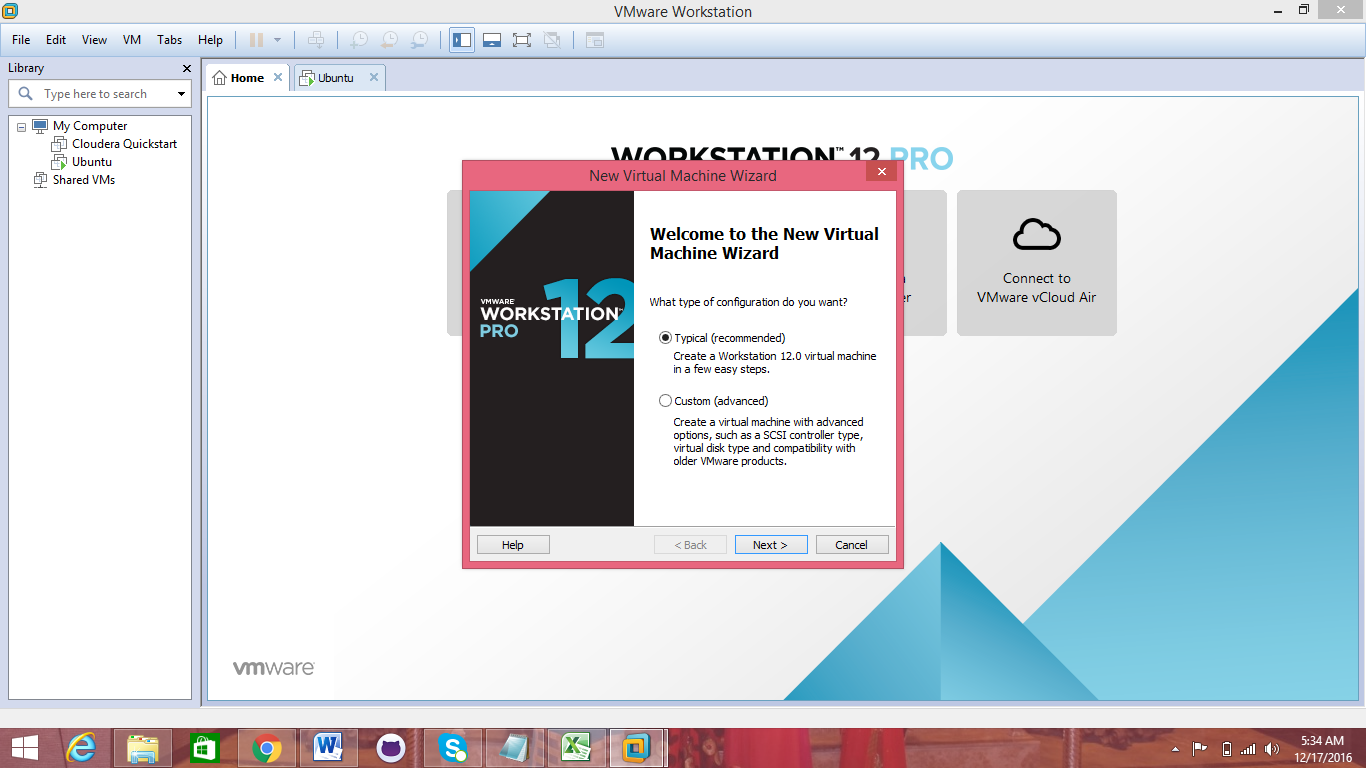
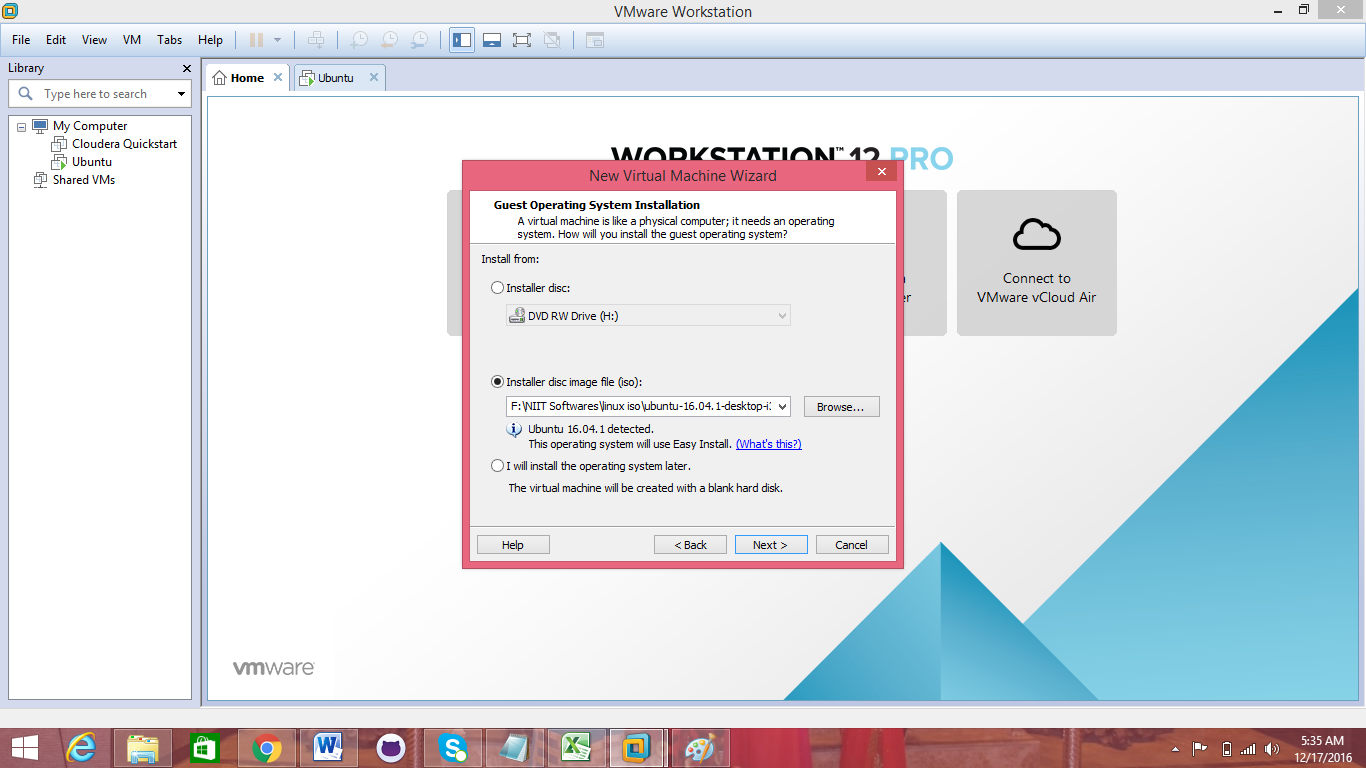
Installation Steps

1. Installation of Ubuntu:

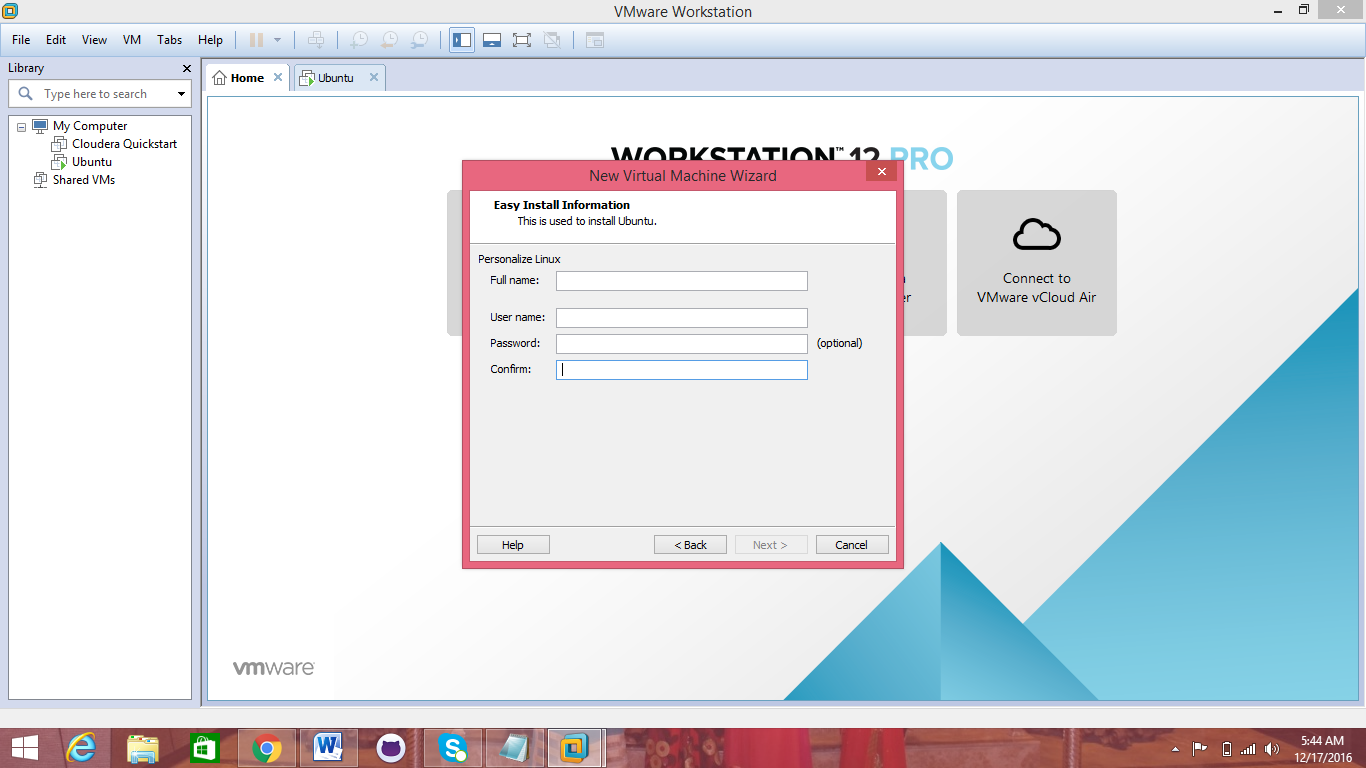
* Firstly you have to download the VMware Workstation 12 in your system.
* Secondly you have to download the Ubuntu operating system from the [www.ubuntu.com/download/desktop](http://www.ubuntu.com/download/desktop) to your system. It will be an iso file consists of size minimum 1GB. After downloading it, move that file into a directory in one of your harddisks other than C drive. This helps to decrease the burden on the C drive.
* Now open VMware workstation and create a new virtual machine. The following window will be shown:-



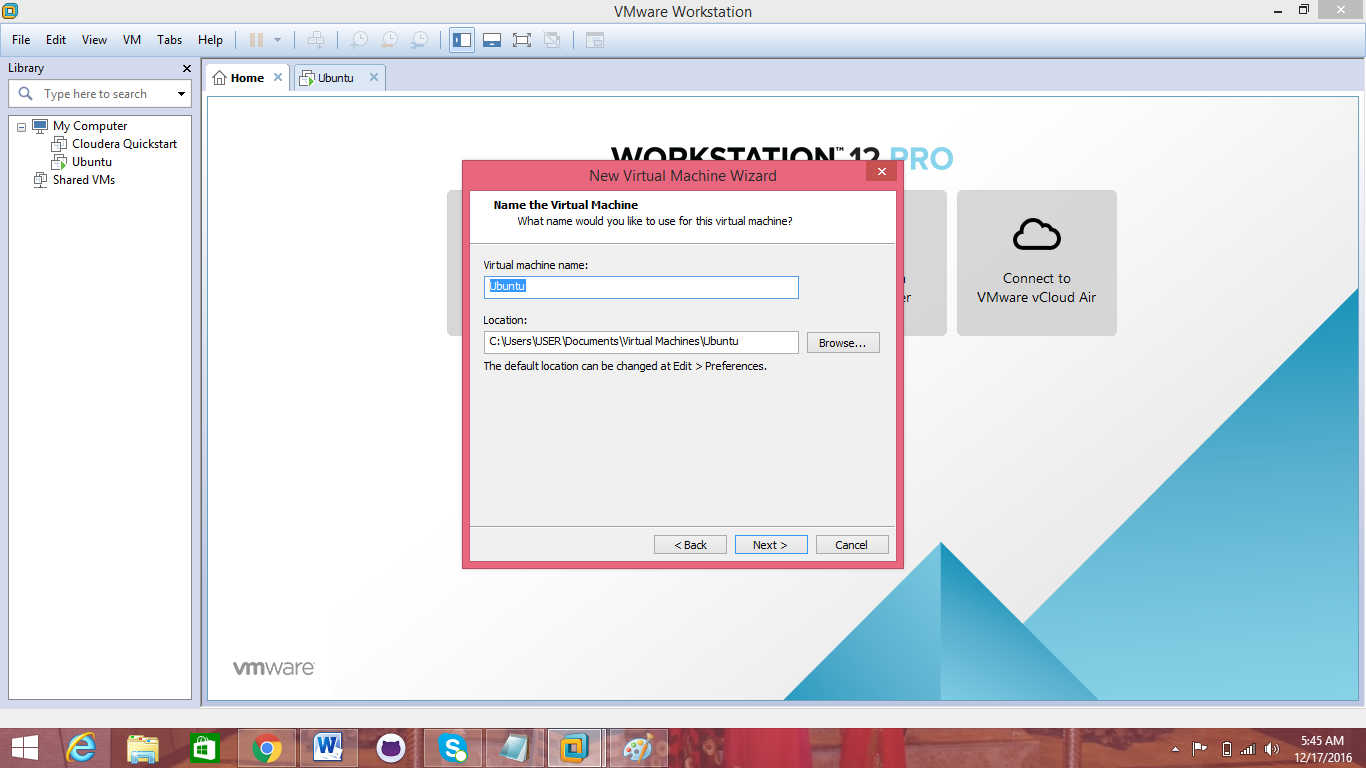
Click next and proceed further. Then you will get another window where you have to browse the folder in which downloaded Ubuntu is present. As your downloaded file is in iso format, select the second radio button and give the path of the directory over there and click next.



After this you will get another window asking your information and username and password for your linux machine. Remember that whatever be the information you will submit here will be finalized and used every time while you log in to your Ubuntu machine. After writing the details, click next and proceed further.

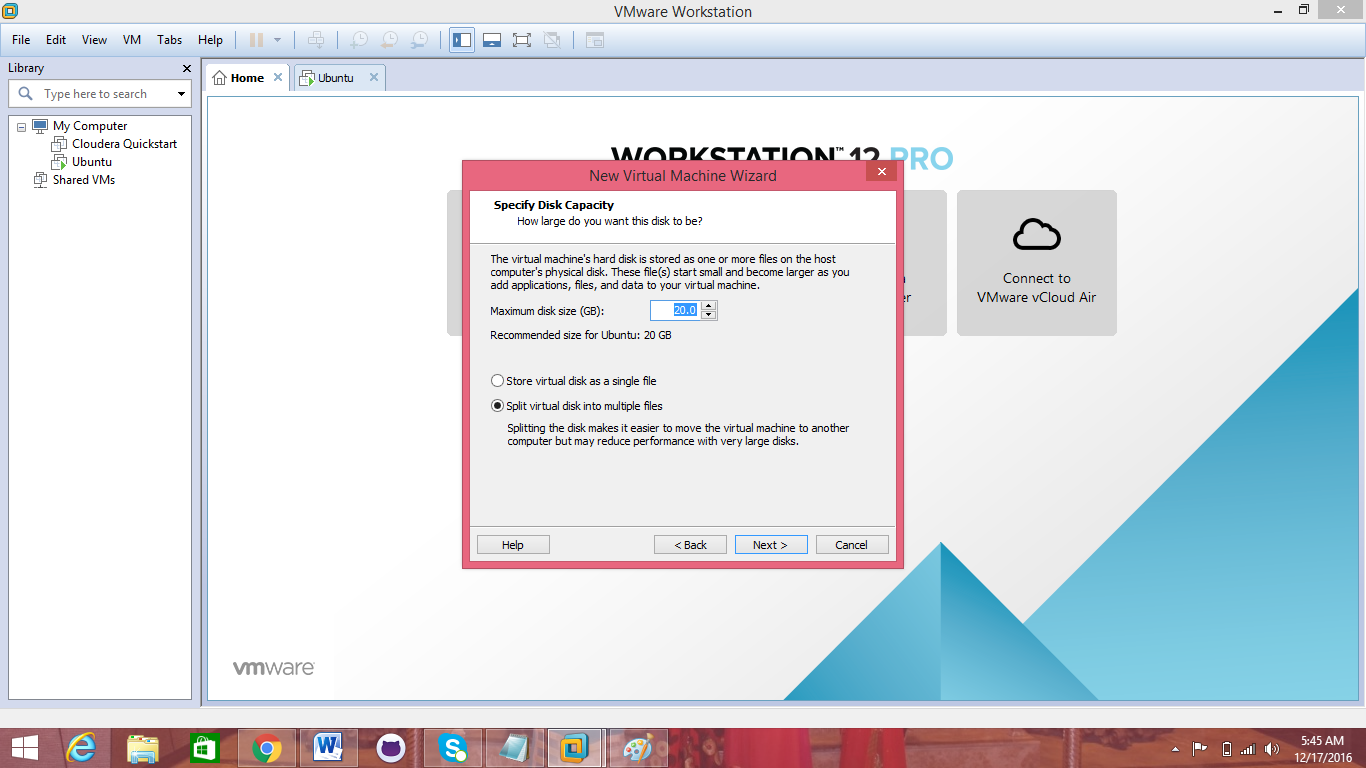


Now you have to browse the directory where you want your virtual machine and its data will be stored. Remember one point that select the place where you have more free space and it can be other than C drive.



Click next.

Last step where you have to specify the capacity which your virtual machine will hold. 20GB is minimum recommended for Ubuntu.



After giving the memory click next and the installation process will be started which will take some time. After importing all the files into Ubuntu machine, your linux operating system will be ready and you can access it by giving your username and password which you created during installation process.

lsb\_release –a =command to know the operating system on which we are currently working on

2. Installation of MySQL in Ubuntu:

* Open your Ubuntu terminal and type:

sudo apt-get update and press enter

sudo apt-get install mysql-server and press enter

Installation process will get started soon after the above command. It will take a while and a window will be displayed asking for change of password for your installing mysql, ignore it and press enter. If it doesn’t accept then keep one password and press enter. After sometime your process will be completed and by using following command you can get your mysql commands:

Mysql –u root –p

Enter password: enter the password and press enter and you will be in mysql mode:

mysql>

To exit from mysql, you can give exit and press enter.

3. Installation of Java:

sudo apt-get update

sudo apt-get install default-jdk

You can check the version of java installed in your machine.

java –version

4. Adding a dedicated Hadoop user:

sudo addgroup hadoop

sudo adduser –ingroup hadoop hduser

5. Installing SSH:

sudo apt-get install ssh

which ssh

which sshd

6. Creating and Setup of SSH Certificates:

su hduser

ssh–keygen –t rsa –P ““

cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys

ssh localhost

7. Installation of Hadoop:

tar xvzf hadoop-2.6.0.tar.gz

sudo mv hadoop /usr/local/

su k

sudo adduser hduser sudo

sudo mv hadoop /usr/local/

sudo chown –R hduser:hadoop /usr/local/hadoop

8. Setup Configuration Files:

The following files will have to be modified to complete the Hadoop setup:

1. ~/.bashrc
2. /usr/local/hadoop/etc/hadoop/hadoop-env.sh
3. /usr/local/hadoop/etc/hadoop/core-site.xml
4. /usr/local/hadoop/etc/hadoop/mapred-site.xml
5. /usr/local/hadoop/etc/hadoop/hdfs-site.xml

1. ~/.bashrc

update-alternatives –config java

sudo nano ~/.bashrc

**#HADOOP VARIABLES START**

**export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-i386**

**export HADOOP\_INSTALL=/usr/local/hadoop**

**export PATH=$PATH:$HADOOP\_INSTALL/bin**

**export PATH=$PATH:$HADOOP\_INSTALL/sbin**

**export HADOOP\_MAPRED\_HOME=$HADOOP\_INSTALL**

**export HADOOP\_COMMON\_HOME=$HADOOP\_INSTALL**

**export HADOOP\_HDFS\_HOME=$HADOOP\_INSTALL**

**export YARN\_HOME=$HADOOP\_INSTALL**

**export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_INSTALL/lib/native**

**export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_INSTALL/lib"**

**#HADOOP VARIABLES END**

source ~/.bashrc

javac –version

which javac

readlink –f /usr/local/javac

2. /usr/local/hadoop/etc/hadoop/hadoop-env.sh

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

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-i386

3. /usr/local/hadoop/etc/hadoop/core-site.xml

sudo mkdir –p /app/hadoop/tmp

sudo chown hduser:hadoop /app/hadoop/tmp

sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml

**<configuration>**

**<property>**

**<name>hadoop.tmp.dir</name>**

**<value>/app/hadoop/tmp</value>**

**<description>A base for other temporary directories.</description>**

**</property>**

**<property>**

**<name>fs.default.name</name>**

**<value>hdfs://localhost:54310</value>**

**<description>The name of the default file system. A URI whose**

**scheme and authority determine the FileSystem implementation. The**

**uri's scheme determines the config property (fs.SCHEME.impl) naming**

**the FileSystem implementation class. The uri's authority is used to**

**determine the host, port, etc. for a filesystem.</description>**

**</property>**

**</configuration>**

4. /usr/local/hadoop/etc/hadoop/mapred-site.xml

sudo nano cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.template /usr/local/hadoop/etc/hadoop/mapred-site.xml

**<configuration>**

**<property>**

**<name>mapred.job.tracker</name>**

**<value>localhost:54311</value>**

**<description>The host and port that the MapReduce job tracker runs**

**at. If "local", then jobs are run in-process as a single map**

**and reduce task.**

**</description>**

**</property>**

**</configuration>**

5. /usr/local/hadoop/etc/hadoop/hdfs-site.xml

sudo mkdir –p /usr/local/hadoop\_store/hdfs/namenode

sudo mkdir –p /usr/local/hadoop\_store/hdfs/datanode

sudo chown –R hduser:hadoop /usr/local/hadoop\_store

sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml

**<configuration>**

**<property>**

**<name>dfs.replication</name>**

**<value>1</value>**

**<description>Default block replication.**

**The actual number of replications can be specified when the file is created.**

**The default is used if replication is not specified in create time.**

**</description>**

**</property>**

**<property>**

**<name>dfs.namenode.name.dir</name>**

**<value>file:/usr/local/hadoop\_store/hdfs/namenode</value>**

**</property>**

**<property>**

**<name>dfs.datanode.data.dir</name>**

**<value>file:/usr/local/hadoop\_store/hdfs/datanode</value>**

**</property>**

**</configuration>**

9. Format the New Hadoop FileSystem:

hadoop namenode –format

10. Starting Hadoop:

sudo su hduser

start-all.sh

jps

netstat -plten | grep java

11. Stopping Hadoop:

stop-all.sh