Apache Hadoop 3.3.0 installation on Ubuntu Part 1

Apache Hadoop 3.3.0 installation on Ubuntu Part 1

With this tutorial, we will learn the complete process to install Hadoop 3.3.1 on Ubuntu 20.

Supported Java Versions

- Apache Hadoop 3.3 and upper supports Java 8 and Java 11 (runtime only)
- Please compile Hadoop with Java 8. Compiling Hadoop with Java 11 is not supported: HADOOP-16795 – Java 11 compile support OPEN
- Apache Hadoop from 3.0.x to 3.2.x now supports only Java 8
- Apache Hadoop from 2.7.x to 2.10.x support both Java 7 and 8

Required software for Linux include:

- Java must be installed. Recommended Java versions are described at HadoopJavaVersions.
- ssh must be installed and sshd must be running to use the Hadoop scripts that manage remote Hadoop daemons if the optional start and stop scripts are to be used.

Steps for Installing JAVA 8 on Ubuntu

Step 1 – Install Java 8 on Ubuntu

The OpenJDK 8 is available under default Apt repositories. You can simply install Java 8 on an Ubuntu system using the following commands.

```
$sudo apt update
$sudo apt install openjdk-8-jdk -y
```

Step 2 – Verify Java Installation

You have successfully installed Java 8 on your system. Let’s verify the installed and current active version using the following command.

```
$java -version
openjdk version "1.8.0_252"
OpenJDK Runtime Environment (build 1.8.0_252-8u252-b09-1ubuntu1-b09)
OpenJDK 64-Bit Server VM (build 25.252-b09, mixed mode)
```

Step 3 – Setup JAVA_HOME and JRE_HOME Variable

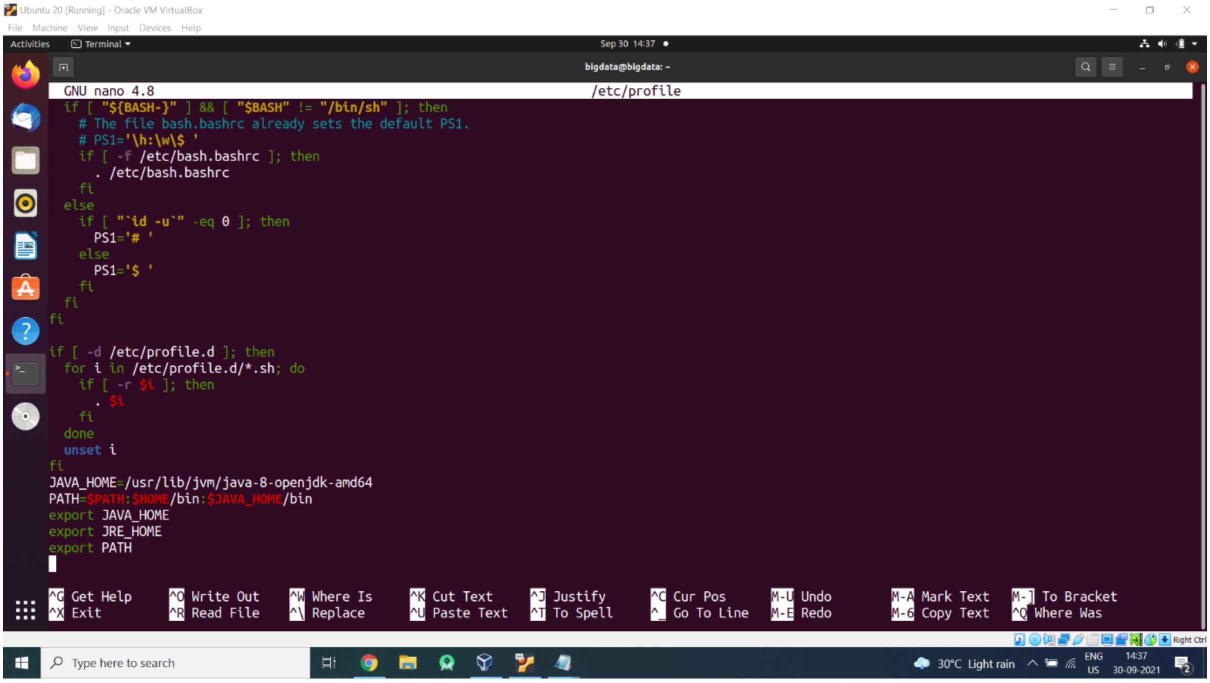
As you have installed Java on your Linux system, You must have to set JAVA_HOME and JRE_HOME environment variables,

Edit the system Path file /etc/profile

```
sudo nano /etc/profile
```

Add the following lines at the end

```
JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
PATH=$PATH:$HOME/bin:$JAVA_HOME/bin
export JAVA_HOME
export JRE_HOME
export PATH
```



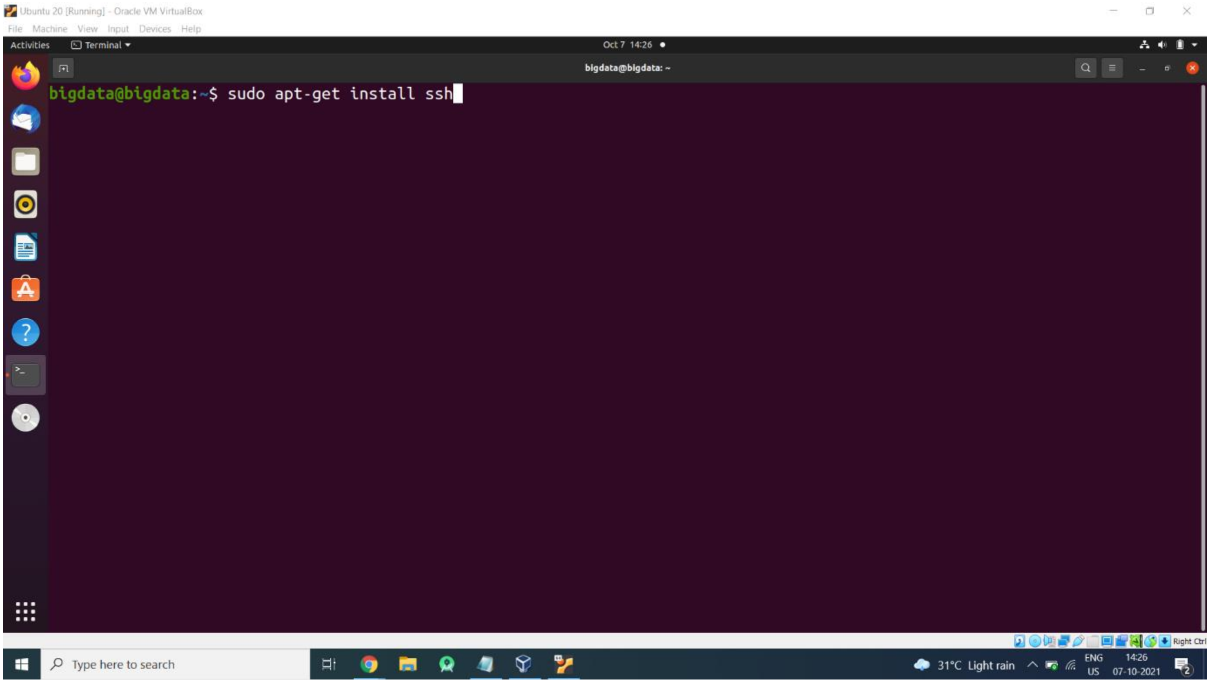
```
  GNU nano 4.8                                  /etc/profile
  if [ "${BASH-}" ] && [ "$BASH" != "/bin/sh" ]; then
    # The file bash.bashrc already sets the default PS1.
    # PS1='\h:\w\$ '
    if [ -f /etc/bash.bashrc ]; then
      . /etc/bash.bashrc
    fi
  else
    if [ "`id -u`" -eq 0 ]; then
      PS1='# '
    else
      PS1='$ '
    fi
  fi
fi

if [ -d /etc/profile.d ]; then
  for i in /etc/profile.d/*.sh; do
    if [ -r $i ]; then
      . $i
    fi
  done
  unset i
fi
JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
PATH=$PATH:$HOME/bin:$JAVA_HOME/bin
export JAVA_HOME
export JRE_HOME
export PATH
█

^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify     ^C Cur Pos     M-U Undo       M-A Mark Text  M-] To Bracket
^X Exit        ^R Read File   ^\ Replace     ^U Paste Text  ^T To Spell    ^_ Go To Line  M-E Redo       M-6 Copy Text  ^Q Where Was
```

Steps for Installing ssh on Ubuntu

Secure Shell (SSH) is a cryptographic network protocol for operating network services securely over an unsecured network. Typical applications include remote command-line, login, and remote command execution, but any network service can be secured with SSH.

Install ssh on your system using the below command:

```
sudo apt-get install ssh
```

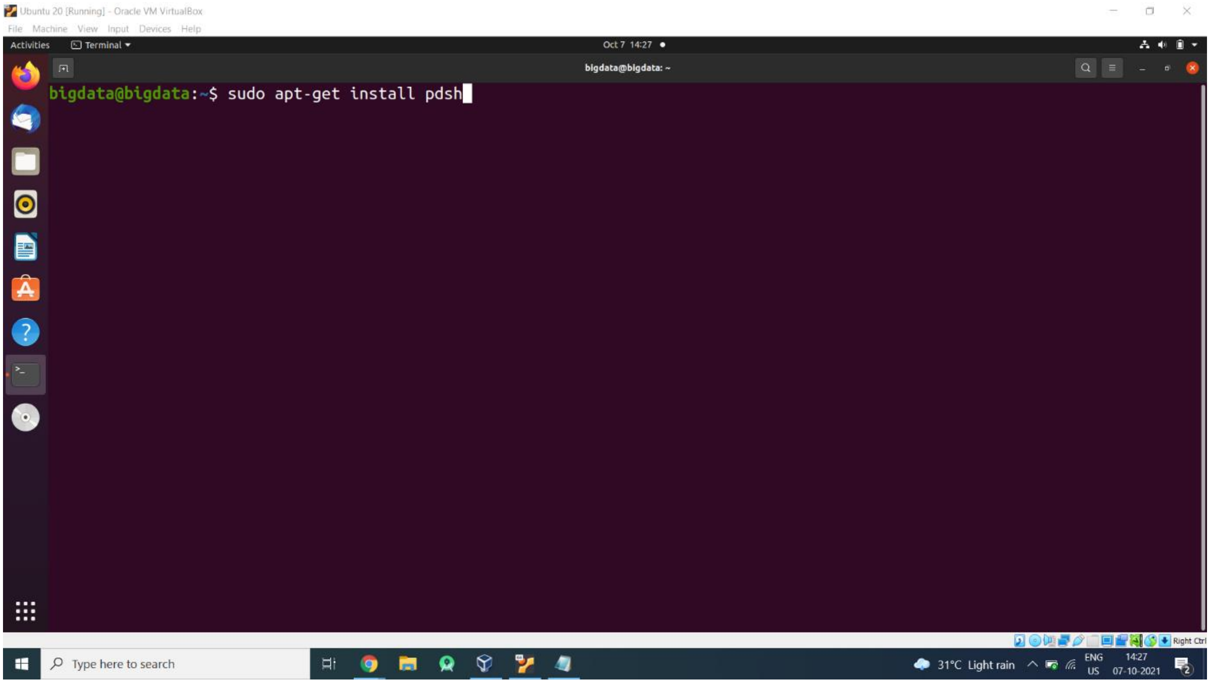


```
bigdata@bigdata:~$ sudo apt-get install ssh
```

Type the password for the sudo user and then press Enter.

Install pdsh on your system using the below command:

```
sudo apt-get install pdsh
```

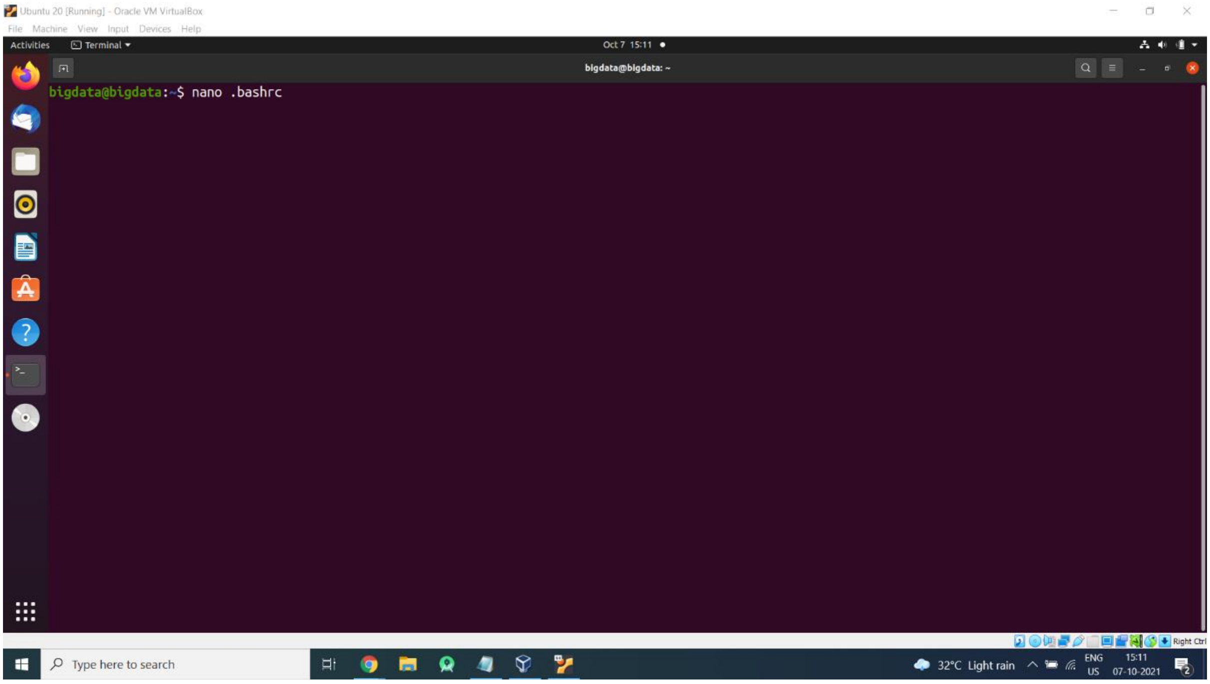


```
bigdata@bigdata:~$ sudo apt-get install pdsh
```

Type ‘Y’ and then press Enter to continue with the installation process.

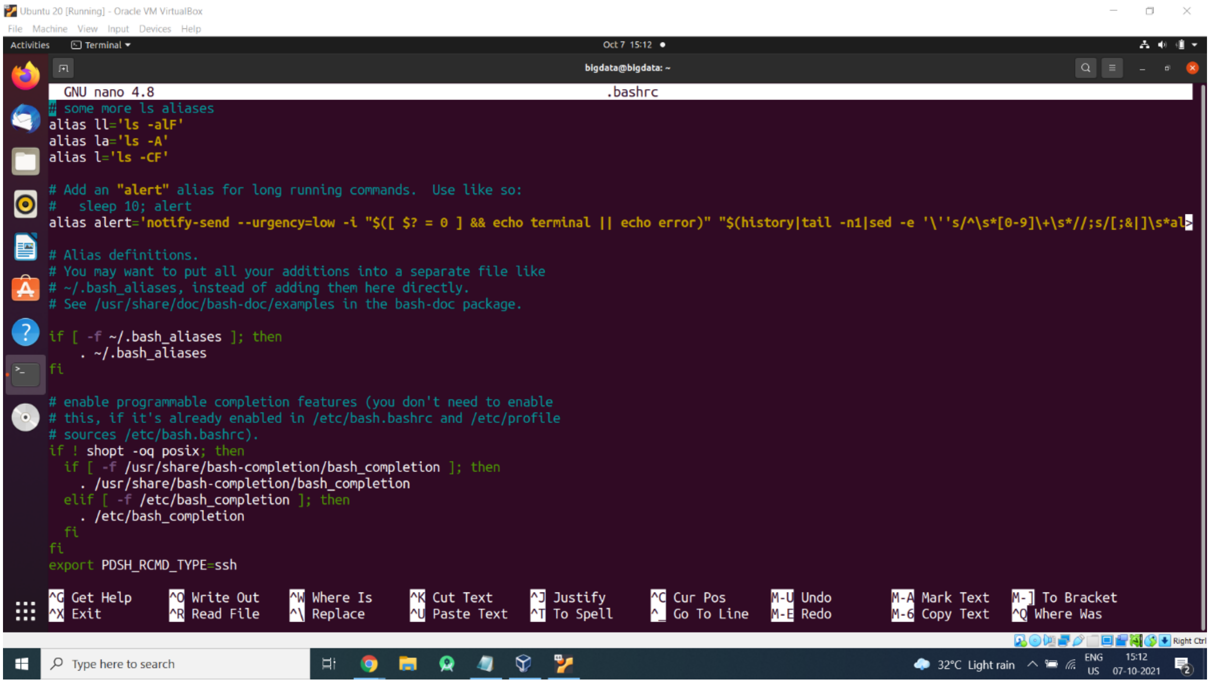
Open the .bashrc file in the nano editor using the following command:

```
nano .bashrc
```



```
bigdata@bigdata:~$ nano .bashrc
```

Now set the PDSH_RCMD_TYPE environment variable to ssh



```
  GNU nano 4.8                                         .bashrc
# some more ls aliases
alias ll='ls -alF'
alias la='ls -A'
alias l='ls -CF'

# Add an "alert" alias for long running commands.  Use like so:
#   sleep 10; alert
alias alert='notify-send --urgency=low -i "$([ $? = 0 ] && echo terminal || echo error)" "$(history|tail -n1|sed -e '\''s/^\s*[0-9]\+\s*//;s/[;&|]\s*al>

# Alias definitions.
# You may want to put all your additions into a separate file like
# ~/.bash_aliases, instead of adding them here directly.
# See /usr/share/doc/bash-doc/examples in the bash-doc package.

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
export PDSH_RCMD_TYPE=ssh

^G Get Help   ^O Write Out  ^W Where Is   ^K Cut Text   ^J Justify    ^C Cur Pos    M-U Undo      M-A Mark Text  M-] To Bracket
^X Exit       ^R Read File  ^\ Replace    ^U Paste Text ^T To Spell   ^_ Go To Line M-E Redo      M-6 Copy Text  ^Q Where Was
```

Steps for Installing Hadoop on Ubuntu

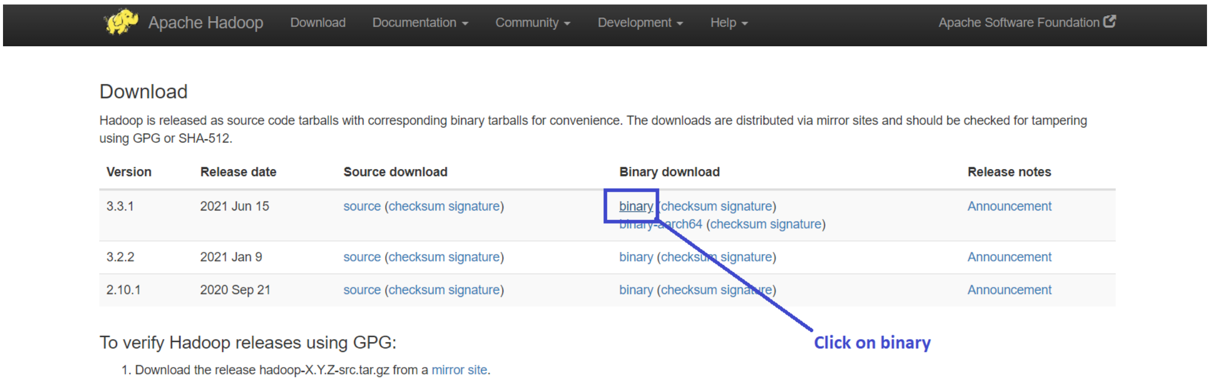
- Create a directory for example

$mkdir /home/bigdata/hadoop

- Move to hadoop directory

$cd /home/bigdata/hadoop

Download Hadoop (Link will change with respect to country so please get the download link from hadoop website ie https://hadoop.apache.org/releases.html



Download

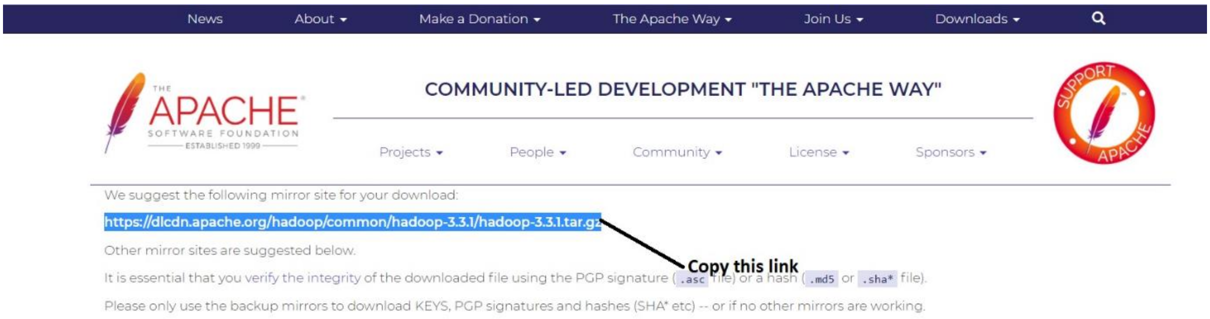
Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-512.

Version	Release date	Source download	Binary download	Release notes
3.3.1	2021 Jun 15	source (checksum signature)	binary (checksum signature) binary-aarch64 (checksum signature)	Announcement
3.2.2	2021 Jan 9	source (checksum signature)	binary (checksum signature)	Announcement
2.10.1	2020 Sep 21	source (checksum signature)	binary (checksum signature)	Announcement

To verify Hadoop releases using GPG:

1. Download the release hadoop-X.Y.Z-src.tar.gz from a mirror site.

A new web page will get open and copy the link



We suggest the following mirror site for your download:

https://dlcdn.apache.org/hadoop/common/hadoop-3.3.1/hadoop-3.3.1.tar.gz

Other mirror sites are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature (.asc file) or a hash (.md5 or .sha* file).

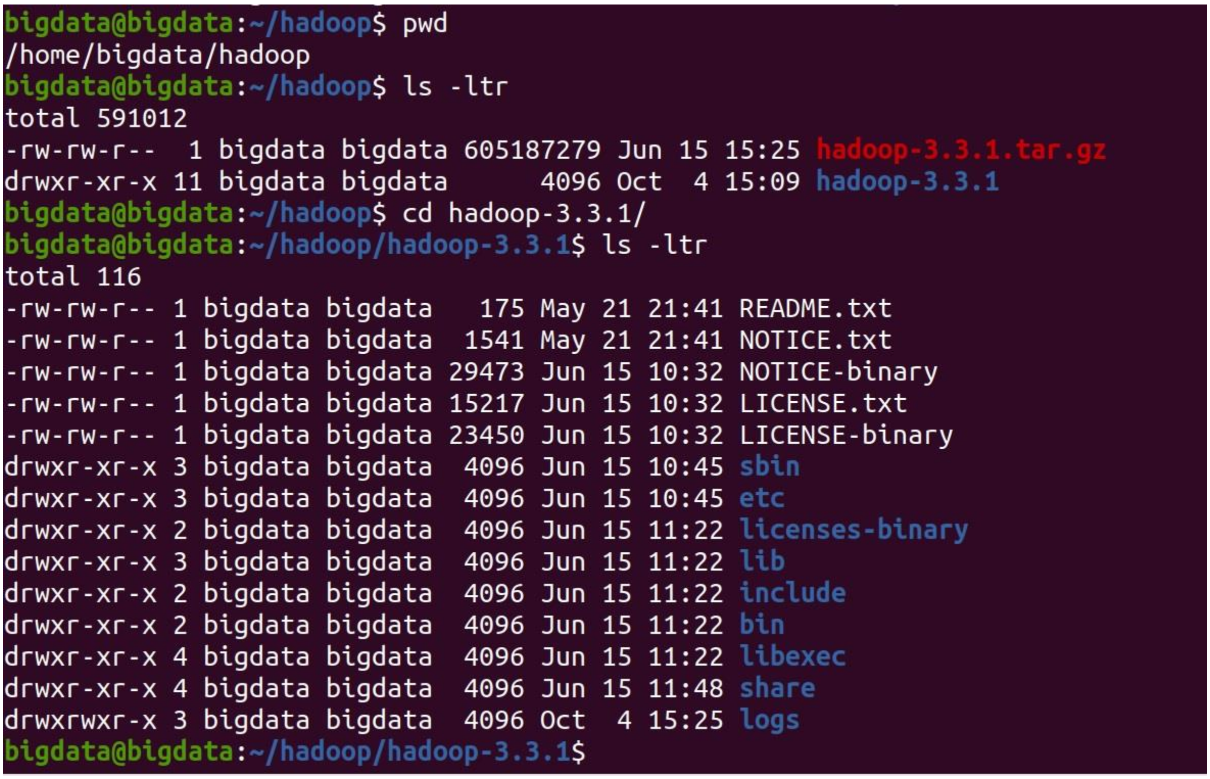
Please only use the backup mirrors to download KEYS, PGP signatures and hashes (SHA* etc) -- or if no other mirrors are working.

In Ubuntu terminal type

```
$wget https://dlcdn.apache.org/hadoop/common/hadoop-3.3.1/hadoop-3.3.1.tar.gz
```

Then type

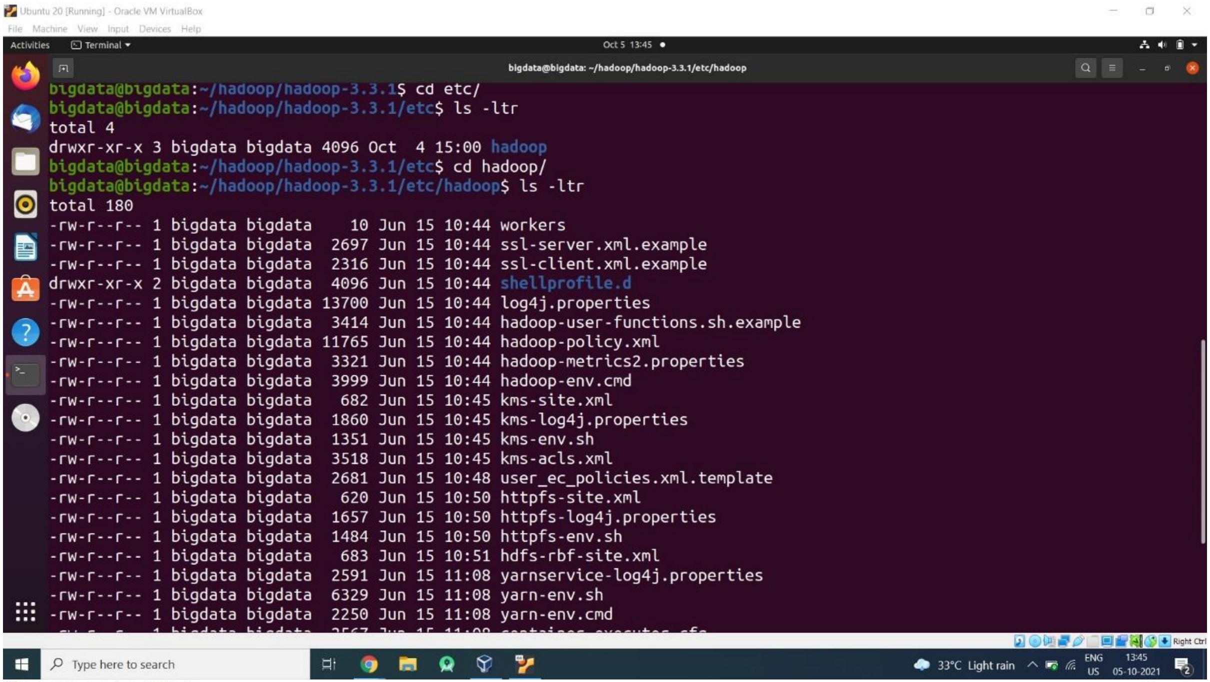
1. `$tar xvf hadoop-3.3.1.tar.gz`
2. `$cd hadoop-3.3.1`



```
bigdata@bigdata:~/hadoop$ pwd
/home/bigdata/hadoop
bigdata@bigdata:~/hadoop$ ls -ltr
total 591012
-rw-rw-r--  1 bigdata bigdata 605187279 Jun 15 15:25 hadoop-3.3.1.tar.gz
drwxr-xr-x 11 bigdata bigdata      4096 Oct  4 15:09 hadoop-3.3.1
bigdata@bigdata:~/hadoop$ cd hadoop-3.3.1/
bigdata@bigdata:~/hadoop/hadoop-3.3.1$ ls -ltr
total 116
-rw-rw-r-- 1 bigdata bigdata   175 May 21 21:41 README.txt
-rw-rw-r-- 1 bigdata bigdata  1541 May 21 21:41 NOTICE.txt
-rw-rw-r-- 1 bigdata bigdata 29473 Jun 15 10:32 NOTICE-binary
-rw-rw-r-- 1 bigdata bigdata 15217 Jun 15 10:32 LICENSE.txt
-rw-rw-r-- 1 bigdata bigdata 23450 Jun 15 10:32 LICENSE-binary
drwxr-xr-x 3 bigdata bigdata  4096 Jun 15 10:45 sbin
drwxr-xr-x 3 bigdata bigdata  4096 Jun 15 10:45 etc
drwxr-xr-x 2 bigdata bigdata  4096 Jun 15 11:22 licenses-binary
drwxr-xr-x 3 bigdata bigdata  4096 Jun 15 11:22 lib
drwxr-xr-x 2 bigdata bigdata  4096 Jun 15 11:22 include
drwxr-xr-x 2 bigdata bigdata  4096 Jun 15 11:22 bin
drwxr-xr-x 4 bigdata bigdata  4096 Jun 15 11:22 libexec
drwxr-xr-x 4 bigdata bigdata  4096 Jun 15 11:48 share
drwxrwxr-x 3 bigdata bigdata  4096 Oct  4 15:25 logs
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
```

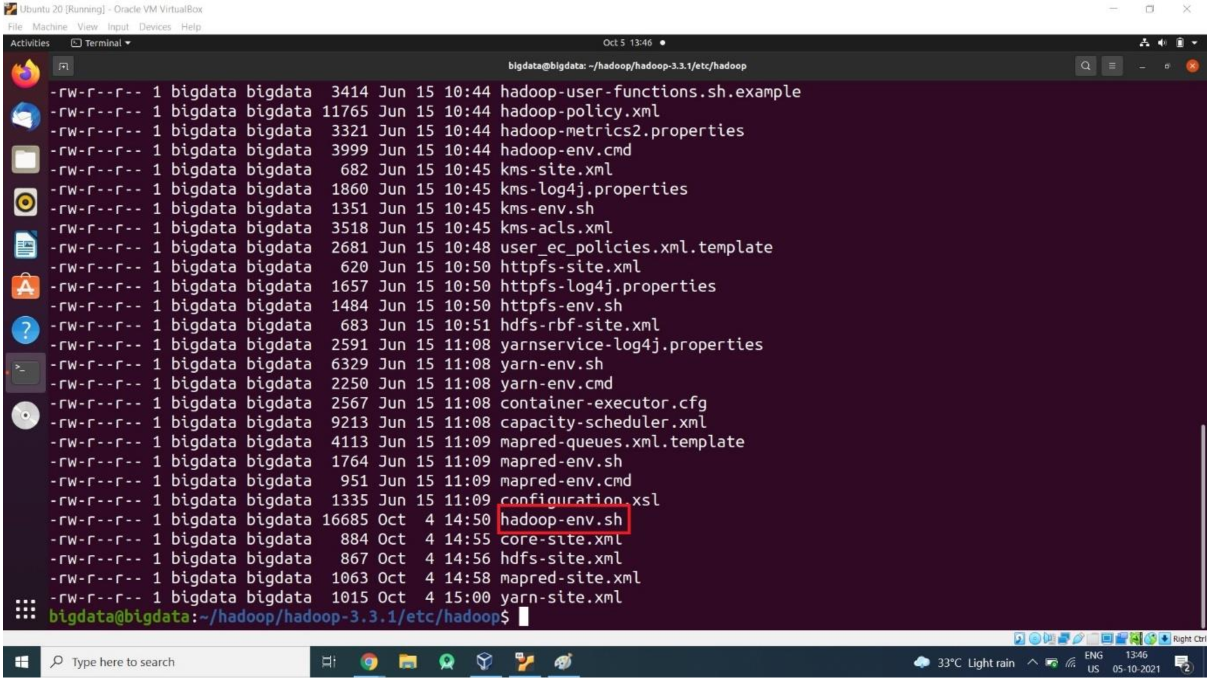
Edit the file etc/hadoop/hadoop-env.sh to define some parameters as follows:

```
$cd etc/
```



Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 5 13:45
bigdata@bigdata: ~/hadoop/hadoop-3.3.1/etc/hadoop
Right Ctrl
Type here to search
33°C Light rain
ENG US
13:45 05-10-2021

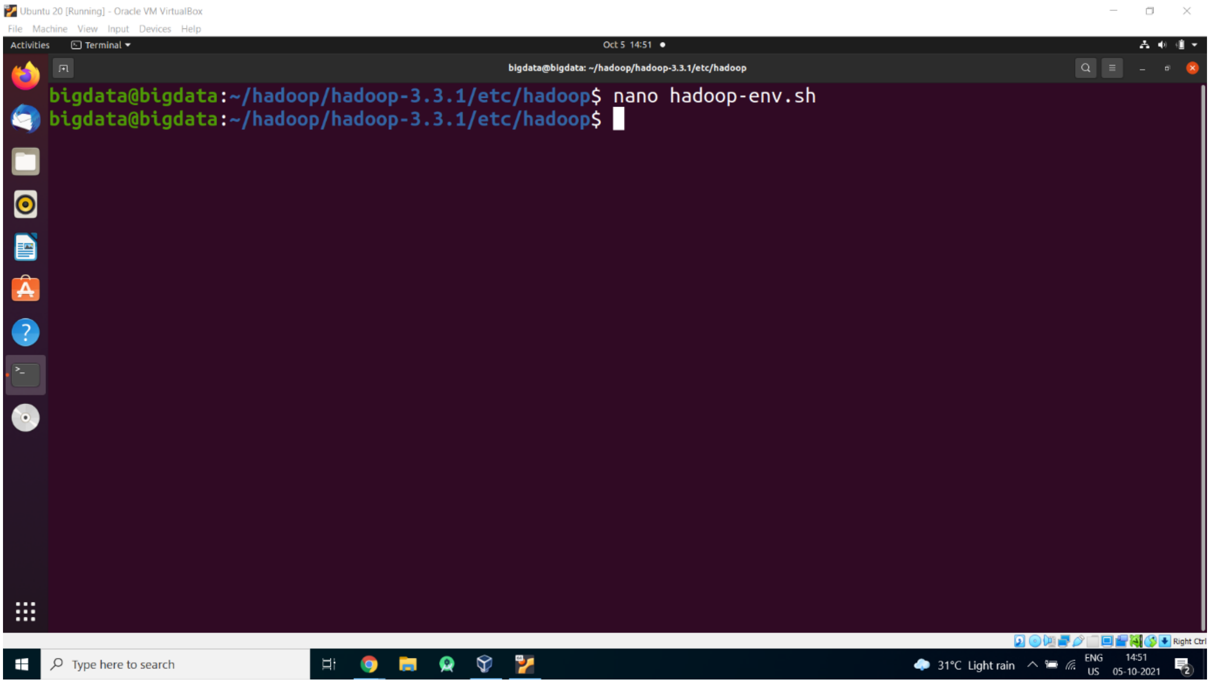
```
bigdata@bigdata:~/hadoop/hadoop-3.3.1$ cd etc/
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc$ ls -ltr
total 4
drwxr-xr-x 3 bigdata bigdata 4096 Oct  4 15:00 hadoop
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc$ cd hadoop/
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc/hadoop$ ls -ltr
total 180
-rw-r--r-- 1 bigdata bigdata    10 Jun 15 10:44 workers
-rw-r--r-- 1 bigdata bigdata  2697 Jun 15 10:44 ssl-server.xml.example
-rw-r--r-- 1 bigdata bigdata  2316 Jun 15 10:44 ssl-client.xml.example
drwxr-xr-x 2 bigdata bigdata  4096 Jun 15 10:44 shellprofile.d
-rw-r--r-- 1 bigdata bigdata 13700 Jun 15 10:44 log4j.properties
-rw-r--r-- 1 bigdata bigdata  3414 Jun 15 10:44 hadoop-user-functions.sh.example
-rw-r--r-- 1 bigdata bigdata 11765 Jun 15 10:44 hadoop-policy.xml
-rw-r--r-- 1 bigdata bigdata  3321 Jun 15 10:44 hadoop-metrics2.properties
-rw-r--r-- 1 bigdata bigdata  3999 Jun 15 10:44 hadoop-env.cmd
-rw-r--r-- 1 bigdata bigdata   682 Jun 15 10:45 kms-site.xml
-rw-r--r-- 1 bigdata bigdata  1860 Jun 15 10:45 kms-log4j.properties
-rw-r--r-- 1 bigdata bigdata  1351 Jun 15 10:45 kms-env.sh
-rw-r--r-- 1 bigdata bigdata  3518 Jun 15 10:45 kms-acls.xml
-rw-r--r-- 1 bigdata bigdata  2681 Jun 15 10:48 user_ec_policies.xml.template
-rw-r--r-- 1 bigdata bigdata   620 Jun 15 10:50 httpfs-site.xml
-rw-r--r-- 1 bigdata bigdata  1657 Jun 15 10:50 httpfs-log4j.properties
-rw-r--r-- 1 bigdata bigdata  1484 Jun 15 10:50 httpfs-env.sh
-rw-r--r-- 1 bigdata bigdata   683 Jun 15 10:51 hdfs-rbf-site.xml
-rw-r--r-- 1 bigdata bigdata  2591 Jun 15 11:08 yarnservice-log4j.properties
-rw-r--r-- 1 bigdata bigdata  6329 Jun 15 11:08 yarn-env.sh
-rw-r--r-- 1 bigdata bigdata  2250 Jun 15 11:08 yarn-env.cmd
-rw-r--r-- 1 bigdata bigdata  2567 Jun 15 11:08 container-executor.cfg
```



Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 5 13:46
bigdata@bigdata: ~/hadoop/hadoop-3.3.1/etc/hadoop
Right Ctrl
Type here to search
33°C Light rain
ENG US
13:46 05-10-2021

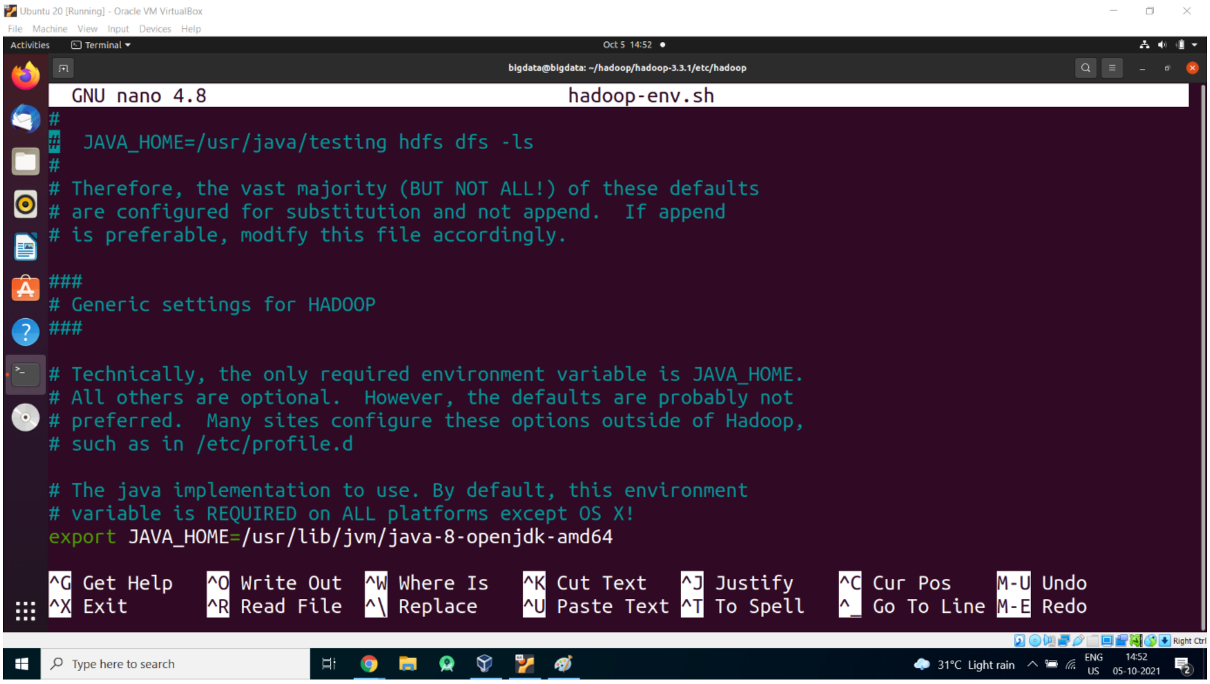
```
-rw-r--r-- 1 bigdata bigdata  3414 Jun 15 10:44 hadoop-user-functions.sh.example
-rw-r--r-- 1 bigdata bigdata 11765 Jun 15 10:44 hadoop-policy.xml
-rw-r--r-- 1 bigdata bigdata  3321 Jun 15 10:44 hadoop-metrics2.properties
-rw-r--r-- 1 bigdata bigdata  3999 Jun 15 10:44 hadoop-env.cmd
-rw-r--r-- 1 bigdata bigdata   682 Jun 15 10:45 kms-site.xml
-rw-r--r-- 1 bigdata bigdata  1860 Jun 15 10:45 kms-log4j.properties
-rw-r--r-- 1 bigdata bigdata  1351 Jun 15 10:45 kms-env.sh
-rw-r--r-- 1 bigdata bigdata  3518 Jun 15 10:45 kms-acls.xml
-rw-r--r-- 1 bigdata bigdata  2681 Jun 15 10:48 user_ec_policies.xml.template
-rw-r--r-- 1 bigdata bigdata   620 Jun 15 10:50 httpfs-site.xml
-rw-r--r-- 1 bigdata bigdata  1657 Jun 15 10:50 httpfs-log4j.properties
-rw-r--r-- 1 bigdata bigdata  1484 Jun 15 10:50 httpfs-env.sh
-rw-r--r-- 1 bigdata bigdata   683 Jun 15 10:51 hdfs-rbf-site.xml
-rw-r--r-- 1 bigdata bigdata  2591 Jun 15 11:08 yarnservice-log4j.properties
-rw-r--r-- 1 bigdata bigdata  6329 Jun 15 11:08 yarn-env.sh
-rw-r--r-- 1 bigdata bigdata  2250 Jun 15 11:08 yarn-env.cmd
-rw-r--r-- 1 bigdata bigdata  2567 Jun 15 11:08 container-executor.cfg
-rw-r--r-- 1 bigdata bigdata  9213 Jun 15 11:08 capacity-scheduler.xml
-rw-r--r-- 1 bigdata bigdata  4113 Jun 15 11:09 mapred-queues.xml.template
-rw-r--r-- 1 bigdata bigdata  1764 Jun 15 11:09 mapred-env.sh
-rw-r--r-- 1 bigdata bigdata   951 Jun 15 11:09 mapred-env.cmd
-rw-r--r-- 1 bigdata bigdata  1335 Jun 15 11:09 configuration.xsl
-rw-r--r-- 1 bigdata bigdata 16685 Oct  4 14:50 hadoop-env.sh
-rw-r--r-- 1 bigdata bigdata   884 Oct  4 14:55 core-site.xml
-rw-r--r-- 1 bigdata bigdata   867 Oct  4 14:56 hdfs-site.xml
-rw-r--r-- 1 bigdata bigdata  1063 Oct  4 14:58 mapred-site.xml
-rw-r--r-- 1 bigdata bigdata  1015 Oct  4 15:00 yarn-site.xml
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc/hadoop$
```

```
$nano hadoop-env.sh
```



```
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc/hadoop$ nano hadoop-env.sh
bigdata@bigdata:~/hadoop/hadoop-3.3.1/etc/hadoop$ 
```

Set the Java Path in hadoop-env.sh as shown in the image.



```
  GNU nano 4.8                        hadoop-env.sh
#
#  JAVA_HOME=/usr/java/testing hdfs dfs -ls
#
# Therefore, the vast majority (BUT NOT ALL!) of these defaults
# are configured for substitution and not append.  If append
# is preferable, modify this file accordingly.

###
# Generic settings for HADOOP
###

# Technically, the only required environment variable is JAVA_HOME.
# All others are optional.  However, the defaults are probably not
# preferred.  Many sites configure these options outside of Hadoop,
# such as in /etc/profile.d

# The java implementation to use. By default, this environment
# variable is REQUIRED on ALL platforms except OS X!
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64

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

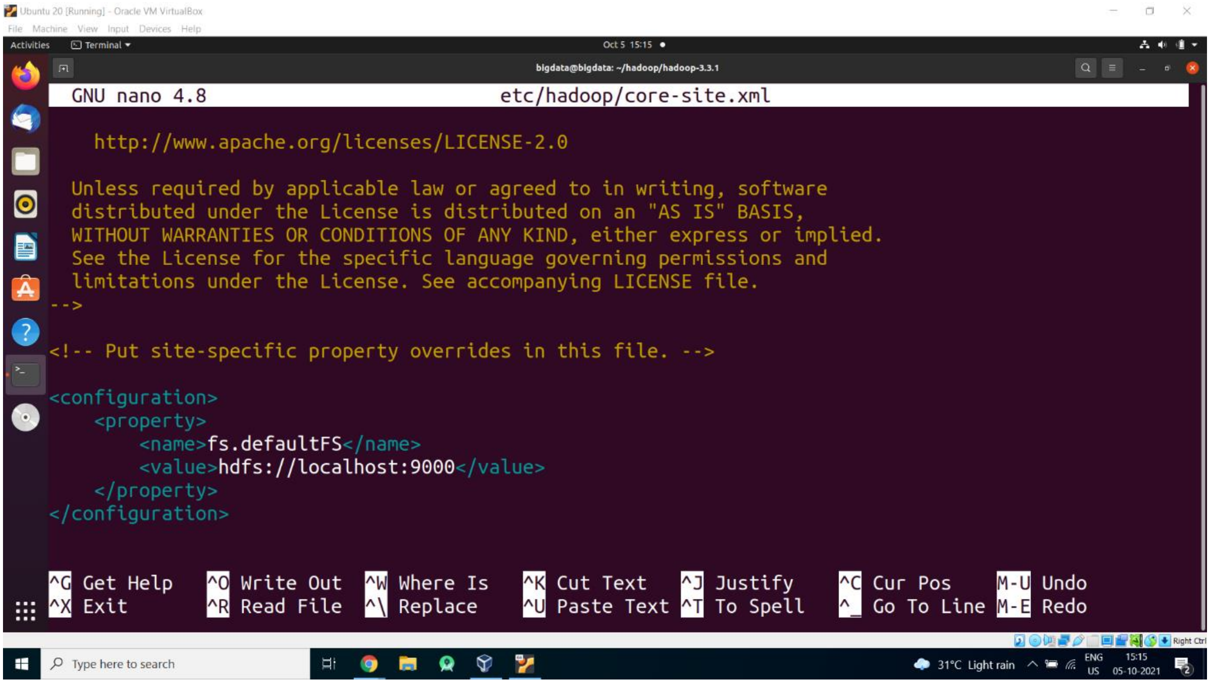
Apache Hadoop 3.3.0 installation on Ubuntu Part 2

Apache Hadoop 3.3.0 installation on Ubuntu Part 2

Use the following property in the respective files

File: nano etc/hadoop/core-site.xml:

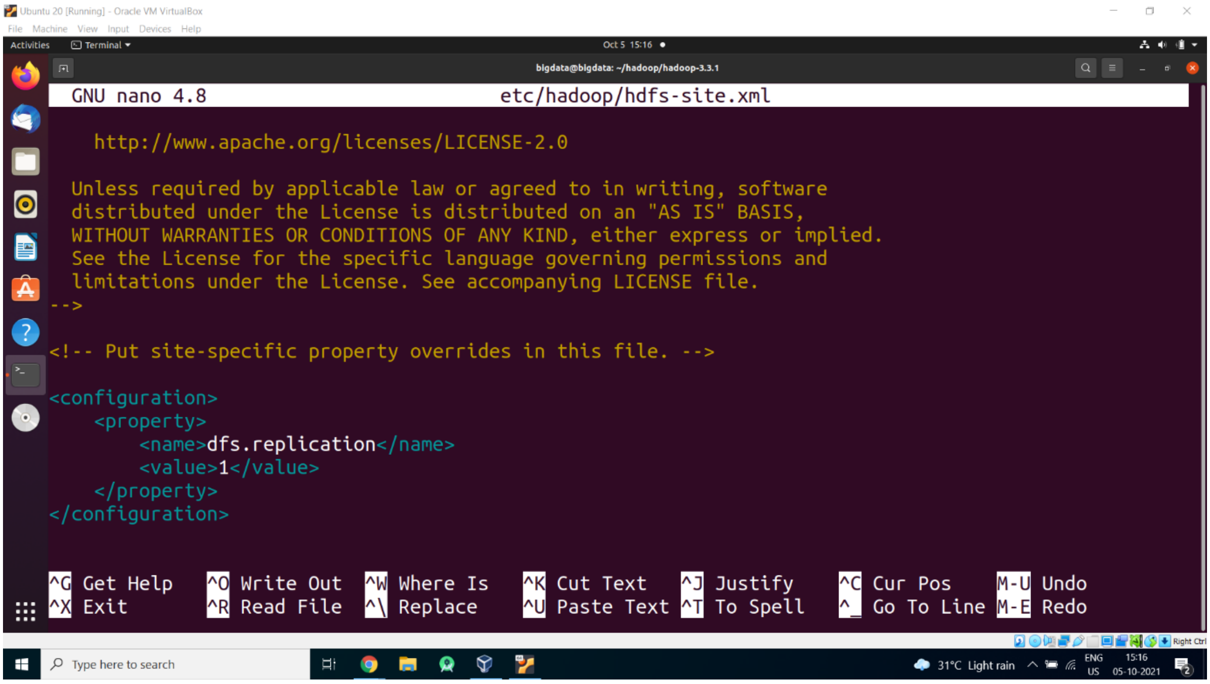
```
<configuration>
<property>
<name>fs.defaultFS</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>
```


Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 5 15:15
bigdata@bigdata: ~/hadoop/hadoop-3.3.1
Type here to search
31°C Light rain
ENG US
15:15 05-10-2021
Right Ctrl

```
  GNU nano 4.8                        etc/hadoop/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.defaultFS</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    M-U Undo
^X Exit       ^R Read File  ^\ Replace    ^U Paste Text ^T To Spell   ^_ Go To Line M-E Redo
```

File: nano etc/hadoop/hdfs-site.xml

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



Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 5 15:16
bigdata@bigdata: ~/hadoop/hadoop-3.3.1
Type here to search
31°C Light rain
ENG US
15:16 05-10-2021
Right Ctrl

```
  GNU nano 4.8                    etc/hadoop/hdfs-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>dfs.replication</name>
        <value>1</value>
    </property>
</configuration>

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

File: nano etc/hadoop/mapred-site.xml

<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

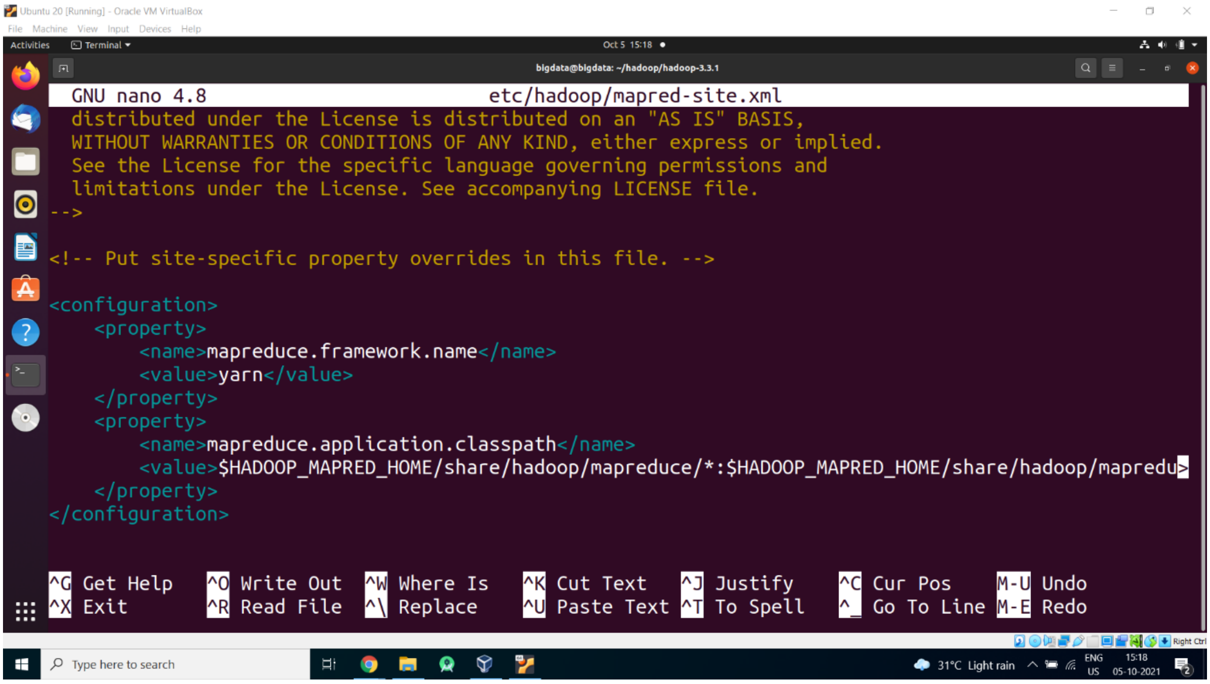
<property>

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

<value>$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*:$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*</value>

</property>

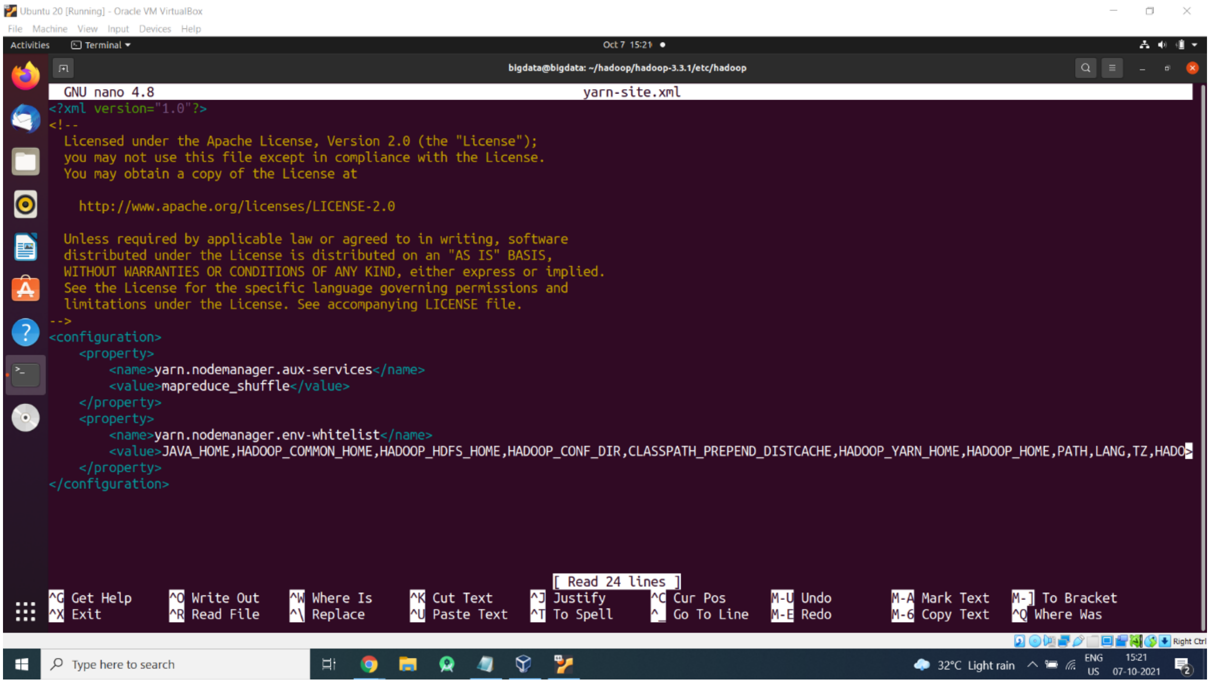
</configuration>


Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 5 15:18
bigdata@bigdata: ~/hadoop/hadoop-3.3.1
Type here to search
31°C Light rain
ENG US
15:18 05-10-2021
Right Ctrl

```
  GNU nano 4.8                    etc/hadoop/mapred-site.xml
  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>
    <property>
        <name>mapreduce.application.classpath</name>
        <value>$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*:$HADOOP_MAPRED_HOME/share/hadoop/mapredu>
    </property>
</configuration>


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

File: nano etc/hadoop/yarn-site.xml

```
<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.env-whitelist</name>
<value>
JAVA_HOME,HADOOP_COMMON_HOME,HADOOP_HDFS_HOME,HADOOP_CONF_DIR,
CLASSPATH_PREPEND_DISTCACHE,
HADOOP_YARN_HOME,HADOOP_HOME,PATH,LANG,TZ,HADOOP_MAPRED_HOME*<
/value>
</property>
</configuration>
```



Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 7 15:21
bigdata@bigdata: ~/hadoop/hadoop-3.3.1/etc/hadoop
GNU nano 4.8
yarn-site.xml
Type here to search
32°C Light rain
ENG US
15:21 07-10-2021

```
<?xml version="1.0"?>
<!--
  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.
-->
<configuration>
    <property>
        <name>yarn.nodemanager.aux-services</name>
        <value>mapreduce_shuffle</value>
    </property>
    <property>
        <name>yarn.nodemanager.env-whitelist</name>
        <value>JAVA_HOME,HADOOP_COMMON_HOME,HADOOP_HDFS_HOME,HADOOP_CONF_DIR,CLASSPATH_PREPEND_DISTCACHE,HADOOP_YARN_HOME,HADOOP_HOME,PATH,LANG,TZ,HADO>
    </property>
</configuration>

                                   [ Read 24 lines ]
```

Now check that you can ssh to the localhost without a passphrase:

```
$ ssh localhost
```

If you cannot ssh to localhost without a passphrase, execute the following commands:

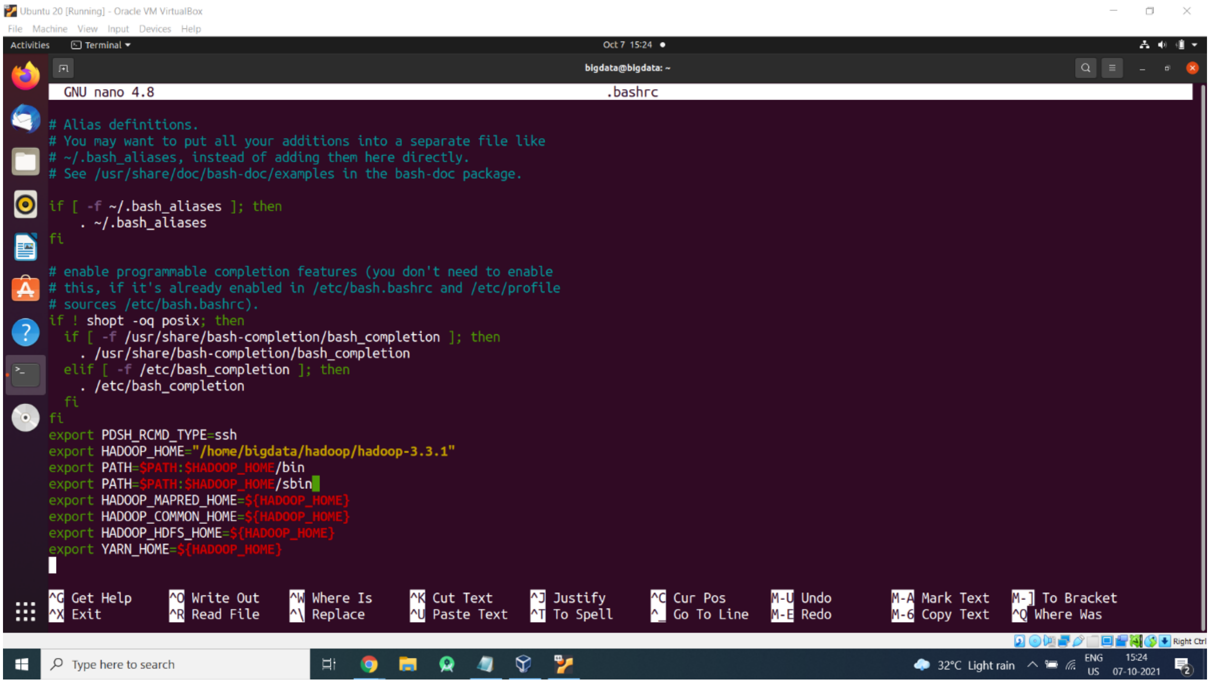
```
1. $ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
2. $ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
3. $ chmod 0600 ~/.ssh/authorized_keys
```

Open the bashrc files in the nano editor using the following command:

```
nano .bashrc
```

edit .bashrc file located in the user’s home directory and add the following parameters:

```
1. export HADOOP_HOME="/home/bigdata/hadoop/hadoop-3.3.1”
2. export PATH=$PATH:$HADOOP_HOME/bin
3. export PATH=$PATH:$HADOOP_HOME/sbin
4. export HADOOP_MAPRED_HOME=${HADOOP_HOME}
5. export HADOOP_COMMON_HOME=${HADOOP_HOME}
6. export HADOOP_HDFS_HOME=${HADOOP_HOME}
7. export YARN_HOME=${HADOOP_HOME}
```



```
  GNU nano 4.8                                 .bashrc

# Alias definitions.
# You may want to put all your additions into a separate file like
# ~/.bash_aliases, instead of adding them here directly.
# See /usr/share/doc/bash-doc/examples in the bash-doc package.

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
export PDSH_RCMD_TYPE=ssh
export HADOOP_HOME="/home/bigdata/hadoop/hadoop-3.3.1"
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}


^G Get Help   ^O Write Out  ^W Where Is   ^K Cut Text   ^J Justify    ^C Cur Pos    M-U Undo      M-A Mark Text  M-] To Bracket
^X Exit       ^R Read File  ^\ Replace    ^U Paste Text ^T To Spell   ^_ Go To Line M-E Redo      M-6 Copy Text  ^Q Where Was
```

To save the changes you've made, press Ctrl+O. To exit the nano editor, press Ctrl+X and then press 'Y' to exit the editor.

Now, source the bashrc file so that the changes will come into effect:

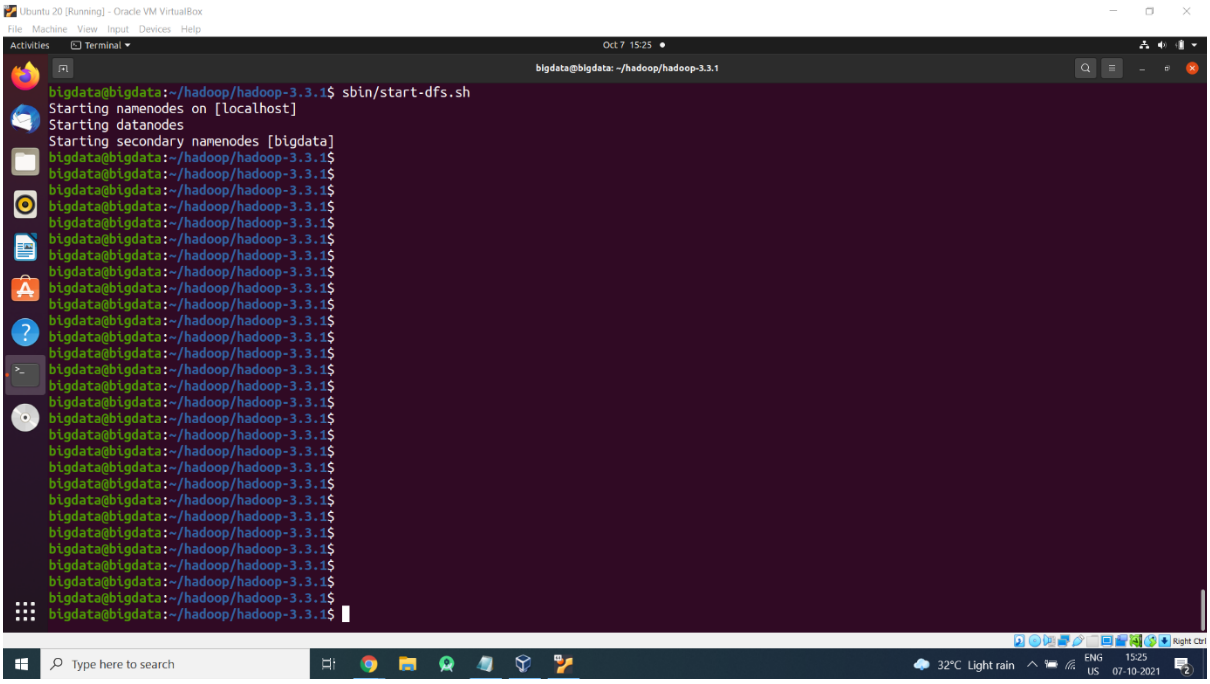
```
source ~/.bashrc
```

Format the filesystem:

```
$ bin/hdfs namenode -format
```

Start NameNode daemon and DataNode daemon:

```
$ sbin/start-dfs.sh
```

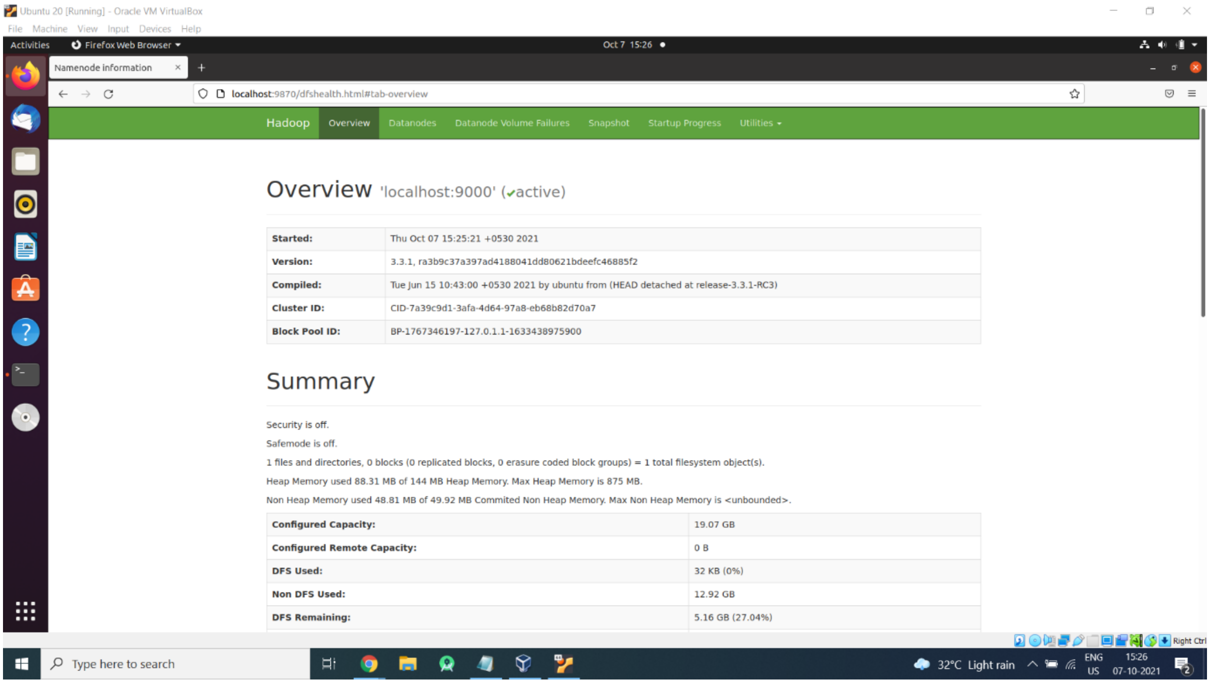


```
bigdata@bigdata:~/hadoop/hadoop-3.3.1$ sbin/start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bigdata]
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
bigdata@bigdata:~/hadoop/hadoop-3.3.1$
```

The hadoop daemon log output is written to the $HADOOP_LOG_DIR directory (defaults to $HADOOP_HOME/logs).

Browse the web interface for the NameNode; by default it is available at:

NameNode – http://localhost:9870/



Overview 'localhost:9000' (✓active)

Started:	Thu Oct 07 15:25:21 +0530 2021
Version:	3.3.1, ra3b9c37a397ad4188041dd80621bdeefc46885f2
Compiled:	Tue Jun 15 10:43:00 +0530 2021 by ubuntu from (HEAD detached at release-3.3.1-RC3)
Cluster ID:	CID-7a39c9d1-3afa-4d64-97a8-eb68b82d70a7
Block Pool ID:	BP-1767346197-127.0.1.1-1633438975900

Summary

Security is off.

Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).

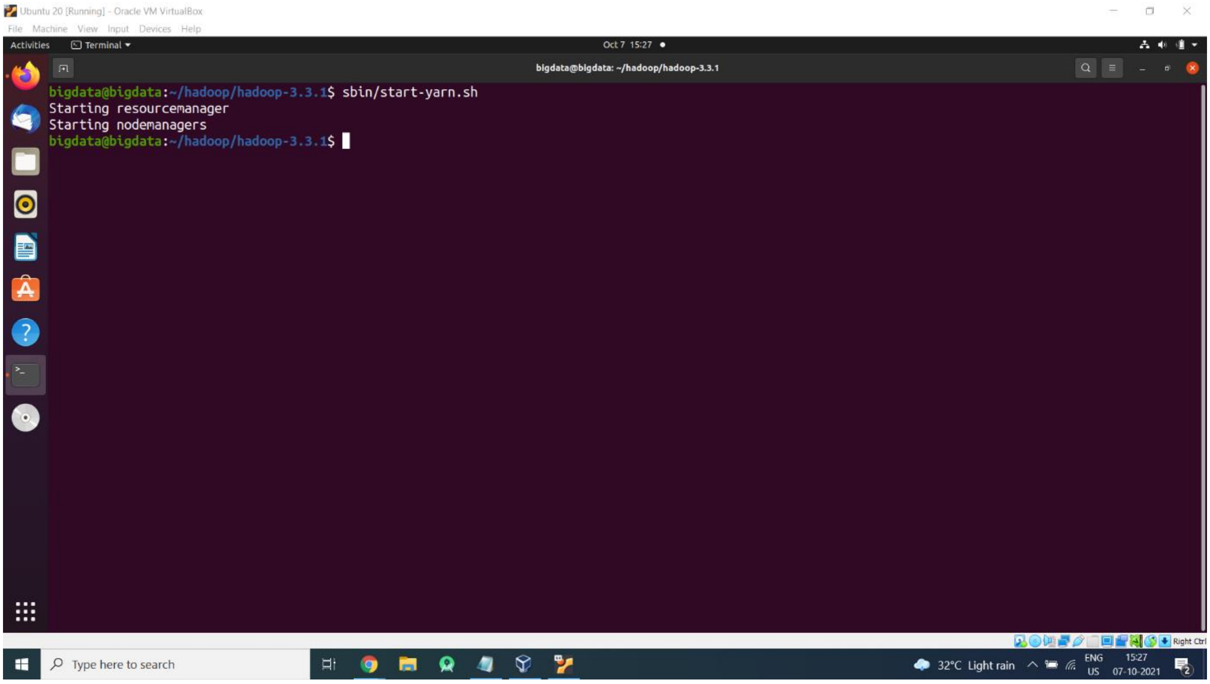
Heap Memory used 88.31 MB of 144 MB Heap Memory. Max Heap Memory is 875 MB.

Non Heap Memory used 48.81 MB of 49.92 MB Commited Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	19.07 GB
Configured Remote Capacity:	0 B
DFS Used:	32 KB (0%)
Non DFS Used:	12.92 GB
DFS Remaining:	5.16 GB (27.04%)

Start ResourceManager daemon and NodeManager daemon:

```
$ sbin/start-yarn.sh
```

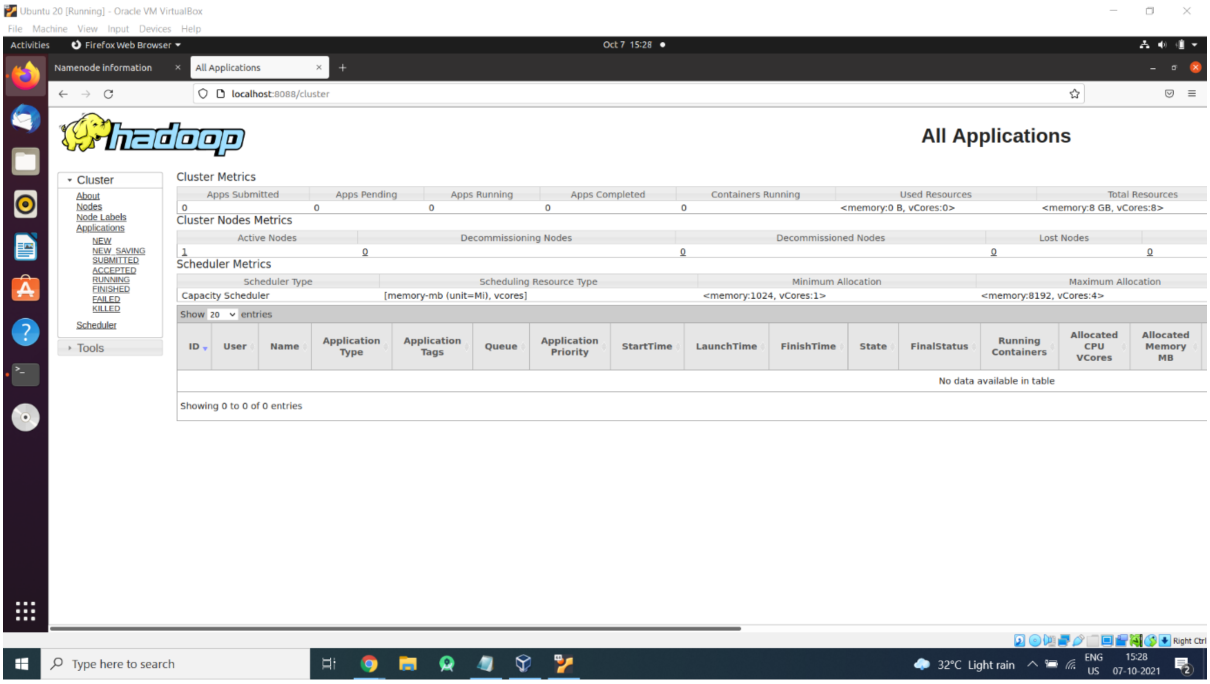


Ubuntu 20 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
Oct 7 15:27
bigdata@bigdata: ~/hadoop/hadoop-3.3.1
Type here to search
32°C Light rain
ENG US
15:27
07-10-2021

```
bigdata@bigdata:~/hadoop/hadoop-3.3.1$ sbin/start-yarn.sh
Starting resourcemanager
Starting nodemanagers
bigdata@bigdata:~/hadoop/hadoop-3.3.1$ 
```

Browse the web interface for the ResourceManager; by default it is available at:

ResourceManager – http://localhost:8088/



All Applications

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Used Resources	Total Resources
0	0	0	0	0	<memory:0 B, vCores:0>	<memory:8 GB, vCores:8>

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	
1	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>

Show 20 entries

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCores	Allocated Memory MB
No data available in table														

Showing 0 to 0 of 0 entries