SHREENIVASA ENGINEERING COLLEGE

(Approved by AICTE and Affiliated to Anna University,Chennai) B.Pallipatti,

Bommidi, Pappireddipatti (Tk), Dharmapuri Dt, Tamilnadu-635 301

DEPARTMENT ARTIFICIAL INTELLIGENCE AND DATA SCIENCE



Bonafide Certificate

This is to certify that this is a Bonafide Record of Practical work done by…….....

…….. of……..Semester of.....

……. Department in the..... Laboratory

During the year……

Head of the Department

Staff Incharge

Submitted for the University Examination held on _____

Internal Examiner

External Examiner

LIST OF EXPERIMENTS

Ex. No.	*Date*	*Name of the Experiment*	*Page no*	*Mark*	*Staff Sign*
1					
2					
3.					
4.					
5.					
6.					
7					
8					
9					

Ex.No.1 **CREATION OF VIRTUAL MACHINES**

Date:

AIM:

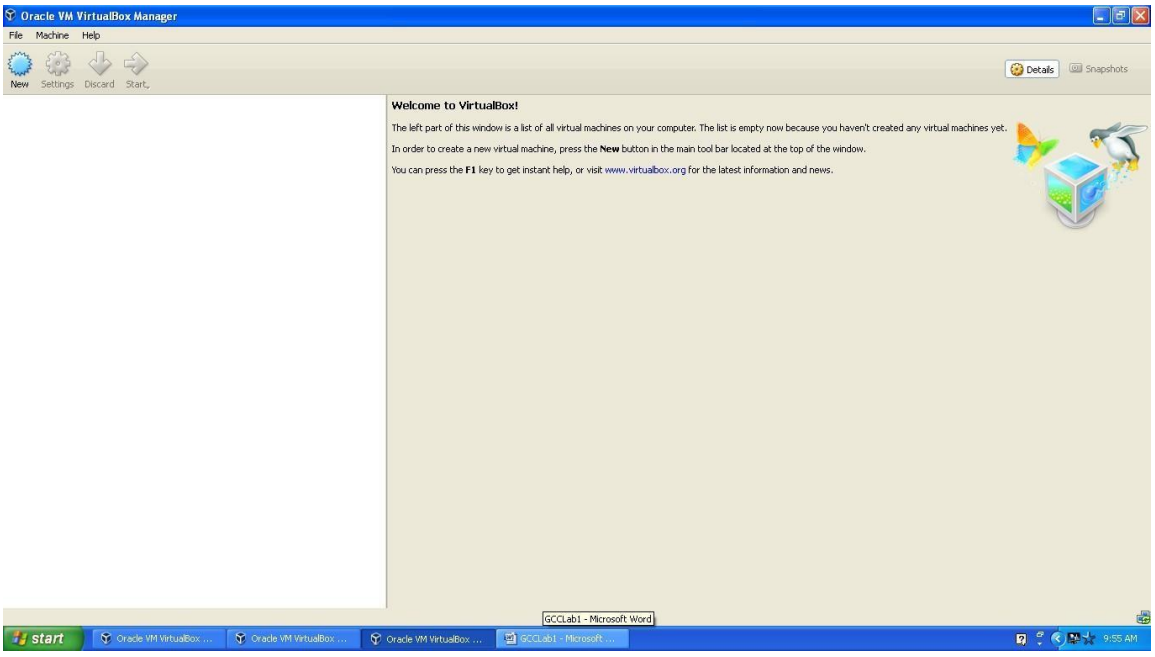
To find procedure to run the virtual machine of different configuration and checkhow many virtual machines can be utilized at particular time.

PROCEDURE:

Installing Ubuntu using Oracle Virtual Box

Step1:

Open Oracle virtual box manager and click create new -> virtual machine.

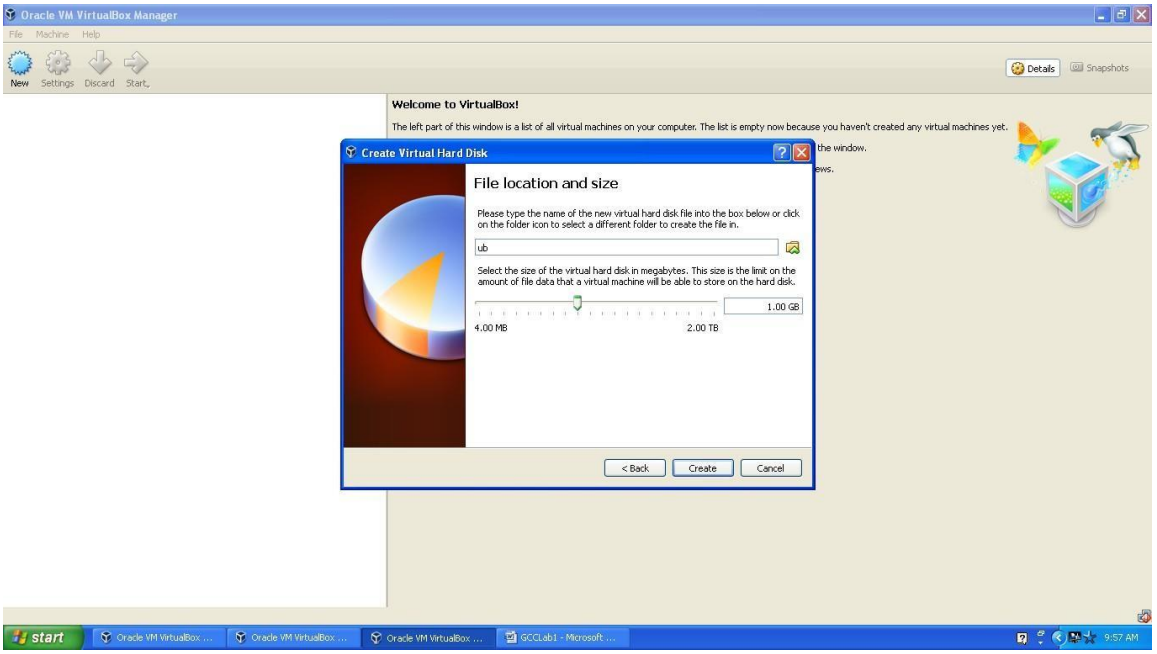


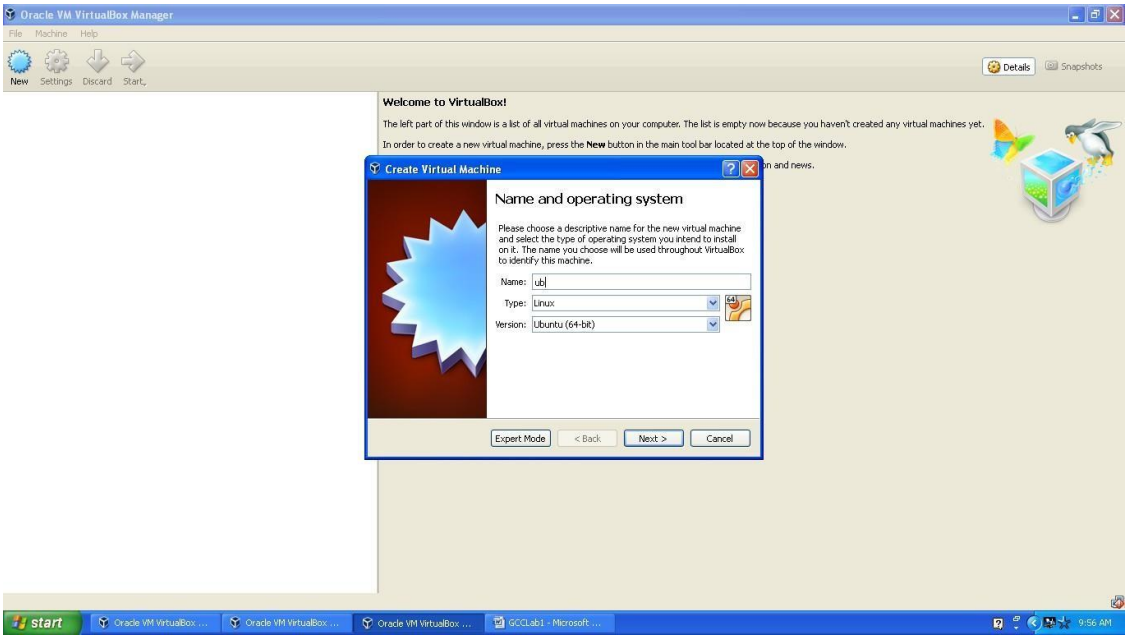
Step 2:

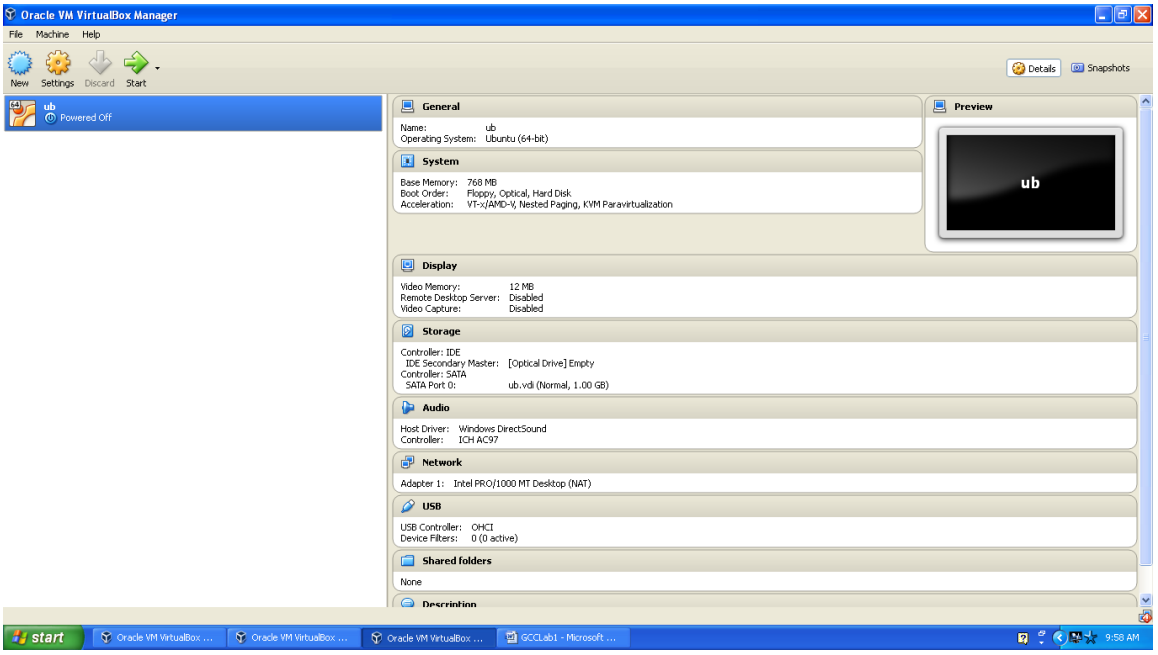
Provide a name for the virtual machine and select the hard disk size for the virtualmachine.

Select the storage size as Dynamically allocated memory size and click OK.The

virtual machine will be created.

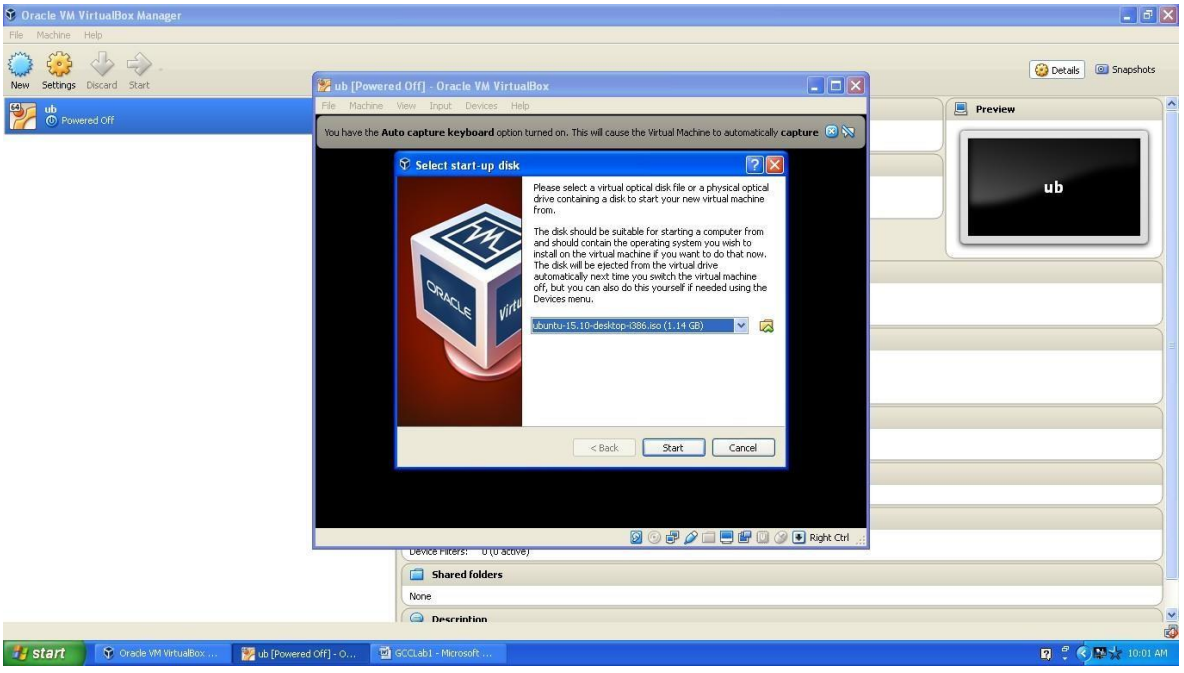

Oracle VM VirtualBox Manager
File Machine Help
New Settings Discard Start
Details Snapshots
Welcome to VirtualBox!
The left part of this window is a list of all virtual machines on your computer. The list is empty now because you haven't created any virtual machines yet.
Create Virtual Hard Disk
File location and size
Please type the name of the new virtual hard disk file into the box below or click on the folder icon to select a different folder to create the file in.
ub
Select the size of the virtual hard disk in megabytes. This size is the limit on the amount of file data that a virtual machine will be able to store on the hard disk.
1.00 GB
4.00 MB
2.00 TB
< Back
Create
Cancel
start
9:57 AM


Oracle VM VirtualBox Manager
File Machine Help
New Settings Discard Start
Details Snapshots
Welcome to VirtualBox!
The left part of this window is a list of all virtual machines on your computer. The list is empty now because you haven't created any virtual machines yet.
In order to create a new virtual machine, press the New button in the main tool bar located at the top of the window.
Create Virtual Machine
Name and operating system
Please choose a descriptive name for the new virtual machine and select the type of operating system you intend to install on it. The name you choose will be used throughout VirtualBox to identify this machine.
Name: ub
Type: Linux
Version: Ubuntu (64-bit)
Expert Mode
< Back
Next >
Cancel
start
9:56 AM



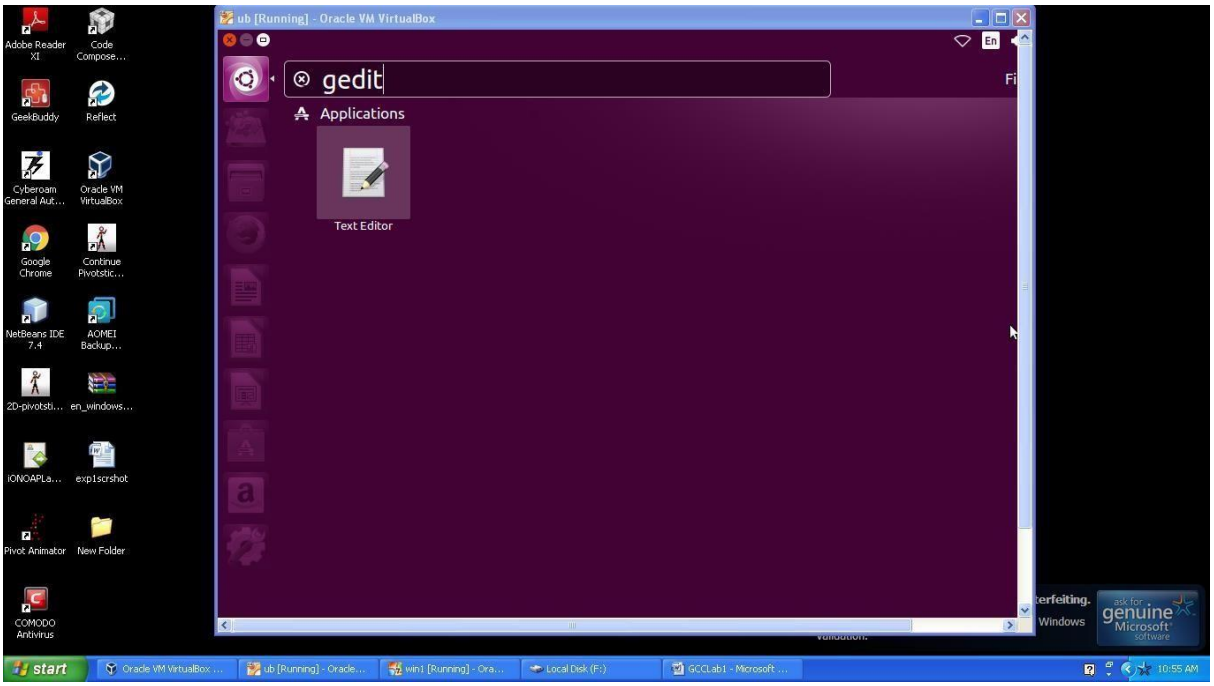
Step 3:

Select the iso file of the virtual OS Ubuntu and click Start.



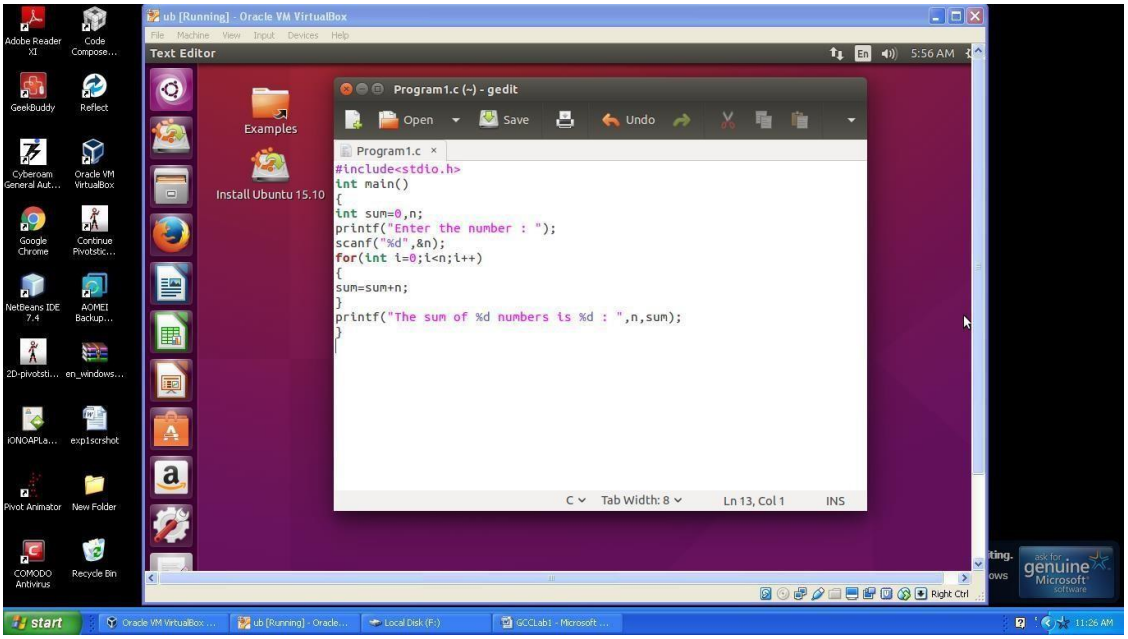
Step 4:

The virtual OS Ubuntu is opened successfully. Now type “gedit” in the search box to open text editor in Ubuntu.



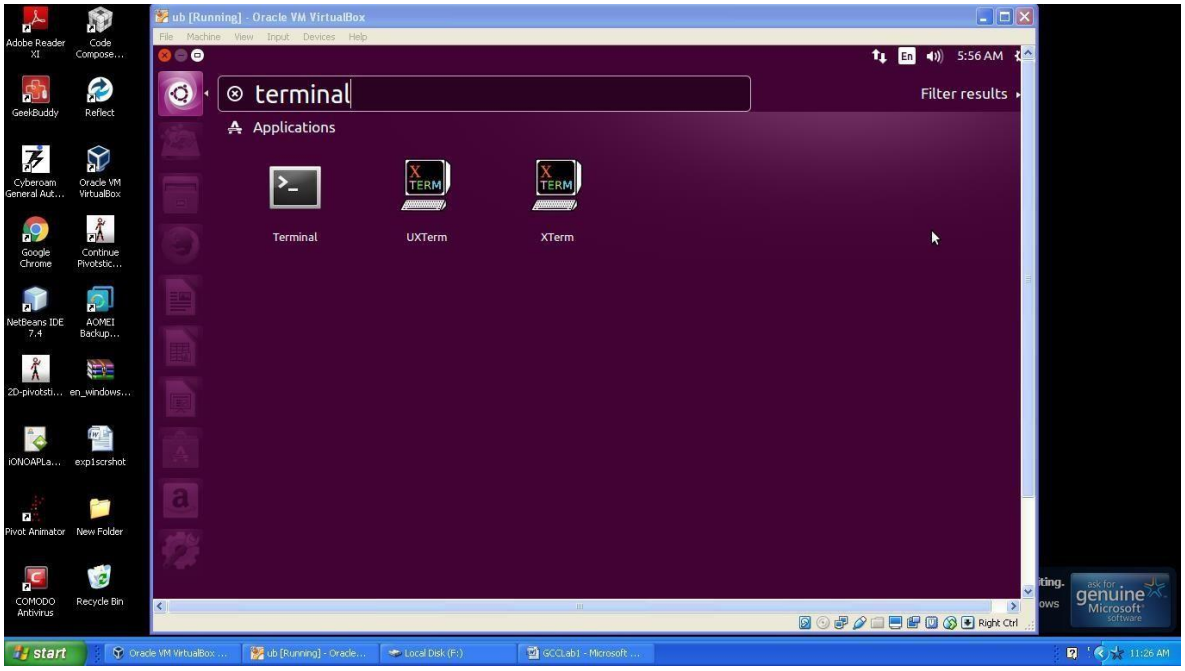
Step 5:

Type your desired C program in text editor and save it with the extension (.c).



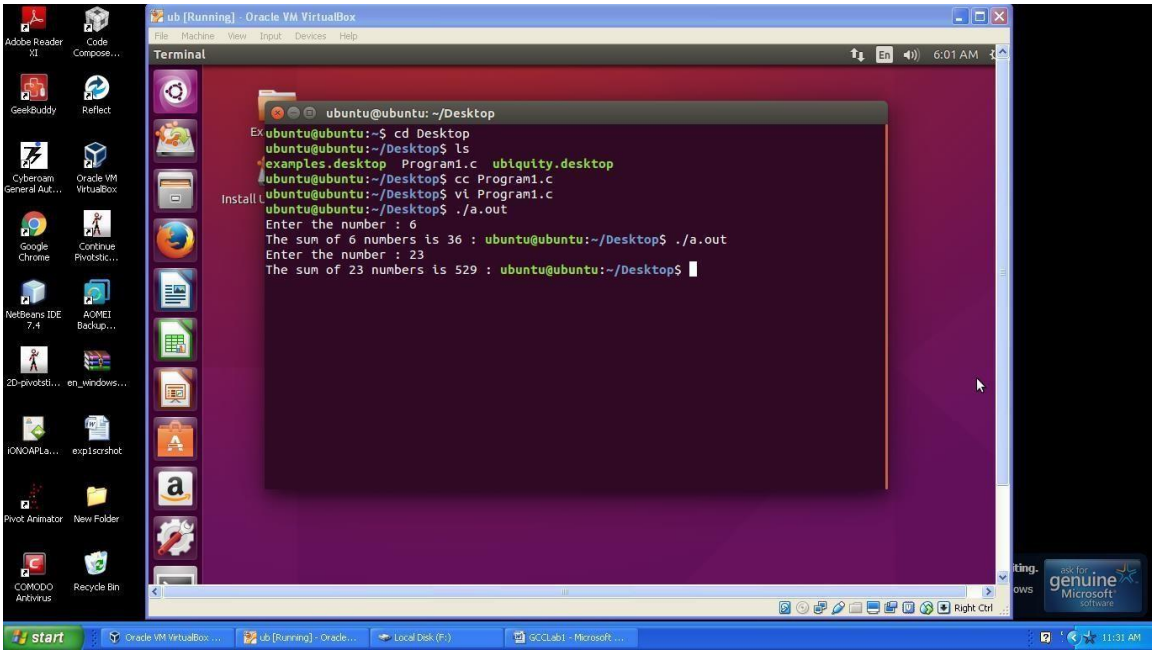
Step 6:

Type “terminal” in the search box to open the command window.



Step 7:

Type the necessary commands to compile and run the C program.

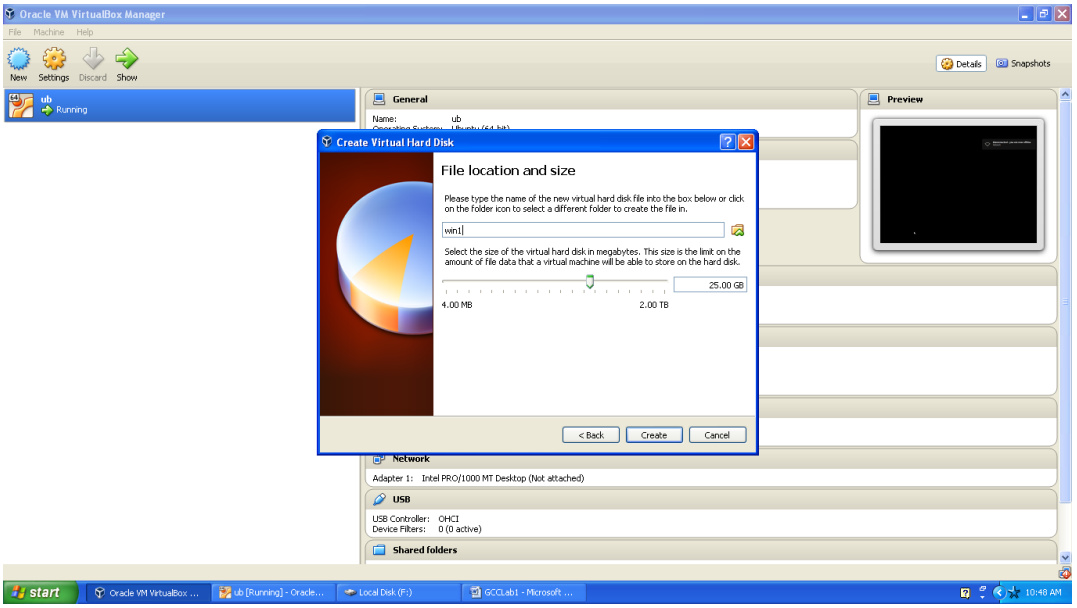


Installing Windows 7 using Oracle Virtual Box

PROCEDURE:

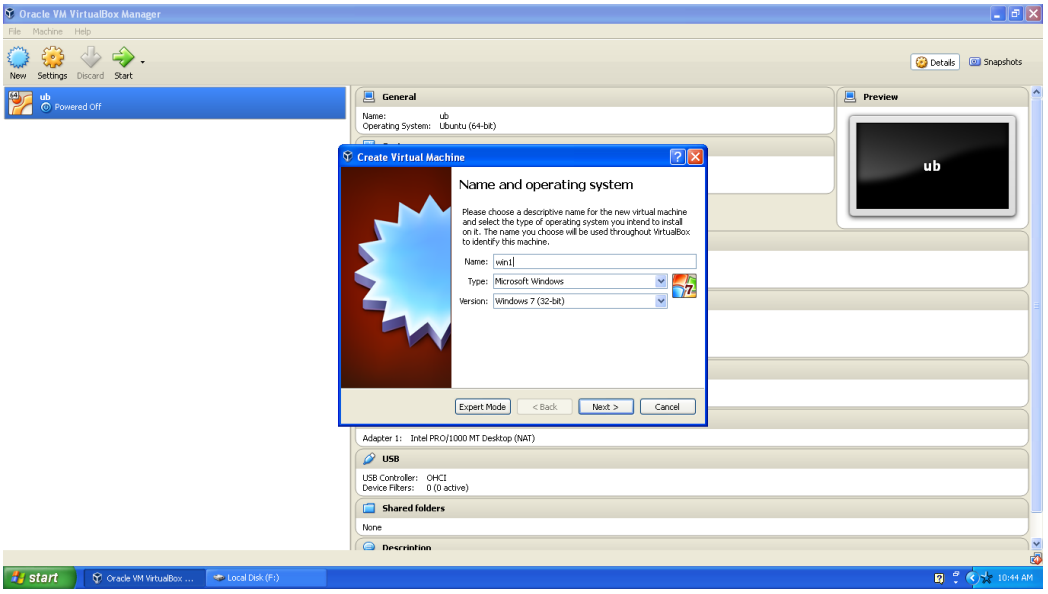
Step1:

Open Oracle virtual box manager and click create new -> virtual machine. Provide and name for the operating system and select the memory size to be occupied in memory.

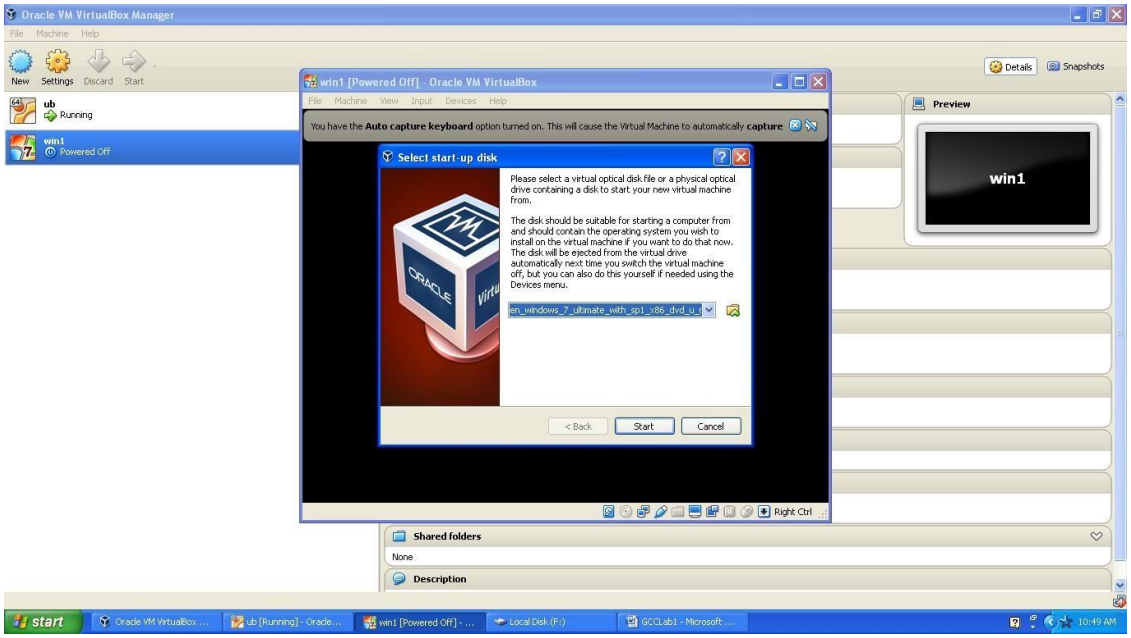


Step 2:

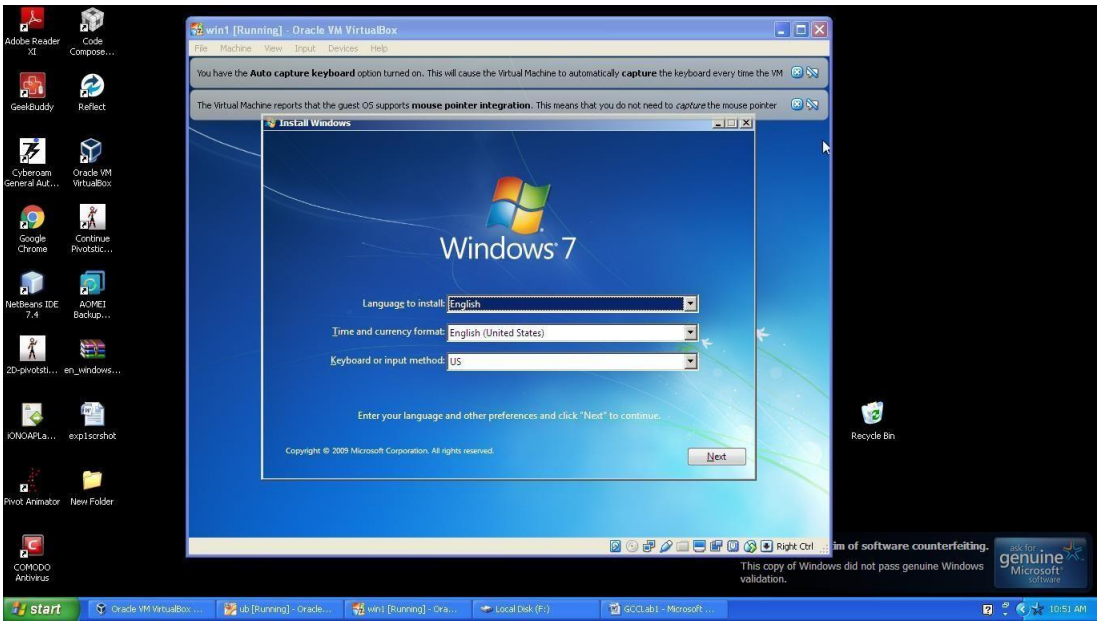
Select the iso file of the virtual OS Windows7 and click Start.

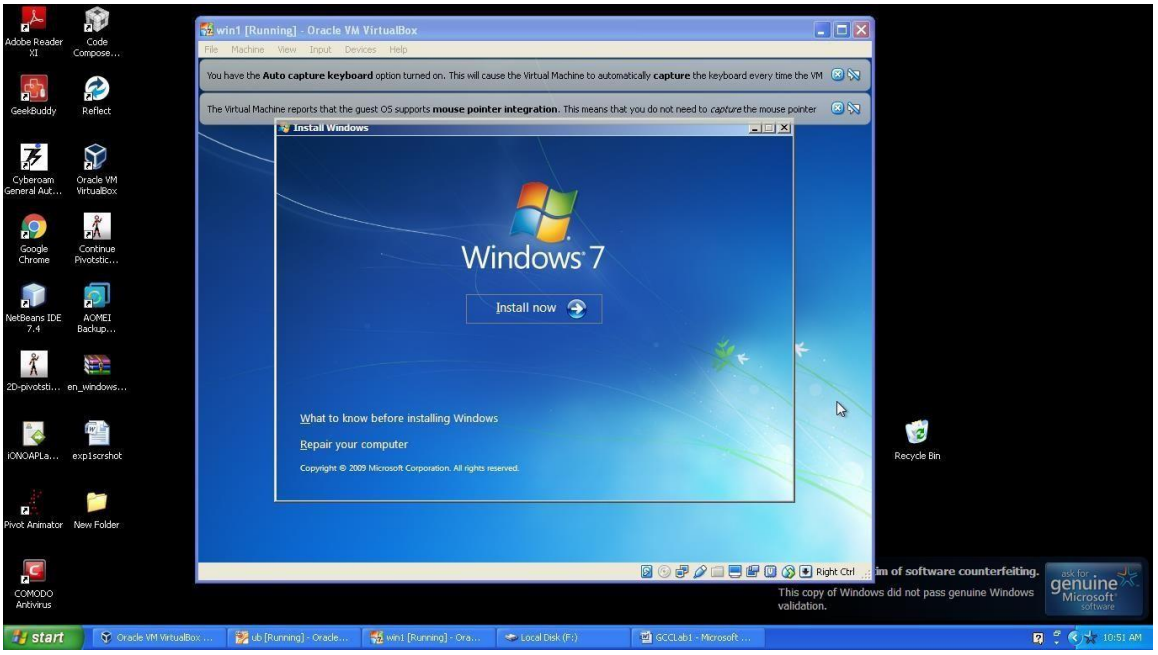


Step 3:



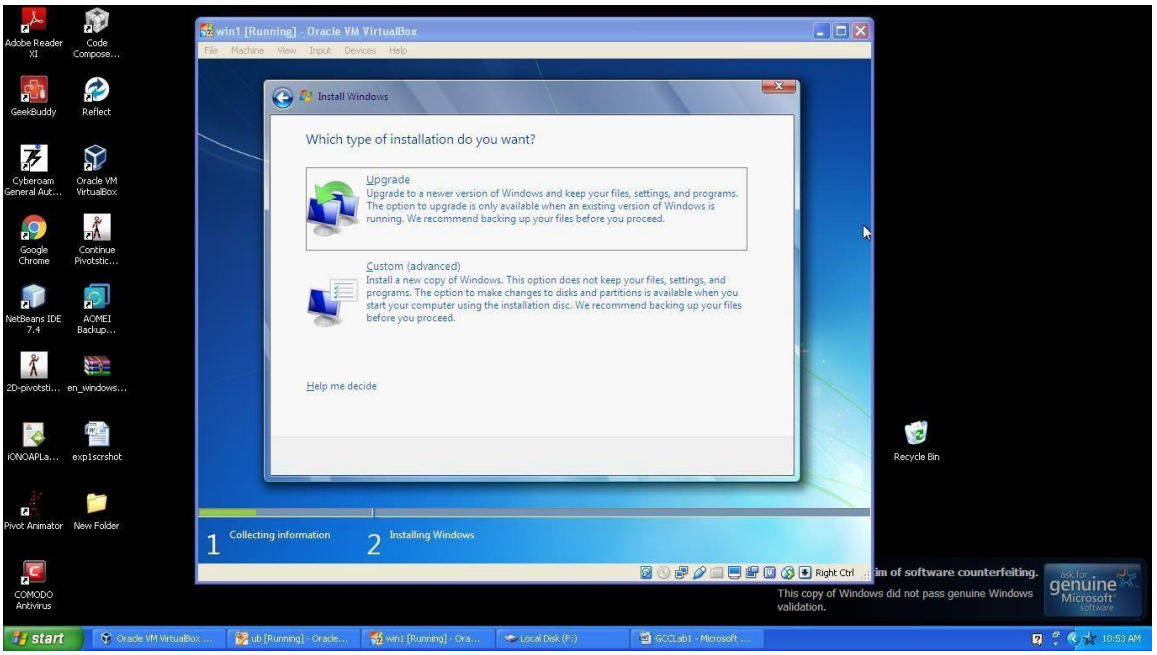
Select the language to use in the Operating System and click Install Now.

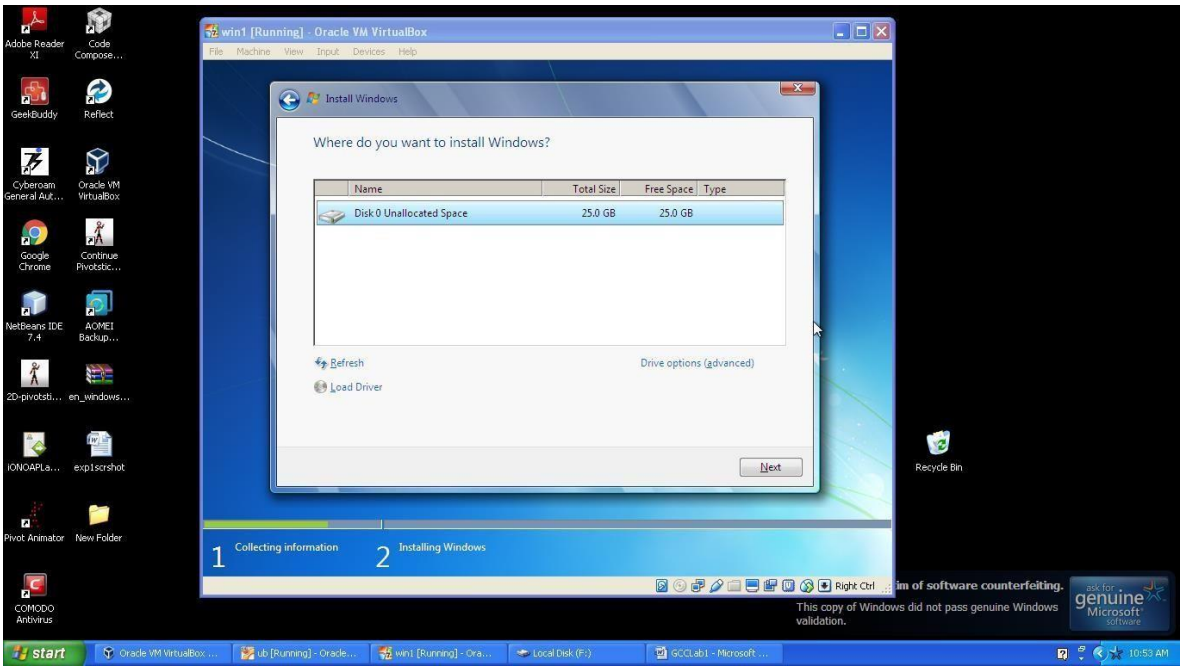




Step 4:

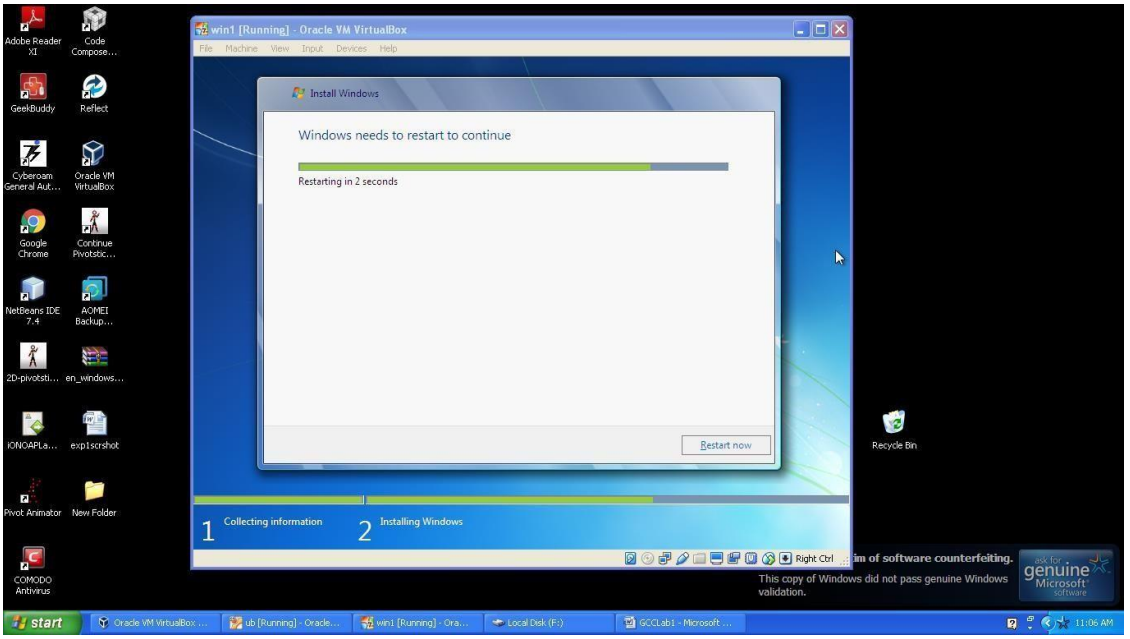
Select the type of installation as Custom for new installation and allocate Diskspace according to your convenience. Click Next to start the installation.

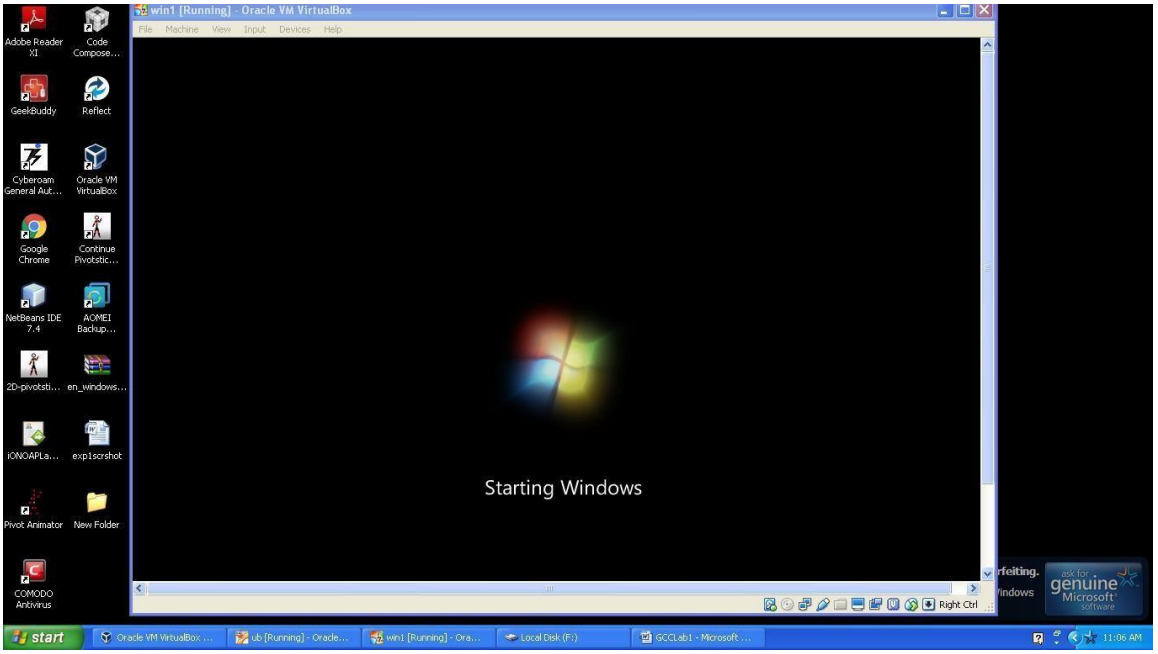




Step 5:

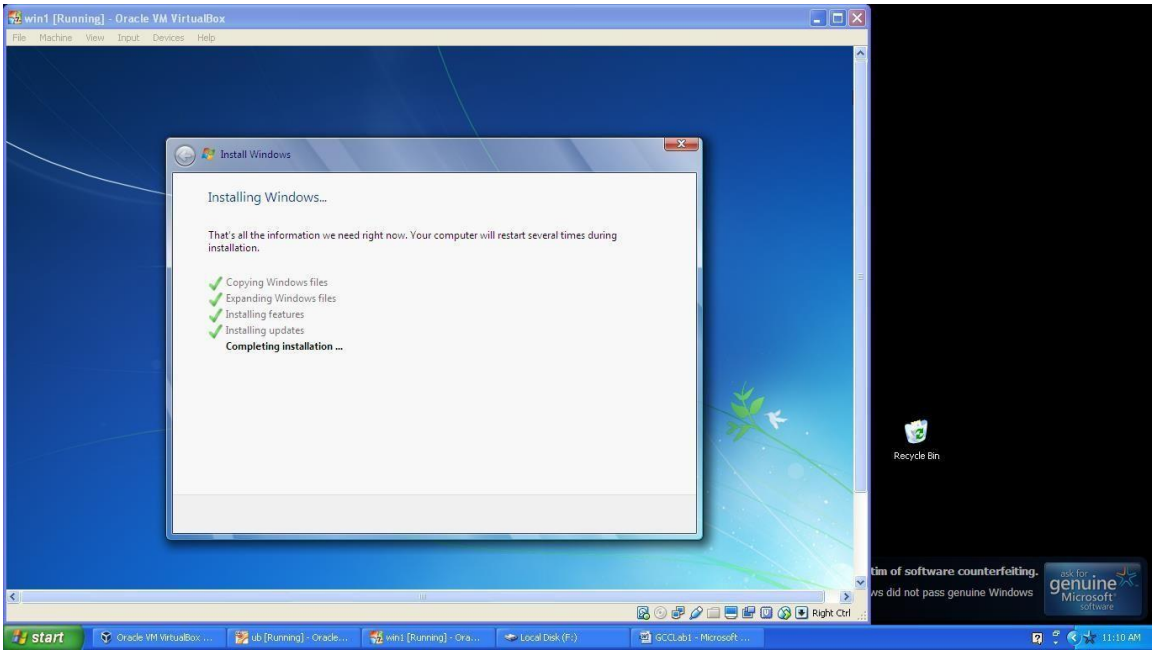
After installation the system will be restarted to complete the installation.

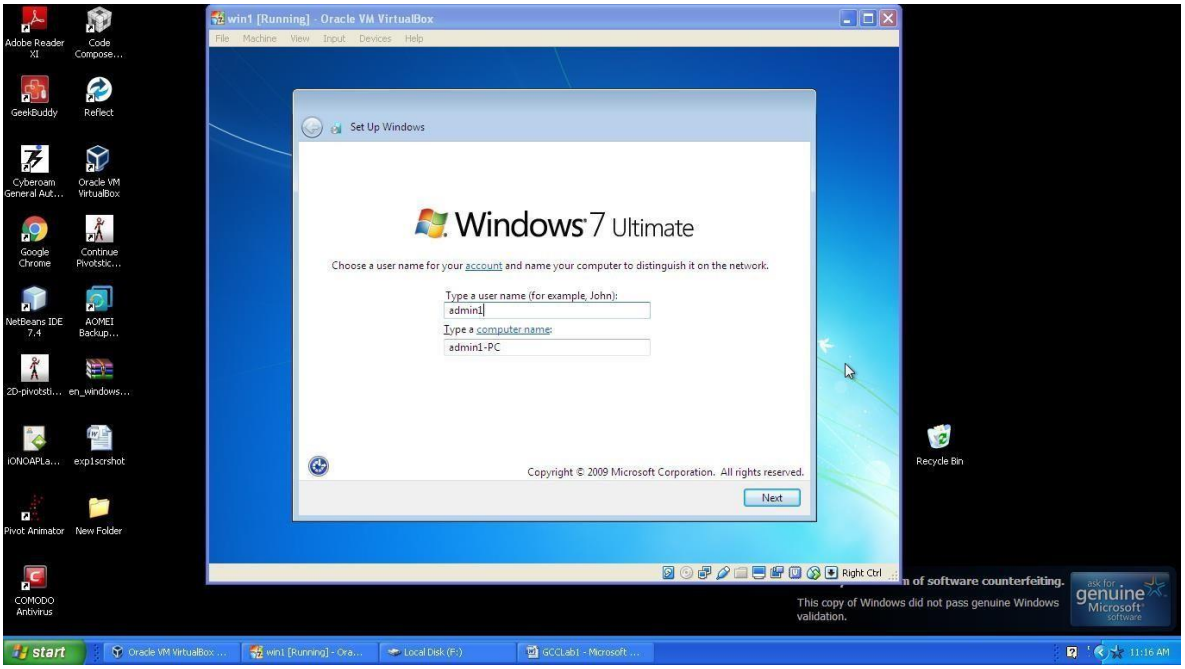




Step 6:

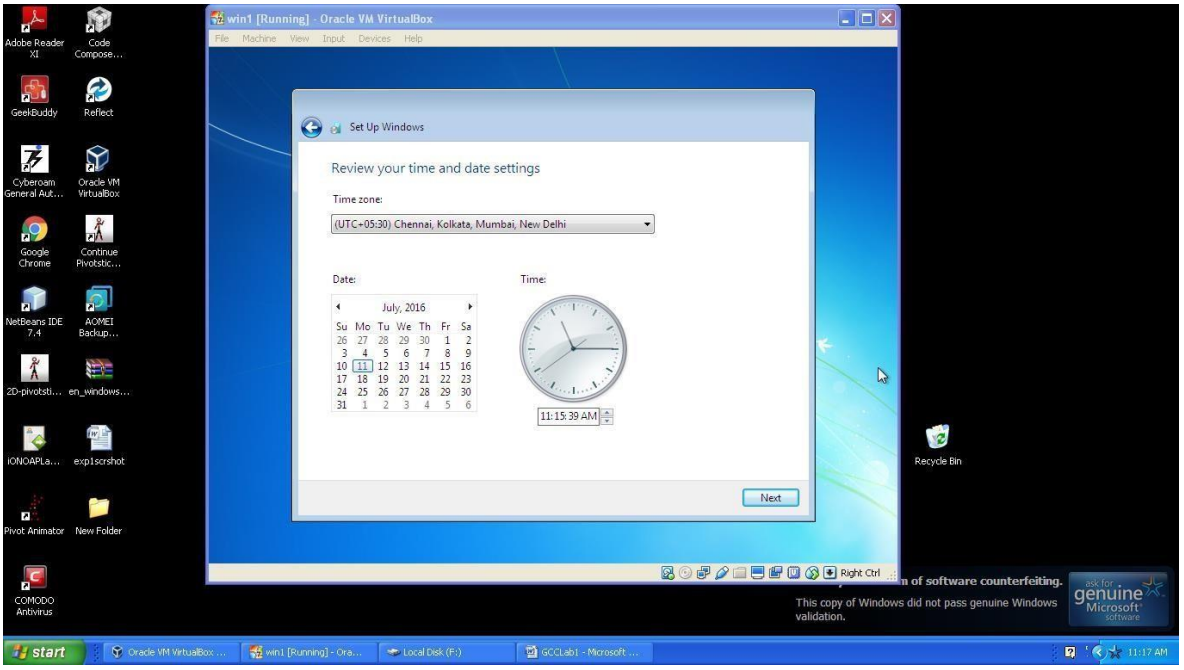
Provide a user name and password(optional) to gain access over the OS.





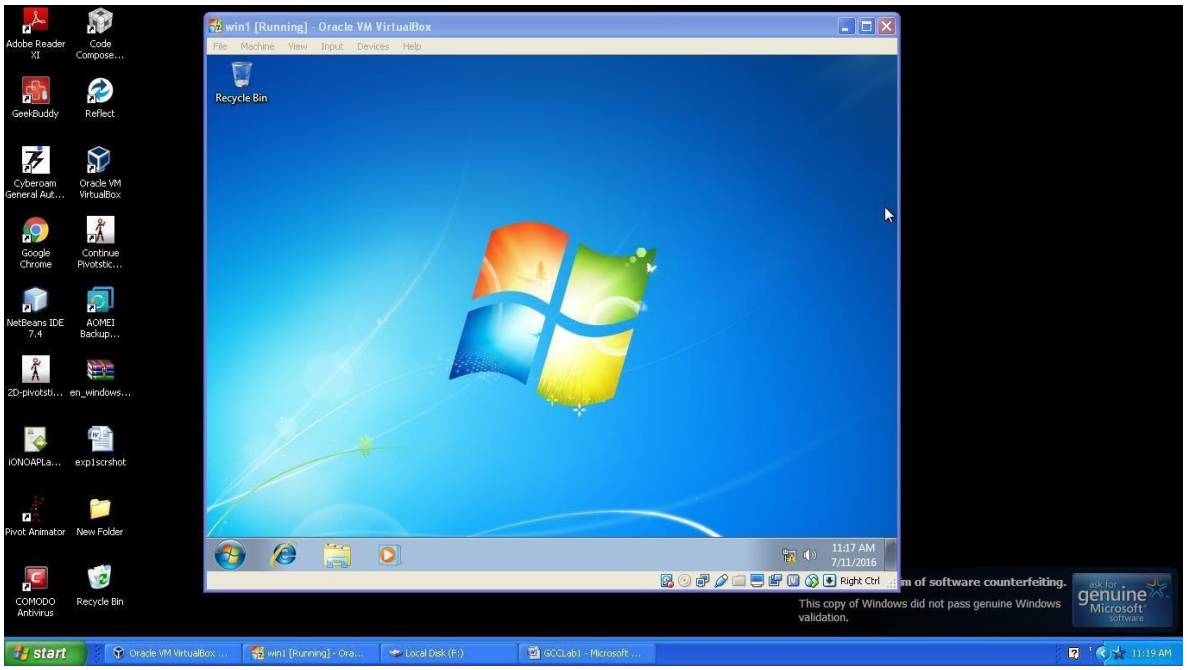
Step 7:

Set the time and date for the new Operating System.



Step 8:

Thus the new Operating System Windows7 will be opened as the virtual machine.



RESULT:

Thus the procedure to run different virtual machines on a single system usingOracle Virtual Box is studied and implemented successfully.

EX NO:2 **TO ATTACH VIRTUAL BLOCK TO THE VIRTUAL MACHINE AND**

DATE: **CHECK WHETHER IT HOLDS THE DATA EVEN AFTER**

THE RELEASE OF THE VIRTUAL MACHINE

AIM:

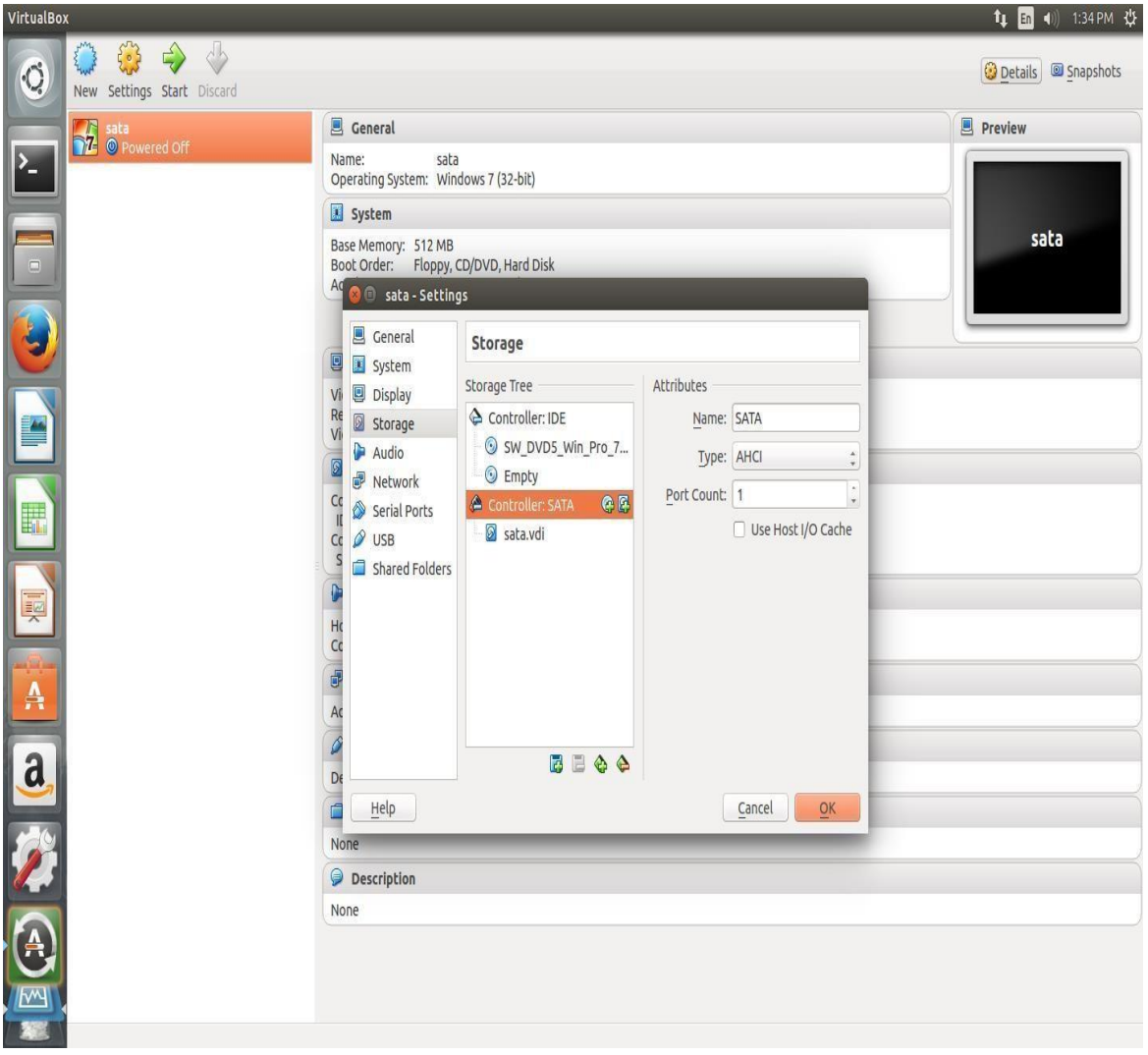
To find a procedure to attach virtual block to the vritual machine and check whether it holds the data even after the release of the virtual machine.

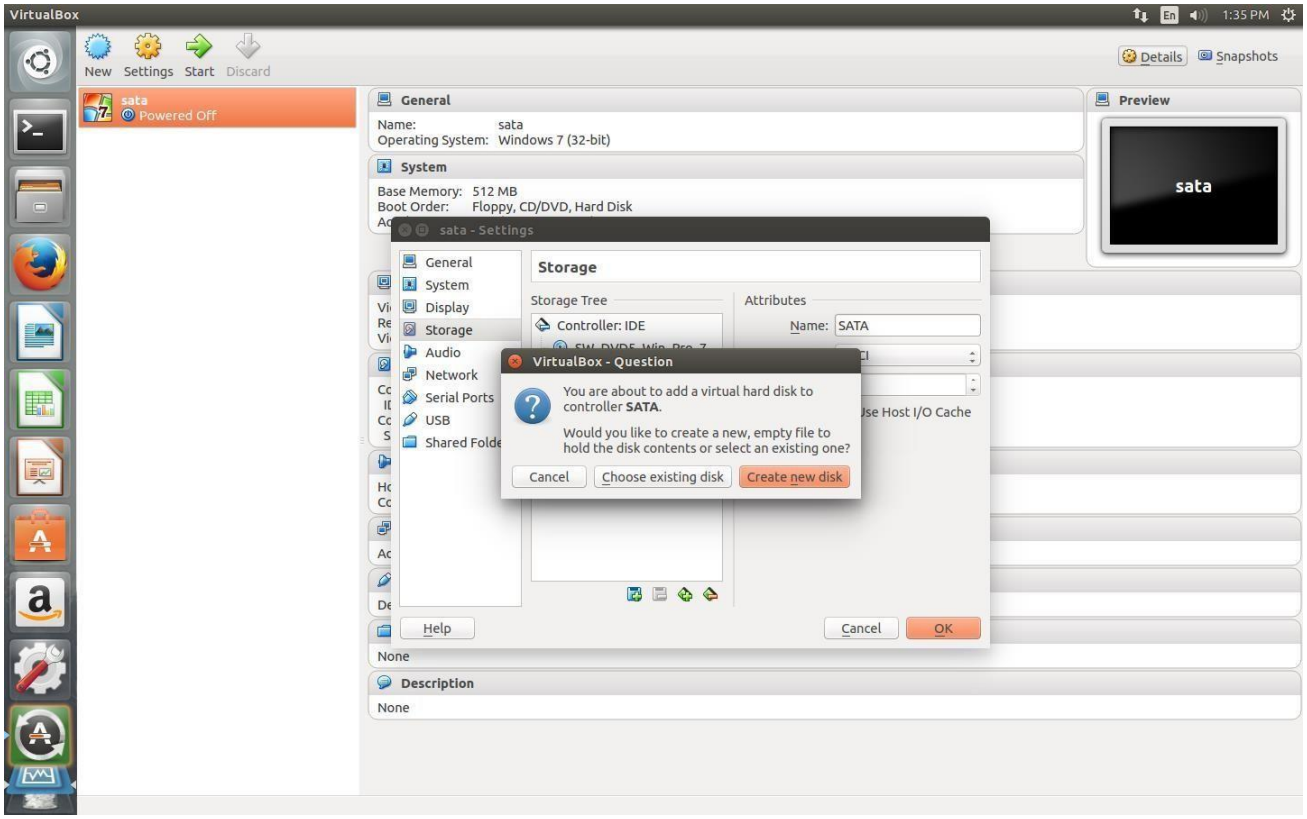
PROCEDURE:

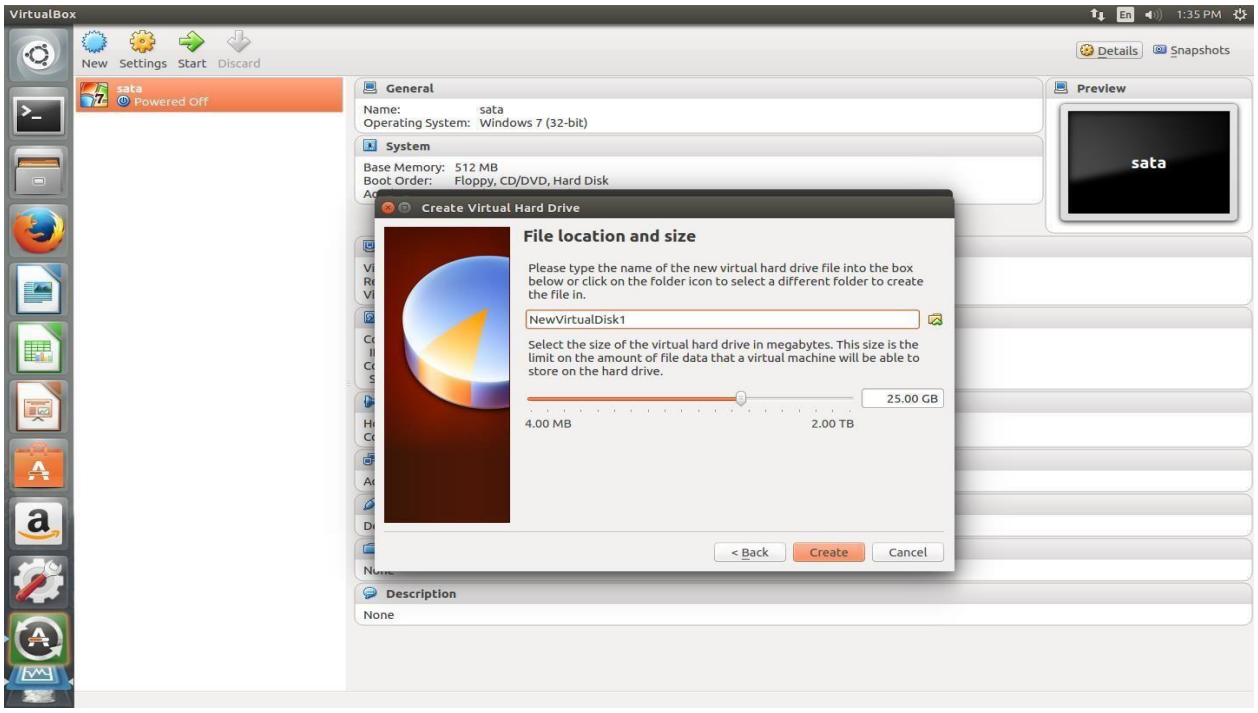
Step1: Open the virtual box and create a new virtual machine - Windows 7(32bit)Step2:

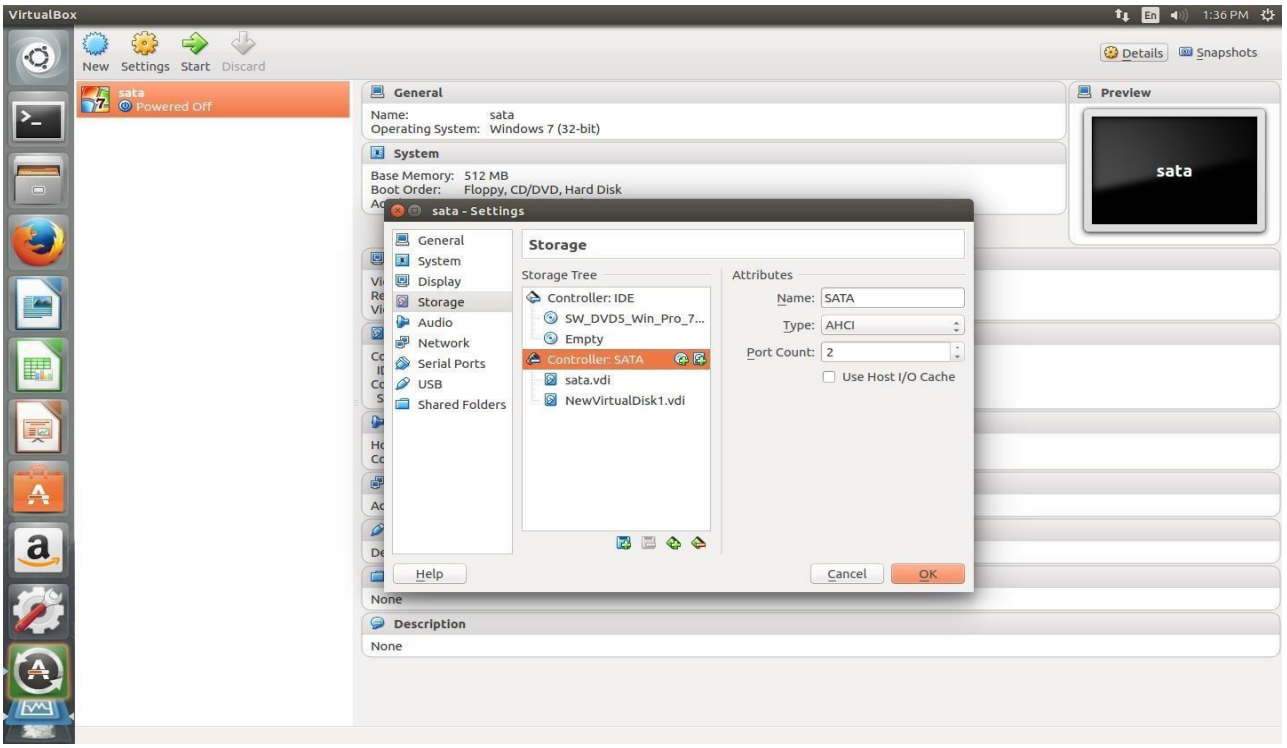
In **Storage tab** select Controller IDE:and then "**choose disk**" is selected

Step3: In **Storage tab** select **Controller :SATA** and then "**choose new disk**" is selected.

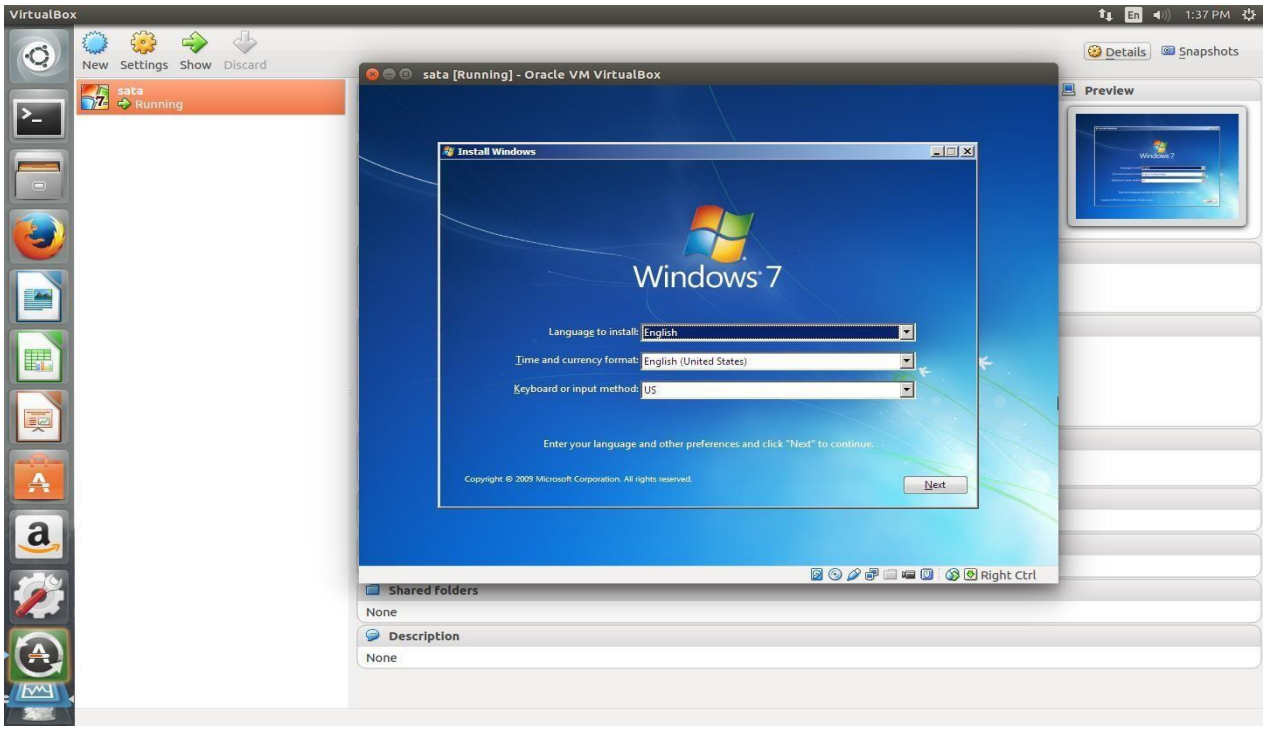


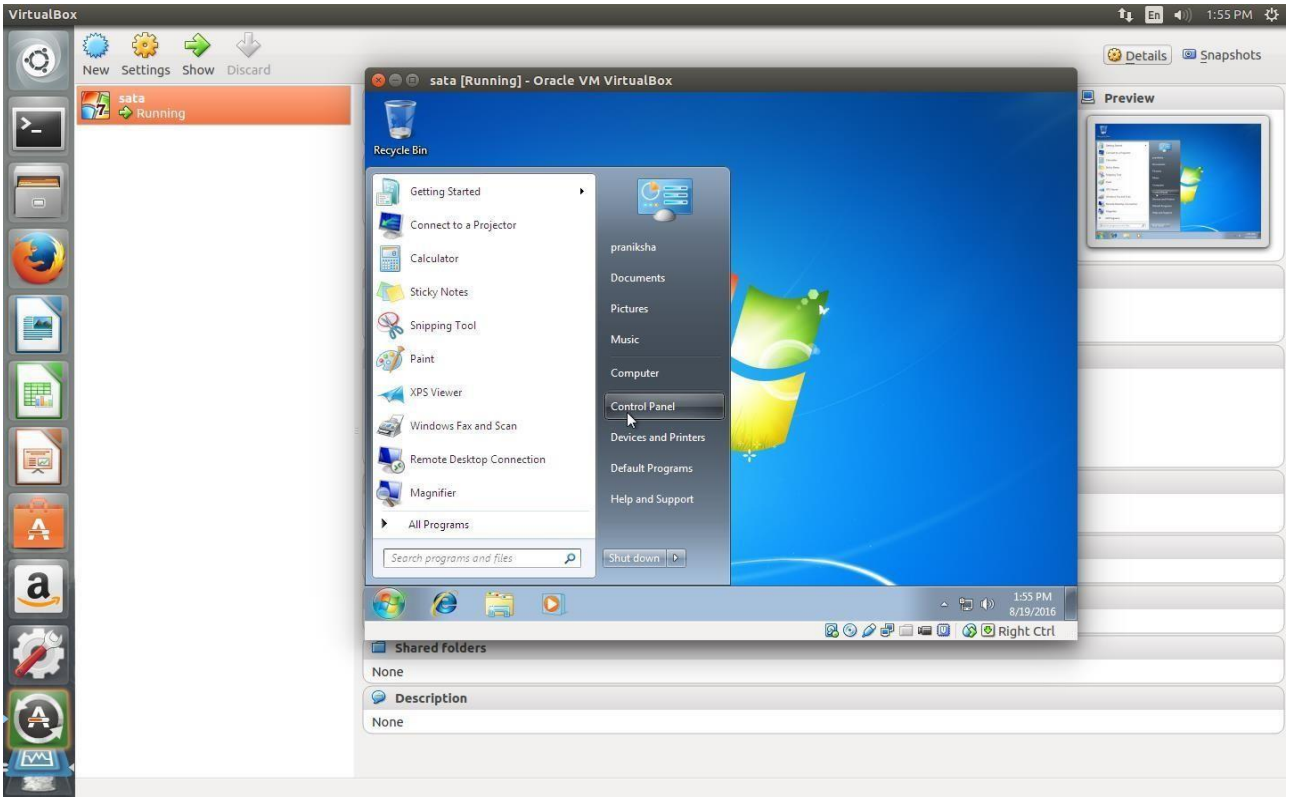

VirtualBox
1:35 PM
New
Settings
Start
Discard
Details
Snapshots
sata
Powered Off
General
Name: sata
Operating System: Windows 7 (32-bit)
System
Base Memory: 512 MB
Boot Order: Floppy, CD/DVD, Hard Disk
Preview
sata
sata - Settings
General
System
Display
Storage
Audio
Network
Serial Ports
USB
Shared Folde
Storage
Storage Tree
Controller: IDE
Attributes
Name: SATA
VirtualBox - Question
You are about to add a virtual hard disk to controller SATA.
Would you like to create a new, empty file to hold the disk contents or select an existing one?
Cancel
Choose existing disk
Create new disk
Use Host I/O Cache
Help
Cancel
OK
None
Description
None


VirtualBox
1:35 PM
New
Settings
Start
Discard
Details
Snapshots
sata
Powered Off
General
Name: sata
Operating System: Windows 7 (32-bit)
System
Base Memory: 512 MB
Boot Order: Floppy, CD/DVD, Hard Disk
Preview
sata
Create Virtual Hard Drive
File location and size
Please type the name of the new virtual hard drive file into the box below or click on the folder icon to select a different folder to create the file in.
NewVirtualDisk1
Select the size of the virtual hard drive in megabytes. This size is the limit on the amount of file data that a virtual machine will be able to store on the hard drive.
25.00 GB
4.00 MB
2.00 TB
< Back
Create
Cancel
Description
None

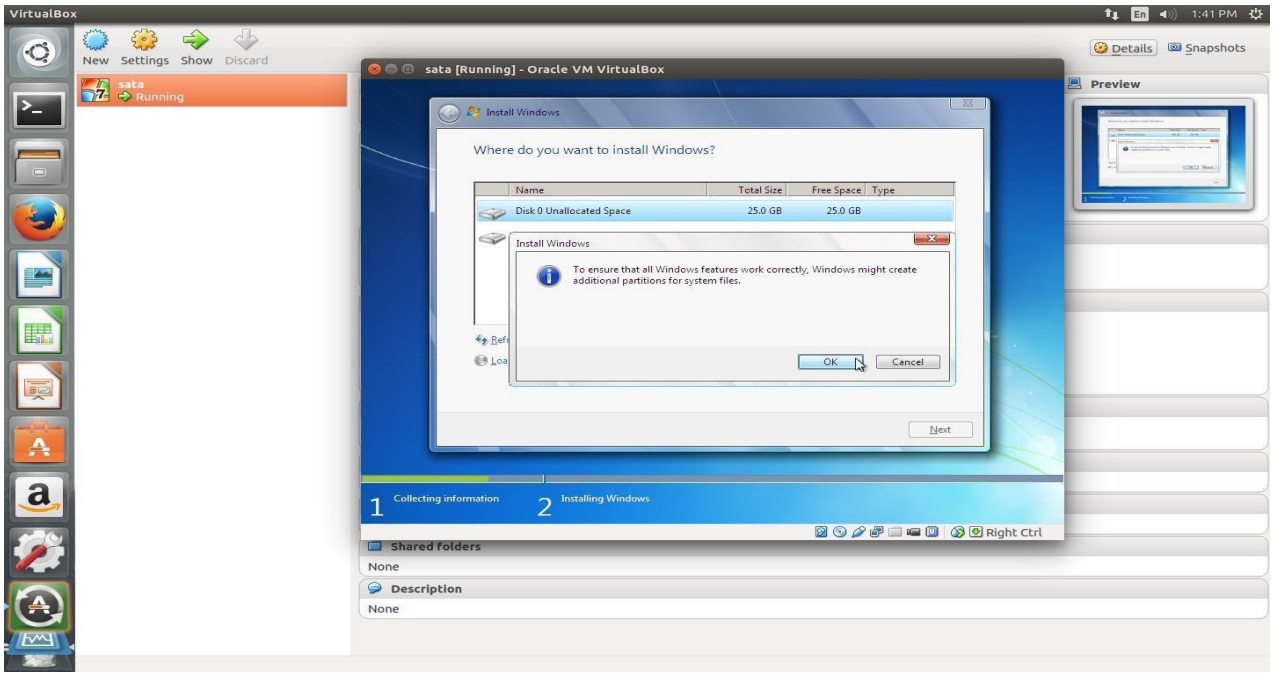


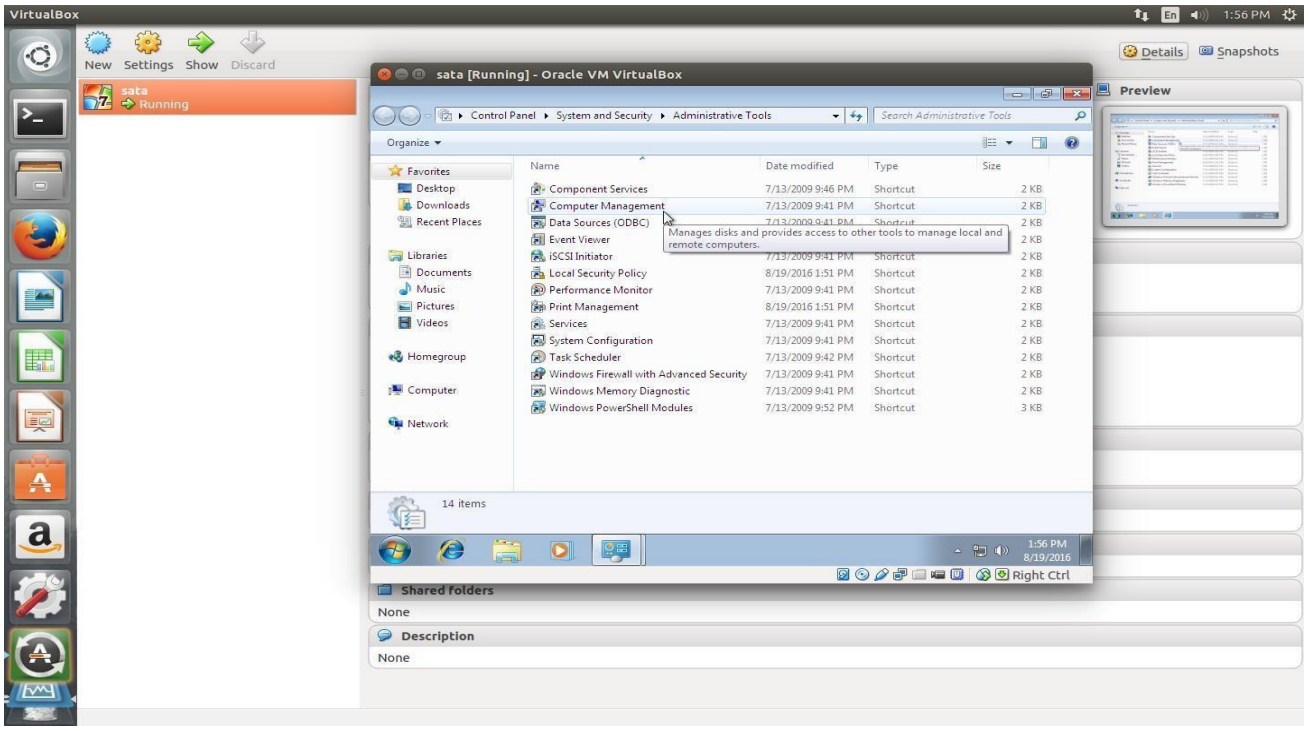
Step4: Now **Start** the virtual machine and **install windows**



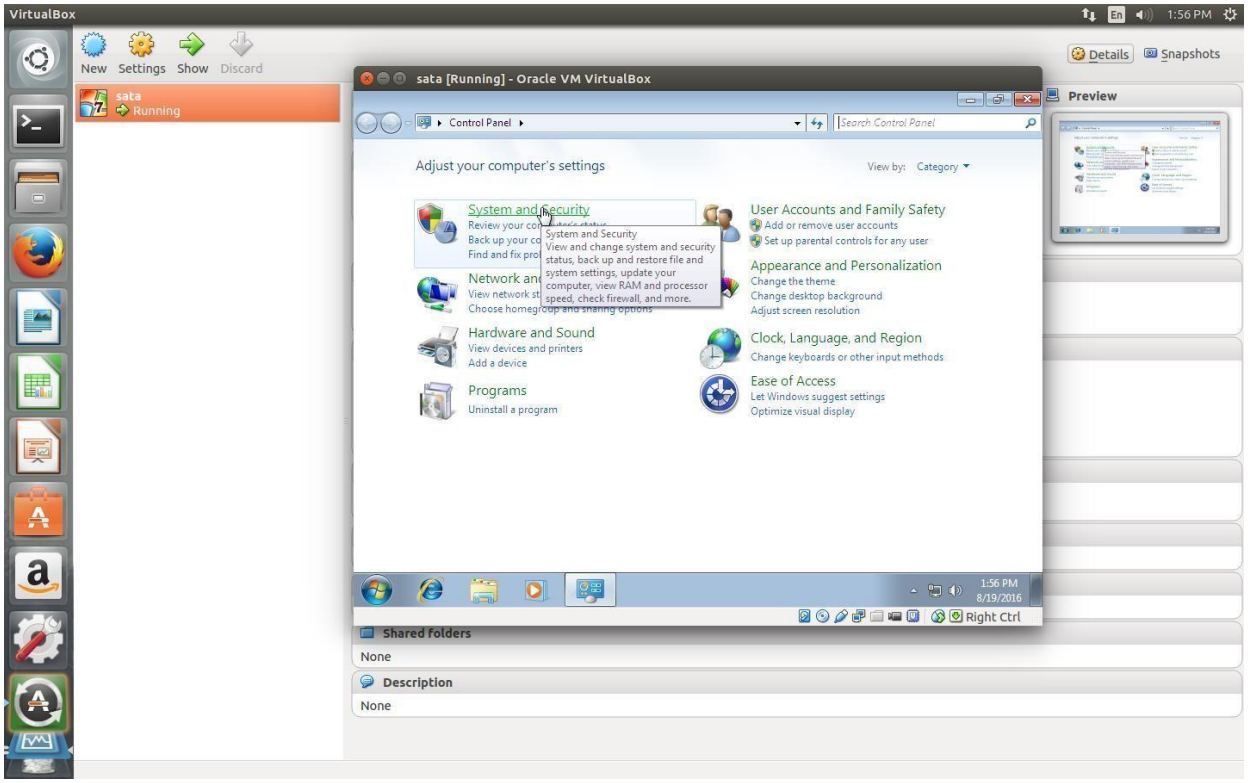


Step5: Go to **Start** in Virtual machine and select **control panelP**

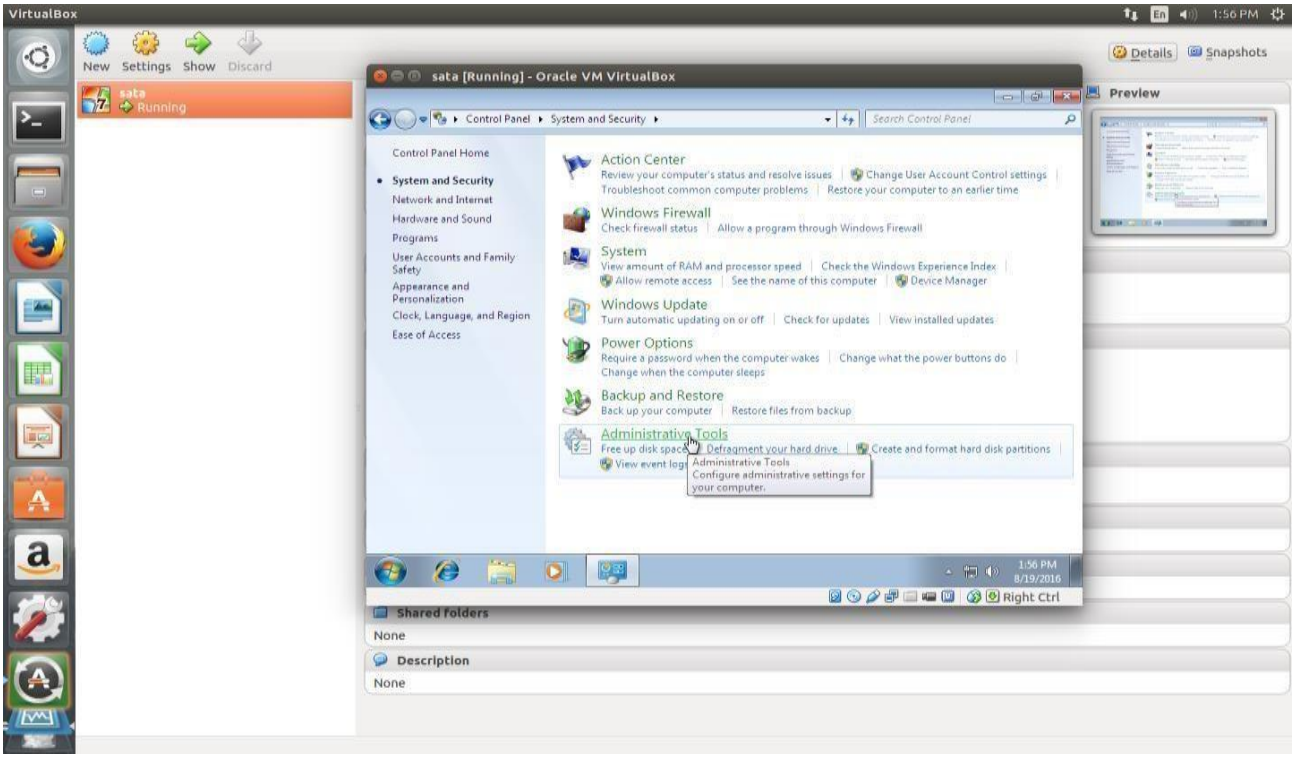


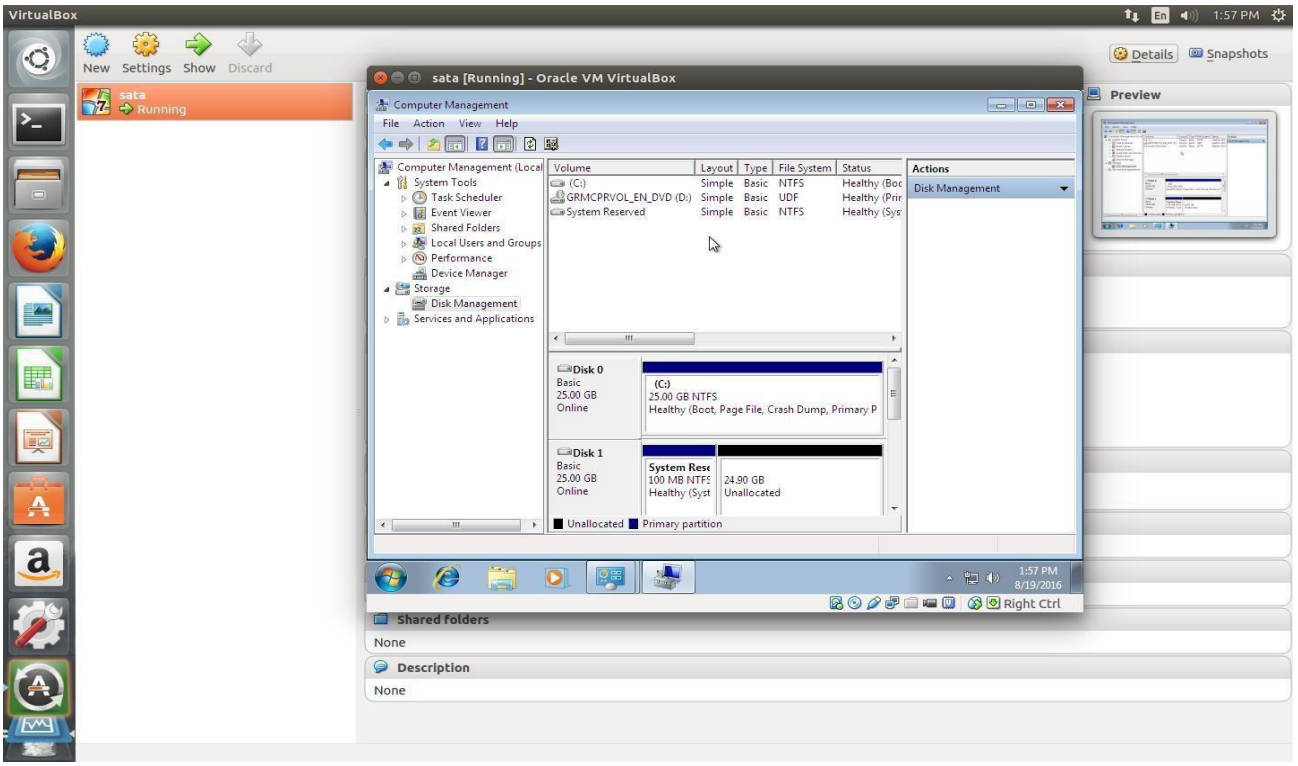


Step6: In Control Panel select **"system security**" and then "**administrative tools" .**

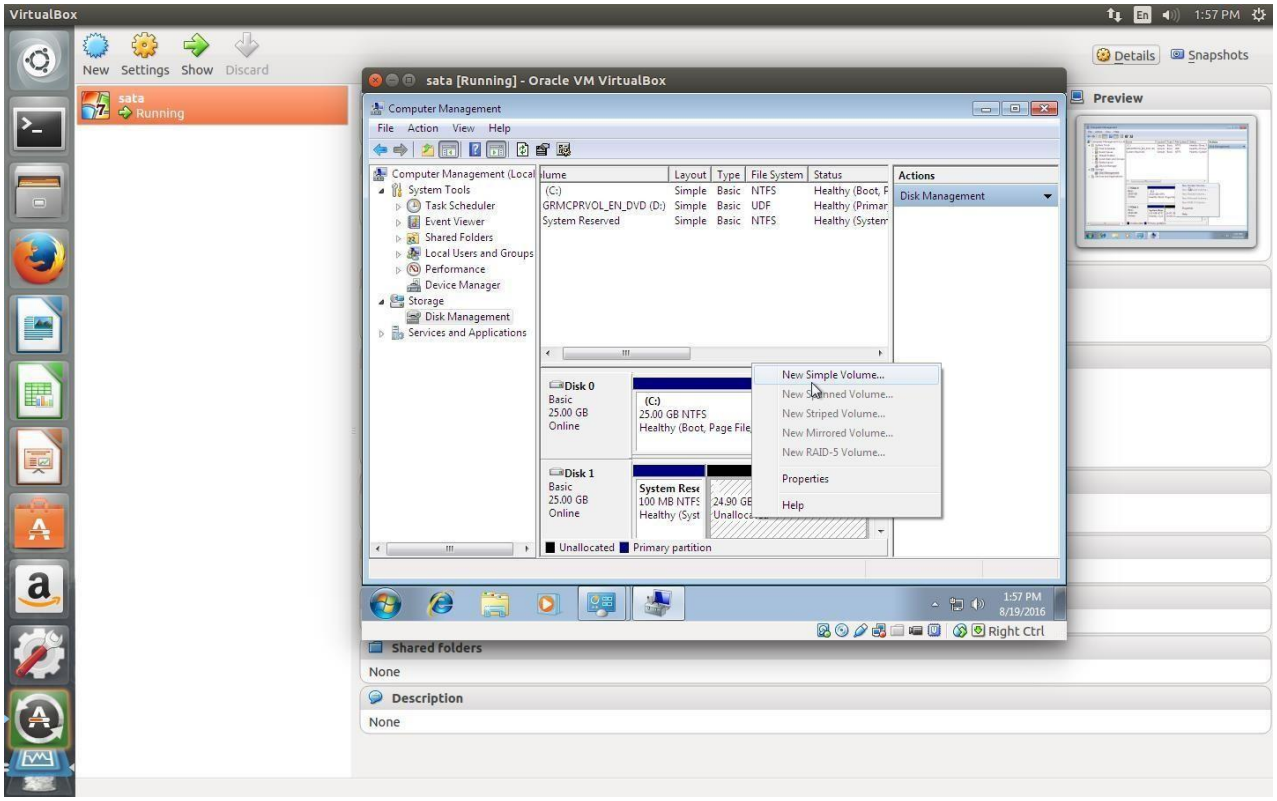


Step8: Select "**Disk Management** " from storage tab

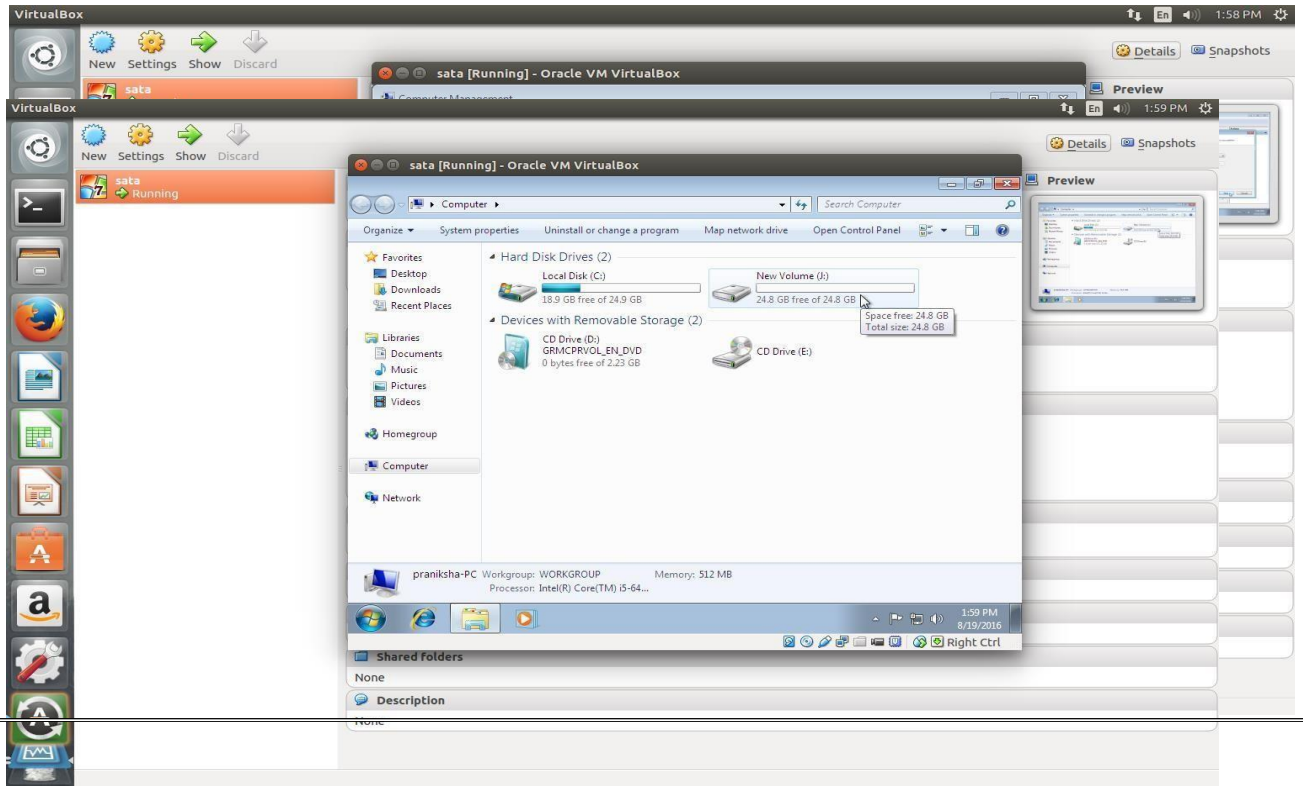




Step9: Select the **disk1** and right click on **unallocated** option and "**select new simple**



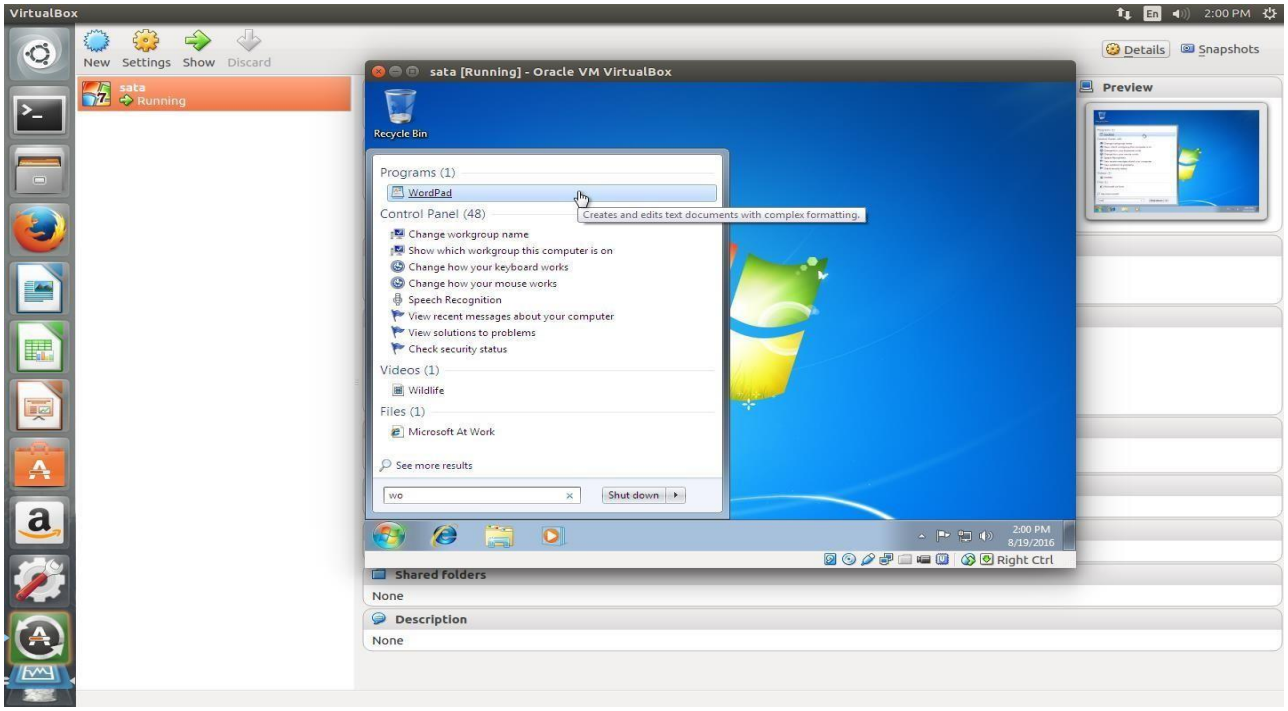
Step10: Select **next(2)** option and also choose "**disk name"** and then give **next(2).** finally **"finish**" the window.

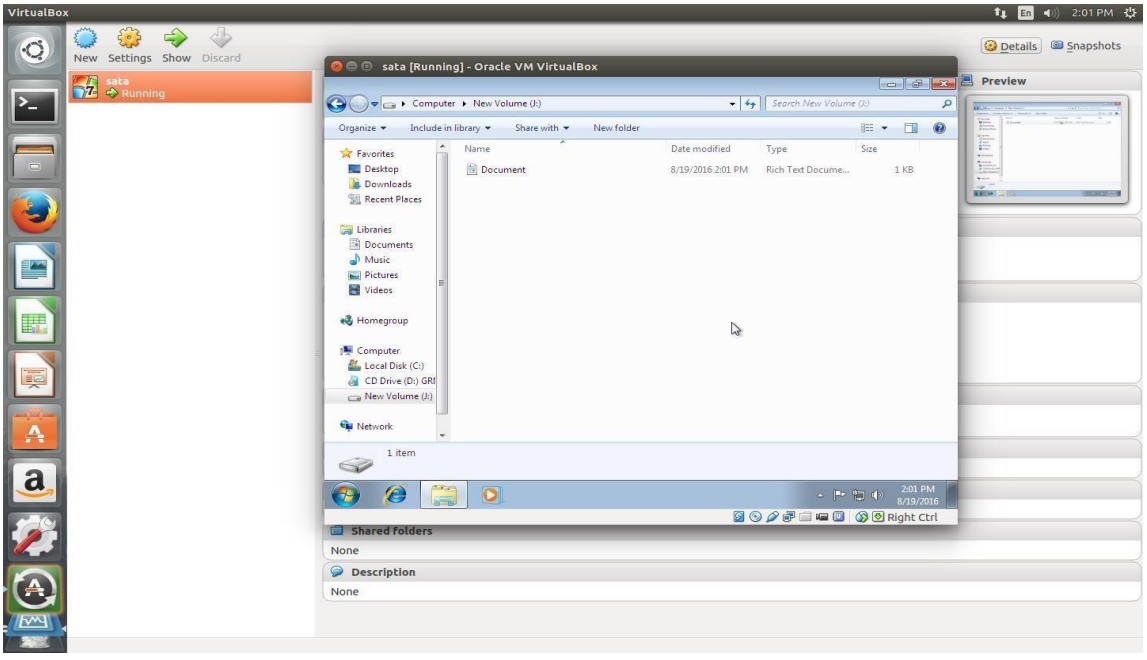


Step

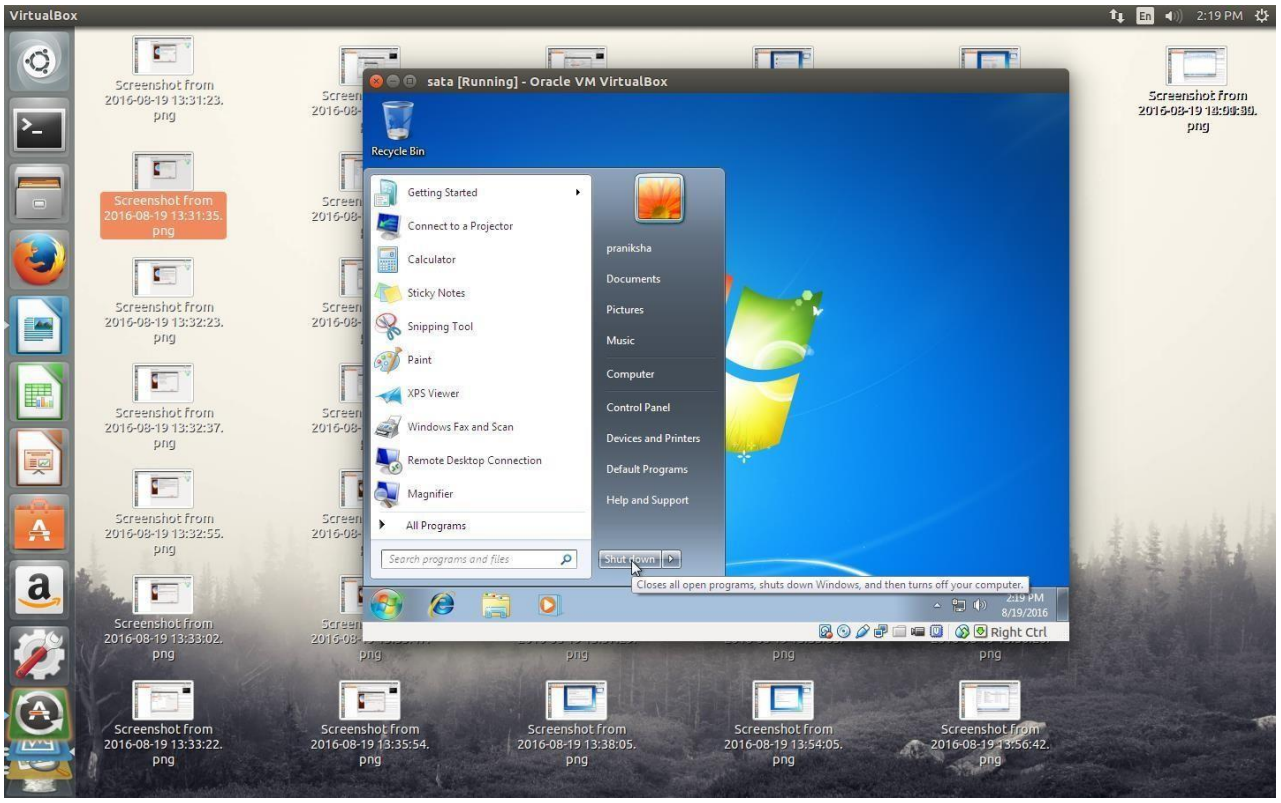
11: Select **Start** menu in virtual machine and choose computer icon and the memorypartition will be displayed.

Type a word file and save in new memory partition.





Step12: **Shut Down** the virtual machine and log in again the file created in the partition willbe displayed.



RESULT:

Thus the virtual block has been attached and the virtual machine has beenchecked.

Ex.No.3 **EXECUTION OF A SAMPLE PROGRAM IN A**

Date: **VIRTUAL MACHINE**

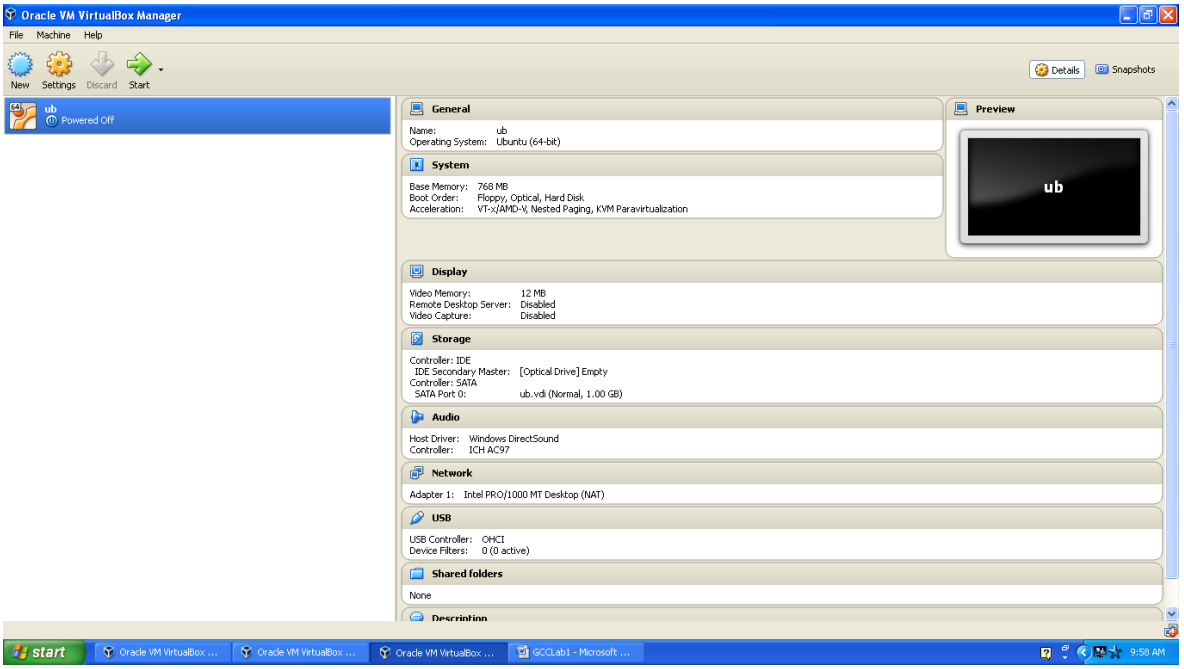
AIM:

To find a procedure to use the C compiler in the virtual machine and execute asample program.

PROCEDURE:

Step 1:

Open the virtual machine in which you want to run the C program.

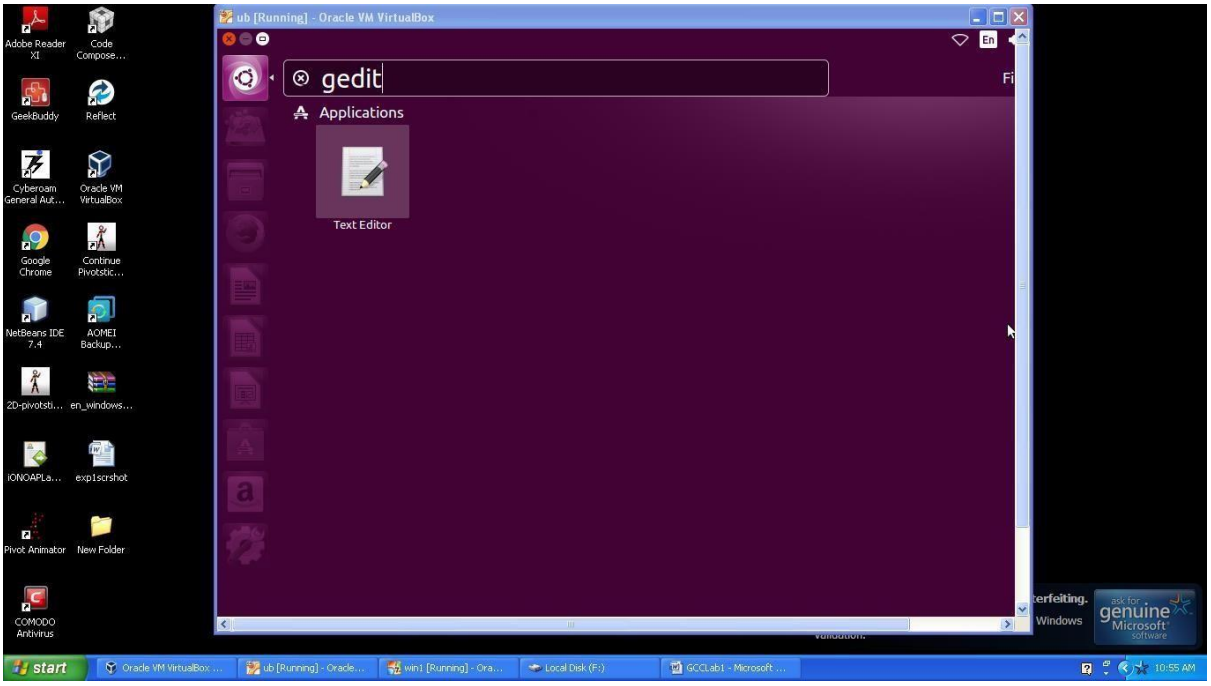


Step 2:

The text editor used by the Ubuntu Operating System is the GEDIT.It can

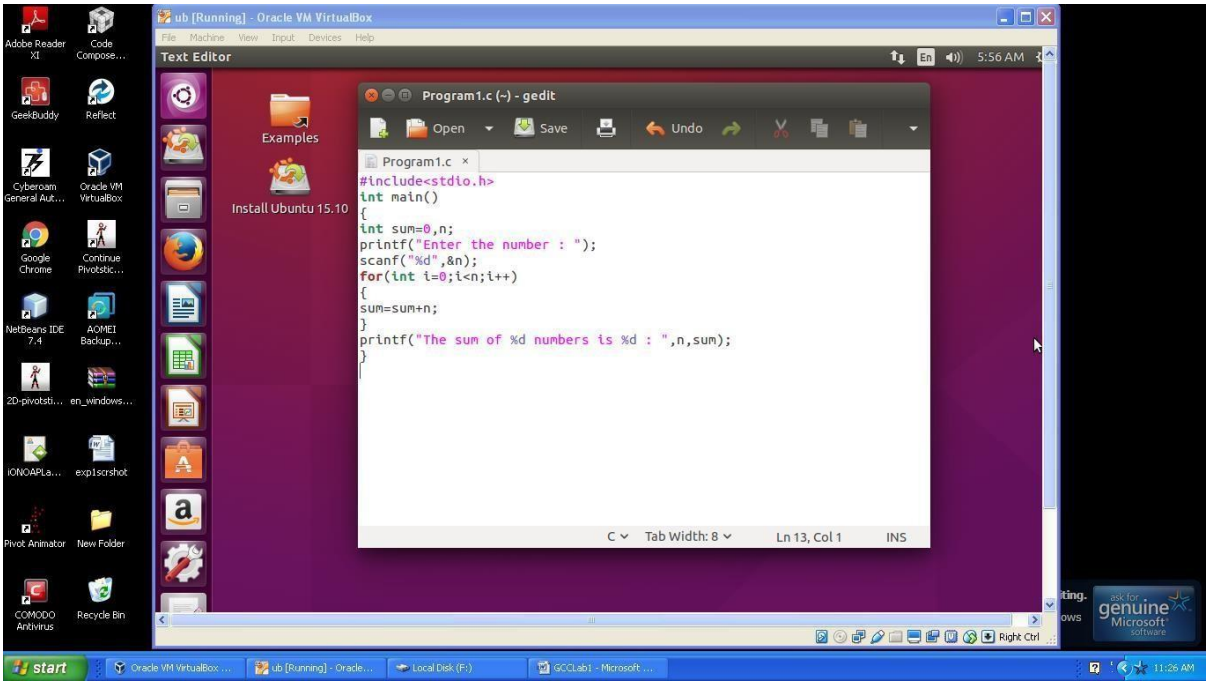
be opened by using the search option by typing gedit in it.

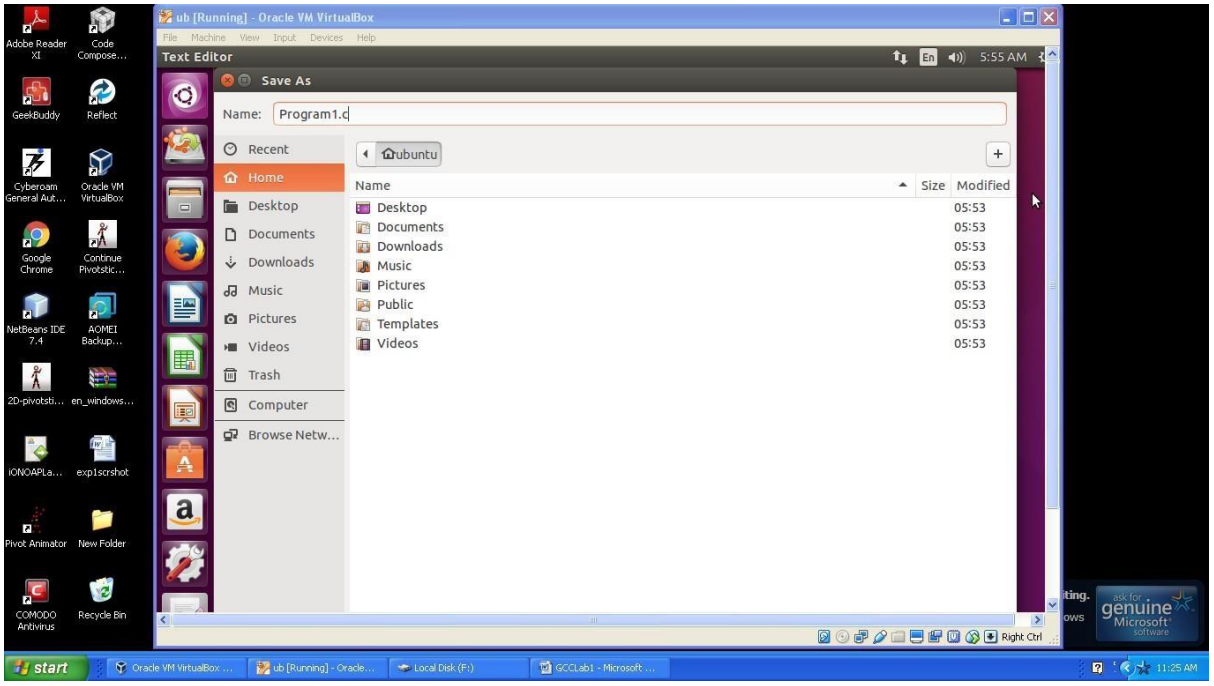
The text editor will be opened now.



3.Step

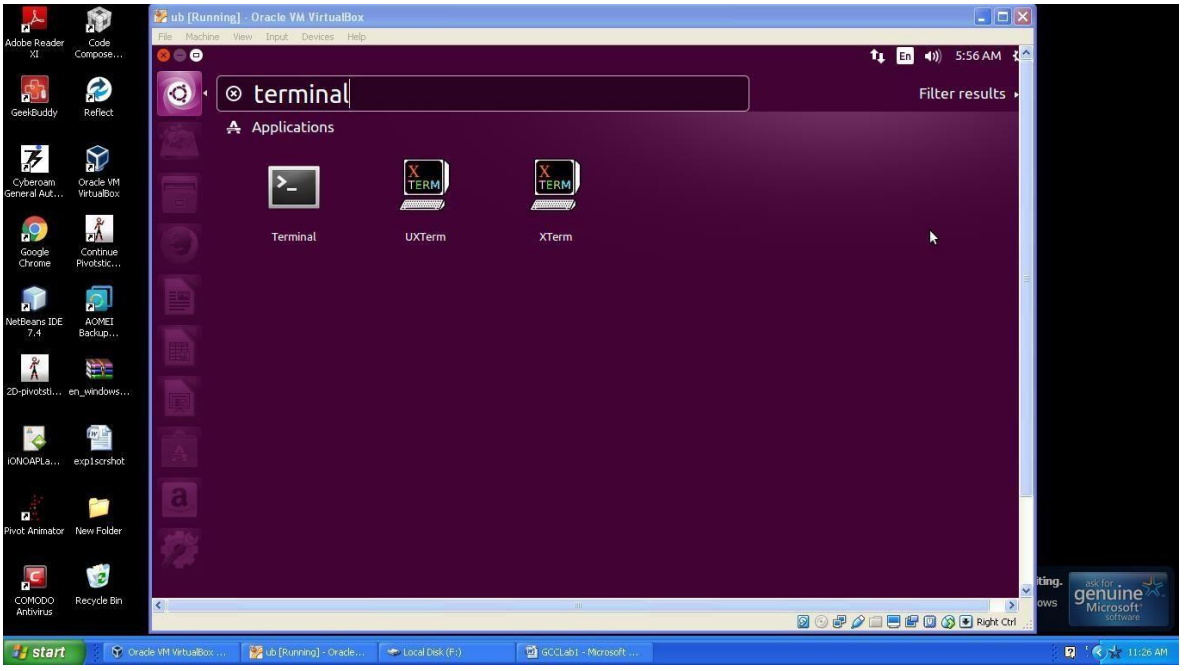
Type your desired C program in the text editor and save it as a C file using theExtension (.c) for C programs.





Step 4:

Type “terminal” in the search box to open the command window.

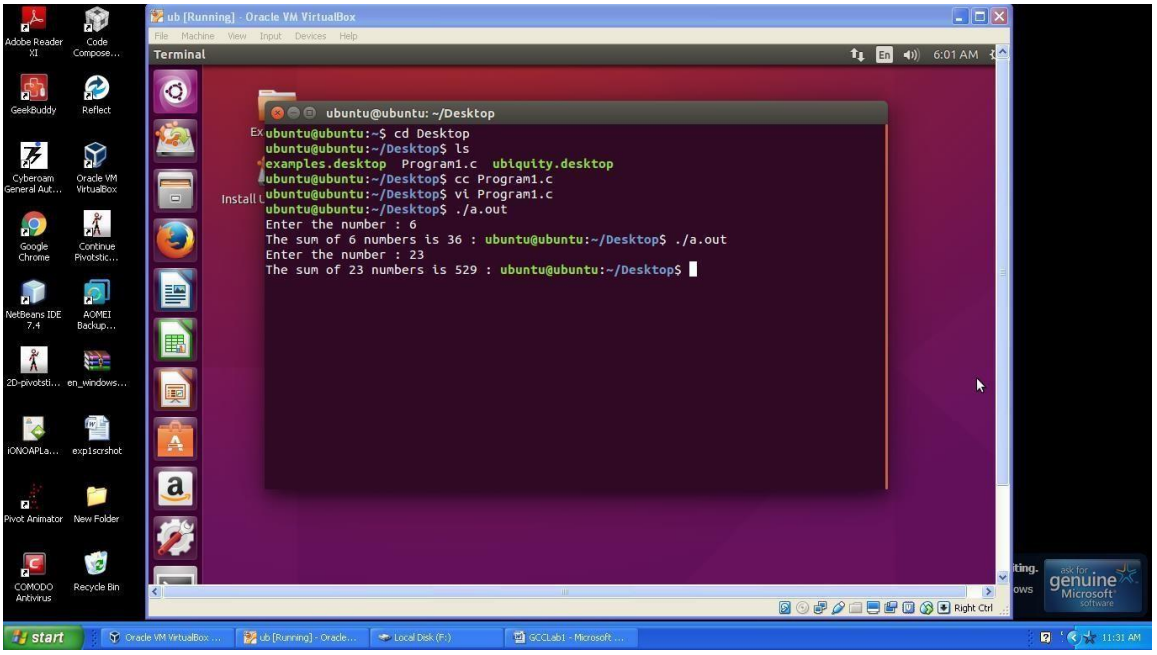


Step 5:

Type the necessary commands to compile and run the C program.(1).

cc_filename to compile the C program.

(2). ./a.out to display the output of the last compiled program.



RESULT:

Thus the procedure to use the C compiler in the virtual machine and execute asample program is implemented successfully.

EX.NO.:4 **VIRTUAL MACHINE MIGRATION**

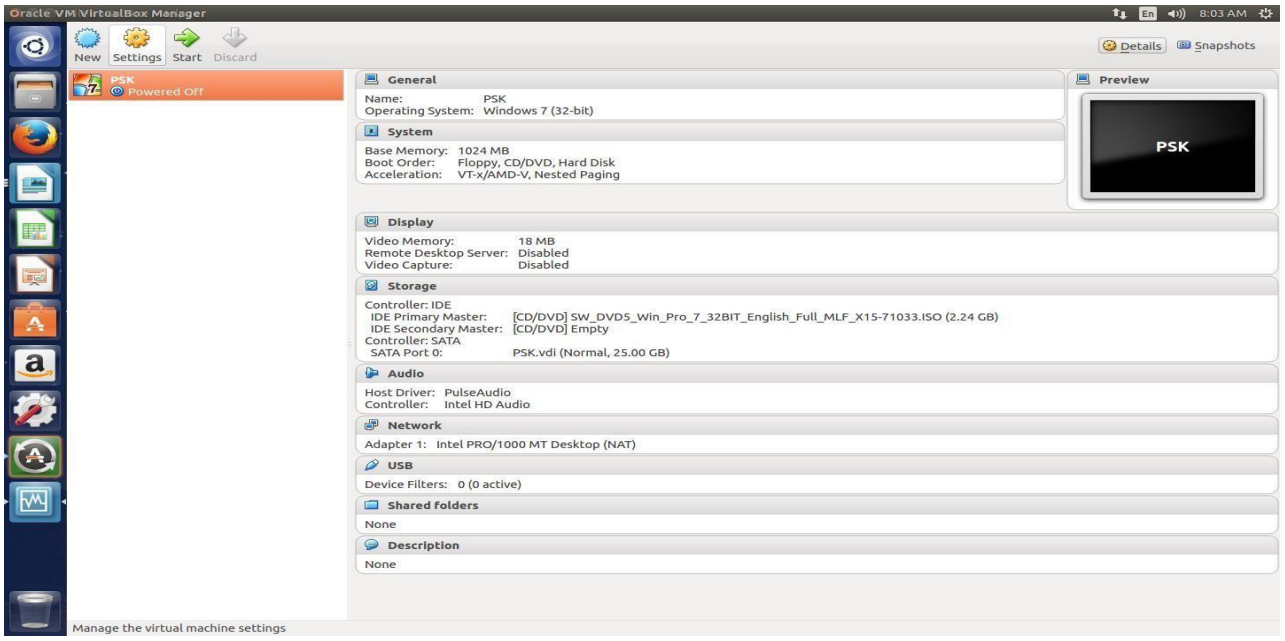
DATE:

AIM:

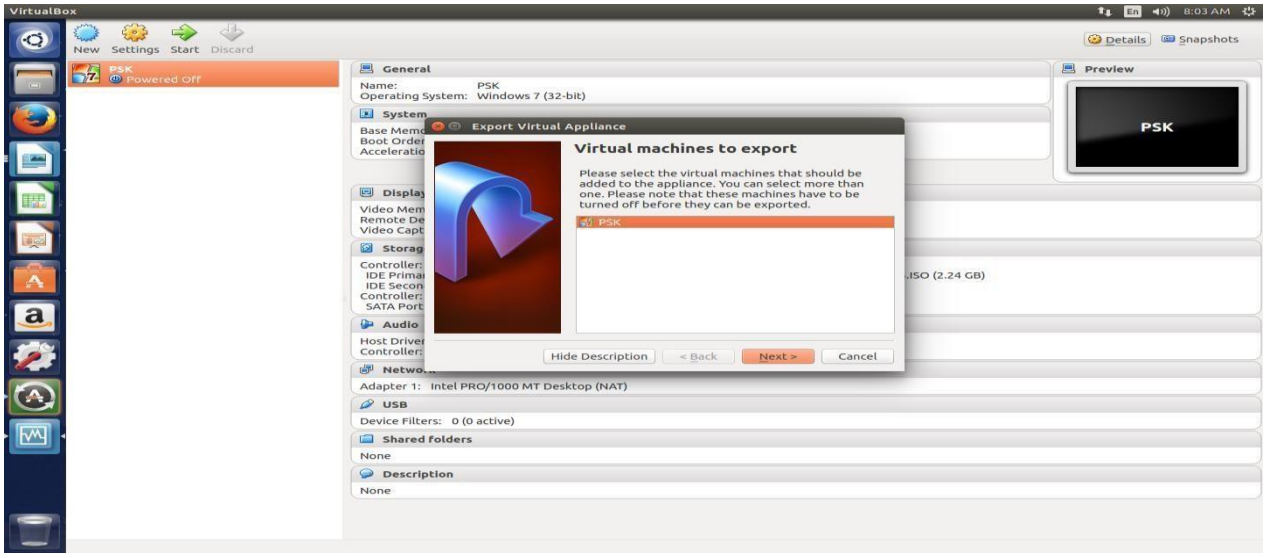
To show the Virtual Machine Migration based on the certain condition from one node to the other.

PROCEDURE:

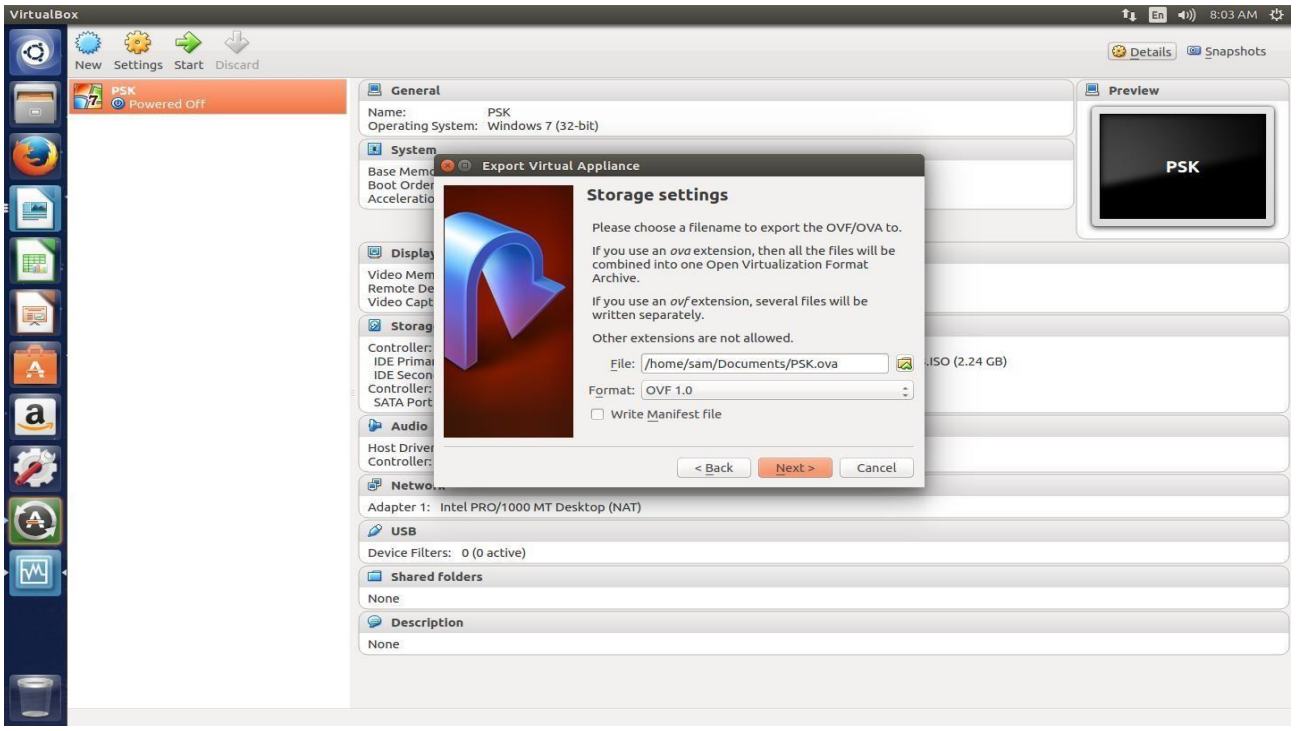
Step1: Select the VM and click File->Export Appliance



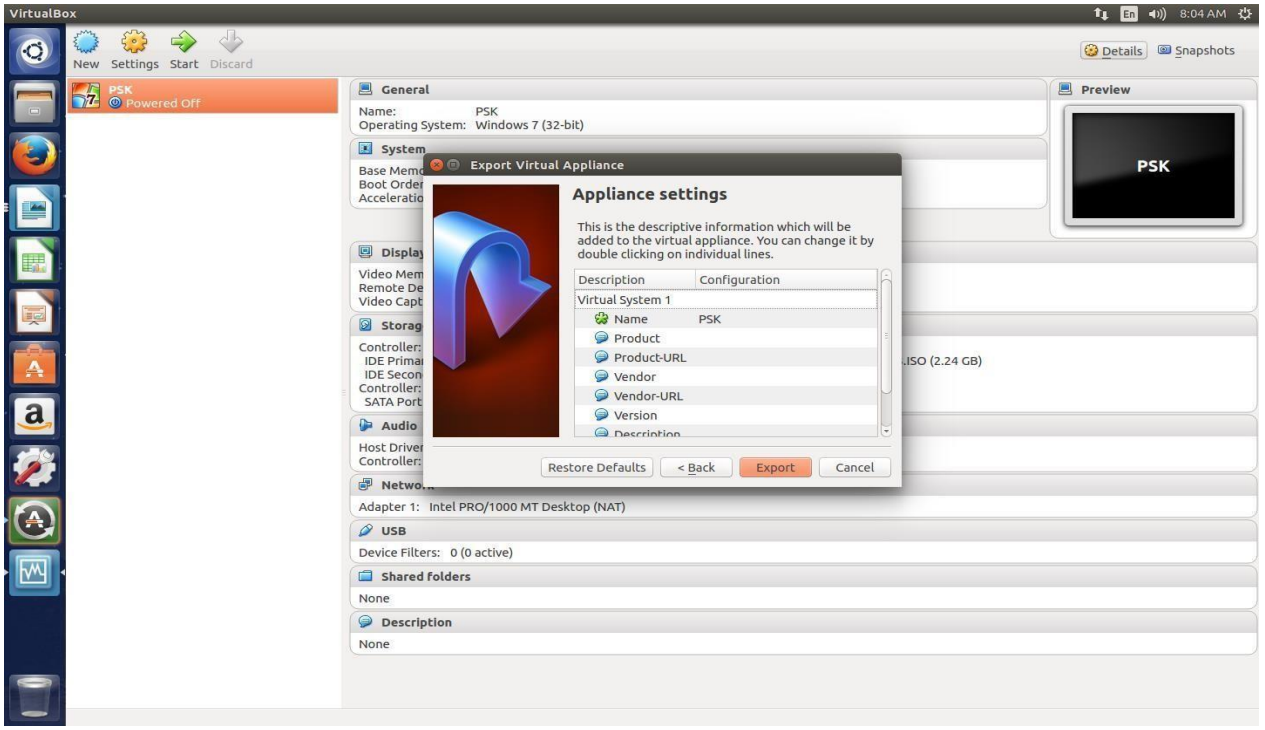
Step2: Select the VM to be exported and click NEXT.



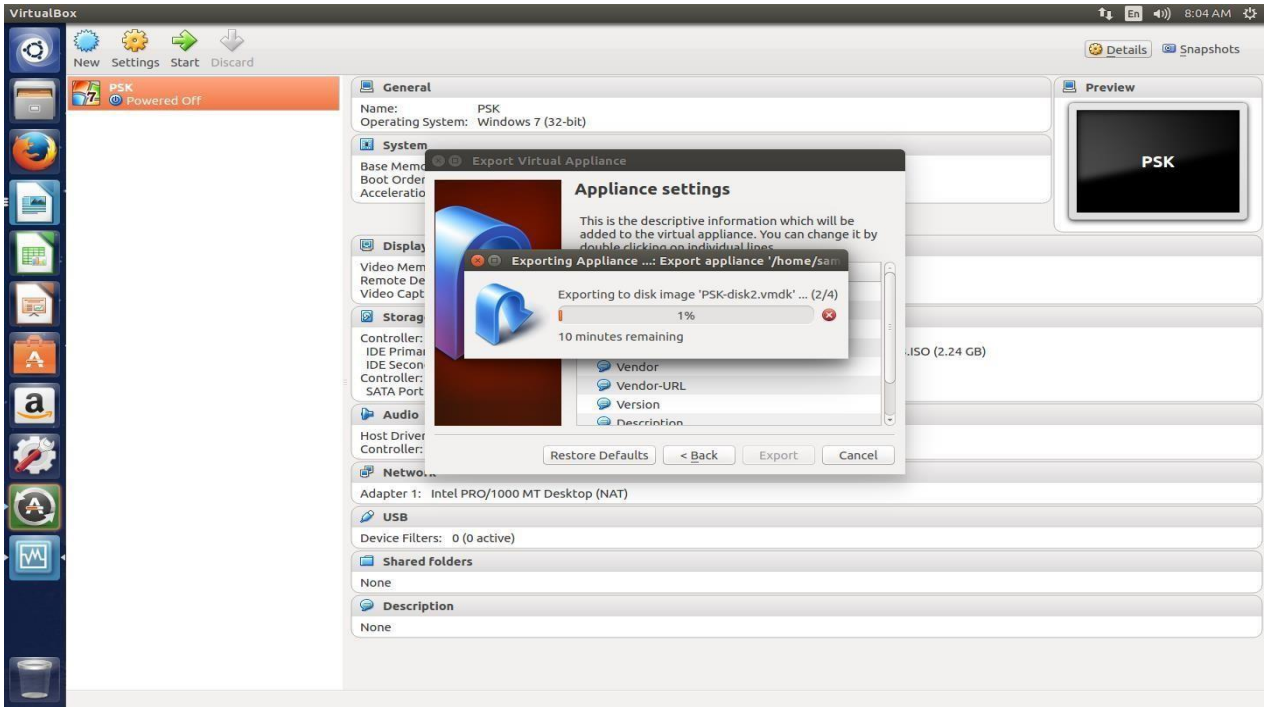
Step3: Note the file path and click “Next”



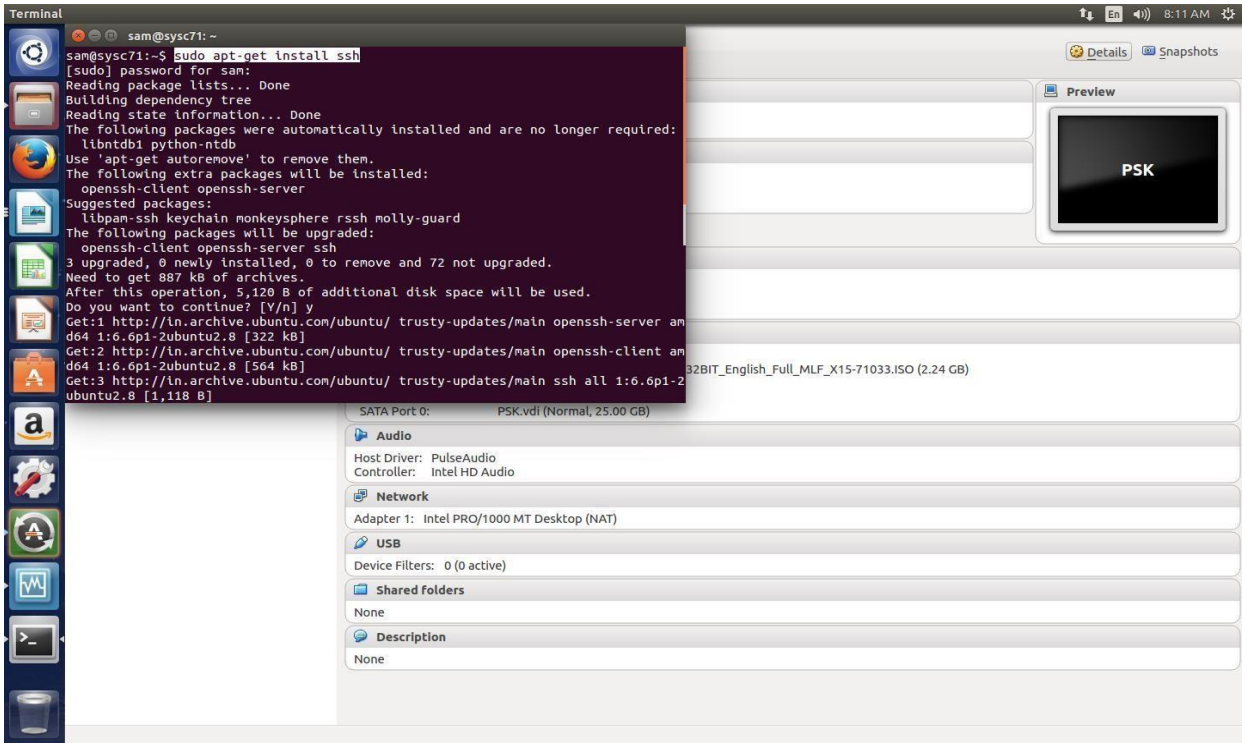
Step4: Click “Export”



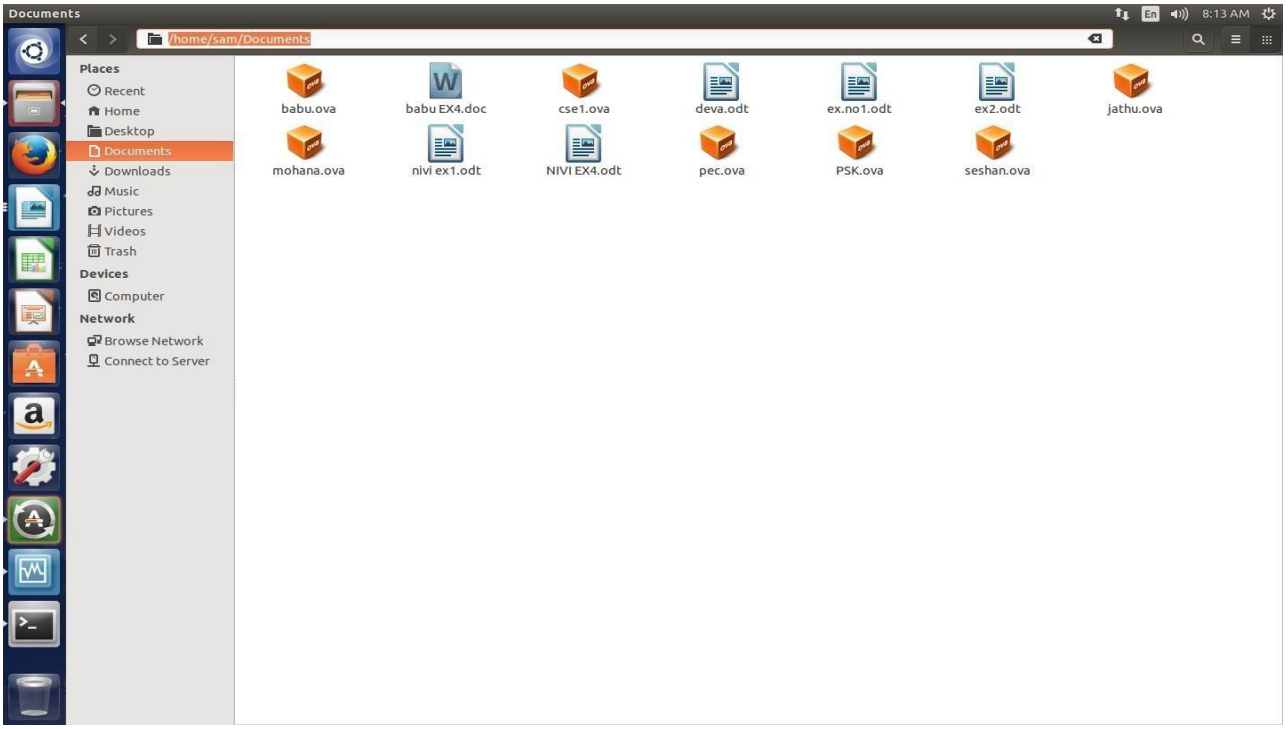
Step5: The Virtual machine is being exported.



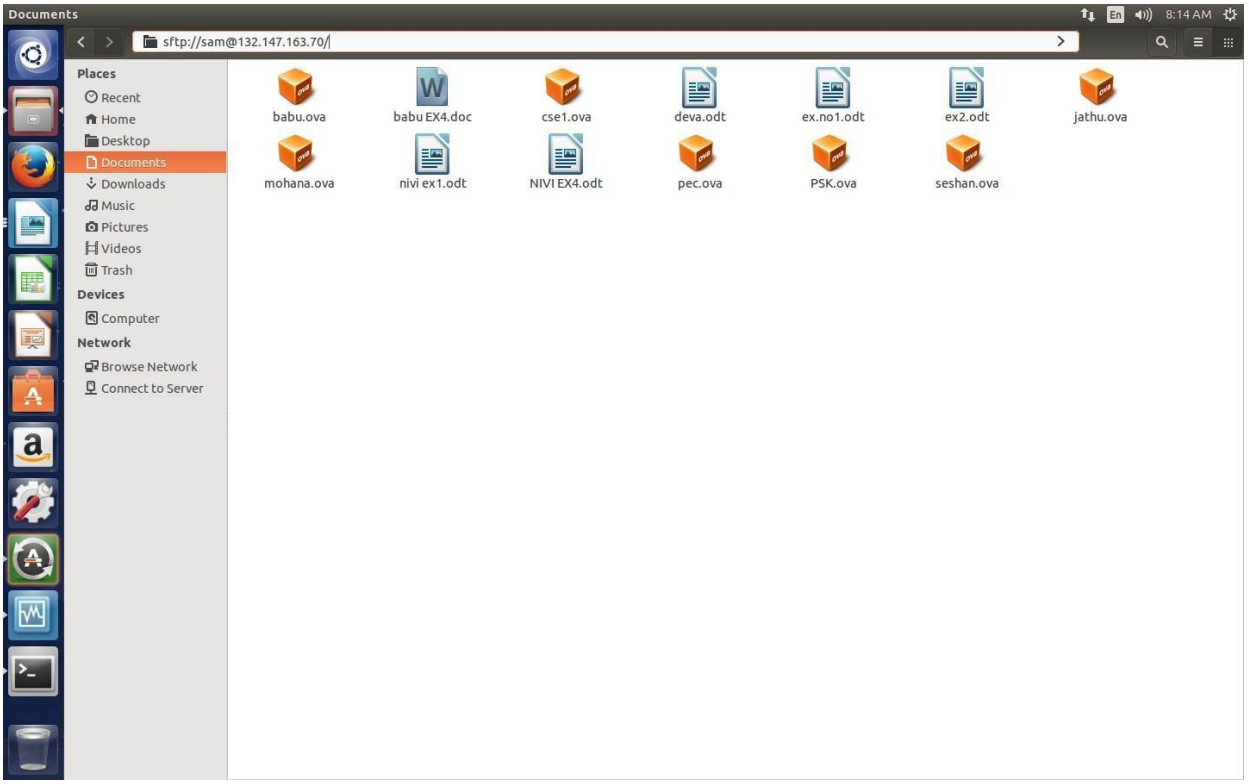
Step6: Install “ssh” to access the neighbour's VM.



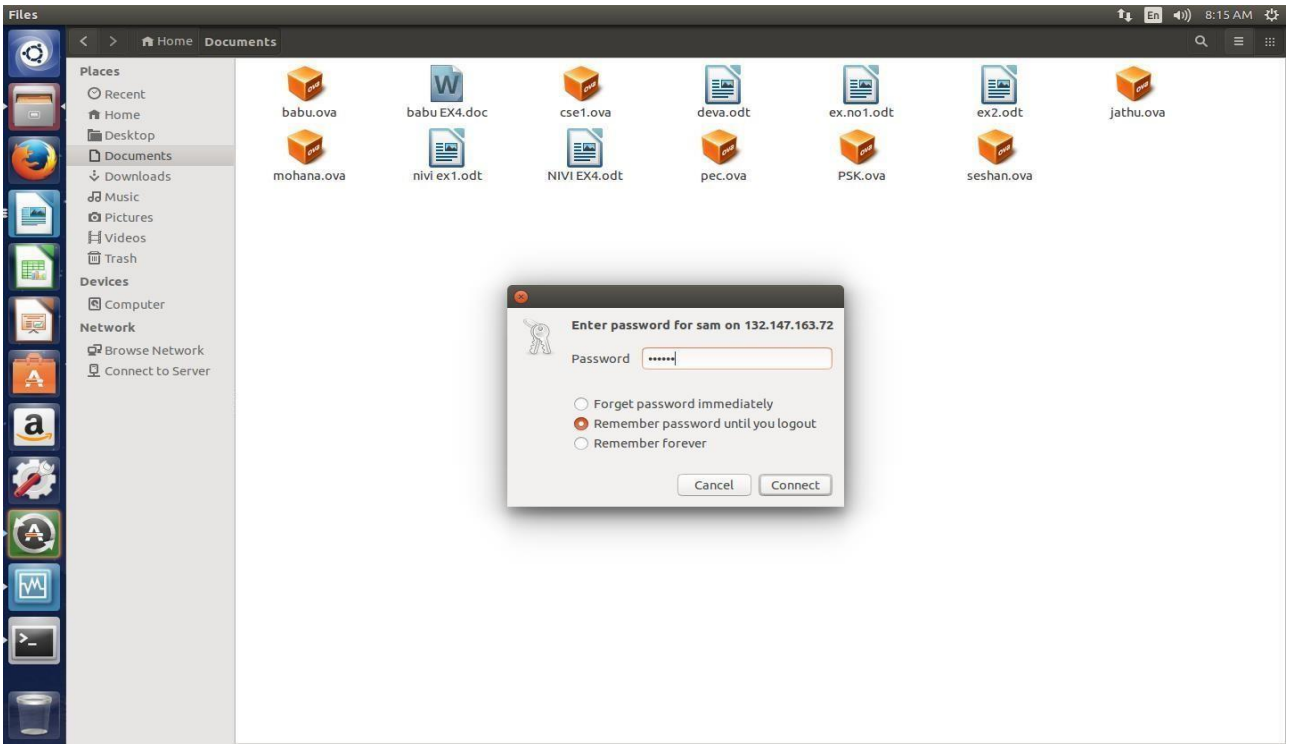
Step7: Go to File->Computer:/home/sam/Documents/



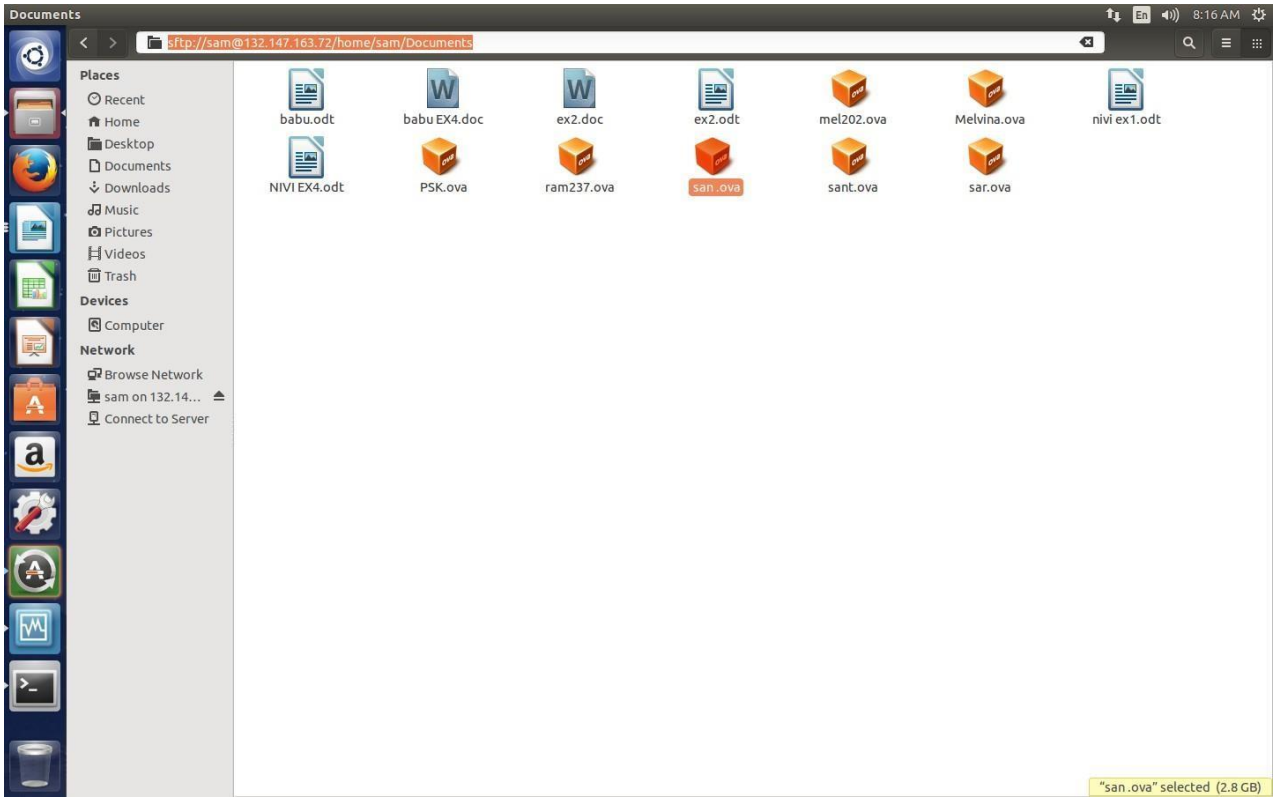
Step8: Type the neighbour's URL: sftp://sam@132.147.163._/



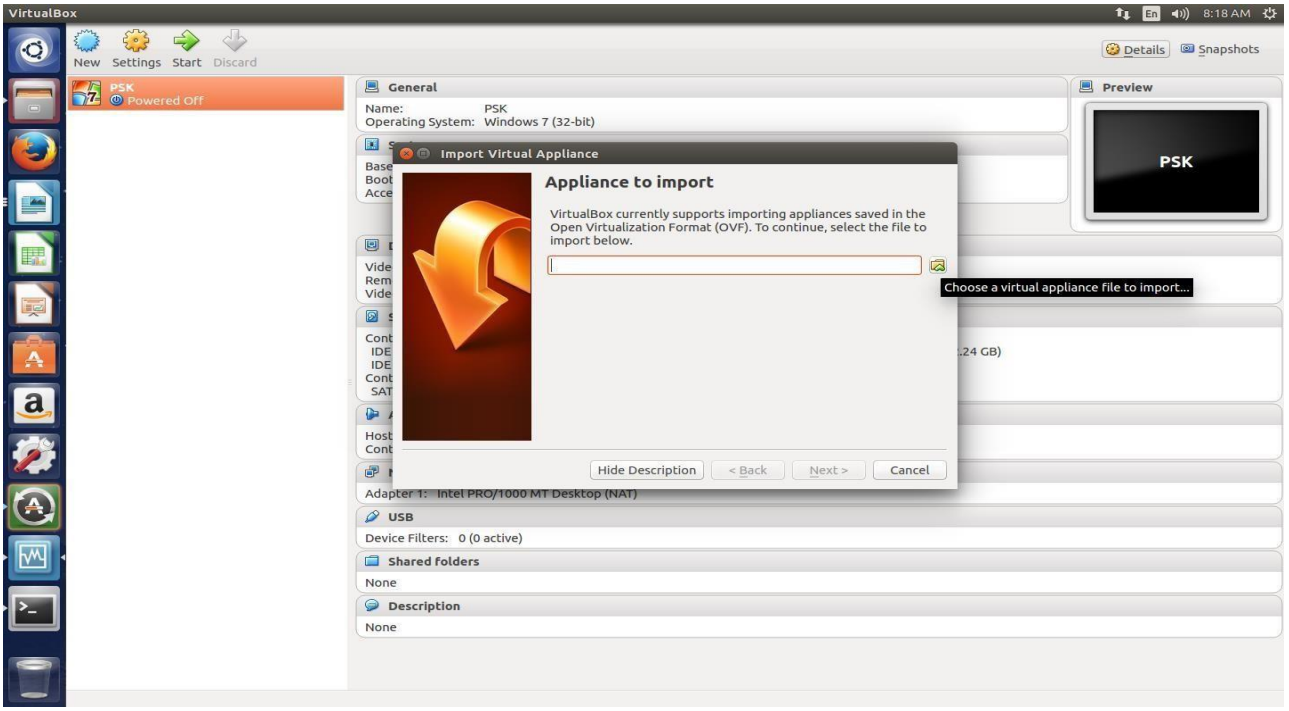
Step9: Give the password(sam123) and get connected.



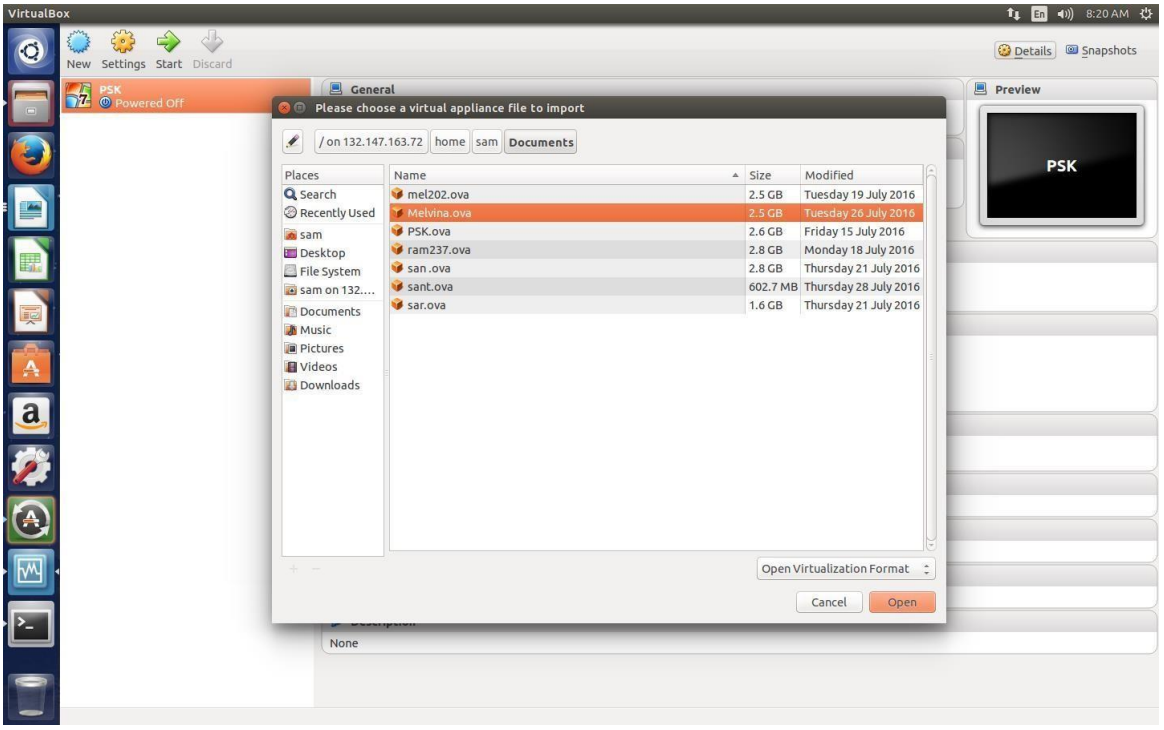
Step10: Select the VM and copy it in desktop.



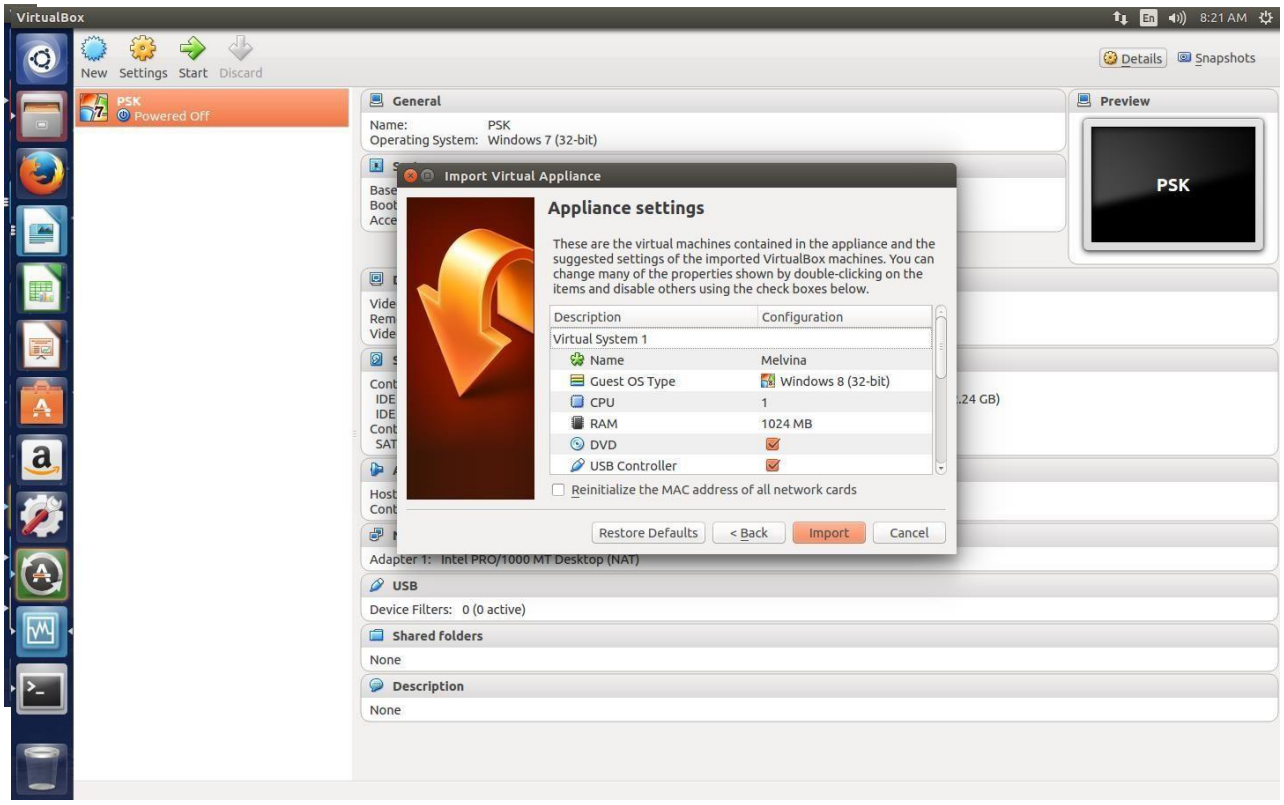
Step11: Open VirtualBox and select File->Import Appliance->Browse



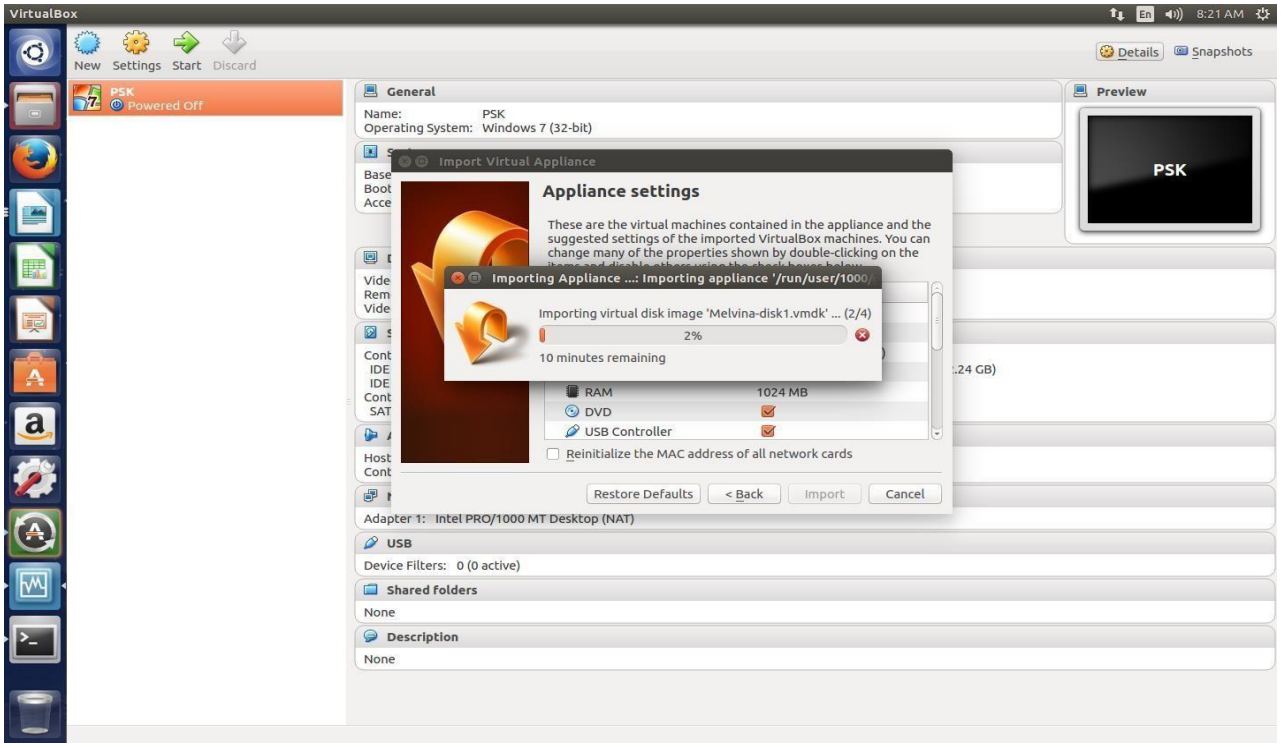
Step12: Select the VM to be imported and click “Open”.



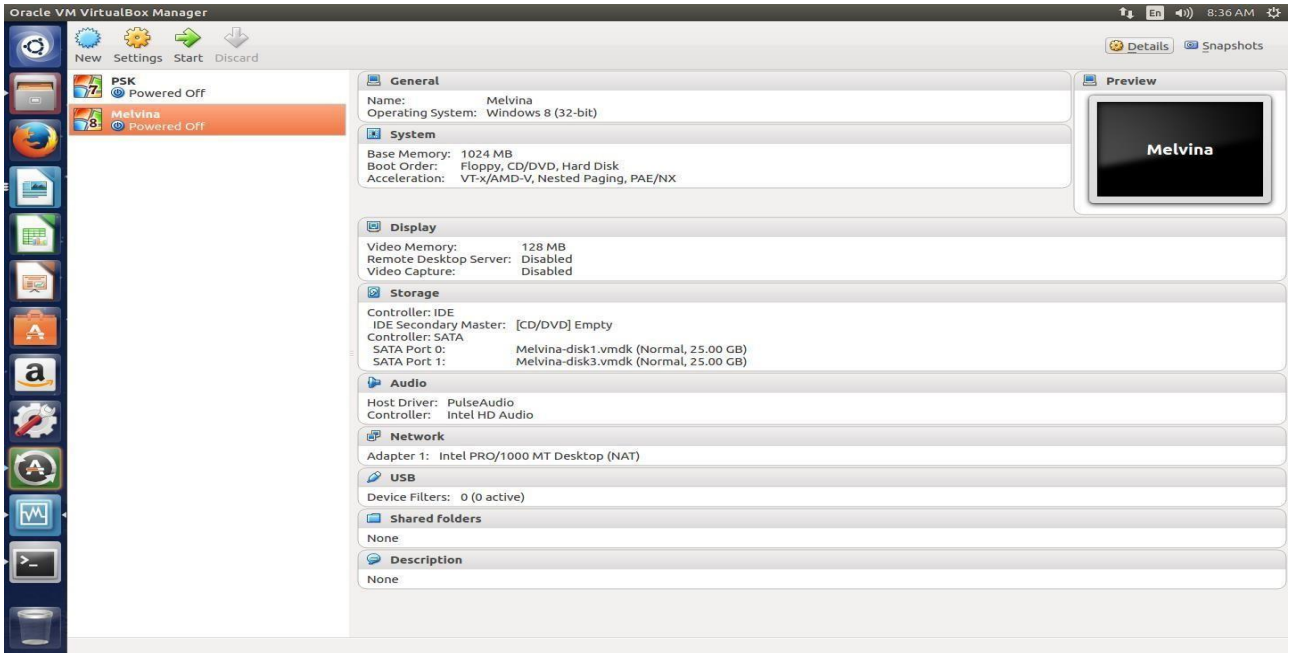
Step13: Click “Next” button.



Step15:TheVirtual Machine starts to get imported.



Step16: Thus the VM is imported.



RESULT:

Thus,Virtual machine migration has been implemented successfully.

EX:NO:5 **INSTALLING STORAGE CONTROLLER**

Date:

AIM:

To find procedure to install storage controller and interact with it.

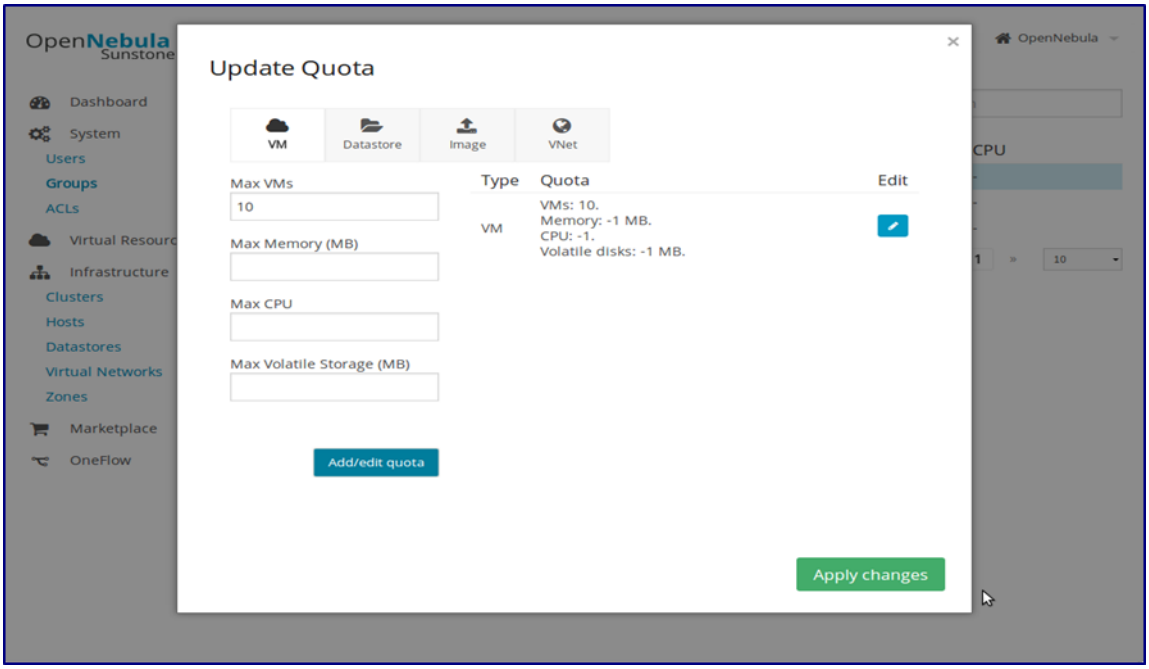
PROCEDURE:

Optionally, Set Quotas

The cloud administrator can set *usage quotas* for the vDC. In this case, we will put a limitof 10 VMs.

$ onegroup show web-dev GROUP
100 INFORMATION
ID 100
NAME : web-dev
GROUP TEMPLATE
GROUP_ADMINS="web-dev-admin"
GROUP_ADMIN_VIEWS="vdcadmin"
SUNSTONE_VIEWS="cloud"
USERS
ID
2
RESOURCE PROVIDERS
ZONE CLUSTER
0 100
RESOURCE USAGE & QUOTAS

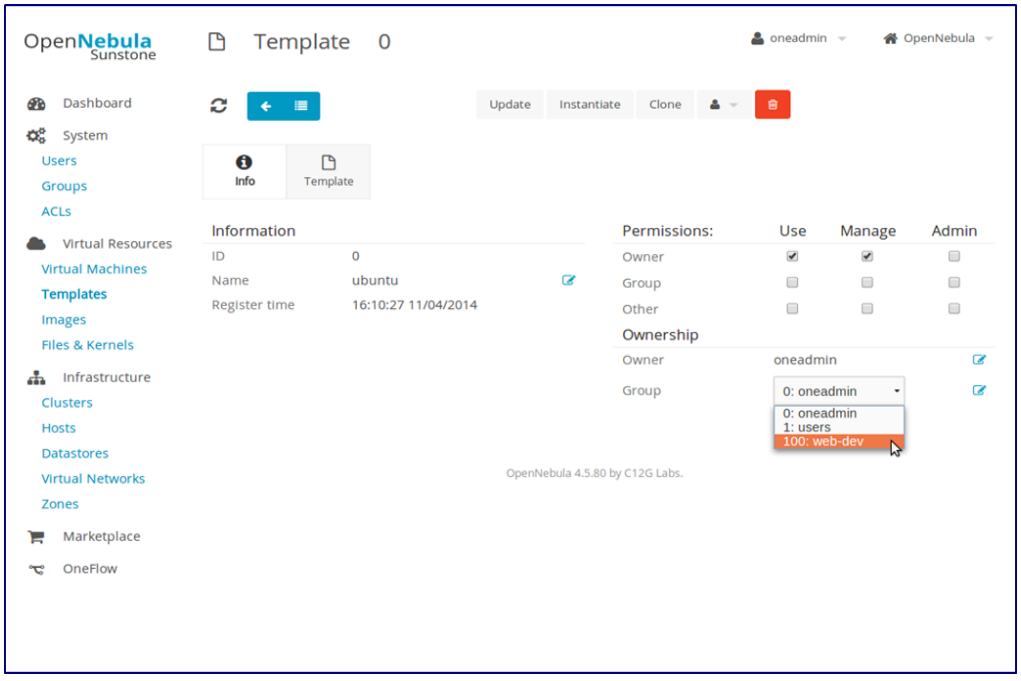
NUMBER OF VMS	MEMORY	CPU	VOLATILE_SIZE
0 / 10	0M / 0M	0.00 / 0.00	0M / 0M



Prepare Virtual Resources for the Users

At this point, the cloud administrator can also prepare working Templates andImages for the vDC users.

$ onetemplate chgrp ubuntu web-dev



RESULT:

Thus the procedure to install storage controller is executed and an interaction isdone with it successfully.

EX NO: 6 **HADOOP 2.6 INSTALLING ON UBUNTU 14.04**

DATE: **(SINGLE-NODE CLUSTER)**

AIM:

The aim of this experiment is to install hadoop in the system.

PROCEDURE:

- Open Ubundu Operating System .
- Open the terminal and follow the steps.

STEP:1

CS117@user:~$ cd ~

Update the source list CS117@user:~$

sudo apt-get update

The OpenJDK project is the default version of Java # that

is provided from a supported Ubuntu repository.

CS117@user:~$ sudo apt-get install default-jdk

CS117@user:~$ java -version

java version "1.7.0_65"

OpenJDK Runtime Environment (IcedTea 2.5.3) (7u71-2.5.3-0ubuntu0.14.04.1)

OpenJDK 64-Bit Server VM (build 24.65-b04, mixed mode)

STEP:2

Adding a dedicated Hadoop user

CS117@user:~$ sudo addgroup hadoop

Adding group `hadoop' (GID 1002) ...

Done.

CS117@user:~$ sudo adduser --ingroup hadoop hduser

Adding user `hduser' ...

Adding new user `hduser' (1001) with group `hadoop' ...

Creating home directory `/home/hduser' ...

Copying files from `/etc/skel' ...

Enter new UNIX password:

Retype new UNIX password:

passwd: password updated successfully

Changing the user information for hduser

Enter the new value, or press ENTER for the defaultFull

Name []:

Room Number []:

Work Phone []:

Home Phone []:

Other []:

Is the information correct? [Y/n] Y

STEP:3

Installing SSH

ssh has two main components:

1. **ssh** : The command we use to connect to remote machines - the client.
2. **sshd** : The daemon that is running on the server and allows clients to connect tothe server.

The **ssh** is pre-enabled on Linux, but in order to start **sshd** daemon, we need to install **ssh** first. Use this command to do that :

CS117@user:~$ sudo apt-get install ssh

This will install ssh on our machine. If we get something similar to the following, we canthink it is setup properly:

CS117@user:~$ which ssh
/usr/bin/ssh

CS117@user:~$ which sshd
/usr/sbin/sshd

Create and Setup SSH Certificates

CS117@user:~$ su hduser
Password:
CS117@user:~$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
Created directory '/home/hduser/.ssh'.
Your identification has been saved in /home/hduser/.ssh/id_rsa. Your
public key has been saved in /home/hduser/.ssh/id_rsa.pub.The key
fingerprint is: 50:6b:f3:fc:0f:32:bf:30:79:c2:41:71:26:cc:7d:e3
hduser@laptop The key's randomart image is:
+--[RSA 2048] ----+

```
|        .oo.o     |
|        . .o=. o  |
|       . + .  o . |
|        o =     E |
|        S +       |
|        . +     |
|         O +    |
|           O o  |
|           o..  |
+_____________+
```

hduser@laptop:/home/k$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys

check if ssh works:

hduser@laptop:/home/k$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established. ECDSA key fingerprint is e1:8b:a0:a5:75:ef:f4:b4:5e:a9:ed:be:64:be:5c:2f.Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 14.04.1 LTS (GNU/Linux 3.13.0-40-generic x86_64)
...

STEP:4

Install Hadoop

hduser@laptop:~$ wget http://mirrors.sonic.net/apache/hadoop/common/hadoop-2.6.0/hadoop-2.6.0.tar.gz

hduser@laptop:~$ tar xvzf hadoop-2.6.0.tar.gz

hduser@laptop:~/hadoop-2.6.0$ su k
Password:

CS117@user:/home/hduser$ sudo adduser hduser sudo[sudo]
password for k:
Adding user `hduser' to group `sudo' ...
Adding user hduser to group sudo Done.

CS117@user:/home/hduser$ sudo su hduser

hduser@laptop:~/hadoop-2.6.0$ sudo mv * /usr/local/hadoop hduser@laptop:~/hadoop-2.6.0$ sudo chown -R hduser:hadoop /usr/local/hadoop

STEP:5

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

~/.bashrc:

hduser@laptop update-alternatives --config java
There is only one alternative in link group java (providing /usr/bin/java): /usr/lib/jvm/java-7-openjdk-amd64/jre/bin/java
Nothing to configure.

hduser@laptop:~$ vi ~/.bashrc

#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64
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

hduser@laptop:~$ source ~/.bashrc

hduser@ubuntu-VirtualBox:~$ javac -version

javac 1.7.0_75

hduser@ubuntu-VirtualBox:~$ which javac
/usr/bin/javac

hduser@ubuntu-VirtualBox:~$ readlink -f /usr/bin/javac
/usr/lib/jvm/java-7-openjdk-amd64/bin/javac

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

We need to set **JAVA_HOME** by modifying **hadoop-env.sh** file.

hduser@laptop:~$ vi /usr/local/hadoop/etc/hadoop/hadoop-env.shexport

JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64

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

   hduser@laptop:~$ sudo mkdir -p /app/hadoop/tmp hduser@laptop:~$
   sudo chown hduser:hadoop /app/hadoop/tmp

hduser@laptop:~$ vi /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) namingthe 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**

hduser@laptop:~$ 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 runsat.  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**

hduser@laptop:~$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
hduser@laptop:~$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode hduser@laptop:~$
sudo chown -R hduser:hadoop /usr/local/hadoop_store

hduser@laptop:~$ vi /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>
```

STEP:6

Format the New Hadoop Filesystem

```
hduser@laptop:~$ hadoop namenode -format
DEPRECATED: Use of this script to execute hdfs command is deprecated.Instead use
the hdfs command for it.

15/04/18 14:43:03 INFO namenode.NameNode: STARTUP_MSG:
/************************************************
************STARTUP_MSG: Starting NameNode
STARTUP_MSG:  host = laptop/192.168.1.1
STARTUP_MSG: args = [-format]
STARTUP_MSG:  version = 2.6.0
STARTUP_MSG:  classpath = /usr/local/hadoop/etc/hadoop
...
STARTUP_MSG:  java = 1.7.0_65
************************************************************/
15/04/18 14:43:03 INFO namenode.NameNode: registered UNIX signal handlers for[TERM,
HUP, INT]
15/04/18 14:43:03 INFO namenode.NameNode: createNameNode [-format]
15/04/18 14:43:07 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor
your platform... using builtin-java classes where applicable
Formatting using clusterid: CID-e2f515ac-33da-45bc-8466-5b1100a2bf7f 15/04/18
14:43:09 INFO namenode.FSNamesystem: No KeyProvider found.15/04/18
14:43:09 INFO namenode.FSNamesystem: fsLock is fair:true 15/04/18 14:43:10
INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit=1000
15/04/18 14:43:10 INFO blockmanagement.DatanodeManager:
dfs.namenode.datanode.registration.ip-hostname-check=true
15/04/18 14:43:10 INFO blockmanagement.BlockManager:
dfs.namenode.startup.delay.block.deletion.sec is set to 000:00:00:00.000
15/04/18 14:43:10 INFO blockmanagement.BlockManager: The block deletion will startaround
2015 Apr 18 14:43:10
15/04/18 14:43:10 INFO util.GSet: Computing capacity for map BlocksMap
15/04/18 14:43:10 INFO util.GSet: VM type           = 64-bit
15/04/18 14:43:10 INFO util.GSet: 2.0% max memory 889 MB = 17.8 MB
15/04/18 14:43:10 INFO util.GSet: capacity          = 2^21 = 2097152 entries
15/04/18 14:43:10 INFO blockmanagement.BlockManager:
dfs.block.access.token.enable=false
15/04/18 14:43:10 INFO blockmanagement.BlockManager: defaultReplication         = 1
15/04/18 14:43:10 INFO blockmanagement.BlockManager: maxReplication             =
512
15/04/18 14:43:10 INFO blockmanagement.BlockManager: minReplication             = 1
15/04/18 14:43:10 INFO blockmanagement.BlockManager: maxReplicationStreams
= 2
```

15/04/18 14:43:10 INFO blockmanagement.BlockManager:
shouldCheckForEnoughRacks = false
15/04/18 14:43:10 INFO blockmanagement.BlockManager: replicationRecheckInterval =3000
15/04/18 14:43:10 INFO blockmanagement.BlockManager: encryptDataTransfer =
false
15/04/18 14:43:10 INFO blockmanagement.BlockManager: maxNumBlocksToLog
= 1000
15/04/18 14:43:10 INFO namenode.FSNamesystem: fsOwner = hduser
(auth:SIMPLE)
15/04/18 14:43:10 INFO namenode.FSNamesystem: supergroup = supergroup
15/04/18 14:43:10 INFO namenode.FSNamesystem: isPermissionEnabled = true 15/04/18
14:43:10 INFO namenode.FSNamesystem: HA Enabled: false
15/04/18 14:43:10 INFO namenode.FSNamesystem: Append Enabled: true
15/04/18 14:43:11 INFO util.GSet: Computing capacity for map INodeMap
15/04/18 14:43:11 INFO util.GSet: VM type = 64-bit
15/04/18 14:43:11 INFO util.GSet: 1.0% max memory 889 MB = 8.9 MB
15/04/18 14:43:11 INFO util.GSet: capacity = 2^20 = 1048576 entries
15/04/18 14:43:11 INFO namenode.NameNode: Caching file names occuring more than10 times
15/04/18 14:43:11 INFO util.GSet: Computing capacity for map cachedBlocks
15/04/18 14:43:11 INFO util.GSet: VM type = 64-bit
15/04/18 14:43:11 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
15/04/18 14:43:11 INFO util.GSet: capacity = 2^18 = 262144 entries
15/04/18 14:43:11 INFO namenode.FSNamesystem:
dfs.namenode.safemode.threshold-pct = 0.9990000128746033
15/04/18 14:43:11 INFO namenode.FSNamesystem:
dfs.namenode.safemode.min.datanodes = 0 15/04/18
14:43:11 INFO namenode.FSNamesystem:
dfs.namenode.safemode.extension = 30000
15/04/18 14:43:11 INFO namenode.FSNamesystem: Retry cache on namenode isenabled
15/04/18 14:43:11 INFO namenode.FSNamesystem: Retry cache will use 0.03 of totalheap and
retry cache entry expiry time is 600000 millis
15/04/18 14:43:11 INFO util.GSet: Computing capacity for map NameNodeRetryCache
15/04/18 14:43:11 INFO util.GSet: VM type = 64-bit
15/04/18 14:43:11 INFO util.GSet: 0.029999999329447746% max memory 889 MB =
273.1 KB
15/04/18 14:43:11 INFO util.GSet: capacity = 2^15 = 32768 entries
15/04/18 14:43:11 INFO namenode.NNConf: ACLs enabled? false 15/04/18
14:43:11 INFO namenode.NNConf: XAttrs enabled? true
15/04/18 14:43:11 INFO namenode.NNConf: Maximum size of an xattr: 16384
15/04/18 14:43:12 INFO namenode.FSImage: Allocated new BlockPoolId: BP-
130729900-192.168.1.1-1429393391595
15/04/18 14:43:12 INFO common.Storage: Storage directory
/usr/local/hadoop_store/hdfs/namenode has been successfully formatted.

15/04/18 14:43:12 INFO namenode.NNStorageRetentionManager: Going to retain 1images with txid >= 0
15/04/18 14:43:12 INFO util.ExitUtil: Exiting with status 0
15/04/18 14:43:12 INFO namenode.NameNode: SHUTDOWN_MSG:
/**
SHUTDOWN_MSG: Shutting down NameNode at laptop/192.168.1.1
**/

STEP:7

Starting Hadoop

@laptop:~$ cd /usr/local/hadoop/sbin

CS117@user:/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

CS117@user:/usr/local/hadoop/sbin$ sudo su hduser

hduser@laptop:/usr/local/hadoop/sbin$ start-all.sh
hduser@laptop:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
15/04/18 16:43:13 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-namenode-laptop.out
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-laptop.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hduser-secondarynamenode-laptop.out
15/04/18 16:43:58 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hduser-resourcemanager-laptop.out

localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduser-nodemanager-laptop.out

STEP:8

hduser@laptop:/usr/local/hadoop/sbin$ jps
9026 NodeManager
7348 NameNode
9766 Jps
8887 ResourceManager
7507 DataNode
Stopping Hadoop

$ pwd

/usr/local/hadoop/sbin

$ ls

distribute-exclude.sh httpfs.sh start-all.sh start-yarn.cmd stop-dfs.cmd yarn-daemon.sh

hadoop-daemon.sh mr-jobhistory-daemon.sh start-balancer.sh start-yarn.sh stop-dfs.sh yarn-daemons.sh

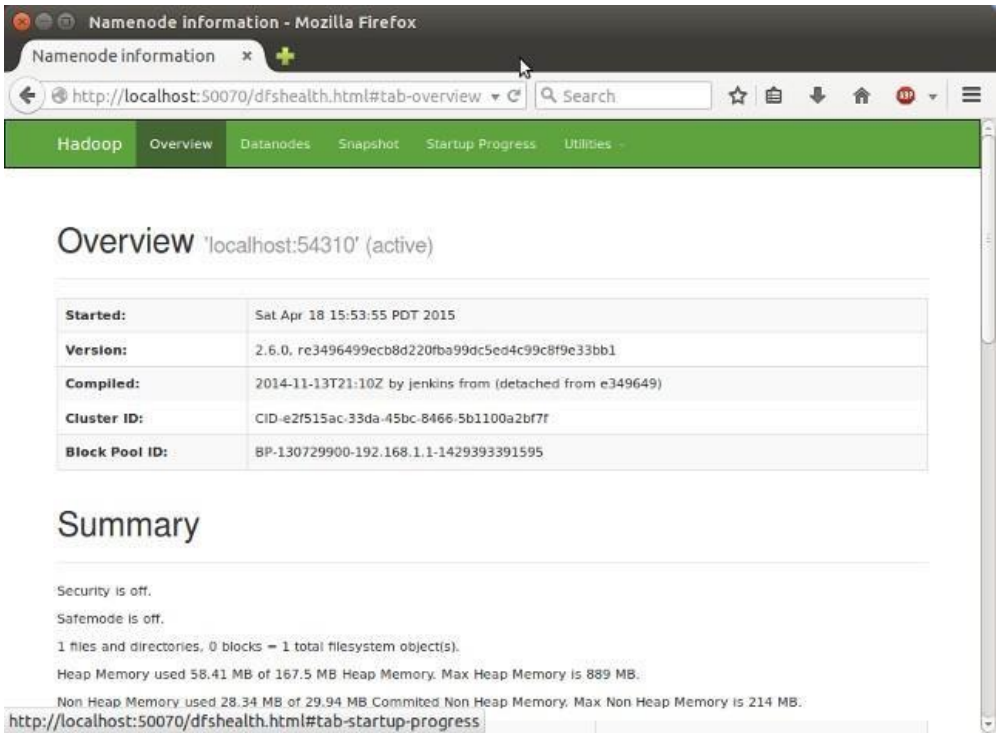
hadoop-daemons.sh refresh-namenodes.sh start-dfs.cmd stop-all.cmd stop-secure-dns.sh

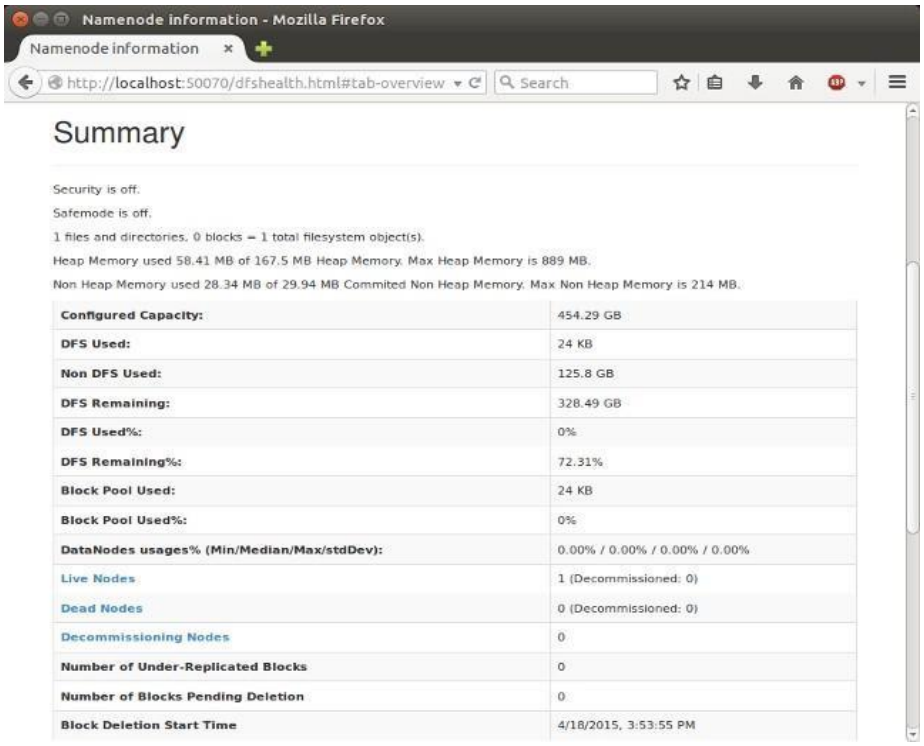
hdfs-config.cmd slaves.sh start-dfs.sh stop-all.sh stop-yarn.cmd

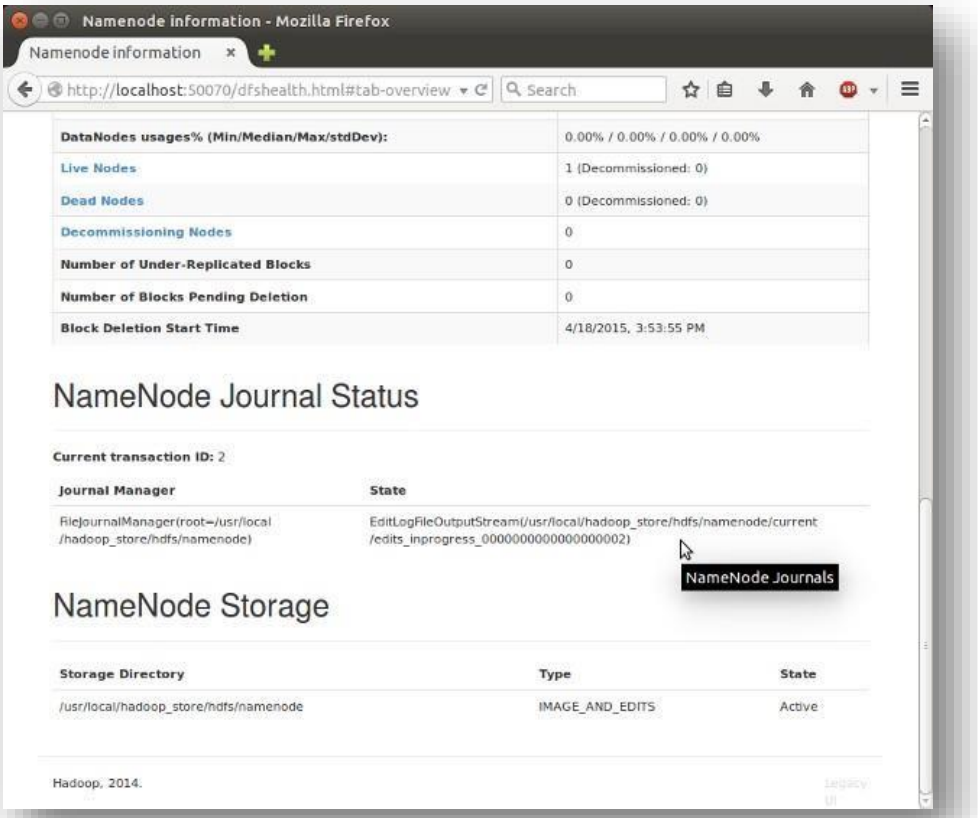
hdfs-config.sh start-all.cmd start-secure-dns.sh stop-balancer.sh stop-yarn.sh

OUTPUT:

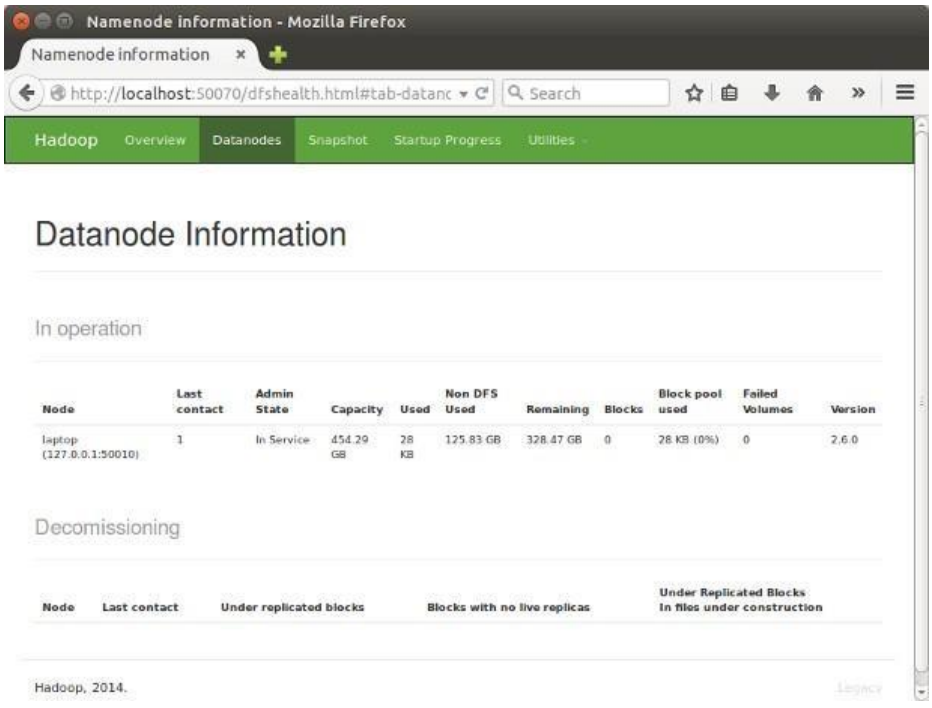
hduser@laptop:/usr/local/hadoop/sbin$ start-all.sh






Namenode information - Mozilla Firefox
Namenode information
http://localhost:50070/dfshealth.html#tab-overview
Search
DataNodes usages% (Min/Median/Max/stdDev): 0.00% / 0.00% / 0.00% / 0.00%
Live Nodes 1 (Decommissioned: 0)
Dead Nodes 0 (Decommissioned: 0)
Decommissioning Nodes 0
Number of Under-Replicated Blocks 0
Number of Blocks Pending Deletion 0
Block Deletion Start Time 4/18/2015, 3:53:55 PM
NameNode Journal Status
Current transaction ID: 2
Journal Manager
State
FileJournalManager(root=/usr/local/hadoop_store/hdfs/namenode)
EditLogFileOutputStream(/usr/local/hadoop_store/hdfs/namenode/current/edits_inprogress_0000000000000000002)
NameNode Journals
NameNode Storage
Storage Directory
Type
State
/usr/local/hadoop_store/hdfs/namenode
IMAGE_AND_EDITS
Active
Hadoop, 2014.

DataNode


Namenode information - Mozilla Firefox
Namenode information
http://localhost:50070/dfshealth.html#tab-datanc
Search
Hadoop
Overview
Datanodes
Snapshot
Startup Progress
Utilities
Datanode Information
In operation
Node
Last contact
Admin State
Capacity
Used
Non DFS Used
Remaining
Blocks
Block pool used
Failed Volumes
Version
laptop (127.0.0.1:50010)
1
In Service
454.29 GB
28 KB
125.83 GB
328.47 GB
0
28 KB (0%)
0
2.6.0
Decomissioning
Node
Last contact
Under replicated blocks
Blocks with no live replicas
Under Replicated Blocks In files under construction
Hadoop, 2014.

Result:

Thus the installation of Hadoop 2.6 on Ubuntu 14.04 which is a single nodeclusted is executed successfully.

Ex.No:7 **MOUNT THE ONE NODE HADOOP CLUSTER USING FUSE.**

Date:

Aim:

To Mount the one node Hadoop cluster using FUSE

Procedure:

However, one can leverage FUSE to write a userland application that exposes HDFS via a traditional filesystem interface. fuse-dfs is one such FUSE-based application which allows you to mount HDFS as if it were a traditional Linux filesystem.If you would like to **mount HDFS on Linux**, you can install fuse-dfs, along with FUSEas follows:

wget http://archive.cloudera.com/one-click-install/maverick/cdh3-repository_1.0_all.deb sudo dpkg -i cdh3-repository_1.0_all.deb

sudo apt-get update

sudo apt-get install hadoop-0.20-fuse

Once fuse-dfs is installed, go ahead and mount HDFS using FUSE as follows. sudo hadoop-fuse-dfs dfs://<name_node_hostname>:<namenode_port> <mount_point>

Once HDFS has been mounted at <mount_point>, you can use most of the traditional filesystem operations (e.g., cp, rm, cat, mv, mkdir, rmdir, more, scp). However, random write operations such as rsync, and permission related operations such as chmod, chown are not supported in FUSE-mounted HDFS.

Result:

Thus the one node Hadoop cluster is mounted successfully using FUSE.

Ex.No.8 **API'S OF HADOOP TO INTERACT WITH IT – TO DISPLAY FILE**

Date: **CONTENT OF A FILE EXIST IN HDFS**

Aim:

To write a program to use the API's of Hadoop to interact with it to display filecontent of a file existing in hdfs.

Procedure:

/home/hduser/HadoopFScat.java:

```
import java.io.InputStream;
import java.net.URI;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem; import
org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IOUtils;
public class HadoopFScat {
        public static void main(String[] args) throws Exception {
                String uri = args[0];
                Configuration conf = new Configuration();
                FileSystem fileSystem = FileSystem.get(URI.create(uri), conf);
                InputStream inputStream = null;
                try{    inputStream = fileSystem.open(new Path(uri));
                        IOUtils.copyBytes(inputStream, System.out, 4096, false);
                }finally{
                        IOUtils.closeStream(inputStream); }
        } }
```

Download the jar file:

Download **Hadoop-core-1.2.1.jar,** which is used to compile and execute the MapReduce program. Visit the following link

http://mvnrepository.com/artifact/org.apache.hadoop/hadoop-core/1.2.1 to download thejar. Let us assume the downloaded folder is /home/hduser/**.**

Creating a direfctory to collect class files:

hduser@nspublin:/usr/local/hadoop/sbin$ mkdir /home/hduser/fscat

Compiling the java file - HadoopFScat.java:

hduser@nspublin:/usr/local/hadoop/sbin$ sudo /usr/lib/jvm/java-8-oracle/bin/javac -classpath /home/hduser/hadoop-core-1.2.1.jar -d /home/hduser/fscat /home/hduser/HadoopFScat.java

hduser@nspublin:/usr/local/hadoop/sbin$ ls /home/hduser/fscat

HadoopFScat.class

Creating jar file for HadoopFScat.java:

hduser@nspublin:/usr/local/hadoop/sbin$ jar -cvf /home/hduser/fscat.jar -C /home/hduser/fscat/ .

added manifest

adding: HadoopFScat.class(in = 1224) (out= 667)(deflated 45%)

OUTPUT:

Executing jar file for HadoopFScat.java:

hduser@nspublin:/usr/local/hadoop/sbin$ hadoop jar /home/hduser/fscat.jar HadoopFScat /user/input/file.txt

16/06/08 15:29:03 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

Alzheimer's virtual reality app simulates dementia2

June 2016 Last updated at 19:13 BST

A virtual reality app has been launched to provide a sense of what it is like to live withdifferent forms of dementia.

A Walk Through Dementia was created by the charity Alzheimer's Research UK.It has been welcomed by other experts in the field.

We will increasingly be asked for help by people with dementia, and having had some insight into what may be happening for them will improve how we can help, said Tula Brannelly from the University of Southampton.

A woman living with the condition and her husband told the Today programme why they supported the Android app's creation.

Visitors to St Pancras International station in London can try out the app until 1700 on Saturday 4 June.

RESULT:

Thus a program to use the API's of Hadoop to interact with it to display file content of a file existing in hdfs is created and executed successfully.

EX:NO:9 **WORD COUNT PROGRAM**

Date:

AIM:

To write a word count program to demonstrate the use of Map and Reduce task.

PROCEDURE:

Step 1:

cs1-17@cs117-HP-Pro-3330-MT:~$ sudo su user

[sudo] password for cs1-17:

user@cs117-HP-Pro-3330-MT:/home/cs1-17$ cd\

user@cs117-HP-Pro-3330-MT:~$ start-all.sh

This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh

16/09/20 10:09:52 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

Starting namenodes on [localhost]

localhost: starting namenode, logging to /usr/local/hadoop1/logs/hadoop-user-namenode-cs117-HP-Pro-3330-MT.out

localhost: starting datanode, logging to /usr/local/hadoop1/logs/hadoop-user-datanode-cs117-HP-Pro-3330-MT.out

Starting secondary namenodes [0.0.0.0]

0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop1/logs/hadoop-user-secondarynamenode-cs117-HP-Pro-3330-MT.out

16/09/20 10:10:08 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

starting yarn daemons

starting resourcemanager, logging to /usr/local/hadoop1/logs/yarn-user-resourcemanager-cs117-HP-Pro-3330-MT.out

localhost: starting nodemanager, logging to /usr/local/hadoop1/logs/yarn-user-nodemanager-cs117-HP-Pro-3330-MT.out

user@cs117-HP-Pro-3330-MT:~$ jps

9551 NodeManager

8924 NameNode

9857 Jps

9076 DataNode

9265 SecondaryNameNode

9420 ResourceManager

Step 2:

create a directory named ip1 on the desktop. in the ip1 directory create a two.txt file forwordcount purpose. create a directory named op1 on the desktop.

user@cs117-HP-Pro-3330-MT:~$ cd /usr/local/hadoop1

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ bin/hdfs dfs -mkdir /user2

16/09/20 10:46:01 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ bin/hdfs dfs -put '/home/cs1-17/Desktop/ip1' /user2

16/09/20 10:48:42 WARN util.NativeCobin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /user1/inputdata output1 deLoader: Unable to load native-hadoop library for yourplatform... using builtin-java classes where applicable

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ bin/hdfs dfs -put '/home/cs1-17/Desktop/op1' /user2

16/09/20 11:02:01 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

Step 3:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ bin/hadoop jar

share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /user2/ip1op1

16/09/20 11:02:12 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

16/09/20 11:02:12 INFO Configuration.deprecation: session.id is deprecated. Instead,use dfs.metrics.session-id

16/09/20 11:02:12 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=

16/09/20 11:02:12 INFO input.FileInputFormat: Total input paths to process : 2

16/09/20 11:02:12 INFO mapreduce.JobSubmitter: number of splits:2

16/09/20 11:02:13 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1146489696_0001

16/09/20 11:02:13 INFO mapreduce.Job: The url to track the job: http://localhost:8080/16/09/20

11:02:13 INFO mapreduce.Job: Running job: job_local1146489696_0001 16/09/20 11:02:13

INFO mapred.LocalJobRunner: OutputCommitter set in config null

16/09/20 11:02:13 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Waiting for map tasks

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Starting task: attempt_local1146489696_0001_m_000000_0

16/09/20 11:02:13 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

16/09/20 11:02:13 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/user2/ip1/two.txt:0+42

16/09/20 11:02:13 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)

16/09/20 11:02:13 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 10016/09/20

11:02:13 INFO mapred.MapTask: soft limit at 83886080

16/09/20 11:02:13 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600 16/09/20

11:02:13 INFO mapred.MapTask: kvstart = 26214396; length = 6553600

16/09/20 11:02:13 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer

16/09/20 11:02:13 INFO mapred.LocalJobRunner:

16/09/20 11:02:13 INFO mapred.MapTask: Starting flush of map output16/09/20

11:02:13 INFO mapred.MapTask: Spilling map output

16/09/20 11:02:13 INFO mapred.MapTask: bufstart = 0; bufend = 69; bufvoid =104857600

16/09/20 11:02:13 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214372(104857488); length = 25/6553600

16/09/20 11:02:13 INFO mapred.MapTask: Finished spill 0

16/09/20 11:02:13 INFO mapred.Task: Task:attempt_local1146489696_0001_m_000000_0 is done. And is in the process of committing

16/09/20 11:02:13 INFO mapred.LocalJobRunner: map

16/09/20 11:02:13 INFO mapred.Task: Task 'attempt_local1146489696_0001_m_000000_0' done.

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Finishing task: attempt_local1146489696_0001_m_000000_0

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Starting task: attempt_local1146489696_0001_m_000001_0

16/09/20 11:02:13 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

16/09/20 11:02:13 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/user2/ip1/one.txt~:0+0

16/09/20 11:02:13 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)

16/09/20 11:02:13 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 10016/09/20

11:02:13 INFO mapred.MapTask: soft limit at 83886080

16/09/20 11:02:13 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600 16/09/20

11:02:13 INFO mapred.MapTask: kvstart = 26214396; length = 6553600

16/09/20 11:02:13 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer

16/09/20 11:02:13 INFO mapred.LocalJobRunner:

16/09/20 11:02:13 INFO mapred.MapTask: Starting flush of map output16/09/20

11:02:13 INFO mapred.Task: Task:attempt_local1146489696_0001_m_000001_0 is done. And is in the process of committing

16/09/20 11:02:13 INFO mapred.LocalJobRunner: map

16/09/20 11:02:13 INFO mapred.Task: Task 'attempt_local1146489696_0001_m_000001_0' done.

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Finishing task: attempt_local1146489696_0001_m_000001_0

16/09/20 11:02:13 INFO mapred.LocalJobRunner: map task executor complete.

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Waiting for reduce tasks

16/09/20 11:02:13 INFO mapred.LocalJobRunner: Starting task: attempt_local1146489696_0001_r_000000_0

16/09/20 11:02:13 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

16/09/20 11:02:13 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@b0a9ac0

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=333971456, maxSingleShuffleLimit=83492864, mergeThreshold=220421168, ioSortFactor=10, memToMemMergeOutputsThreshold=10

16/09/20 11:02:13 INFO reduce.EventFetcher: attempt_local1146489696_0001_r_000000_0 Thread started: EventFetcher for fetchingMap Completion Events

16/09/20 11:02:13 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output ofmap attempt_local1146489696_0001_m_000000_0 decomp: 37 len: 41 to MEMORY

16/09/20 11:02:13 INFO reduce.InMemoryMapOutput: Read 37 bytes from map-outputfor attempt_local1146489696_0001_m_000000_0

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-outputof size: 37, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->37

16/09/20 11:02:13 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output ofmap attempt_local1146489696_0001_m_000001_0 decomp: 2 len: 6 to MEMORY

16/09/20 11:02:13 INFO reduce.InMemoryMapOutput: Read 2 bytes from map-outputfor attempt_local1146489696_0001_m_000001_0

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-outputof size: 2, inMemoryMapOutputs.size() -> 2, commitMemory -> 37, usedMemory ->39

16/09/20 11:02:13 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning16/09/20

11:02:13 INFO mapred.LocalJobRunner: 2 / 2 copied.

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: finalMerge called with 2 in-memory map-outputs and 0 on-disk map-outputs

16/09/20 11:02:13 INFO mapred.Merger: Merging 2 sorted segments

16/09/20 11:02:13 INFO mapred.Merger: Down to the last merge-pass, with 1 segmentsleft of total size: 29 bytes

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: Merged 2 segments, 39 bytes todisk to satisfy reduce memory limit

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: Merging 1 files, 41 bytes from disk

16/09/20 11:02:13 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes frommemory into reduce

16/09/20 11:02:13 INFO mapred.Merger: Merging 1 sorted segments

16/09/20 11:02:13 INFO mapred.Merger: Down to the last merge-pass, with 1 segmentsleft of total size: 29 bytes

16/09/20 11:02:13 INFO mapred.LocalJobRunner: 2 / 2 copied.

16/09/20 11:02:14 INFO Configuration.deprecation: mapred.skip.on is deprecated.Instead, use mapreduce.job.skiprecords

16/09/20 11:02:14 INFO mapred.Task: Task:attempt_local1146489696_0001_r_000000_0 is done. And is in the process ofcommitting

16/09/20 11:02:14 INFO mapreduce.Job: Job job_local1146489696_0001 running inuber mode : false

16/09/20 11:02:14 INFO mapred.LocalJobRunner: 2 / 2 copied.

16/09/20 11:02:14 INFO mapred.Task: Task attempt_local1146489696_0001_r_000000_0 is allowed to commit now

16/09/20 11:02:14 INFO mapreduce.Job: map 100% reduce 0%

16/09/20 11:02:14 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1146489696_0001_r_000000_0' to hdfs://localhost:54310/user/user/op1/_temporary/0/task_local1146489696_0001_r_000 000

16/09/20 11:02:14 INFO mapred.LocalJobRunner: reduce > reduce

16/09/20 11:02:14 INFO mapred.Task: Task 'attempt_local1146489696_0001_r_000000_0' done.

16/09/20 11:02:14 INFO mapred.LocalJobRunner: Finishing task: attempt_local1146489696_0001_r_000000_0

16/09/20 11:02:14 INFO mapred.LocalJobRunner: reduce task executor complete.16/09/20

11:02:15 INFO mapreduce.Job: map 100% reduce 100%

16/09/20 11:02:15 INFO mapreduce.Job: Job job_local1146489696_0001 completedsuccessfully

16/09/20 11:02:15 INFO mapreduce.Job: Counters: 38File

System Counters

FILE: Number of bytes read=812415 FILE:

Number of bytes written=1575498

FILE: Number of read operations=0 FILE:

Number of large read operations=0FILE:

Number of write operations=0 HDFS:

Number of bytes read=126 HDFS: Number

of bytes written=23 HDFS: Number of read

operations=25

HDFS: Number of large read operations=0

HDFS: Number of write operations=5

Map-Reduce Framework Map

input records=5 Map

output records=7Map

output bytes=69

Map output materialized bytes=47

Input split bytes=211

Combine input records=7

Combine output records=3

Reduce input groups=3

Reduce shuffle bytes=47

Reduce input records=3

Reduce output records=3

Spilled Records=6 Shuffled

Maps =2

Failed Shuffles=0 Merged

Map outputs=2 GC time

elapsed (ms)=14

CPU time spent (ms)=0

Physical memory (bytes) snapshot=0

Virtual memory (bytes) snapshot=0

Total committed heap usage (bytes)=925368320

Shuffle Errors

BAD_ID=0

CONNECTION=0

IO_ERROR=0

WRONG_LENGTH=0

WRONG_MAP=0

WRONG_REDUCE=0

File Input Format Counters

Bytes Read=42

File Output Format Counters

Bytes Written=23

Step 4:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fs -lsop1

16/09/20 11:03:59 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

Found 2 items

-rw-r--r-- 1 user supergroup 0 2016-09-20 11:02 op1/_SUCCESS

-rw-r--r-- 1 user supergroup 23 2016-09-20 11:02 op1/part-r-00000

Step 5:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ usr/local/hadoop1/bin/hadoop fs -cat op1/result.txt

bash: usr/local/hadoop1/bin/hadoop: No such file or directory

Step 6:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ usr/local/hadoop1/bin/hadoop fs -catop1/*

bash: usr/local/hadoop1/bin/hadoop: No such file or directory

Step 7:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fs -cat op1/*

16/09/20 11:05:58 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

hello 3

helo 1

world 3

Step 8:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fs -cat op1/result.txt

16/09/20 11:06:14 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

cat: `op1/result.txt': No such file or directory

Step 9:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fsop1/result.txt

op1/result.txt: Unknown command

Step 10:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fs -cat op1/>>result.txt

16/09/20 11:06:50 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

cat: `op1': Is a directory

Step 11:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ /usr/local/hadoop1/bin/hadoop fs -cat >> op1/result.txt

bash: op1/result.txt: No such file or directory

Step 12:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ stop-all.sh

This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh

16/09/20 11:11:57 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

Stopping namenodes on [localhost]

localhost: stopping namenode

localhost: stopping datanode

Stopping secondary namenodes [0.0.0.0]

0.0.0.0: stopping secondarynamenode

16/09/20 11:12:21 WARN util.NativeCodeLoader: Unable to load native-hadoop libraryfor your platform... using builtin-java classes where applicable

stopping yarn daemons

stopping resourcemanager

localhost: stopping nodemanagerno

proxyserver to stop

Step 13:

user@cs117-HP-Pro-3330-MT:/usr/local/hadoop1$ cd\

>

user@cs117-HP-Pro-3330-MT:~$

Wordcount.java

```
//package org.myorg;
    import java.io.IOException;
   import java.util.StringTokenizer;
   import org.apache.hadoop.fs.Path;
   import org.apache.hadoop.conf.*;
   import org.apache.hadoop.io.*;
   import org.apache.hadoop.mapreduce.*;
   import org.apache.hadoop.mapreduce.lib.input.FileInputFormat; import
   org.apache.hadoop.mapreduce.lib.input.TextInputFormat; import
   org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
```

```
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;

public class WordCount {

 public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {private
    final static IntWritable one = new IntWritable(1);
    private Text word = new Text();

    public void map(LongWritable key, Text value, Context context) throws
IOException, InterruptedException {
      String line = value.toString();
      StringTokenizer tokenizer = new StringTokenizer(line);while
       (tokenizer.hasMoreTokens()) {
          word.set(tokenizer.nextToken());
          context.write(word, one);
       }
    }
 }

 public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {

    public void reduce(Text key, Iterable<IntWritable> values, Context context)throws
      IOException, InterruptedException {
       int sum = 0;
       for (IntWritable val : values) {
```

```
            sum += val.get();
        }
        context.write(key, new IntWritable(sum));
    }
}

public static void main(String[] args) throws Exception {
    Configuration conf = new Configuration();

        Job job = new Job(conf, "wordcount");

    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    job.setMapperClass(Map.class);
    job.setReducerClass(Reduce.class);
    job.setInputFormatClass(TextInputFormat.class);
    job.setOutputFormatClass(TextOutputFormat.class);
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));

    job.waitForCompletion(true);
} }
```

Result:

Thus a word count program to demonstrate the use of Map and Reduce task iscreated and executed successfully.