deb ...
Unpacking openssh-server (1:7.6p1-4ubuntu0.3) ...
Selecting previously unselected package ssh-import-id.
Preparing to unpack .../ssh-import-id_5.7-0ubuntu1.1_all.deb ...
Unpacking ssh-import-id (5.7-0ubuntu1.1) ...
Setting up ncurses-term (6.1-1ubuntu1.18.04) ...
Setting up openssh-client (1:7.6p1-4ubuntu0.3) ...
Setting up ssh-import-id (5.7-0ubuntu1.1) ...
Setting up openssh-sftp-server (1:7.6p1-4ubuntu0.3) ...
Setting up openssh-server (1:7.6p1-4ubuntu0.3) ...

Creating config file /etc/ssh/sshd_config with new version
Creating SSH2 RSA key; this may take some time ...
2048 SHA256:q30DYfb+dtwBFjRzMfpzMTKs8aIaylZ+f4hq90qi2Y root@ubuntu (RSA)
Creating SSH2 ECDSA key; this may take some time ...
256 SHA256:Ph5jgpQjQnS0wZa7H3V43CmGzbpaQLFWQBUFKYH0+4 root@ubuntu (ECDSA)
Creating SSH2 ED25519 key; this may take some time ...
256 SHA256:BjxgJWJzP1ya0dVGqHLzNSkrrWg3rZN6Kqo5SKdxZEE root@ubuntu (ED25519)
Created symlink /etc/systemd/system/sshd.service → /lib/systemd/system/ssh.service
.
Created symlink /etc/systemd/system/multi-user.target.wants/ssh.service → /lib/systemd/system/ssh.service.
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
Processing triggers for ufw (0.36-0ubuntu0.18.04.1) ...
Processing triggers for ureadahead (0.100.0-21) ...
Processing triggers for systemd (237-3ubuntu10.33) ...
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ cd
rafiulamin@ubuntu:~$ hdfs namenode -format
hadoop-2.8.0-src.tar.gz 2017-03-25 00:33 33M
```

```
Index of /dist/hadoop/common/hadoop-2.8.0 - Mozilla Firefox
rafiulamin@ubuntu: /usr/local/hadoop/sbin

File Edit View Search Terminal Help
20/03/19 15:42:30 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1928706493-127.0.1.1-1584657750352
20/03/19 15:42:30 INFO common.Storage: Storage directory /usr/local/hadoop/data/namenode has been successfully formatted.
20/03/19 15:42:30 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
20/03/19 15:42:30 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop/data/namenode/current/fsimage.ckpt_00000000000000000000 of size 327 bytes saved in 0 seconds.
20/03/19 15:42:30 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
20/03/19 15:42:30 INFO util.ExitUtil: Exiting with status 0
20/03/19 15:42:30 INFO namenode.NameNode: SHUTDOWN_MSG:
*****SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1*****
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/
rafiulamin@ubuntu:~$ cd /usr/local/hadoop/sbin/
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ ls
distribute-exclude.sh    start-all.cmd      stop-balancer.sh
hadoop-daemon.sh        start-all.sh       stop-dfs.cmd
hadoop-daemons.sh       start-balancer.sh  stop-dfs.sh
hdfs-config.cmd         start-dfs.cmd     stop-secure-dns.sh
hdfs-config.sh          start-dfs.sh      stop-yarn.cmd
httpfs.sh                start-secure-dns.sh stop-yarn.sh
kms.sh                  start-yarn.cmd   yarn-daemon.sh
mr-jobhistory-daemon.sh start-yarn.sh    yarn-daemons.sh
refresh-namenodes.sh    stop-all.cmd
slaves.sh                stop-all.sh
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ ./start-all.sh
hadoop-2.8.0-src.tar.gz 2017-03-25 00:33 33M
```

It asks for passwords to log into different nodes. Finally all connections are established.

The screenshot shows a terminal window titled "rafiulamin@ubuntu: /usr/local/hadoop/sbin". The window contains the following text output:

```
Are you sure you want to continue connecting (yes/no)? y
Please type 'yes' or 'no': yes
localhost: Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
rafiulamin@localhost's password:
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-rafiulamin-namenode-ubuntu.out
rafiulamin@localhost's password:
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-rafiulamin-datanode-ubuntu.out
Starting secondary namenodes [0.0.0.0]
The authenticity of host '0.0.0.0 (0.0.0.0)' can't be established.
ECDSA key fingerprint is SHA256:Ph5jgpQjQnS0wZa7H3V43CmGzJbpAqlFWQBUFKYH0+4.
Are you sure you want to continue connecting (yes/no)? yes
0.0.0.0: Warning: Permanently added '0.0.0.0' (ECDSA) to the list of known hosts.
rafiulamin@0.0.0.0's password:
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-rafiulamin-secondarynamenode-ubuntu.out
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-2.8.0.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
starting yarn daemons
resource manager running as process 12523. Stop it first.
rafiulamin@localhost's password:
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-rafiulamin-nodemanager-ubuntu.out
rafiulamin@ubuntu:/usr/local/hadoop/sbin$
```

At the bottom of the terminal window, there is a status bar with the text "hadoop-2.8.0-src.tar.gz 2017-03-25 00:33 33M".

We check the sites as well.

Index of /dist/hadoop/com ... | Namenode information | + | localhost:50070/dfshealth.html#tab-overview

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Overview 'localhost:9000' (active)

Started:	Thu Mar 19 15:44:13 -0700 2020
Version:	2.8.0, r91f2b7a13d1e97be65db92ddabc627cc29ac0009
Compiled:	Thu Mar 19 21:12:00 -0700 2017 by jdu from branch-2.8.0
Cluster ID:	CID-42cf5fc-f4f9-4131-8f44-64d7d4fe14d8
Block Pool ID:	BP-1928706493-127.0.1.1-1584657750352

Summary

Security is off.
Safemode is off.
1 files and directories, 0 blocks = 1 total filesystem object(s).
Heap Memory used 32.99 MB of 79.2 MB Heap Memory. Max Heap Memory is 966.69 MB.
Non Heap Memory used 42.99 MB of 46.75 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	19.56 GB
DFS Used:	24 KB (0%)
Non DFS Used:	8.68 GB
DFS Remaining:	9.86 GB (50.4%)

All Applications - Mozilla Firefox

index of /dist/hadoop/com ... | Namenode information | All Applications | + | localhost:8088/cluster

Logged in as: dr.who

All Applications

hadoop

Cluster Metrics

Apps Submitted	0	Apps Pending	0	Apps Running	0	Apps Completed	0	Containers Running	0	Memory Used	0 B	Memory Total	8 GB	Memory Reserved	0 B	Vcores Used	0	Vcores Total	8	Vcores Reserved	0
----------------	---	--------------	---	--------------	---	----------------	---	--------------------	---	-------------	-----	--------------	------	-----------------	-----	-------------	---	--------------	---	-----------------	---

Cluster Nodes Metrics

Active Nodes	1	Decommissioning Nodes	0	Decommissioned Nodes	0	Lost Nodes	0	Unhealthy Nodes	0	Rebooted Nodes	0	Shutdown Nodes	0
--------------	---	-----------------------	---	----------------------	---	------------	---	-----------------	---	----------------	---	----------------	---

Scheduler Metrics

Scheduler Type	Capacity Scheduler	Scheduling Resource Type	[MEMORY]	Minimum Allocation	<memory:1024, vCores:1>	Maximum Allocation	<memory:8192, vCores:4>	Maximum Cluster Application Priority	0
----------------	--------------------	--------------------------	----------	--------------------	-------------------------	--------------------	-------------------------	--------------------------------------	---

Show: 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU	Allocated Memory	% of Queue	% of Cluster	Progress	Tracking UI	Blacklisted Nodes	
No data available in table																		

Showing 0 to 0 of 0 entries

First Previous Next Last

Now we install git in order to clone kmean-mapreduce code:

```
rafiulamin@ubuntu: ~
File Edit View Search Terminal Help
rafiulamin@localhost's password:
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-rafiulamin-
datanode-ubuntu.out
Starting secondary namenodes [0.0.0.0]
The authenticity of host '0.0.0.0 (0.0.0.0)' can't be established.
ECDSA key fingerprint is SHA256:Ph5jgpQjQnS0wZa7H3V43CmGzJbpAQLFWQBUFKYH0+4.
Are you sure you want to continue connecting (yes/no)? yes
0.0.0.0: Warning: Permanently added '0.0.0.0' (ECDSA) to the list of known hosts.
rafiulamin@0.0.0.0's password:
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-rafi
ulamin-secondarynamenode-ubuntu.out
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.ut
il.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.8.0.
jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.se
curity.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflectiv
e access operations
WARNING: All illegal access operations will be denied in a future release
starting yarn daemons
resourcemanager running as process 12523. Stop it first.
rafiulamin@localhost's password:
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-rafiulamin
-nodemanager-ubuntu.out
rafiulamin@ubuntu:/usr/local/hadoop/sbin$ cd
rafiulamin@ubuntu:~$ git clone https://github.com/Maki94/kmeans_mapreduce.git

Command 'git' not found, but can be installed with:

sudo apt install git

rafiulamin@ubuntu:~$
```

We use the following command to install git:

```
$ sudo apt install git
```

Then we clone it from the specified git-hub link:

```
Get:2 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 git amd64 1:2.17.1-1ubuntu0.5 [803 kB]
Get:3 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 git-amd64 1:2.17.1-1ubuntu0.5 [3,912 kB]
Fetched 4,738 kB in 4s (1,097 kB/s)
Selecting previously unselected package liberror-perl.
(Reading database ... 133891 files and directories currently installed.)
Preparing to unpack .../liberror-perl_0.17025-1_all.deb ...
Unpacking liberror-perl (0.17025-1) ...
Selecting previously unselected package git-man.
Preparing to unpack .../git-man_1%3a2.17.1-1ubuntu0.5_all.deb ...
Unpacking git-man (1:2.17.1-1ubuntu0.5) ...
Selecting previously unselected package git.
Preparing to unpack .../git_1%3a2.17.1-1ubuntu0.5_amd64.deb ...
Unpacking git (1:2.17.1-1ubuntu0.5) ...
Setting up git-man (1:2.17.1-1ubuntu0.5) ...
Setting up liberror-perl (0.17025-1) ...
Setting up git (1:2.17.1-1ubuntu0.5) ...
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
rafiulamin@ubuntu:~$ git clone https://github.com/Maki94/kmeans_mapreduce.git
Cloning into 'kmeans_mapreduce'...
remote: Enumerating objects: 153, done.
remote: Total 153 (delta 0), reused 0 (delta 0), pack-reused 153
Receiving objects: 100% (153/153), 52.03 MiB | 23.65 MiB/s, done.
Resolving deltas: 100% (32/32), done.
rafiulamin@ubuntu:~$ █
```

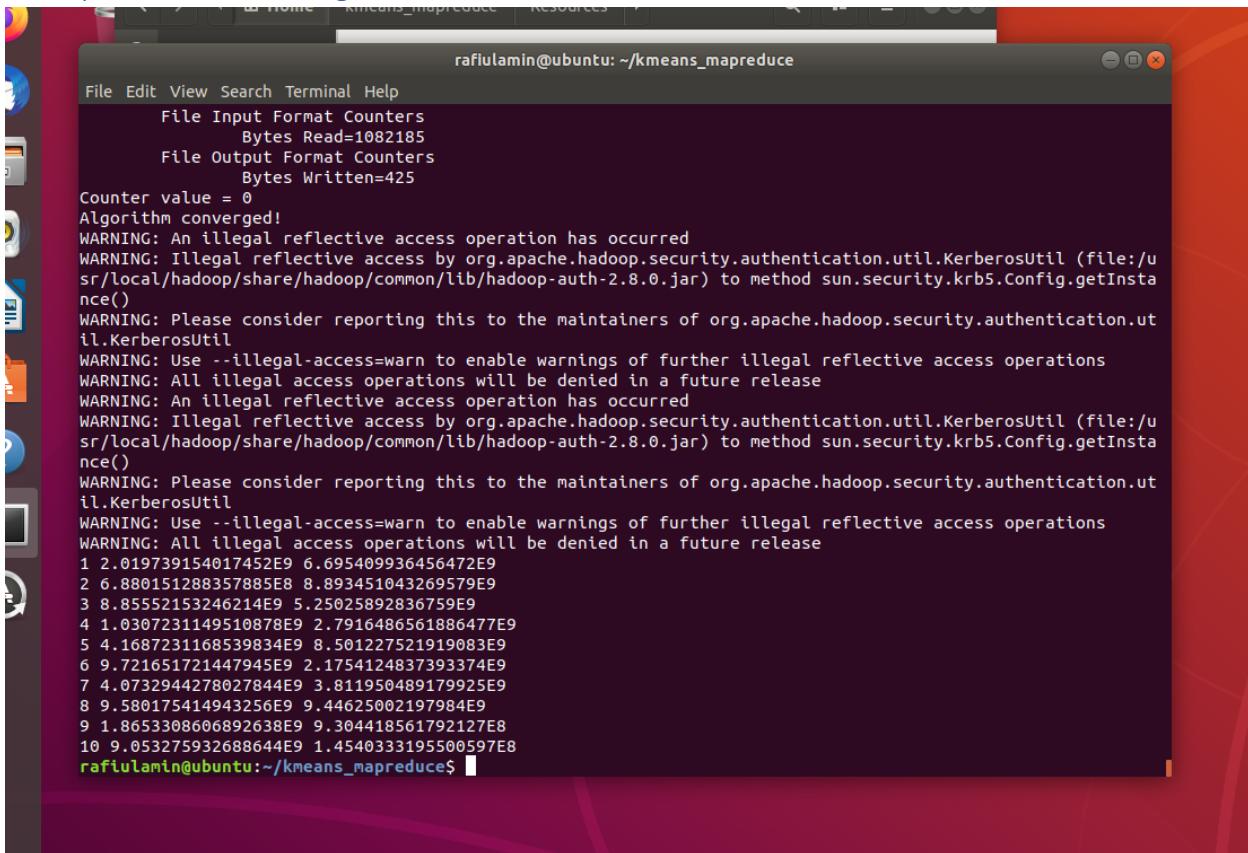
Now is the time to run it!

```
rafiulamin@ubuntu: ~/kmeans_mapreduce
File Edit View Search Terminal Help
0 upgraded, 3 newly installed, 0 to remove and 78 not upgraded.
Need to get 4,738 kB of archives.
After this operation, 33.9 MB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://us.archive.ubuntu.com/ubuntu bionic/main amd64 liberror-perl all 0.17
025-1 [22.8 kB]
Get:2 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 git-man all 1:
2.17.1-1ubuntu0.5 [803 kB]
Get:3 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 git amd64 1:2.
17.1-1ubuntu0.5 [3,912 kB]
Fetched 4,738 kB in 4s (1,097 kB/s)
Selecting previously unselected package liberror-perl.
(Reading database ... 133891 files and directories currently installed.)
Preparing to unpack .../liberror-perl_0.17025-1_all.deb ...
Unpacking liberror-perl (0.17025-1) ...
Selecting previously unselected package git-man.
Preparing to unpack .../git-man_1%3a2.17.1-1ubuntu0.5_all.deb ...
Unpacking git-man (1:2.17.1-1ubuntu0.5) ...
Selecting previously unselected package git.
Preparing to unpack .../git_1%3a2.17.1-1ubuntu0.5_amd64.deb ...
Unpacking git (1:2.17.1-1ubuntu0.5) ...
Setting up git-man (1:2.17.1-1ubuntu0.5) ...
Setting up liberror-perl (0.17025-1) ...
Setting up git (1:2.17.1-1ubuntu0.5) ...
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
rafiulamin@ubuntu:~$ git clone https://github.com/Maki94/kmeans_mapreduce.git
Cloning into 'kmeans_mapreduce'...
remote: Enumerating objects: 153, done.
remote: Total 153 (delta 0), reused 0 (delta 0), pack-reused 153
Receiving objects: 100% (153/153), 52.03 MiB | 23.65 MiB/s, done.
Resolving deltas: 100% (32/32), done.
rafiulamin@ubuntu:~$ cd kmeans_mapreduce
rafiulamin@ubuntu:~/kmeans_mapreduce$ ./run.sh
```

Kmeans is running in Hadoop:

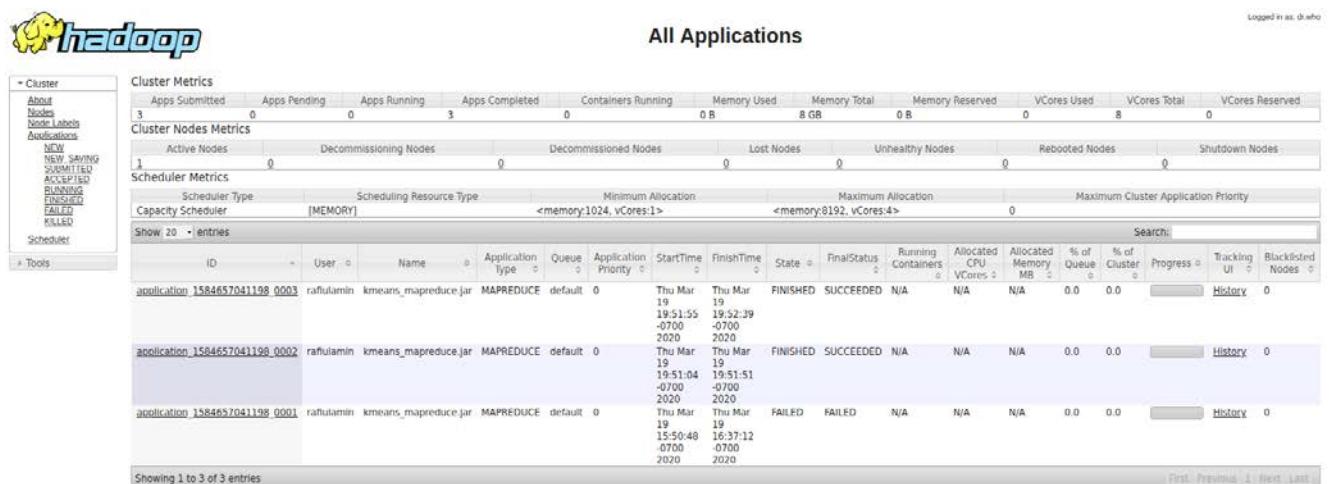
```
Index of /dist/hadoop/comn X | Namenode information X All Applications X +  
raflulamin@ubuntu: ~/kmeans_mapreduce  
File Edit View Search Terminal Help  
state=/KMeans/Resources/Input/clusters.txt  
output=/KMeans/Resources/Output  
max=10  
number=3  
delta=1000000000.0  
distance=eucl  
#####  
WARNING: An illegal reflective access operation has occurred  
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.ut  
il.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.8.0.  
jar) to method sun.security.krb5.Config.getInstance()  
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.se  
curity.authentication.util.KerberosUtil  
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflectiv  
e access operations  
WARNING: All illegal access operations will be denied in a future release  
First iteration adding /KMeans/Resources/Input/clusters.txt  
20/03/19 15:50:44 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8  
032  
20/03/19 15:50:46 INFO input.FileInputFormat: Total input files to process : 1  
20/03/19 15:50:47 INFO mapreduce.JobSubmitter: number of splits:1  
20/03/19 15:50:47 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1584  
657041198_0001  
20/03/19 15:50:48 INFO impl.YarnClientImpl: Submitted application application_1584  
657041198_0001  
20/03/19 15:50:48 INFO mapreduce.Job: The url to track the job: http://ubuntu:8088  
/proxy/application_1584657041198_0001/  
20/03/19 15:50:48 INFO mapreduce.Job: Running job: job_1584657041198_0001  
20/03/19 15:56:22 INFO mapreduce.Job: Job job_1584657041198_0001 running in uber m  
ode : false  
20/03/19 15:56:22 INFO mapreduce.Job: map 0% reduce 0%  
20/03/19 15:56:41 INFO mapreduce.Job: map 100% reduce 0%
```

Finally it finishes running:



```
rafiulamin@ubuntu: ~/kmeans_mapreduce
File Edit View Search Terminal Help
  File Input Format Counters
    Bytes Read=1082185
  File Output Format Counters
    Bytes Written=425
Counter value = 0
Algorithm converged!
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-2.8.0.jar) to method sun.security.krb5.Config.getInsta
nce()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.ut
il.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
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/u
sr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.8.0.jar) to method sun.security.krb5.Config.getInsta
nce()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.ut
il.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
1 2.019739154017452E9 6.695409936456472E9
2 6.880151288357885E8 8.893451043269579E9
3 8.85552153246214E9 5.25025892836759E9
4 1.0307231149510878E9 2.7916486561886477E9
5 4.1687231168539834E9 8.501227521919083E9
6 9.721651721447945E9 2.1754124837393374E9
7 4.0732944278027844E9 3.811950489179925E9
8 9.580175414943256E9 9.44625002197984E9
9 1.8653308606892638E9 9.304418561792127E8
10 9.053275932688644E9 1.4540333195500597E8
rafiulamin@ubuntu:~/kmeans_mapreduce$
```

We also check the site:



The screenshot shows the Hadoop Web UI's "All Applications" page. The left sidebar has sections for Cluster (About Nodes, Node Labels, Applications), Tools, and a log viewer. The main area displays cluster metrics, node metrics, and a detailed table of application logs.

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved
3	0	0	3	0	0 B	8 GB	0 B	0	8	0

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes	Shutdown Nodes
1	0	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

All Applications

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCores	Allocated Memory MB	% of Queue	% of Cluster	Progress	Tracking UI	Blacklisted Nodes
application_1584657041198_0001	rafiulamin	kmeans_mapreduce.jar	MAPREDUCE	default	0	Thu Mar 19 19:51:55 2020	Thu Mar 19 19:52:39 2020	FINISHED	SUCCEEDED	N/A	N/A	0.0	0.0	0	History	0	
application_1584657041198_0002	rafiulamin	kmeans_mapreduce.jar	MAPREDUCE	default	0	Thu Mar 19 19:51:04 2020	Thu Mar 19 19:51:51 2020	FINISHED	SUCCEEDED	N/A	N/A	0.0	0.0	0	History	0	
application_1584657041198_0003	rafiulamin	kmeans_mapreduce.jar	MAPREDUCE	default	0	Thu Mar 19 19:50:48 2020	Thu Mar 19 16:37:12 2020	FAILED	FAILED	N/A	N/A	0.0	0.0	0	History	0	

Showing 1 to 3 of 3 entries