

# Experiment No.1

## Hadoop Installation In Ubuntu-16.04

Mohit Badve

BE COMPS

2017130006

Date of Experiment - 06/08/2020

### Hadoop Installation In Ubuntu-16.04:

#### Step 1:

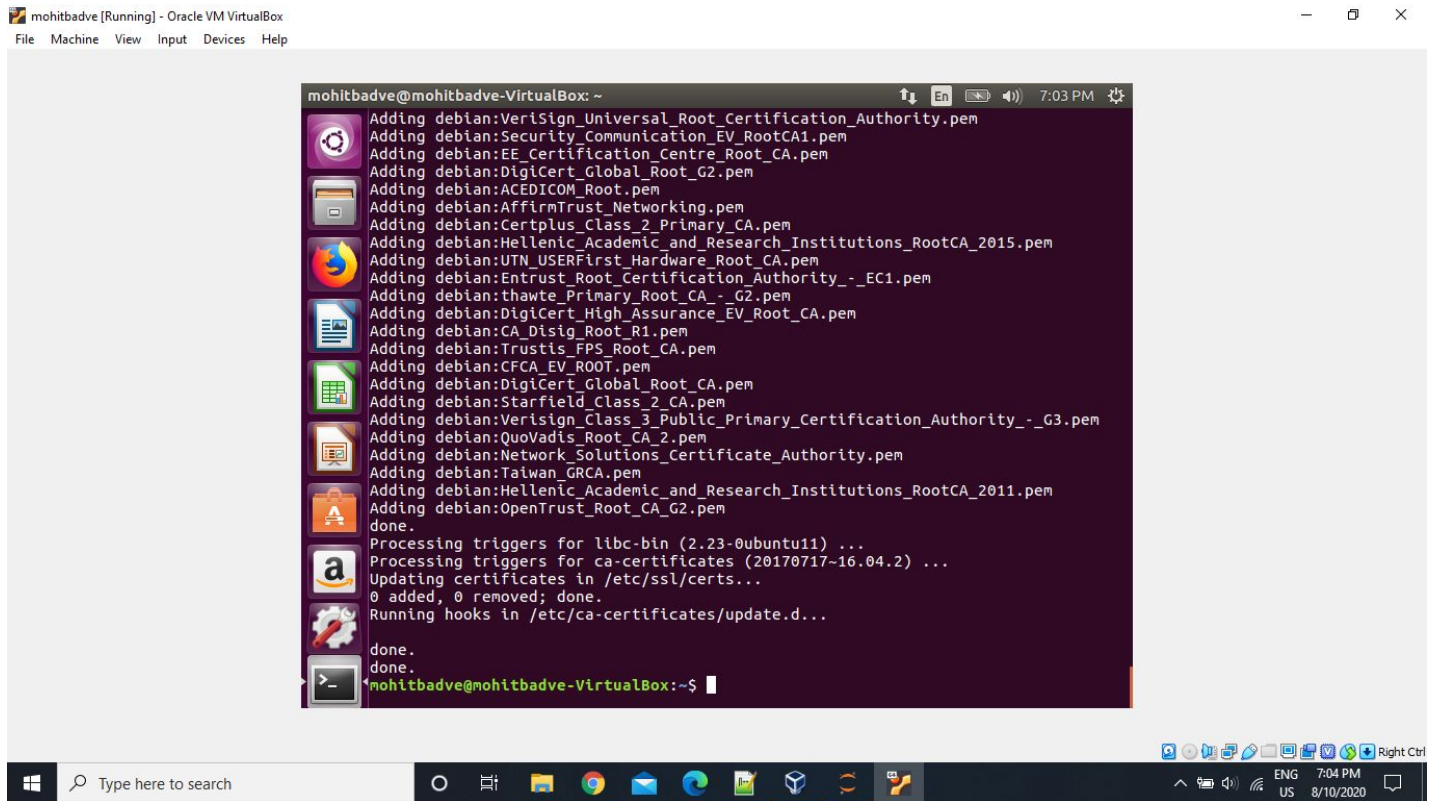
**sudo apt-get update**

mohitbadve [Running] - Oracle VM VirtualBox  
File Machine View Input Devices Help

```
mohitbadve@mohitbadve-VirtualBox: ~  
To run a command as administrator (user "root"), use "sudo <command>".  
See "man sudo_root" for details.  
  
mohitbadve@mohitbadve-VirtualBox:~$ sudo apt-get update  
[sudo] password for mohitbadve:  
Hit:1 http://in.archive.ubuntu.com/ubuntu xenial InRelease  
Hit:2 http://in.archive.ubuntu.com/ubuntu xenial-updates InRelease  
Hit:3 http://in.archive.ubuntu.com/ubuntu xenial-backports InRelease  
Get:4 http://security.ubuntu.com/ubuntu xenial-security InRelease [109 kB]  
Get:5 http://security.ubuntu.com/ubuntu xenial-security/main amd64 DEP-11 Metada  
ta [78.6 kB]  
Get:6 http://security.ubuntu.com/ubuntu xenial-security/universe amd64 DEP-11 Me  
tadata [124 kB]  
Get:7 http://security.ubuntu.com/ubuntu xenial-security/multiverse amd64 DEP-11  
Metadata [2,468 B]  
Fetched 314 kB in 2s (127 kB/s)  
  
Reading package lists... Done  
mohitbadve@mohitbadve-VirtualBox:~$  
mohitbadve@mohitbadve-VirtualBox:~$
```

## Step 2:

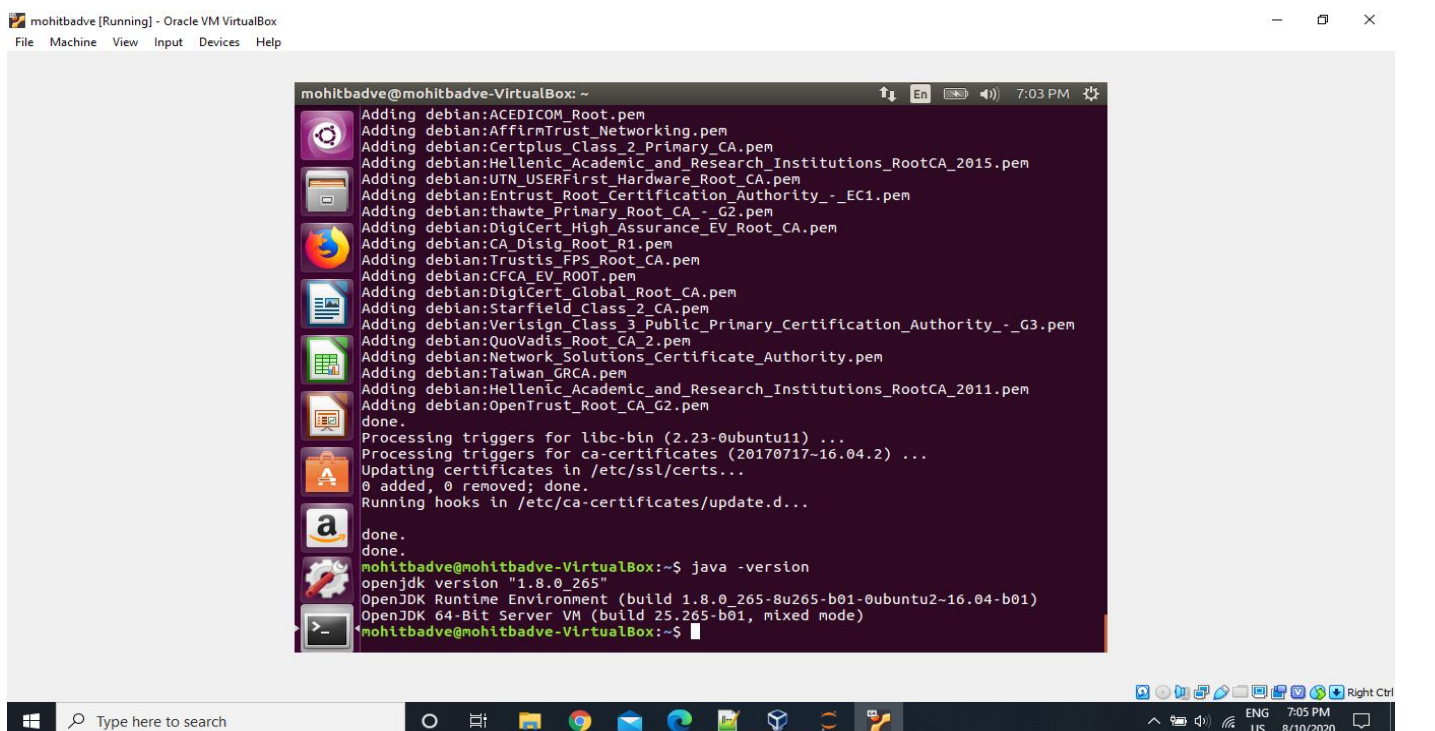
### sudo apt-get install default-jdk



```
mohitbadve@mohitbadve-VirtualBox: ~  
Adding debian:VeriSign_Universal_Root_Certification_Authority.pem  
Adding debian:Security_Communication_EV_RootCA1.pem  
Adding debian:EE_Certification_Centre_Root_CA.pem  
Adding debian:DigiCert_Global_Root_G2.pem  
Adding debian:ACEDICOM_Root.pem  
Adding debian:AffirmTrust_Networking.pem  
Adding debian:Certplus_Class_2_Primary_CA.pem  
Adding debian:Hellenic_Academic_and_Research_Institutions_RootCA_2015.pem  
Adding debian:UTN_USERFirst_Hardware_Root_CA.pem  
Adding debian:Entrust_Root_Certification_Authority_-_EC1.pem  
Adding debian:thawte_Primary_Root_CA_-_G2.pem  
Adding debian:DigiCert_High_Assurance_EV_Root_CA.pem  
Adding debian:CA_Disig_Root_R1.pem  
Adding debian:Trustis_FPS_Root_CA.pem  
Adding debian:CFCA_EV_ROOT.pem  
Adding debian:DigiCert_Global_Root_CA.pem  
Adding debian:Starfield_Class_2_CA.pem  
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority_-_G3.pem  
Adding debian:QuoVadis_Root_CA_2.pem  
Adding debian:Network_Solutions_Certificate_Authority.pem  
Adding debian:Taiwan_GRCA.pem  
Adding debian:Hellenic_Academic_and_Research_Institutions_RootCA_2011.pem  
Adding debian:OpenTrust_Root_CA_G2.pem  
done.  
Processing triggers for libc-bin (2.23-0ubuntu11) ...  
Processing triggers for ca-certificates (20170717-16.04.2) ...  
Updating certificates in /etc/ssl/certs...  
0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d...  
done.  
done.  
mohitbadve@mohitbadve-VirtualBox:~$
```

## Step 3:

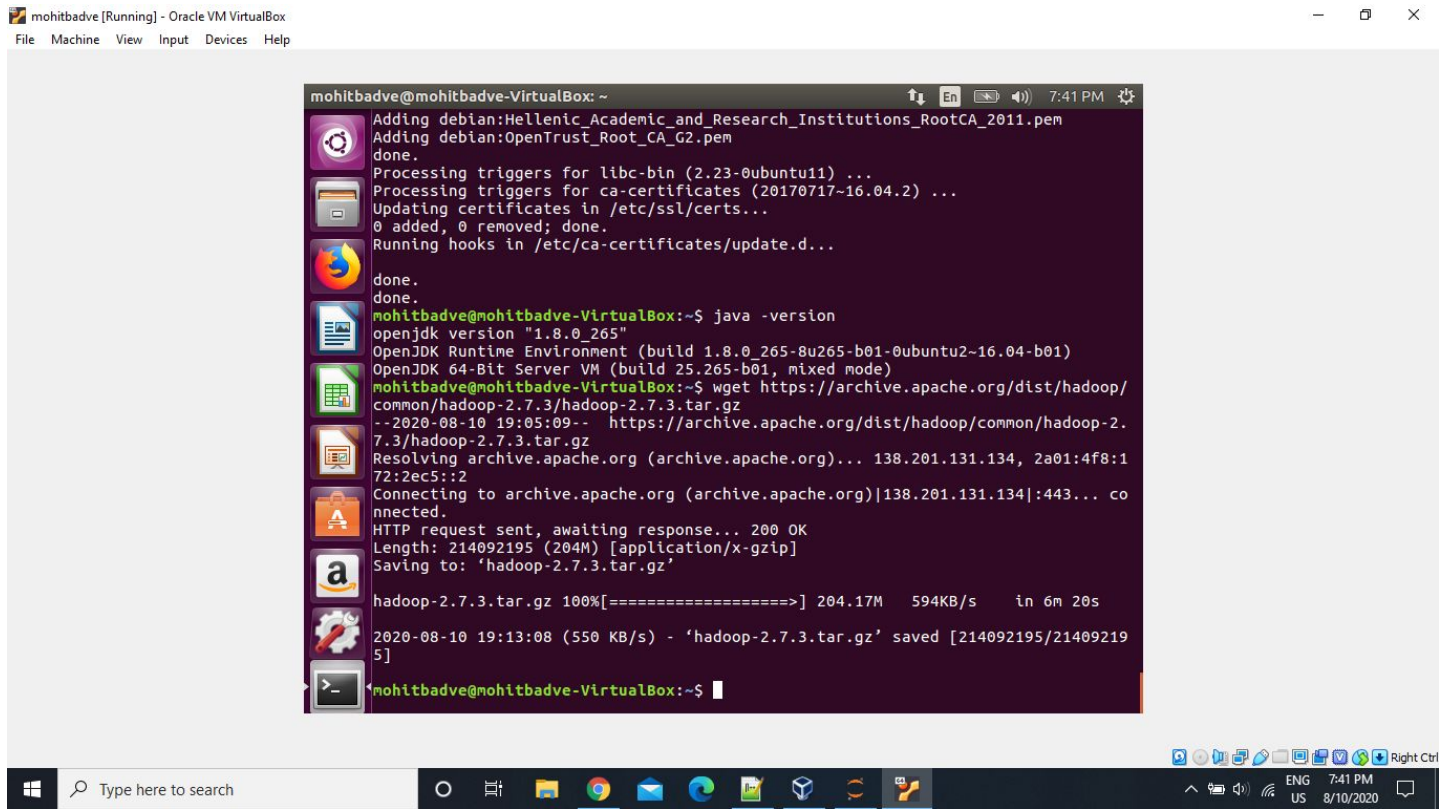
### java -version



```
mohitbadve@mohitbadve-VirtualBox: ~  
Adding debian:ACEDICOM_Root.pem  
Adding debian:AffirmTrust_Networking.pem  
Adding debian:Certplus_Class_2_Primary_CA.pem  
Adding debian:Hellenic_Academic_and_Research_Institutions_RootCA_2015.pem  
Adding debian:UTN_USERFirst_Hardware_Root_CA.pem  
Adding debian:Entrust_Root_Certification_Authority_-_EC1.pem  
Adding debian:thawte_Primary_Root_CA_-_G2.pem  
Adding debian:DigiCert_High_Assurance_EV_Root_CA.pem  
Adding debian:CA_Disig_Root_R1.pem  
Adding debian:Trustis_FPS_Root_CA.pem  
Adding debian:CFCA_EV_ROOT.pem  
Adding debian:DigiCert_Global_Root_CA.pem  
Adding debian:Starfield_Class_2_CA.pem  
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority_-_G3.pem  
Adding debian:QuoVadis_Root_CA_2.pem  
Adding debian:Network_Solutions_Certificate_Authority.pem  
Adding debian:Taiwan_GRCA.pem  
Adding debian:Hellenic_Academic_and_Research_Institutions_RootCA_2011.pem  
Adding debian:OpenTrust_Root_CA_G2.pem  
done.  
Processing triggers for libc-bin (2.23-0ubuntu11) ...  
Processing triggers for ca-certificates (20170717-16.04.2) ...  
Updating certificates in /etc/ssl/certs...  
0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d...  
done.  
done.  
mohitbadve@mohitbadve-VirtualBox:~$ java -version  
openjdk version "1.8.0_265"  
OpenJDK Runtime Environment (build 1.8.0_265-8u265-b01-0ubuntu2-16.04-b01)  
OpenJDK 64-Bit Server VM (build 25.265-b01, mixed mode)  
mohitbadve@mohitbadve-VirtualBox:~$
```

## Step: 4

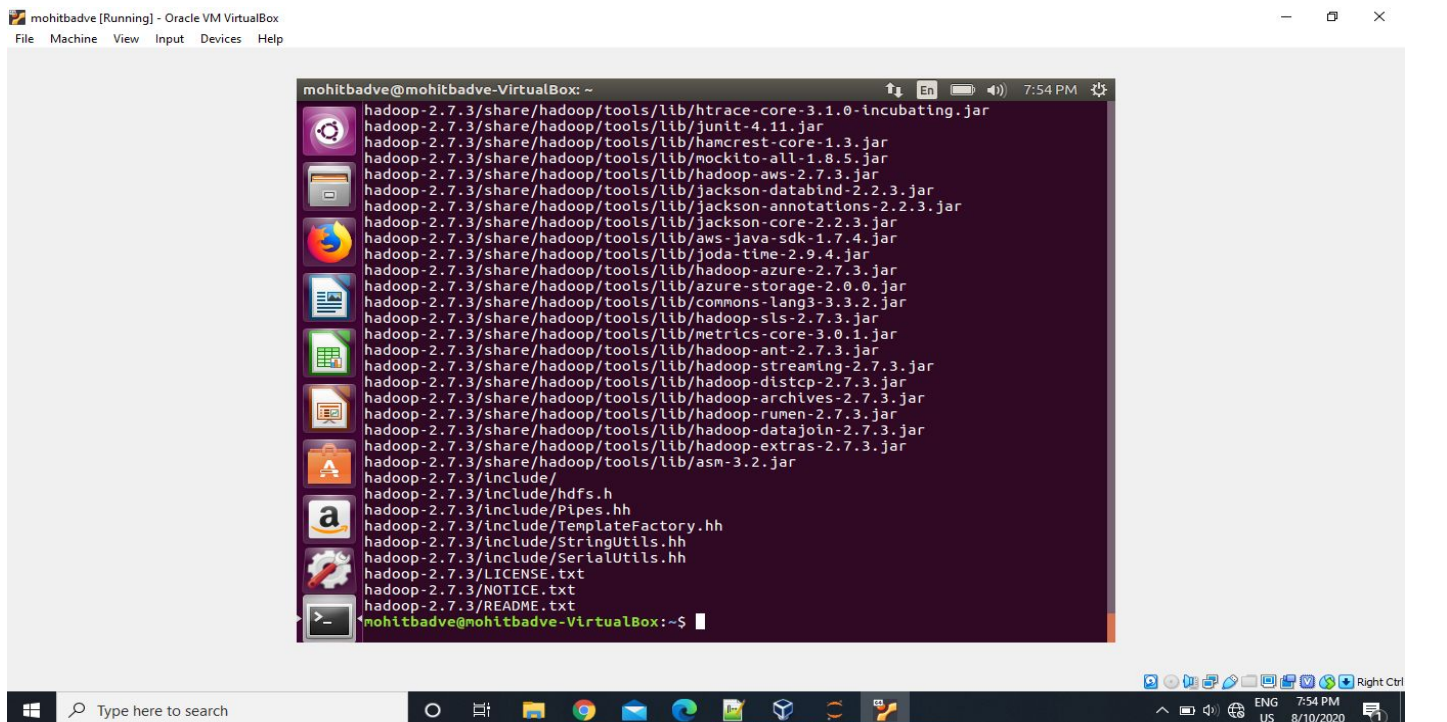
**wget https://archive.apache.org/dist/hadoop/common/hadoop-2.7.3/hadoop-2.7.3.tar.gz**



```
mohitbadve@Oracle VM VirtualBox: ~  
File Machine View Input Devices Help  
mohitbadve@mohitbadve-VirtualBox: ~  
Adding debian:Hellenic_Academic_and_Research_Institutions_RootCA_2011.pem  
Adding debian:OpenTrust_Root_CA_G2.pem  
done.  
Processing triggers for libc-bin (2.23-0ubuntu11) ...  
Processing triggers for ca-certificates (20170717~16.04.2) ...  
Updating certificates in /etc/ssl/certs...  
0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d...  
done.  
done.  
mohitbadve@mohitbadve-VirtualBox:~$ java -version  
openjdk version "1.8.0_265"  
OpenJDK Runtime Environment (build 1.8.0_265-8u265-b01-0ubuntu2~16.04-b01)  
OpenJDK 64-Bit Server VM (build 25.265-b01, mixed mode)  
mohitbadve@mohitbadve-VirtualBox:~$ wget https://archive.apache.org/dist/hadoop/  
common/hadoop-2.7.3/hadoop-2.7.3.tar.gz  
--2020-08-10 19:05:09-- https://archive.apache.org/dist/hadoop/common/hadoop-2.  
7.3/hadoop-2.7.3.tar.gz  
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:1  
72:2ec5::2  
Connecting to archive.apache.org (archive.apache.org)|138.201.131.134|:443... co  
nnected.  
HTTP request sent, awaiting response... 200 OK  
Length: 214092195 (204M) [application/x-gzip]  
Saving to: 'hadoop-2.7.3.tar.gz'  
  
hadoop-2.7.3.tar.gz 100%[=====] 204.17M 594KB/s in 6m 20s  
  
2020-08-10 19:13:08 (550 KB/s) - 'hadoop-2.7.3.tar.gz' saved [214092195/21409219  
5]  
mohitbadve@mohitbadve-VirtualBox:~$
```

## Step: 5

**tar -xzf hadoop-2.7.3.tar.gz**



```
mohitbadve@Oracle VM VirtualBox: ~  
File Machine View Input Devices Help  
mohitbadve@mohitbadve-VirtualBox: ~  
hadoop-2.7.3/share/hadoop/tools/lib/htrace-core-3.1.0-incubating.jar  
hadoop-2.7.3/share/hadoop/tools/lib/junit-4.11.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hamcrest-core-1.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/mockito-all-1.8.5.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-aws-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-databind-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-annotations-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-core-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/aws-java-sdk-1.7.4.jar  
hadoop-2.7.3/share/hadoop/tools/lib/joda-time-2.9.4.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-azure-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/azure-storage-2.0.0.jar  
hadoop-2.7.3/share/hadoop/tools/lib/commons-lang3-3.3.2.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-sls-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/metrics-core-3.0.1.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-ant-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-streaming-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-distcp-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-archives-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-ruven-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-datajoin-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-extras-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/asm-3.2.jar  
hadoop-2.7.3/include/  
hadoop-2.7.3/include/hdfs.h  
hadoop-2.7.3/include/Pipes.hh  
hadoop-2.7.3/include/TemplateFactory.hh  
hadoop-2.7.3/include/StringUtils.hh  
hadoop-2.7.3/include/SerialUtils.hh  
hadoop-2.7.3/LICENSE.txt  
hadoop-2.7.3/NOTICE.txt  
hadoop-2.7.3/README.txt  
mohitbadve@mohitbadve-VirtualBox:~$
```



Step: 6

`sudo mv hadoop-2.7.3 /usr/local/hadoop`

Step: 7

`readlink -f /usr/bin/java | sed "s:bin/java::"`

mohitbadve [Running] - Oracle VM VirtualBox  
File Machine View Input Devices Help

```
mohitbadve@mohitbadve-VirtualBox: ~  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-databind-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-annotations-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/jackson-core-2.2.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/aws-java-sdk-1.7.4.jar  
hadoop-2.7.3/share/hadoop/tools/lib/joda-time-2.9.4.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-azure-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/azure-storage-2.0.0.jar  
hadoop-2.7.3/share/hadoop/tools/lib/commons-lang3-3.3.2.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-sls-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/metrics-core-3.0.1.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-ant-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-streaming-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-distcp-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-archives-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-rumen-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-datajoin-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/hadoop-extras-2.7.3.jar  
hadoop-2.7.3/share/hadoop/tools/lib/asm-3.2.jar  
hadoop-2.7.3/include/  
hadoop-2.7.3/include/hdfs.h  
hadoop-2.7.3/include/Pipes.hh  
hadoop-2.7.3/include/TemplateFactory.hh  
hadoop-2.7.3/include/StringUtils.hh  
hadoop-2.7.3/include/SerialUtils.hh  
hadoop-2.7.3/LICENSE.txt  
hadoop-2.7.3/NOTICE.txt  
hadoop-2.7.3/README.txt  
mohitbadve@mohitbadve-VirtualBox:~$ sudo mv hadoop-2.7.3 /usr/local/hadoop  
[sudo] password for mohitbadve:  
mohitbadve@mohitbadve-VirtualBox:~$ readlink -f /usr/bin/java | sed "s:bin/java::"  
:/usr/lib/jvm/java-8-openjdk-amd64/jre/  
mohitbadve@mohitbadve-VirtualBox:~$
```

ENG 7:58 PM  
US 8/10/2020

## Step: 8

`sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh`

Write the following command

`export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre/`

mohitbadve [Running] - Oracle VM VirtualBox  
File Machine View Input Devices Help

```
mohitbadve@mohitbadve-VirtualBox: ~
GNU nano 2.5.3 File: /usr/local/hadoop/etc/hadoop/hadoop-env.sh Modified

# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements. See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership. The ASF licenses this file
# to you 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.
#
# 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-8-openjdk-amd64/jre/
# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
#
^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit ^R Read File ^\ Replace ^U Uncut Text ^T To Linter ^_ Go To Line
```

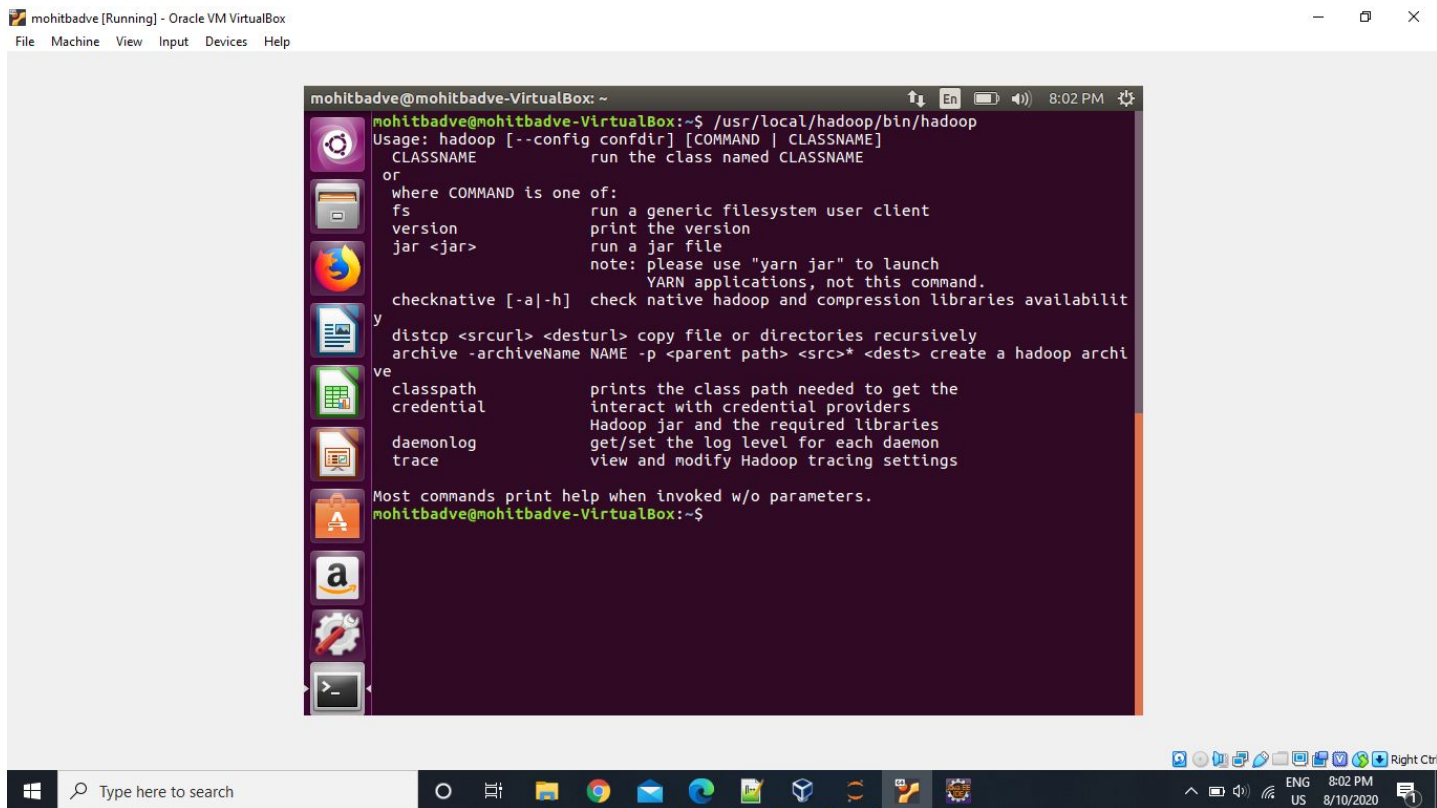
Right Ctrl

Type here to search

ENG 8:00 PM  
US 8/10/2020

## Step: 9

/usr/local/hadoop/bin/hadoop



The screenshot shows a terminal window titled "mohitbadve [Running] - Oracle VM VirtualBox". The terminal prompt is "mohitbadve@mohitbadve-VirtualBox: ~". The user has entered the command "/usr/local/hadoop/bin/hadoop". The terminal output displays the usage of the Hadoop command, including options like "--config", "COMMAND", and "CLASSNAME". It also lists several sub-commands: fs, version, jar, checknative, distcp, archive, classpath, credential, daemonlog, and trace. The terminal window is running on a Windows host, as indicated by the taskbar at the bottom.

```
mohitbadve@mohitbadve-VirtualBox: ~  
mohitbadve@mohitbadve-VirtualBox:~$ /usr/local/hadoop/bin/hadoop  
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]  
CLASSNAME          run the class named CLASSNAME  
or  
where COMMAND is one of:  
fs                  run a generic filesystem user client  
version             print the version  
jar <jar>           run a jar file  
                    note: please use "yarn jar" to launch  
                    YARN applications, not this command.  
checknative [-a|-h] check native hadoop and compression libraries availabilit  
y  
distcp <srcurl> <desturl> copy file or directories recursively  
archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archi  
ve  
classpath           prints the class path needed to get the  
credential           interact with credential providers  
                    Hadoop jar and the required libraries  
daemonlog           get/set the log level for each daemon  
trace              view and modify Hadoop tracing settings  
  
Most commands print help when invoked w/o parameters.  
mohitbadve@mohitbadve-VirtualBox:~$
```

## Conclusion:

For installing Hadoop in ubuntu, the above steps need to be followed and the main step in the installation process is setting up the JAVA\_HOME path in hadoop-env.sh file.