ST.JOHNS

COLLEGE OF ENGINEERING AND TECHNOLOGYYERRAKOTA, YEMMIGANUR–518360

DEPARTMENTOFCOMPUTERSCIENCEANDENGINEERING

LABMANUAL

CLOUDCOMPUTINGLABORATORY

Regulation R20

Year/ Semester:III/II

2022–2023

PREPAREDBY

Mr.G.Viswanath,M.Tech,

AssociateProfessor/CSE

LISTOFEXPERIMENTS

1. InstallVirtualbox/VMwareWorkstationwithdifferentflavoursoflinuxorwindows OSontopof windows7or8.
2. InstallaCcompilerinthevirtualmachinecreatedusingvirtualboxandexecuteSim ple Programs
3. InstallGoogleAppEngine.Create*helloworld*appandothersimplewebapplicationsu singpython/java.
4. UseGAElaunchertolaunchthewebapplications.
5. SimulateacloudscenariousingCloudSimandrunaschedulingalgorithmthat isnotpresentinCloudSim.
6. Findaproceduretotransferthefilesfromonevirtualmachinetoanothervirtualmachine.
7. Findaproceduretolaunchvirtualmachineusingtrystack(OnlineOpenstack DemoVersion)
8. InstallHadoopsinglenodeclusterandrunsimpleapplicationslikewordcount.

TABLEOFCONTENTS

6.		Find aprocedureto transferthefiles fromone virtualmachine toanothervirtualmachine.		

7.		Findaproceduretolaunchvirtualmachineusingtrystack(OnlineOpenstackDemoVersion)		
8.		Install Hadoop singlenodeclusterandrunsimpleapplications likewordcount.		

EX.NO:1

DATE:

InstallVirtualbox/VMwareWorkstationwithdifferentflavoursoflinuxorwindows OSon topofwindows7 or8.

Aim:

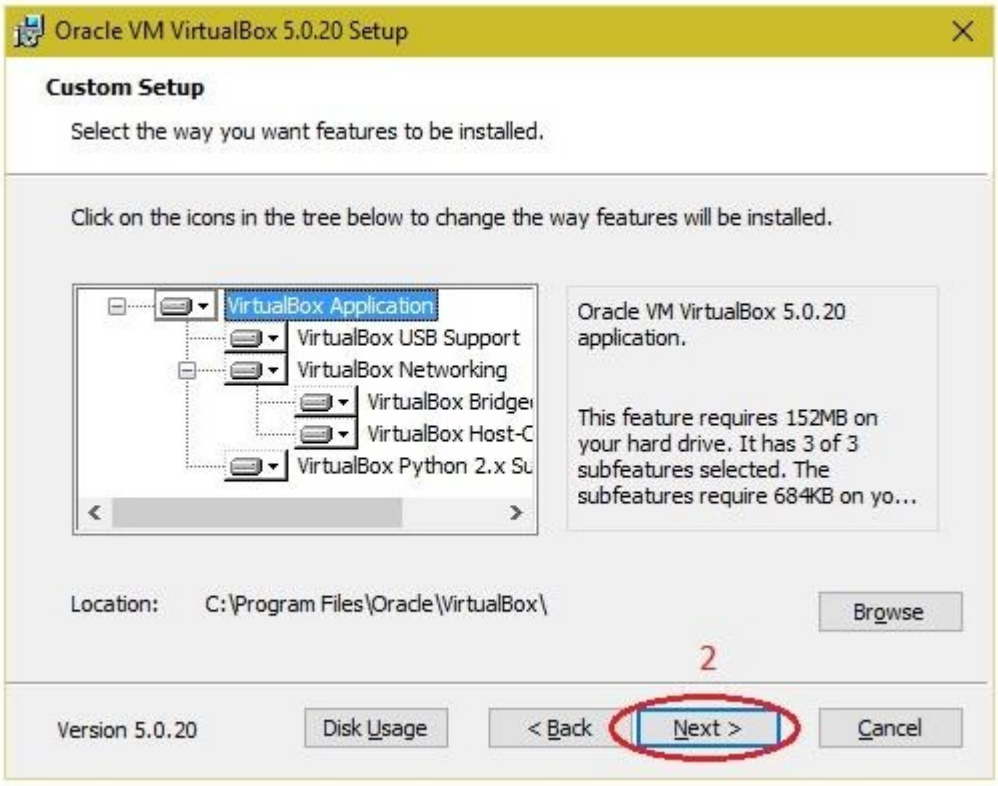
ToInstall Virtualbox / VMwareWorkstation with different flavours of linux or windowsOSontopofwindows7or8.

PROCEDURE:

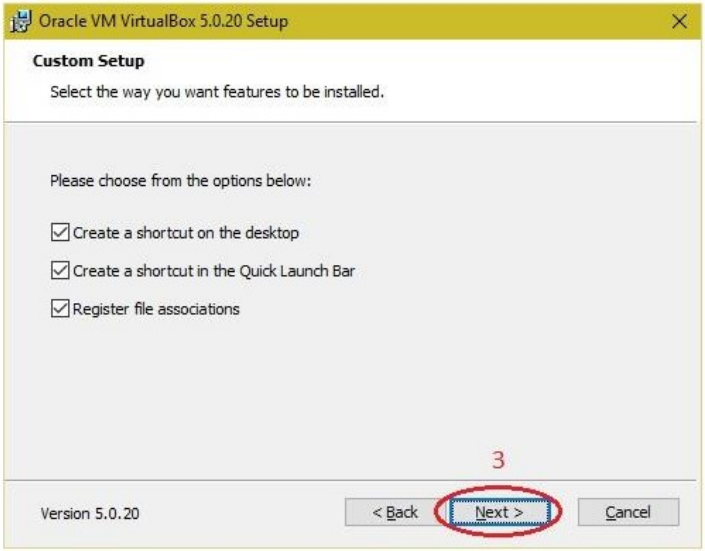
StepstoinstallVirtualBox:

1. DownloadtheVirtualboxexeand clicktheexefile…andselectnextbutton..

2. Clickthenextbutton..

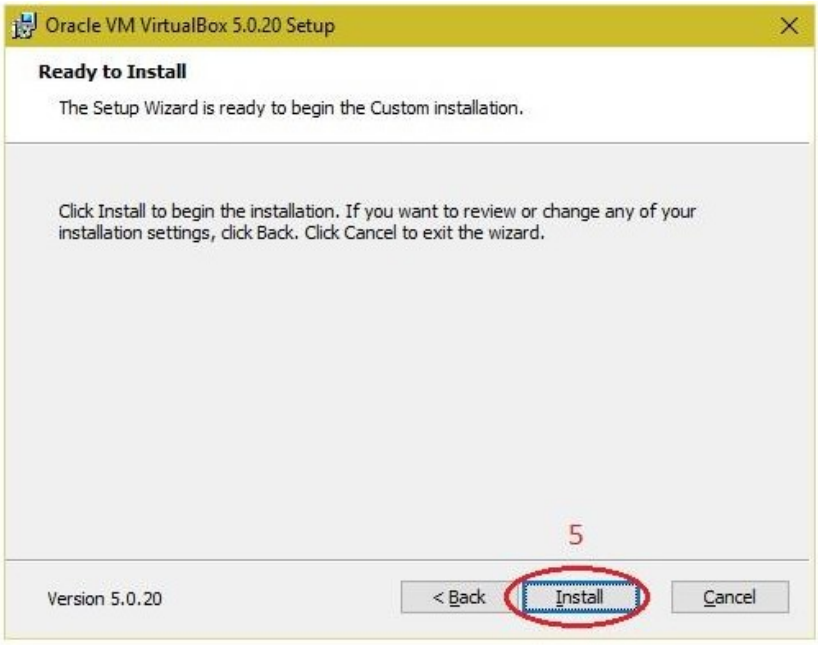


3. Clickthenextbutton



4. ClicktheYESbutton..

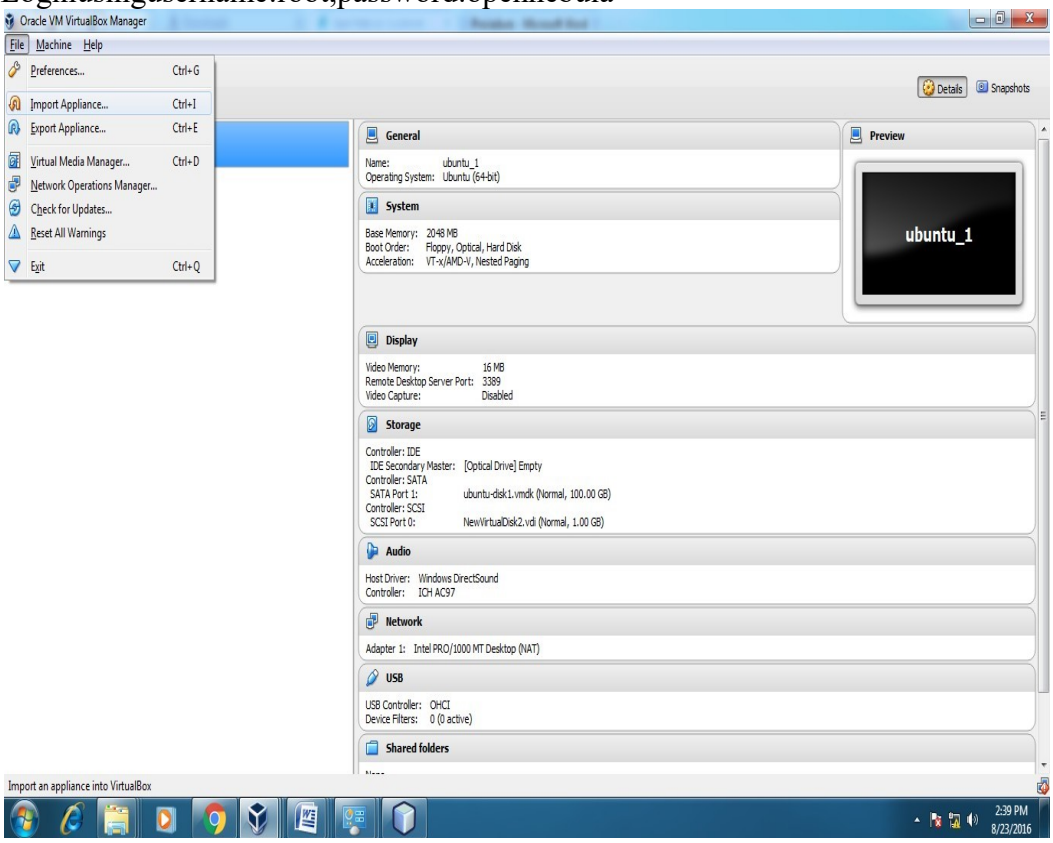
5. Clicktheinstallbutton…

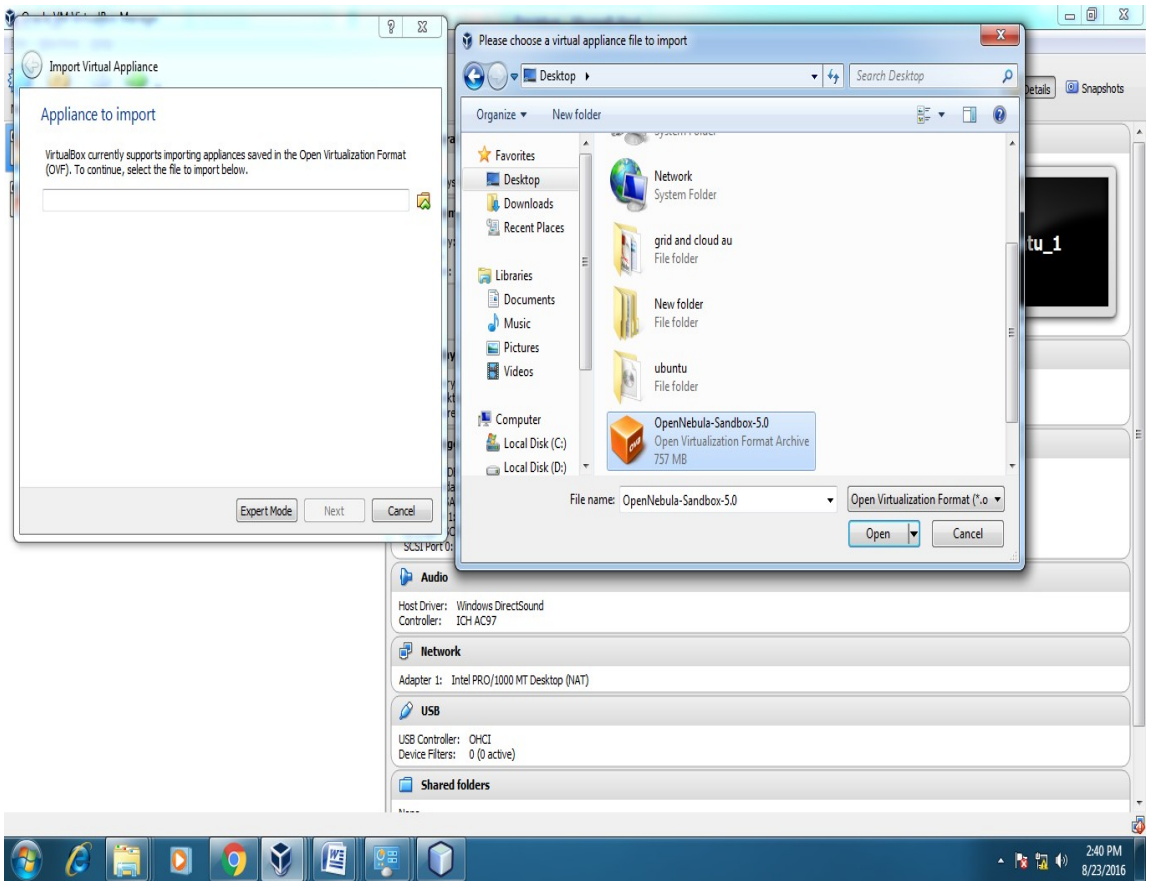


6. Theninstallationwascompleted..theshowvirtualboxiconondesktopscreen….

StepstoimportOpennebulasandbox:

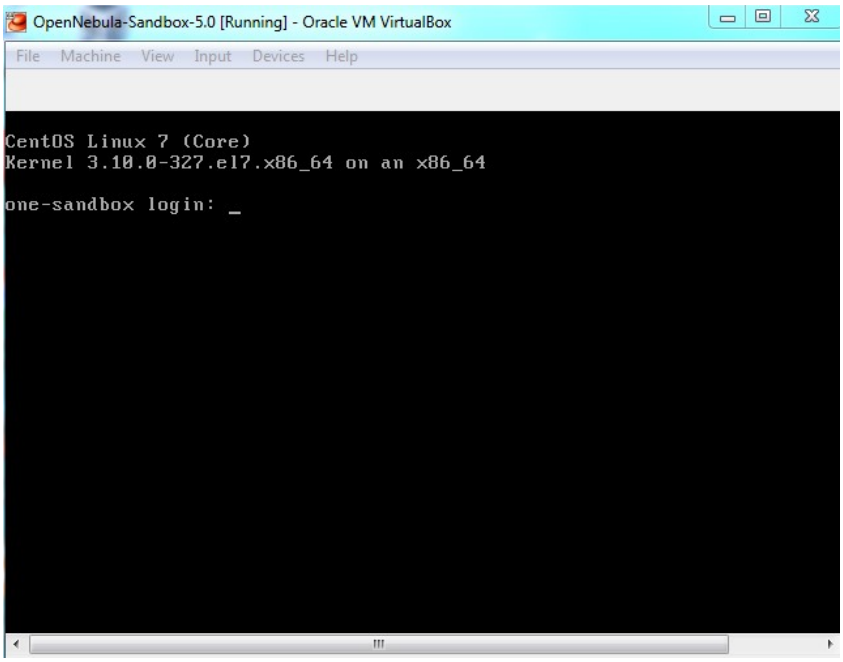
1. OpenVirtualbox
2. File→importAppliance
3. BrowseOpenNebula-Sandbox-5.0.ovafile
4. Thengotosetting, selectUsband chooseUSB1.1
5. ThenStarttheOpenNebula
6. Loginusingusername:root,password:opennebula

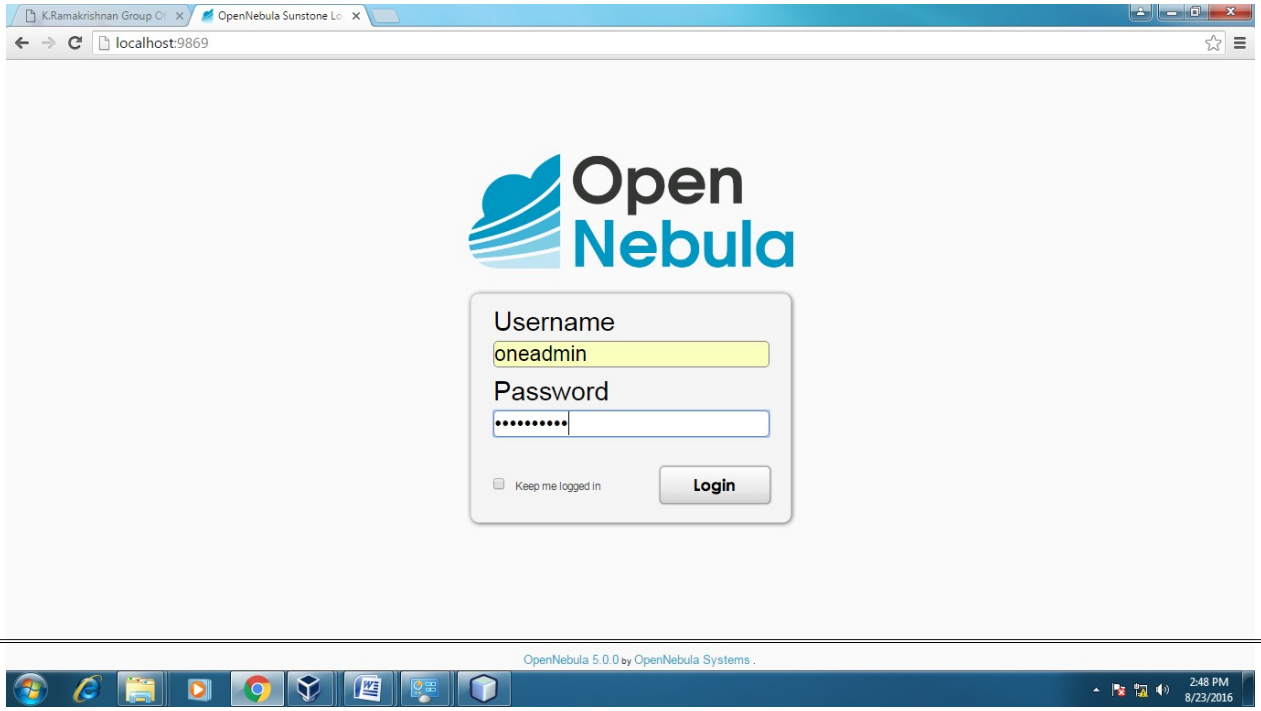


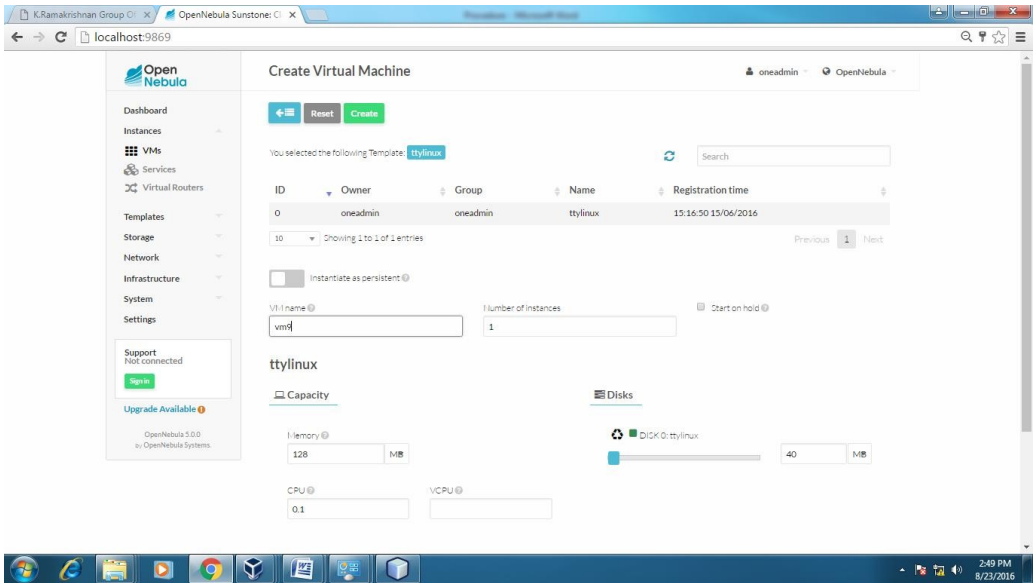


StepstocreateVirtualMachinethroughopennebula

1. OpenBrowser,typelocalhost:9869
2. Loginusingusername:oneadmin,password:opennebula
3. Clickoninstances,select VMsthenfollowthestepstocreate Virtaulmachine
    a. Expandthe+symbol
    b. Selectuseroneadmin
    c. ThenentertheVMname,no.ofinstance,cpu.
    d. Thenclickoncreatebutton.
    e. Repeatthestepsthe C,Dforcreating morethanoneVMs.


OpenNebula-Sandbox-5.0 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
CentOS Linux 7 (Core)
Kernel 3.10.0-327.el7.x86_64 on an x86_64
one-sandbox login: _


K.Ramakrishnan Group Of
OpenNebula Sunstone Lo
localhost:9869
Open Nebula
Username
oneadmin
Password
Keep me logged in
Login
OpenNebula 5.0.0 by OpenNebula Systems .
2:48 PM
8/23/2016



APPLICATIONS:

There are various applications of cloud computing in today's network world. Many search engines and socialwebsitesareusingtheconceptofcloudcomputinglikewww.amazon.com,hotmail.com,facebook.com,linkedln.cometc.theadvantagesofcloudcomputingincontexttoscalabilityislikereducedrisk,lowcosttesting,abilitytosegmentthecustomerbaseandauto-scalingbasedonapplicationload.

RESULT:

Thustheproceduretorunthevirtualmachineofdifferentconfiguration.

EX.NO:2

DATE:

Install a C compiler in the virtual machine created usingvirtual box andexecuteSimplePrograms

---

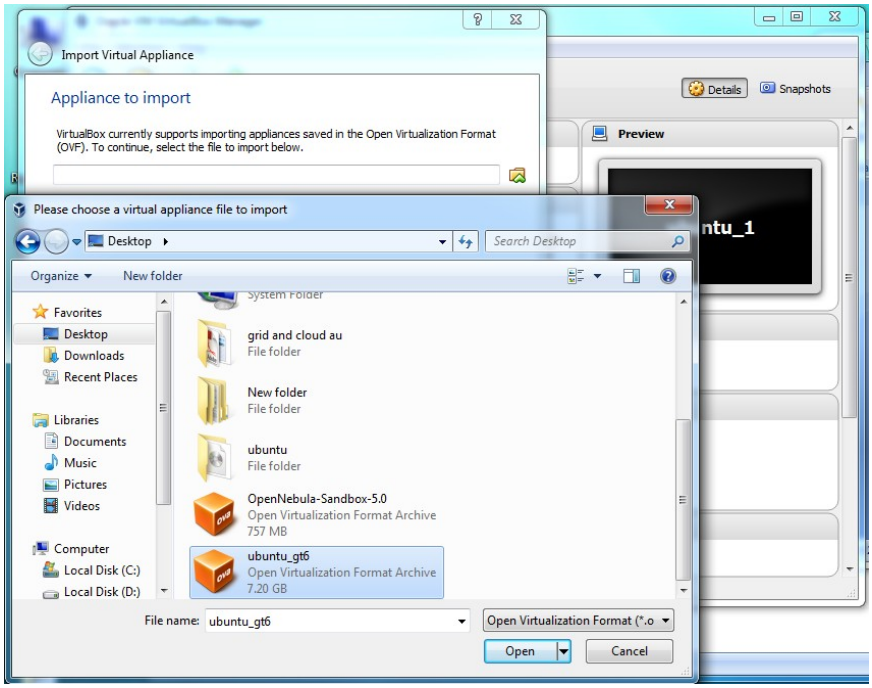
Aim:

ToInstallaCcompilerinthevirtual machinecreatedusing virtualbox and executeSimplePrograms`

PROCEDURE:

<u>Stepsto import.ovafile:</u>

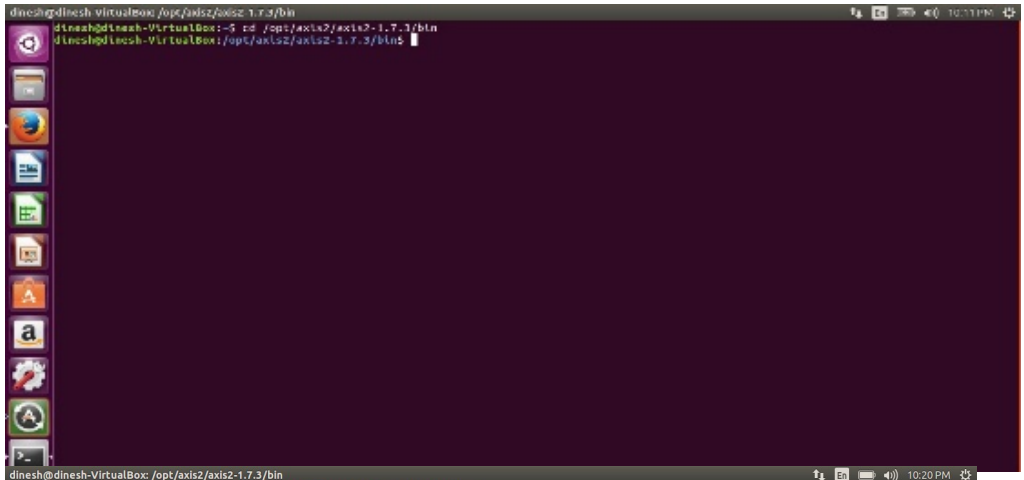
1. OpenVirtualbox
2. File→importAppliance
3. Browseubuntu_gt6.ovafile
4. Thengotosetting, selectUsband chooseUSB1.1
5. ThenStarttheubuntu_gt6
6. Loginusingusername:dinesh,password:99425.



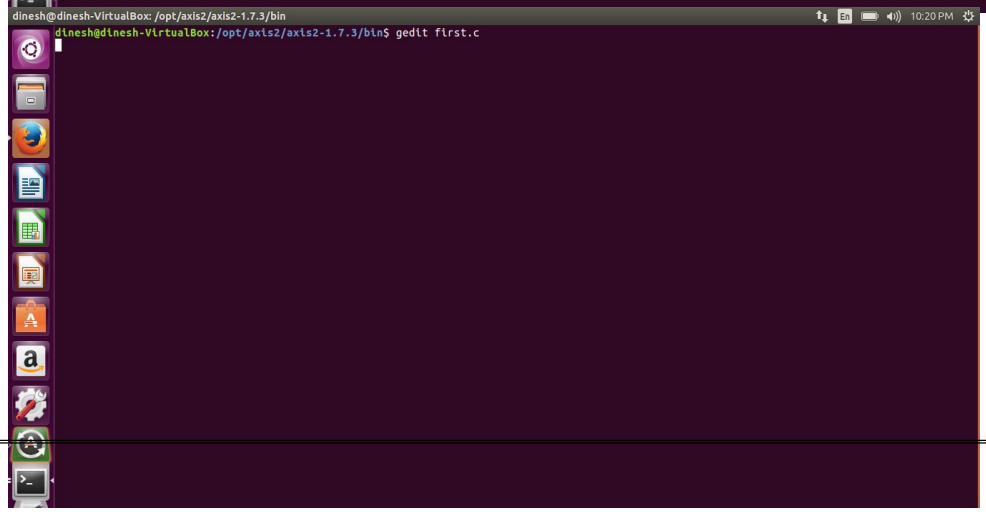
Stepsto runcprogram:

1. Opentheterminal
2. Typecd/opt/axis2/axis2-1.7.3/binthenpressenter
3. gedithello.c
4. gcchello.c
5. ./a.out

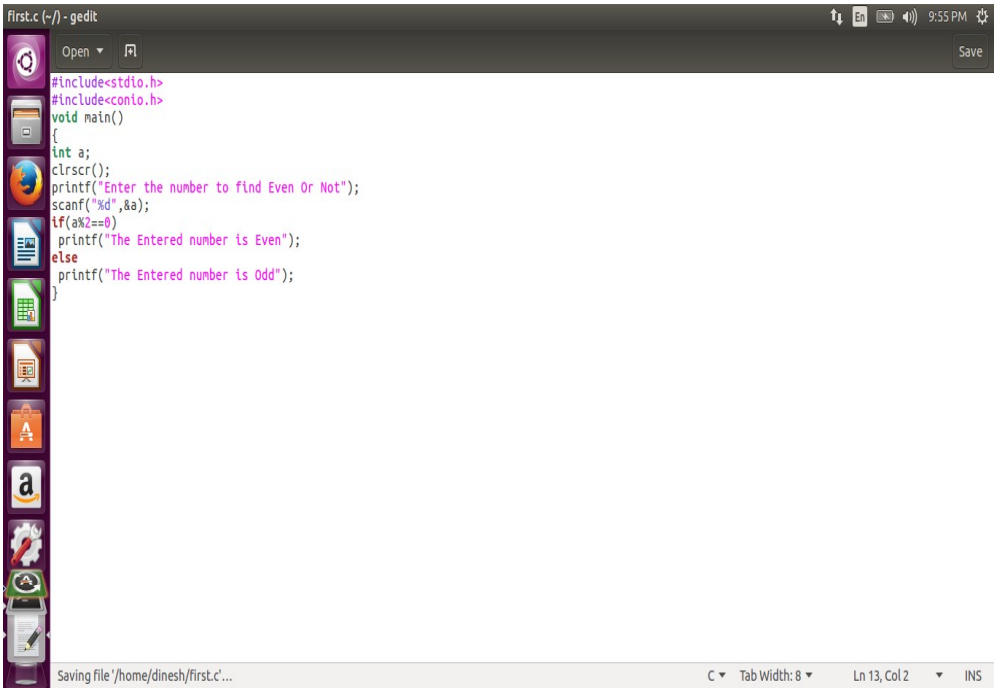
1. **Typecd/opt/axis2/axis2-1.7.3/binthenpressenter**



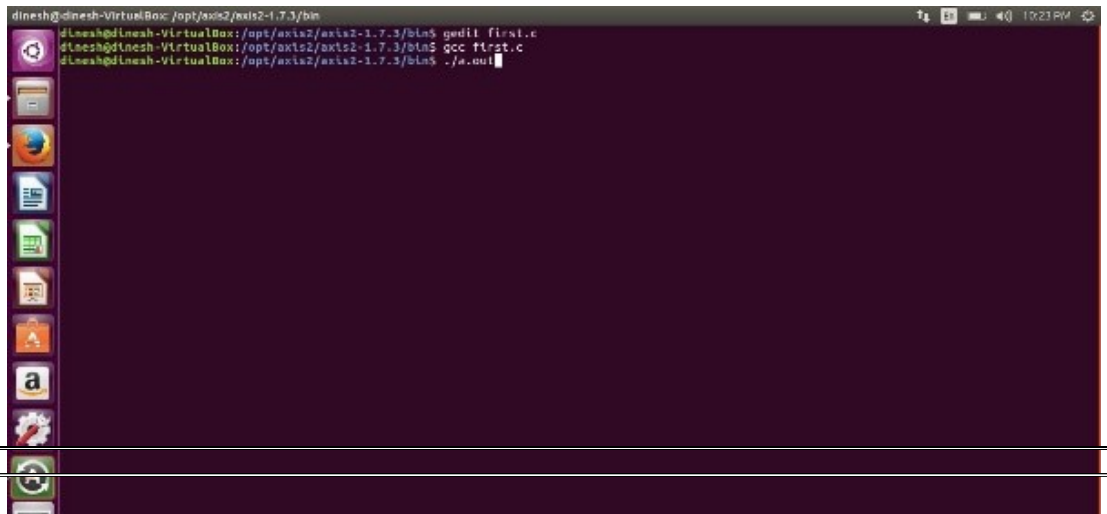
2. **Typegeditfirst.c**



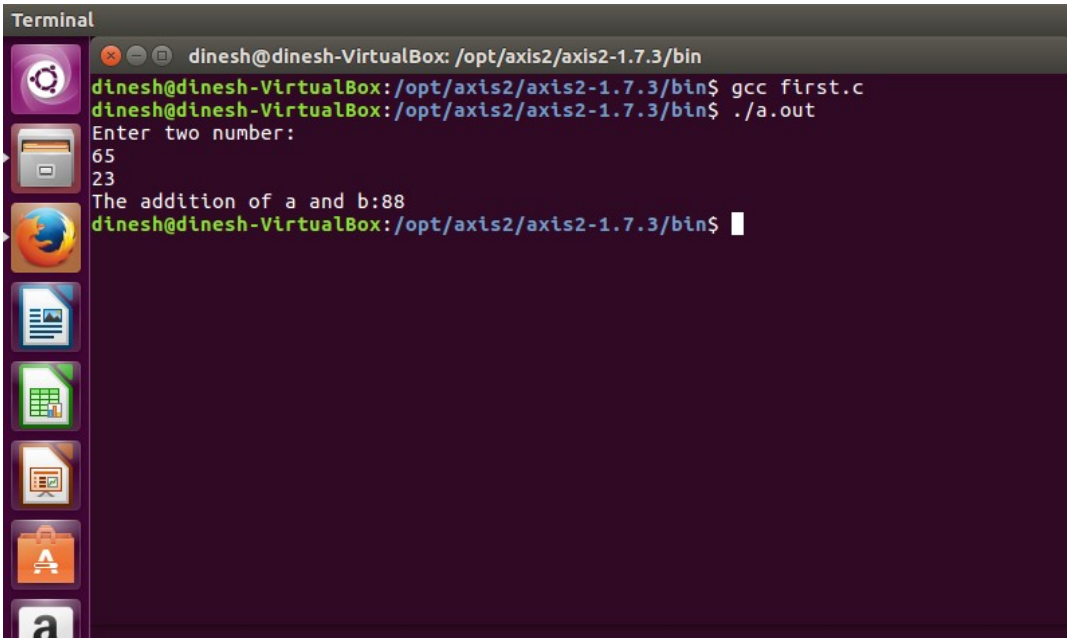
3. Typethec program



4. **Runningthe**Cprogram



4. **Displaytheoutput:**



APPLICATIONS:

Simplyrunningallprogramsingridenvironment.

RESULT:

ThusthesimpleCprograms executedsuccessfully.

EX.NO:3 **DATE:**

Install Google App Engine. Create *hello world* app and other simple webapplicationsusingpython/java.

Aim:

To Install Google App Engine. Create *hello world* app and other simple webapplicationsusingpython/java.

Procedure:

1. InstallGooglePluginforEclipse

Read this guide – how to install Google Plugin for Eclipse. If you install the Google App EngineJava SDK together with "**Google Plugin for Eclipse**", then go to step 2, Otherwise, get the GoogleAppEngineJava SDKandextractit.

2. CreateNewWebApplicationProject

InEclipsetoolbar,clickontheGoogleicon,andselect"**NewWebApplicationProject…**"

Figure–NewWebApplicationProject

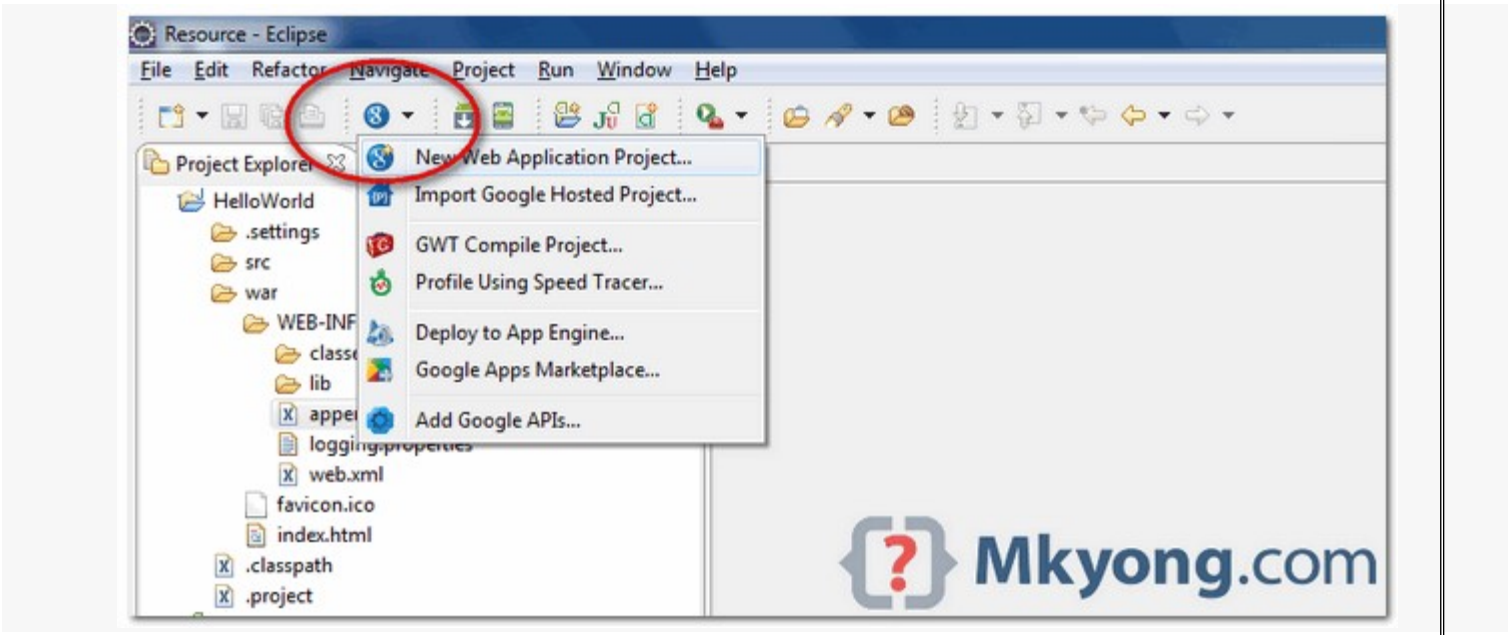
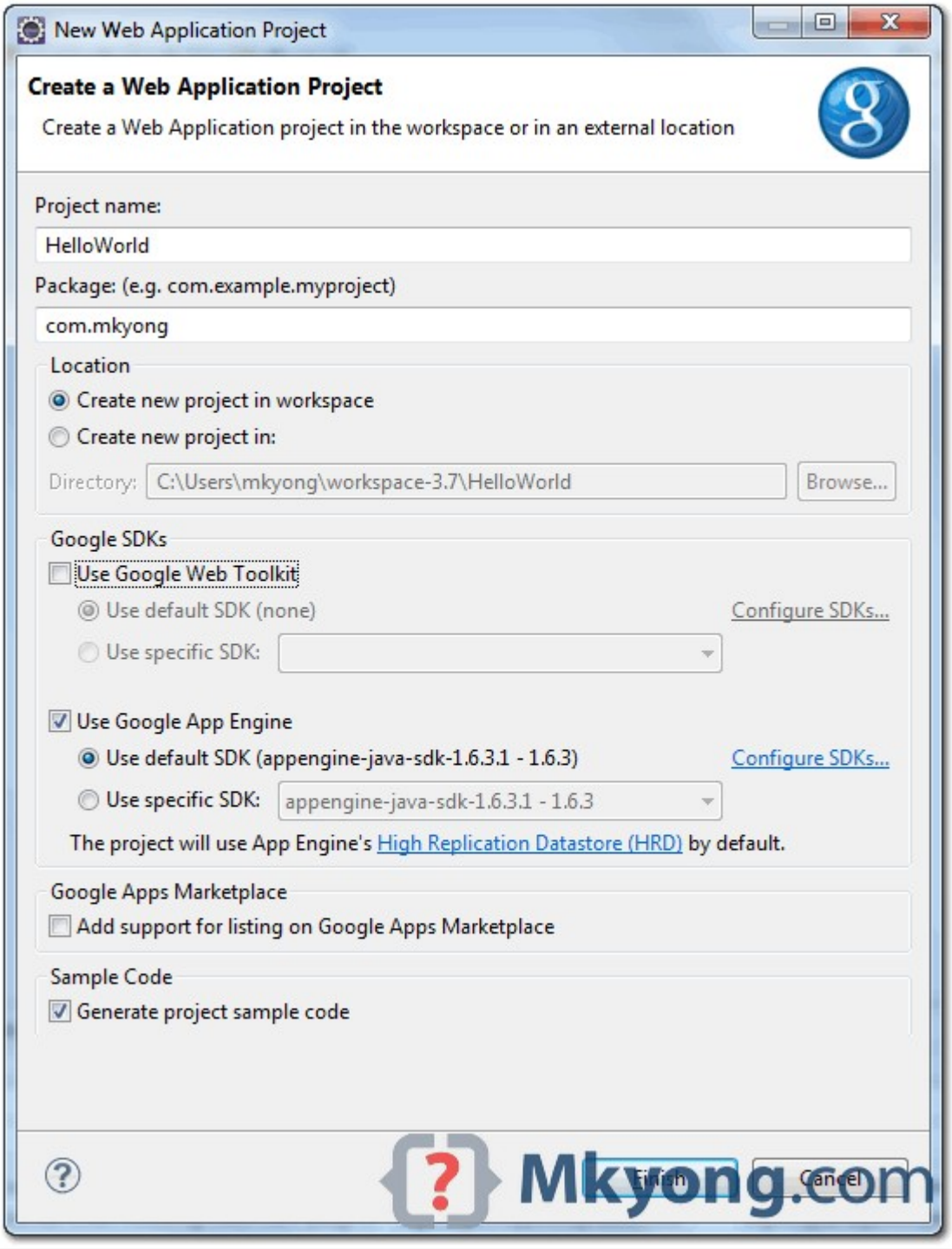


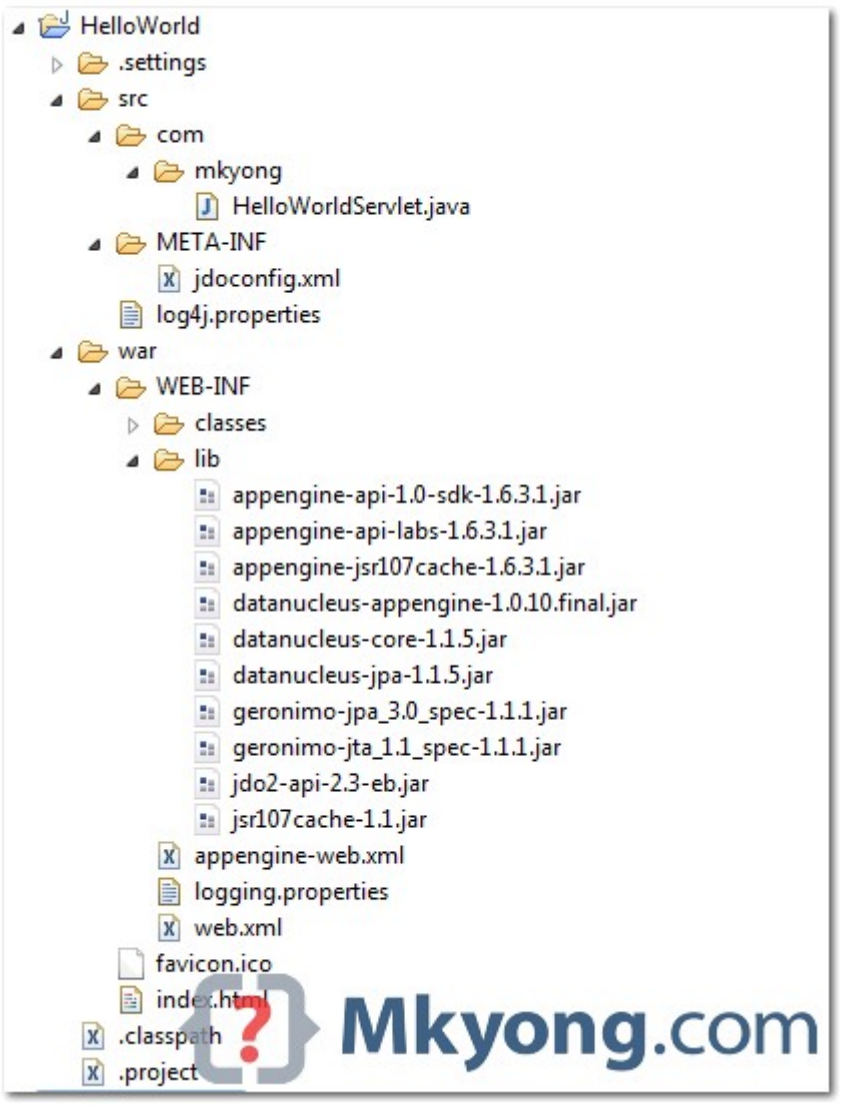
*Figure – Deselect the "**Google Web ToolKit**", and link your GAE Java SDK via the "**configureSDK**"link.*



Clickfinished,GooglePluginforEclipsewillgenerateasampleprojectautomatically.

3. HelloWorld

Reviewthegeneratedprojectdirectory.



Nothingspecial,astandardJavawebproject structure.

```
HelloWorld/
  src/
    ...Java source
    code...META-INF/
      ...other
  configuration...war/
    ...JSPs, images, data
```

files...WEB-INF/
  ...appconfiguration...

```
    lib/
      ...JARs for
    libraries...classes/
      ...compiledclasses...
```
Copy
The extra is this file “appengine-web.xml“, Google App Engine need this to run and deploy theapplication.

File:appengine-web.xml

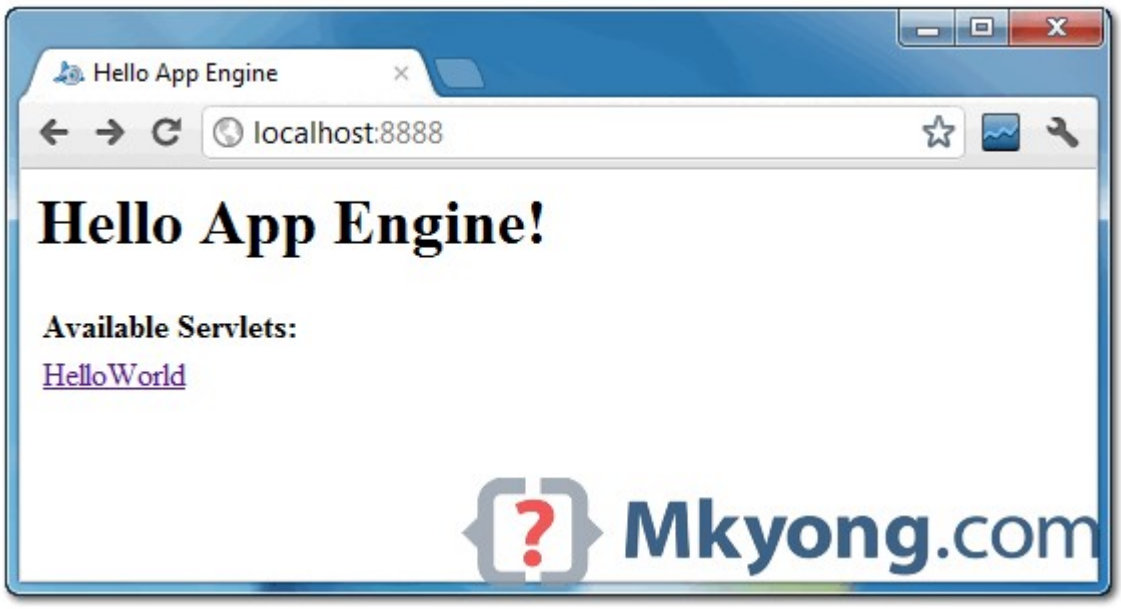
```
<?xmlversion="1.0"encoding="utf-8"?>
<appengine-web-appxmlns="http://appengine.google.com/ns/1.0">
 <application></application>
 <version>1</version>

 <!--Configurejava.util.logging-->
 <system-properties>
  <propertyname="java.util.logging.config.file"value="WEB-INF/logging.properties"/>
 </system-properties>

</appengine-web-
app>Copy
```

4. Runitlocal

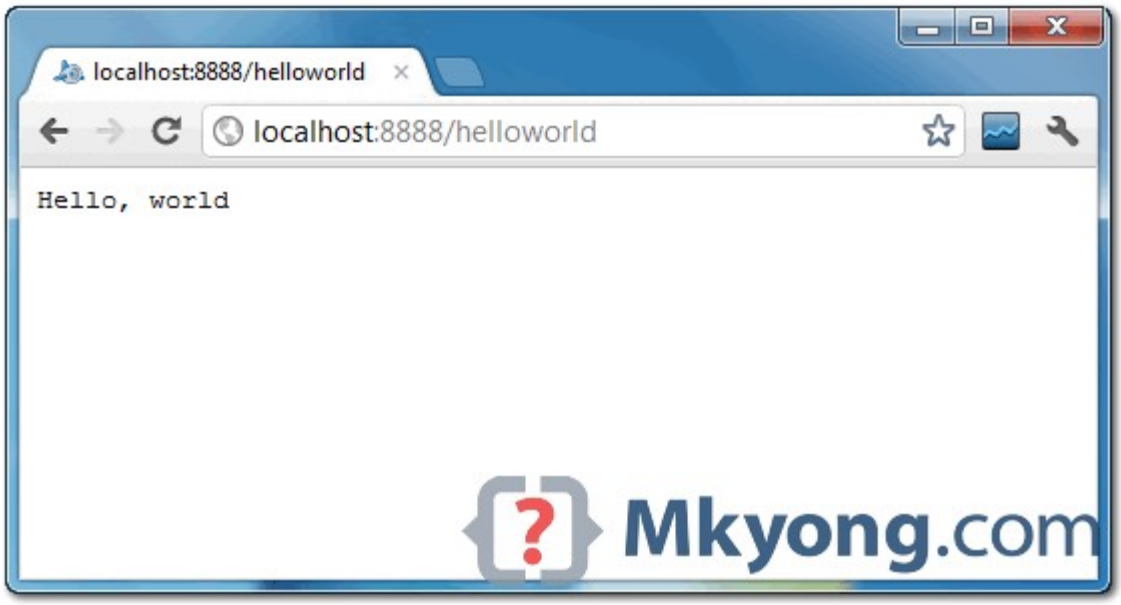
Rightclickontheprojectandrunas“**WebApplication**“.

Eclipseconsole:

```
//...
INFO:Theserverisrunningathttp://localhost:8888/
30 Mac 2012 11:13:01 PM com.google.appengine.tools.development.DevAppServerImpl
startINFO:Theadminconsoleisrunningathttp://localhost:8888/_ah/admin
```
Copy
AccessURLhttp://localhost:8888/,seeoutput



andalsothehelloworldservlet–http://localhost:8888/helloworld



5. DeploytoGoogleAppEngine

Register an account on https://appengine.google.com/, and create an application ID for your webapplication.

In this demonstration, I created an application ID, named “mkyong123”, and put it in appengine-web.xml.

File:appengine-web.xml

```
<?xmlversion="1.0"encoding="utf-8"?>
<appengine-web-appxmlns="http://appengine.google.com/ns/1.0">
  <application>mkyong123</application>
  <version>1</version>

  <!--Configurejava.util.logging-->
  <system-properties>
    <propertyname="java.util.logging.config.file"value="WEB-INF/logging.properties"/>
  </system-properties>

</appengine-web-
app>Copy
```

Todeploy,seefollowingsteps:

Figure1.1–ClickonGAEdeploybuttonon thetoolbar.

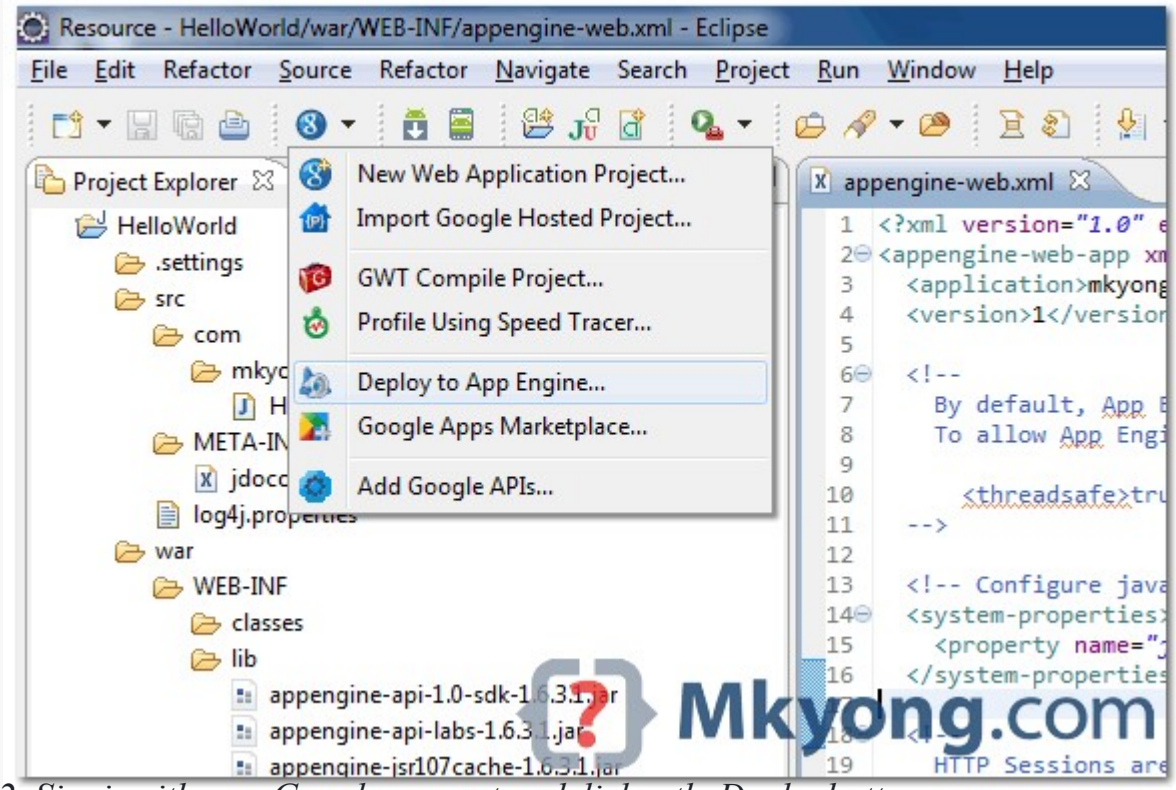


Figure1.2–Signinwith yourGoogleaccount andclickontheDeploybutton.

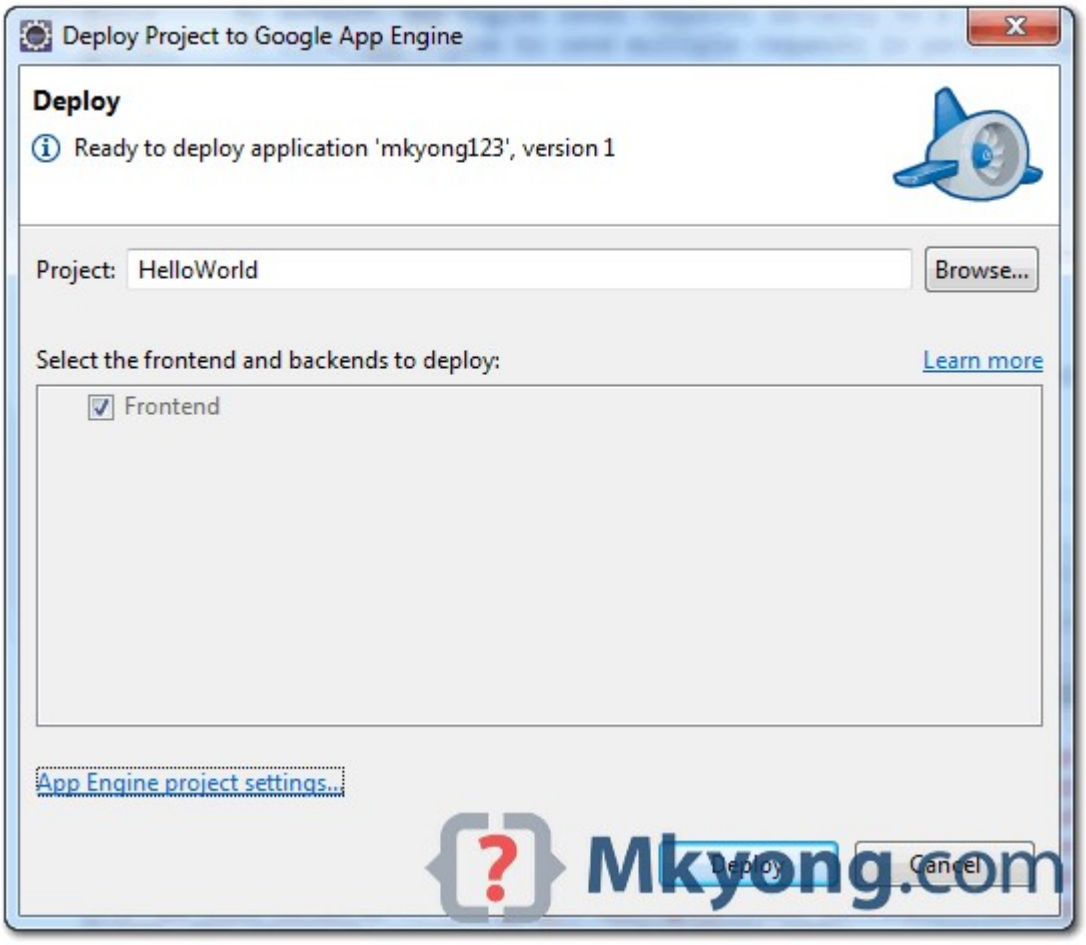


Figure 1.3 – If everything is fine, the hello world web application will be deployed to this URL – http://mkyong123.appspot.com/

Result:

Thusthesimpleapplicationwascreatedsuccessfully.

EX.NO:5 **DATE:**

5.Simulate a cloud scenario using CloudSim and run a schedulingalgorithmthatisnotpresentin CloudSim.

Aim:

ToSimulateacloudscenariousingCloudSimandrunaschedulingalgorithmthatisnotpresent inCloudSim.

Steps:

HowtouseCloudSiminEclipse

CloudSim is written in Java. The knowledge you need to use CloudSim is basic Java programmingand some basics about cloud computing. Knowledge of programming IDEs such as Eclipse orNetBeansisalsohelpful. Itisalibrary and, hence,CloudSimdoes nothavetobeinstalled.Normally, you can unpack the downloaded package in any directory, add it to the Java classpath anditisreadyto be used.Please verifywhetherJavaisavailableonyoursystem.

TouseCloudSiminEclipse:

1. DownloadCloudSiminstallablefiles from*https://code.google.com/p/cloudsim/downloads/listandunzip*
2. OpenEclipse
3. Createanew JavaProject: File-> New
4. ImportanunpackedCloudSimprojectintothenewJavaProject

   The first step is to initialise the CloudSim package by initialising the CloudSim library, asfollows

   CloudSim.init(num_user,calendar,trace_flag)
5. Data centres are the resource providers in CloudSim; hence, creation of datacentres is a second step. To create Datacenter, you need theDatacenterCharacteristicsobjectthatstorestheproperties ofadatacentresuchasarc hitecture, OS, list of machines, allocation policy that covers the time orspaceshared,the timezoneanditsprice:

   Datacenter datacenter9883 = new Datacenter(name, characteristics, newVmAllocationPolicySimple(hostList),s
6. Thethirdstepistocreateabroker:

   DatacenterBrokerbroker=createBroker();
7. The fourth step is to create one virtual machine unique ID of the VM, userIdID of the VM's owner, mips, number Of Pes amount of CPUs, amount of RAM,amount of bandwidth, amount of storage, virtual machine monitor, andcloudletSchedulerpolicyforcloudlets:

   Vmvm =newVm(vmid,brokerId,mips,pesNumber,ram,bw,size,vmm,new CloudletSchedulerTimeShared())
8. Submit theVMlisttothebroker:

   broker.submitVmList(vmlist)
9. Createacloudletwithlength,filesize,outputsize,andutilisationmodel:

   Cloudletcloudlet=newCloudlet(id,length,pesNumber,fileSize,outputSize,utilizationModel,utiliza tionMode
10. Submit the cloudlet list to the broker:broker.submitCloudletList(cloudlet List)Sample Output from the Existing Example:Starting CloudSimExample1...

Initialising...
Starting
CloudSimversion 3.0
Datacenter_0isstarting...
>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>null
Broker
isstarting...Enti
tiesstarted.

: Broker: Cloud Resource List received with
1resource(s)0.0:Broker:TryingtoCreateVM#0inDat
acenter_0
: Broker: VM #0 has been created in Datacenter #2, Host
#00.1: Broker:Sending cloudlet0 toVM #0
400.1:Broker:Cloudlet0received
400.1 : Broker: All Cloudlets
executed.Finishing
.... 400.1:Broker:DestroyingVM
#0
Brokerisshuttingdown...
Simulation: No more
futureevents
CloudInformationService:NotifyallCloudSimentitiesforshutting
down.Datacenter_0isshutting down...
Broker is
shuttingdown
............Simulation
completed.Simulation
completed.
==========OUTPUT==========
CloudletIDSTATUSDatacenterID VM ID Time Start
TimeFinishTime 0 SUCCESS 2 0 400
0.1 400.1
*****Datacenter:
Datacenter_0*****User id
Debt
3 35.6

CloudSimExample1finished!

RESULT:

Thesimulationwassuccessfullyexecuted.

4. UseGAElaunchertolaunch the webapplications.

Aim:

ToUseGAElaunchertolaunchthewebapplications.

Steps:

MakingyourFirstApplication

Nowyouneedtocreateasimpleapplication.Wecouldusethe“+”optiontohavethelaunchermakeusanapplication-butinsteadwewilldoitbyhandtogetabettersenseofwhatisgoingon.

MakeafolderforyourGoogleAppEngineapplications.IamgoingtomaketheFolderonmyDesktopcalled“**apps**”-thepathtothisfolderis:

C:\DocumentsandSettings\csev\Desktop\apps

Andthenmakeasub-folderinwithinappscalled“ae-01-trivial”—thepathtothis folderwouldbe:

C:\DocumentsandSettings\csev\Desktop\apps\ae-01-trivial

Using a text editor such as JEdit (www.jedit.org), create a file called app.yaml in theae-01-trivialfolderwiththefollowingcontents:

```
application: ae-01-
    trivialversion:1
runtime:pythonapi_version:1h
andlers:-url:/.*
script:index.py
```

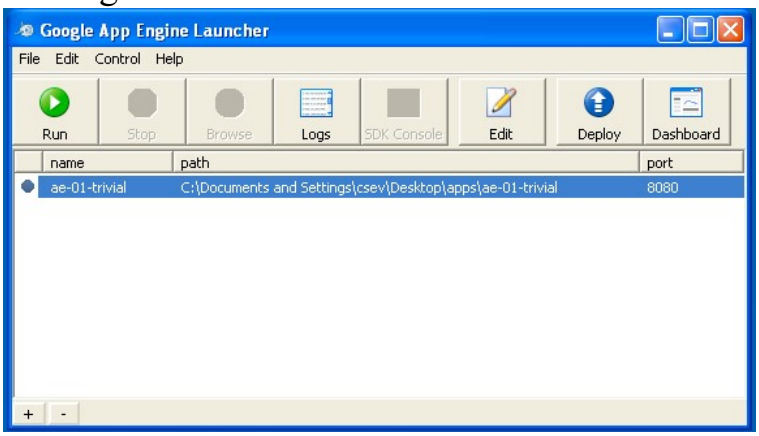
Note:Pleasedonotcopyandpastetheselinesintoyourtexteditor–youmightendupwithstrangecharacters–simplytypethemintoyoureditor.

Thencreateafileintheae-01-trivialfoldercalledindex.pywiththreelinesinit:

```
print 'Content-Type:
```

text/plain'print''
print'HellothereChuck'

Then start the GoogleAppEngineLauncher program that can be foundunderApplications.UsetheFile->Add ExistingApplication command andnavigateintotheappsdirectoryandselecttheae-01-trivialfolder.Once youhaveaddedtheapplication,selectitsothatyoucancontroltheapplicationusin gthelauncher.



Onceyouhaveselectedyourapplicationandpress Run. Afterafewmomentsyour

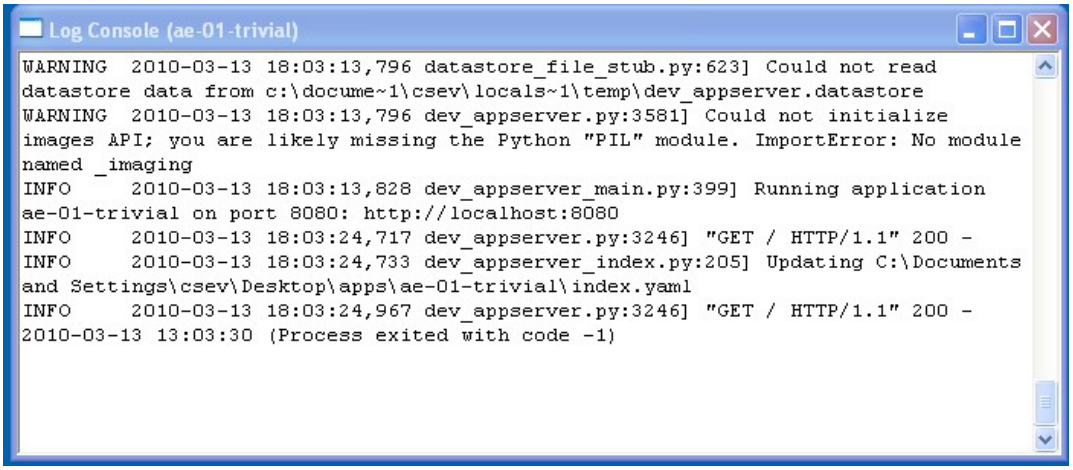
application will start and the launcherwill show a little green icon nextto yourapplication. Then press Browse to open a browser pointing at your applicationwhichisrunningathttp://localhost:8080/

Pastehttp://localhost:8080intoyourbrowserandyoushouldseeyourapplication asfollows:

Justforfun,edittheindex.pytochangethename“Chuck”toyourownnamean dpressRefreshinthebrowsertoverifyyourupdates.

WatchingtheLog

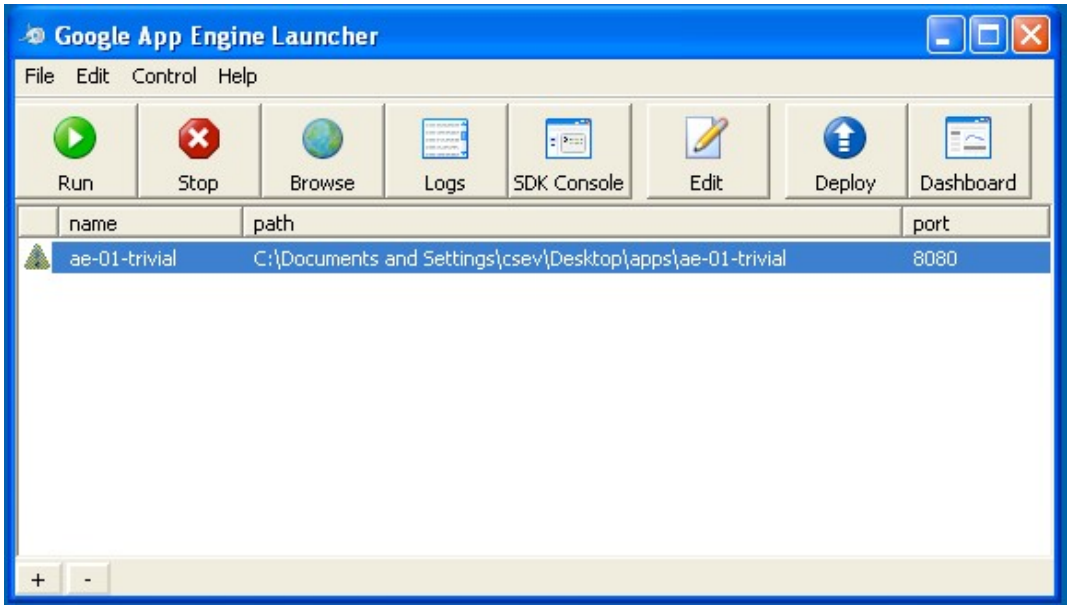
Youcanwatchtheinternallogoftheactionsthatthewebserverisperformingwhen youareinteractingwithyourapplication in thebrowser. SelectyourapplicationintheLauncherandpresstheLogsbuttontobringupalogwindo w:



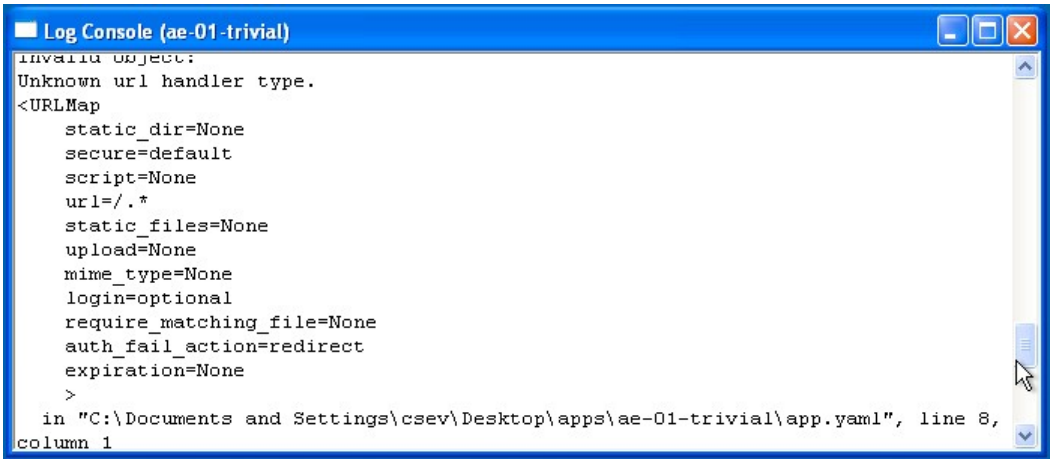
EachtimeyoupressRefreshinyourbrowser–
youcanseeitretrievingtheoutputwith aGETrequest.

DealingWithErrors

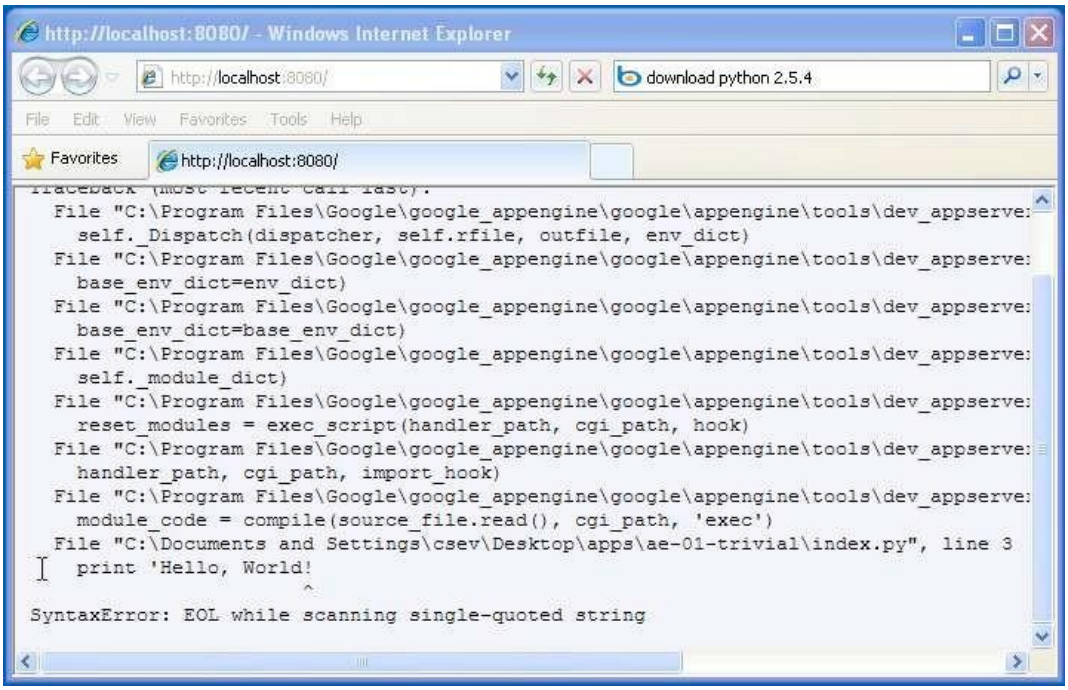
With two files to edit, there are two general categories of errors that you mayencounter. Ifyoumakeamistake on theapp.yamlfile, the App Enginewillnotstartandyourlauncherwillshowayellowiconnearyourapplication:



Togetmoredetailonwhatisgoingwrong,takealookat thelogfortheapplication:



Inthisinstance—themistakeismis-indentingthelastlineintheapp.yaml(line8). Ifyoumakeasyntaxerrorintheindex.pyfile,aPythontracebackerrorwillappearin yourbrowser.



Theerroryouneedtoseeislikelytobethelastfewlinesoftheoutput–inthiscase ImadeaPythonsyntaxerroronlineoneofourone-lineapplication. Reference:http://en.wikipedia.org/wiki/Stack_trace

Whenyoumakeamistakeintheapp.yamlfile– youmustthefixthemistakeandattempttostarttheapplicationagain.

Ifyoumakeamistakeinafilelikeindex.py,youcansimplyfixthefileandpressr efreshinyourbrowser–thereisnoneedtorestarttheserver.

ShuttingDowntheServer

Toshutdowntheserver,usetheLauncher,selectyourapplicationandpresstheStopbutton .

Result:

ThustheGAEwebapplicationswascreated.

6.Findaprocedureto transferthefilesfrom onevirtualmachinetoanothervirtualmachine.

Aim:

ToFindaproceduretotransferthefilesfromonevirtualmachine toanothervirtualmachine.

Steps:

1. Youcancopyfew(ormore)lineswith*copy&paste*mechanism.
   For this you need to share clipboard between host OS and guest OS, installingGuest Addition on both the virtual machines (probably setting *bidirectional*and restarting them). You *copy* from *guest OS* in the clipboard that is sharedwith the *hostOS*.
   Thenyou*paste*fromthe*host OS*tothesecond*guest OS*.
2. You can enable drag and drop too with the same method (Click on themachine,settings,general,advanced,draganddrop:setto*bidirectional*)
3. You can have common *Shared Folders* on both virtual machines anduseone ofthedirectorysharedas buffer tocopy.
   Installing Guest Additions you have the possibility to set Shared Folders too.As you put a file in a shared folder from *host OS* or from *guest OS*, isimmediatelyvisibletotheother.(Keepinmindthatcanarisesomeproblemsfor date/time of the files when there are different clock settings on thedifferentvirtualmachines).
   If you use the same folder shared on more machines you can exchange filesdirectlycopyingtheminthisfolder.
4. You can use usual method to copy files between 2 different computer withclient-server application. (e.g. scp with sshd active for linux, winscp...youcangetsome infoabout SSHserverse.g.here)
   Youneedanactiveserver(sshd)onthereceivingmachineandaclientonthe sending machine. Of course you need to have the authorization setted(viapassword or,better,viaanautomaticauthenticationmethod).
   Note: many Linux/Ubuntu distribution install sshd by default: you can see ifit is running with pgrep sshd from a shell. You can install with sudo apt-getinstall openssh-server.
5. You can mount part of the file system of a virtual machine via NFS

orSSHFS on the other, or you can share file and directory with Samba.YoumayfindinterestingthearticleSharingfilesbetweenguestand

host without VirtualBox shared folders with detailed step by stepinstructions.

Youshouldrememberthatyouarediallingwithalittlenetworkofmachineswith differentoperativesystems,andinparticular:

- Eachvirtualmachinehasitsownoperativesystemrunningonandactsas aphysicalmachine.
- Each virtual machine is an instance of a program *owned* by an *user* in thehostingoperativesystemandshouldundergotherestrictionsofthe*user*inthe*host ingOS*.
  E.g Let we say that Hastur and Meow are users of the hosting machine, butthey did not allow each other to see their directories (no read/write/executeauthorization). When each of them run a virtual machine, for the hosting OSthose virtual machine are two normal programs owned by Hastur and Meowand cannot see the private directory of the other user. This is a restriction dueto the *hosting OS*. It's easy to overcame it: it's enough to give authorization toread/write/executetoadirectoryortochoseadifferentdirectoryinwhichbothusers canread/write/execute.
- WindowslikesmouseandLinuxfingers.:-) ImeanIsuggestyouto enable*Drag&drop*tobecosywiththeWindowsmachines and the*Sharedfolders* ortobecosywithLinux.
  Whenyouwill needto befast withLinuxyouwillfeeltheneed ofssh-keygen and toGenerateonceSSHKeystocopyfileson/fromaremotemachinewithoutwritingpasswordanymo re.Inthiswayitfunctions bashauto-completionremotelytoo!

PROCEDURE:

Steps:

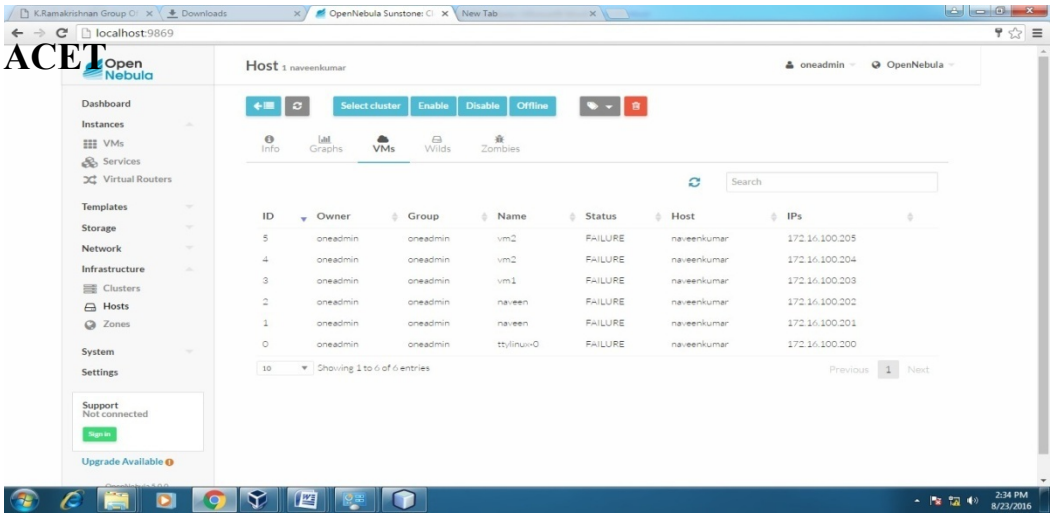
1. OpenBrowser,typelocalhost:9869
2. Loginusingusername:oneadmin,password:opennebula
3. ThenfollowthestepstomigrateVMs
   a. Clickoninfrastructure
   b. Selectclustersandenterthe clustername
   c. Thenselecthosttab,and selectallhost
   d. ThenselectVnetstab,andselectallvnet
   e. Thenselectdatastorestab,andselectalldatastores
   f. Andthenchoosehostunderinfrastructuretab
   g. Clickon+symboltoadd newhost,namethe hostthen clickoncreate.

4. oninstances,select VMstomigratethenfollowthestpes
    a. Clickon8th icon,thedropdownlistdisplay
    b. Selectmigrateonthat,thepopupwindowdisplay

c. Onthatselectthetargethosttomigratethenclickonmigrate.

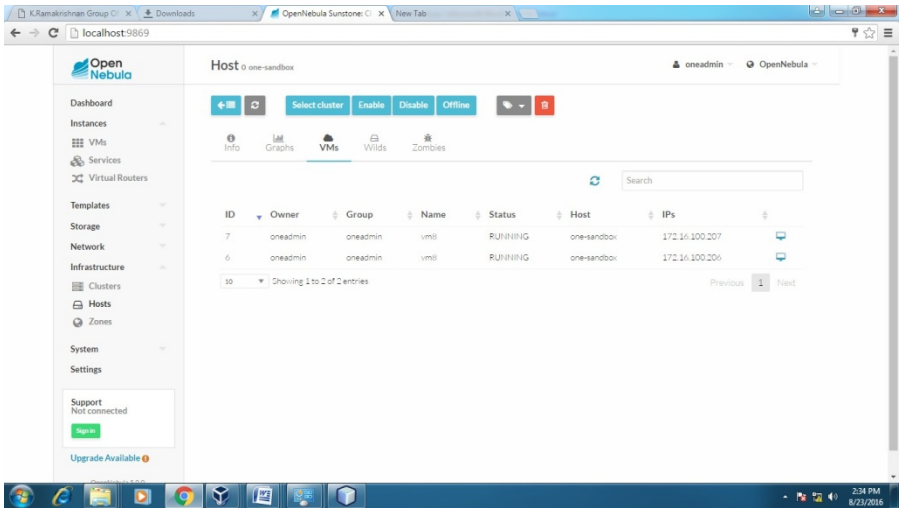
Before

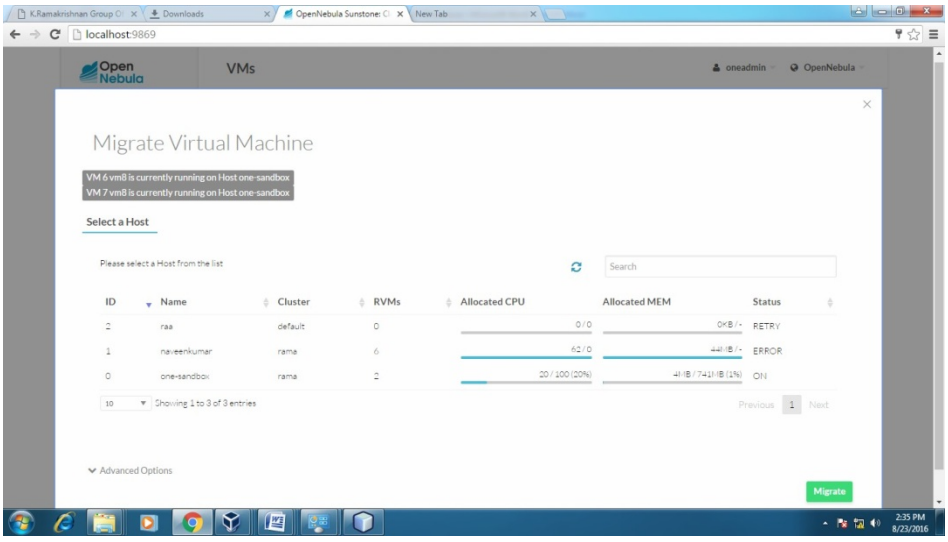
migrationHost:S

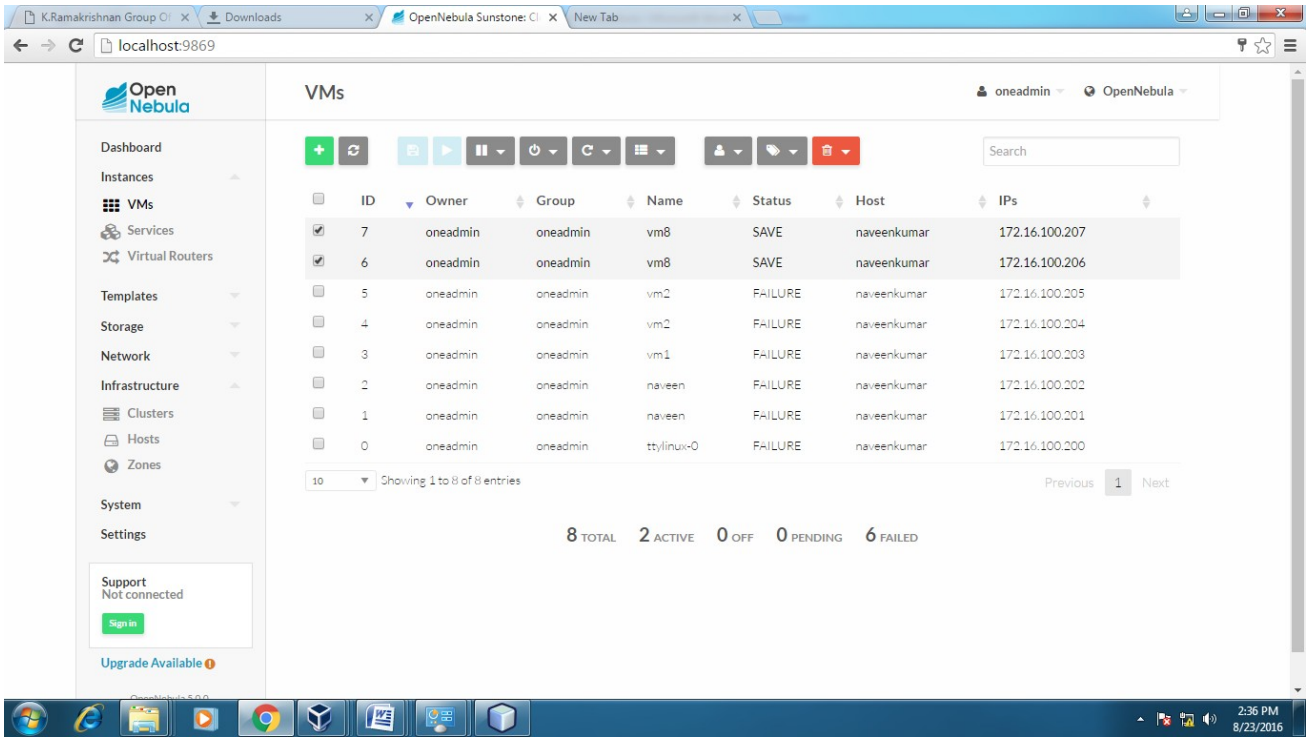


ACET

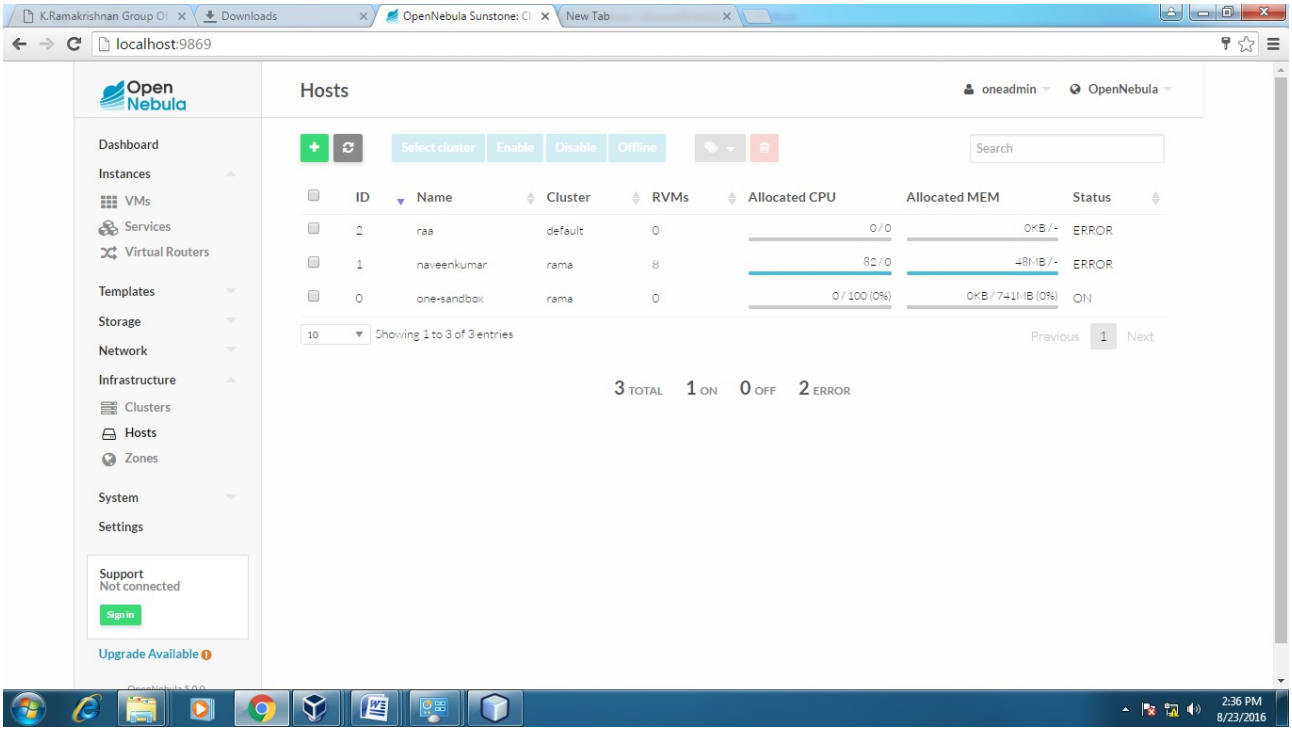
Host:one-sandbox


localhost:9869
Open Nebula
Host 0 one-sandbox
oneadmin
OpenNebula
Dashboard
Instances
VMs
Services
Virtual Routers
Templates
Storage
Network
Infrastructure
Clusters
Hosts
Zones
System
Settings
Support
Not connected
Sign in!
Upgrade Available
Select cluster
Enable
Disable
Offline
Info
Graphs
VMs
Wilds
Zombies
Search
ID
Owner
Group
Name
Status
Host
IPs
7 oneadmin oneadmin vm8 RUNNING one-sandbox 172.16.100.207
6 oneadmin oneadmin vm8 RUNNING one-sandbox 172.16.100.206
Showing 1 to 2 of 2 entries
Previous
1
Next
2:34 PM
8/23/2016

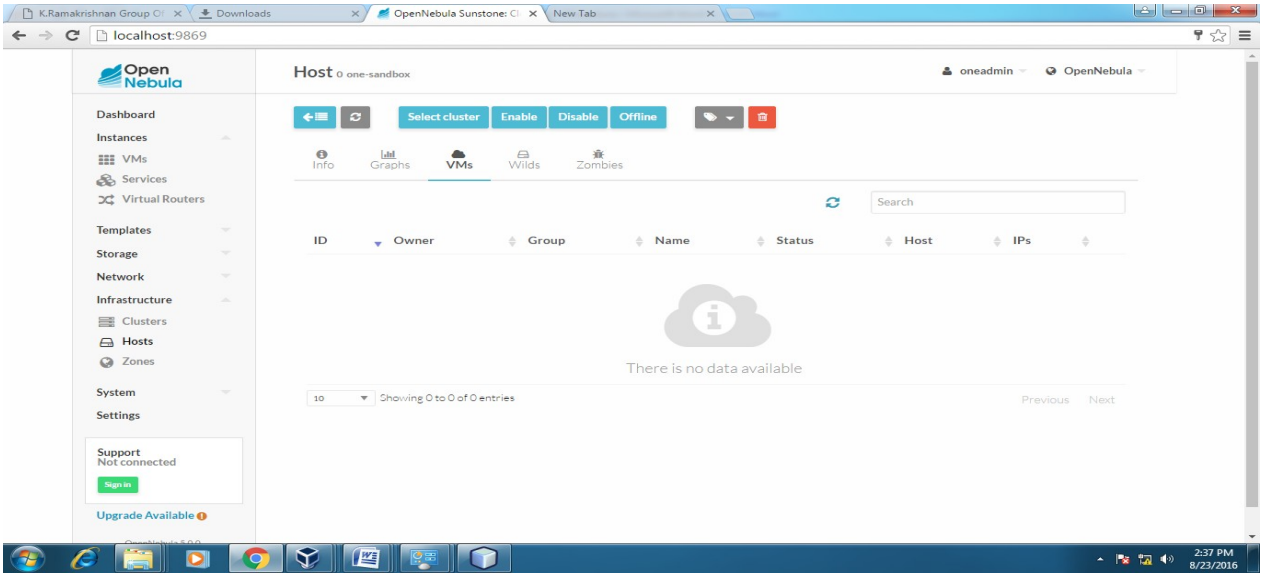

localhost:9869
Open Nebula
VMs
oneadmin
OpenNebula
Migrate Virtual Machine
VM 6 vm8 is currently running on Host one-sandbox
VM 7 vm8 is currently running on Host one-sandbox
Select a Host
Please select a Host from the list
Search
ID
Name
Cluster
RVMs
Allocated CPU
Allocated MEM
Status
2 raa default 0 0 / 0 0KB / - RETRY
1 naveenkumar rama 6 62 / 0 4MB / - ERROR
0 one-sandbox rama 2 20 / 100 (20%) 4MB / 7.41MB (1%) ON
Showing 1 to 3 of 3 entries
Previous
1
Next
Advanced Options
Migrate
2:35 PM
8/23/2016



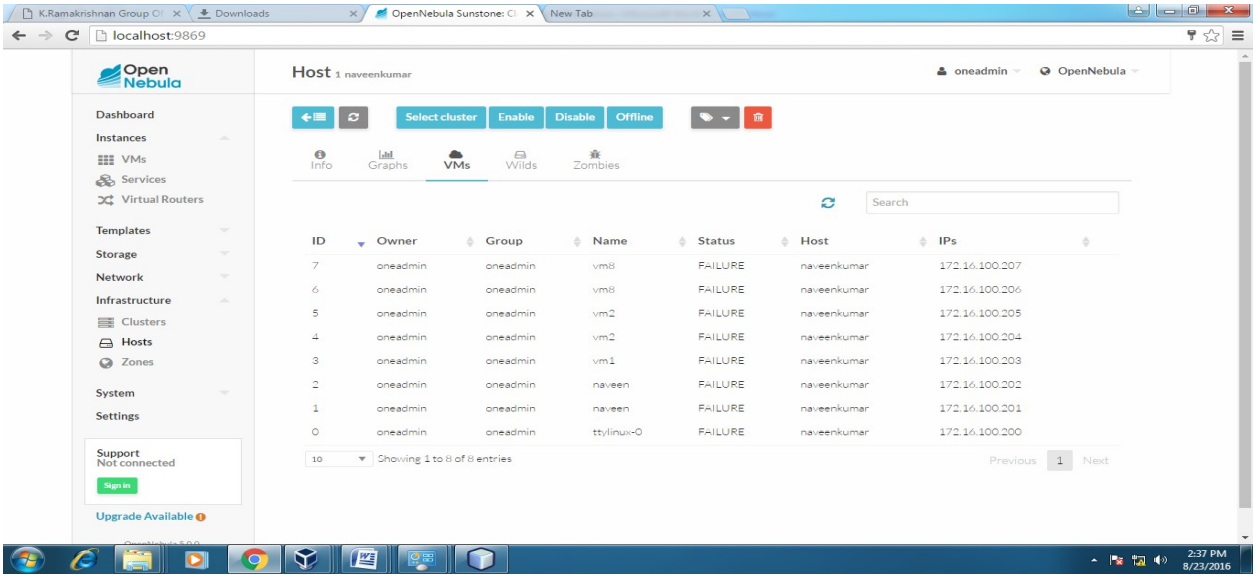
AfterMigration:



Host:one-sandbox



Host:SACET



APPLICATIONS:

Easilymigrateyourvirtualmachinefromonepctoanother.

Result:

ThusthefiletransferbetweenVMwassuccessfullycompleted…..

7.Find a procedure to launch virtual machine using trystack (OnlineOpenstackDemoVersion)

Aim:

ToFindaproceduretolaunchvirtual machineusingtrystack.

Steps:

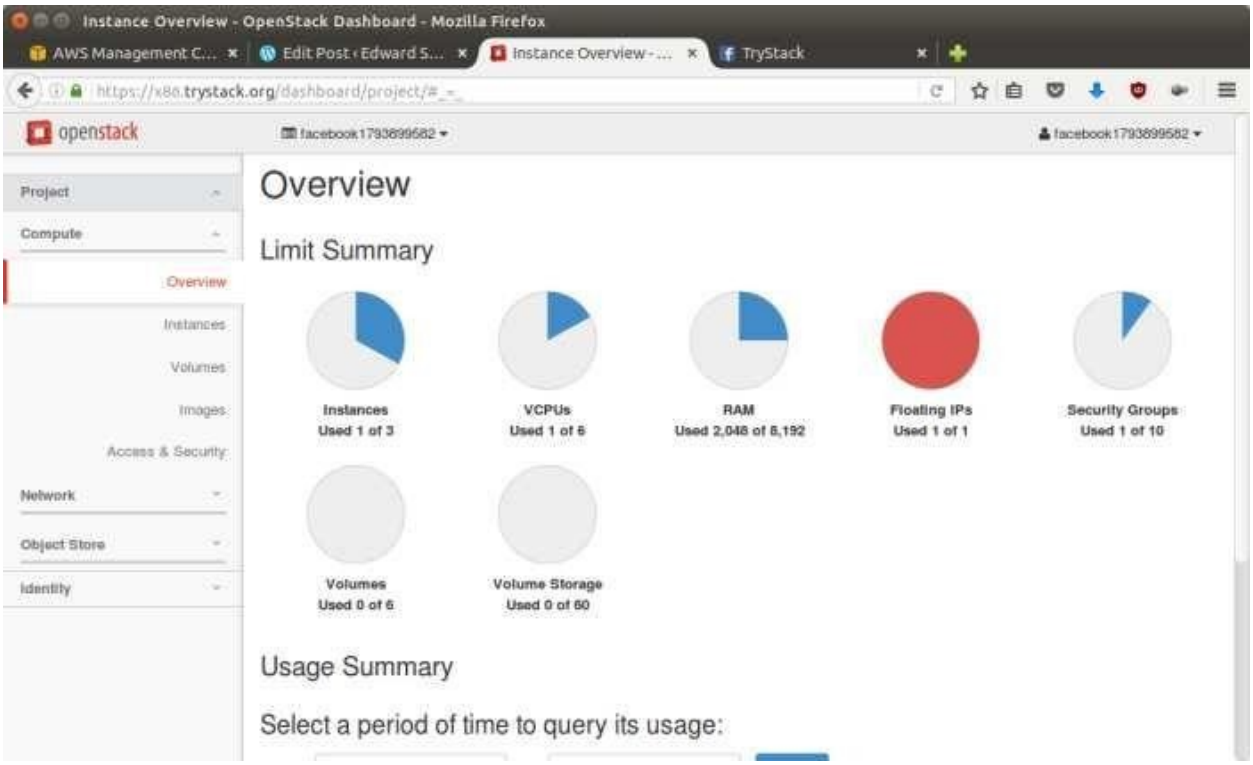
OpenStackisanopen-sourcesoftwarecloudcomputingplatform.

OpenStack is primarily used for deploying an infrastructure as a service (IaaS)solution like Amazon Web Service (AWS). In other words, you can *make yourown AWS* by using OpenStack. If you want to try out OpenStack, **TryStack** is theeasiest andfree waytodoit.

In order to try OpenStack in TryStack, you must register yourself by joiningTryStack Facebook Group. The acceptance of group needs a couple days becauseit's approved manually. After you have been accepted in the TryStack Group, youcanloginTryStack.

TryStack.orgHomepage

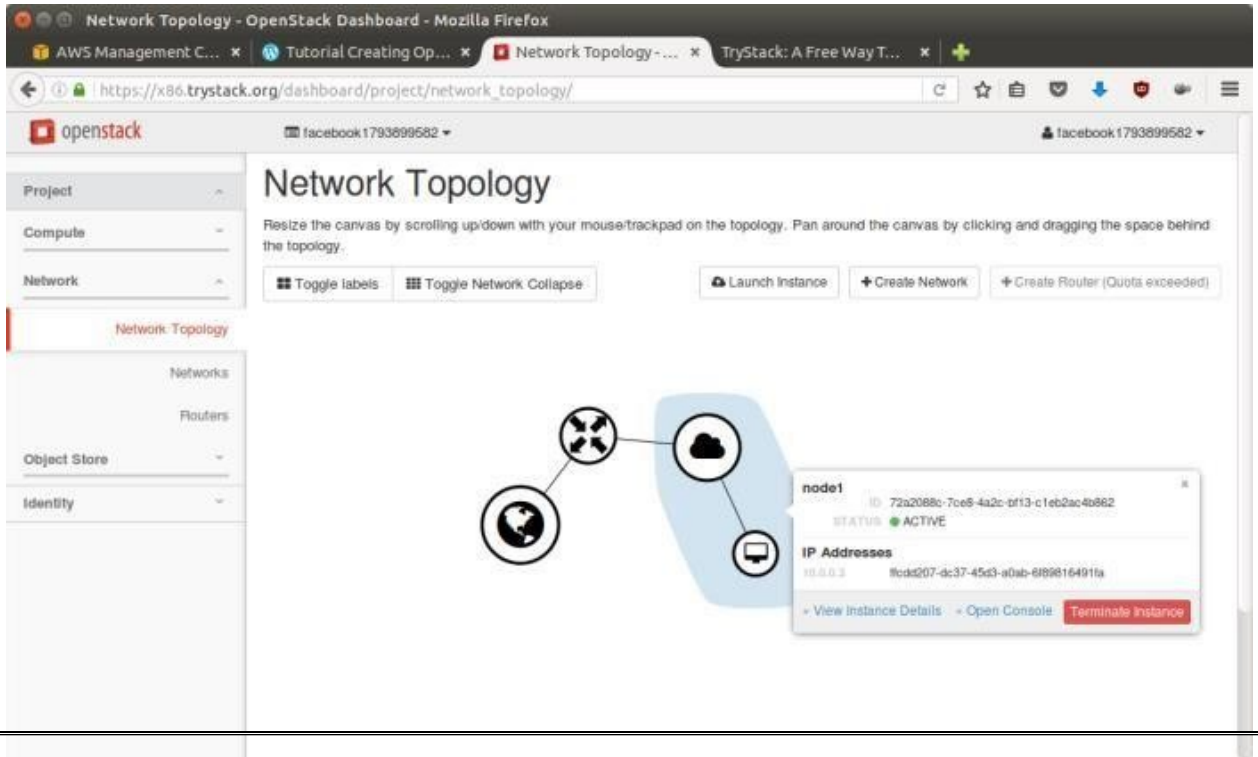
I assume that you already join to the Facebook Group and login to the dashboard.AfteryoulogintotheTryStack,youwillsee the ComputeDashboardlike:



OpenStackComputeDashboard

Overview:Whatwewilldo?

In this post, I will show you how to run an OpenStack instance. The instance willbe accessible through the internet (have a public IP address). The final topologywilllike:



Networktopology

As you see from the image above, the instance will be connected to a localnetwork and thelocalnetwork willbeconnected to internet.

Step 1:CreateNetwork

Network?Yes,thenetworkinhereisourownlocalnetwork.So,yourinstanceswill be not mixed up with the others. You can imagine this as your own LAN(Local Area Network) inthe cloud.

1. Goto**Network>Networks**andthen click**CreateNetwork**.
2. In**Network**tab,fill**NetworkName**forexampleinternalandthenclick**Next**.
3. In **Subnet**tab,
    1. Fill **Network Address** with appropriate CIDR, for example192.168.1.0/24.UseprivatenetworkCIDRblock asthebestpractice.
    2. Select**IPVersion**withappropriateIPversion,inthiscaseIPv4.
    3. Click **Next**.
4. In **Subnet Details** tab, fill **DNS Name Servers** with 8.8.8.8 (GoogleDNS)andthenclick**Create**.

Step 2:CreateInstance

Now, we will create an instance. The instance is a virtual machine in the cloud,like AWS EC2. You need the instance to connect to the network that we justcreatedinthe previousstep.

1. Goto**Compute>Instances**andthen click **LaunchInstance**.
2. In**Details**tab,
    1. Fill**InstanceName**,forexampleUbuntu1.
    2. Select**Flavor**,forexample m1.medium.
    3. Fill**InstanceCount**with**1**.
    4. Select**InstanceBootSource**with**BootfromImage**.
    5. Select**ImageName**with**Ubuntu14.04amd64(243.7MB)**ifyouwantinstallUbuntu 14.04inyourvirtualmachine.
3. In**Access& Security**tab,
    1. Click[+]button of **KeyPair**to import keypair.This keypairisapublicandprivatekeythatwewilluse toconnecttothe instance fromourmachine.
    2. In**ImportKeyPair**dialog,
        1. Fill**KeyPairName**withyourmachinename(forexampleEdward-Key).
        2. Fill **Public Key** with your **SSH public key** (usually is in ~/.ssh/id_rsa.pub). See description in Import Key Pair dialog box formore information. If you are using Windows, you can use **Puttygen**togenerate keypair.
        3. Click**Import key pair**.
    3. In**SecurityGroups**,mark/check**default**.
4. In**Networking**tab,

1. In**SelectedNetworks**,selectnetworkthathavebeencreatedinStep 1,forexampleinternal.

5. Click **Launch**.
6. If you want to create multiple instances, you can repeat step 1-5. I created one more instance with instance name Ubuntu2.

Step3: Create Router

I guess you already know what router is. In the step 1, we created our network, but it is isolated. It doesn't connect to the internet. To make our network has an internet connection, we need a router that running as the gateway to the internet.

1. Go to **Network > Routers** and then click **Create Router**.
2. Fill **Router Name** for example router1 and then click **Create router**.
3. Click on your **router name link**, for example router1, **Router Details** page.
4. Click **Set Gateway** button on in upper right:
   1. Select **External networks** with **external**.
   2. Then **OK**.
5. Click **Add Interface** button.
   1. Select **Subnet** with the network that you have been created in Step1.
   2. Click **Add interface**.
6. Go to **Network > Network Topology**. You will see the network topology. In the example, there are two network, i.e. external and internal, those are bridged by a router. There are instances those are joined to internal network.

Step4: Configure Floating IP Address

Floating IP address is public IP address. It makes your instance is accessible from the internet. When you launch your instance, the instance will have a private network IP, but no public IP. In OpenStack, the public Ips is collected in a pool and managed by admin (in our case is TryStack). You need to request a public (floating) IP address to be assigned to your instance.

1. Go to **Compute > Instance**.
2. In one of your instances, click **More > Associate Floating IP**.
3. In **IP Address**, click Plus [+].
4. Select **Pool** to **external** and then click **Allocate IP**.
5. Click **Associate**.
6. Now you will get a public IP, e.g. 8.21.28.120, for your instance.

Step5: Configure Access & Security

OpenStack has a feature like a firewall. It can whitelist/blacklist your in/out connection. It is called *Security Group*.

1. Go to **Compute > Access & Security** and then open **Security Groups** tab.
2. In **default** row, click **Manage Rules**.
3. Click **Add Rule**, choose **ALL ICMP** rule to enable ping into your instance, and then click **Add**.
4. Click **Add Rule**, choose **HTTP** rule to open HTTP port (port 80), and then click **Add**.

5. Click**AddRule**,choose**SSH**ruletoopenSSH port (port 22),andthen click**Add**.
6. Youcanopenotherportsbycreatingnewrules.

Step 6: SSH toYourInstance

Now,youcanSSH yourinstancestothefloatingIPaddressthatyougotinthestep 4.If you are using Ubuntuimage,theSSHuserwill beubuntu.

Result:

Thustheopenstackdemoworkedsuccessfully.

8.InstallHadoopsinglenodeclusterandrunsimpleapplicationslikewordcount.

Aim:

ToInstallHadoopsinglenodeclusterandrunsimple applicationslikewordcount.

Steps:

Install Hadoop

Step 1: Click hereto download the Java 8 Package. Save this file in yourhomedirectory.

Step2: ExtracttheJava TarFile.

***Command*:**tar-xvfjdk-8u101-linux-i586.tar.gz

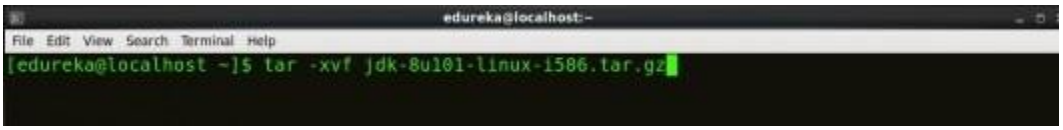


Fig:HadoopInstallation–ExtractingJavaFiles

Step3:**DownloadtheHadoop2.7.3Package.**

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

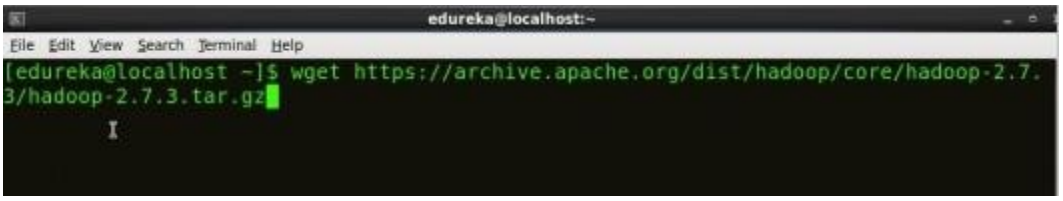


Fig:HadoopInstallation–DownloadingHadoop

Step4: ExtracttheHadooptarFile.

Command:tar-xvfhadoop-2.7.3.tar.gz

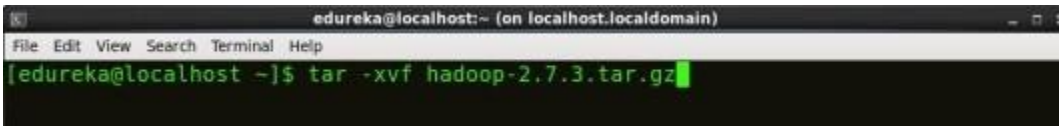
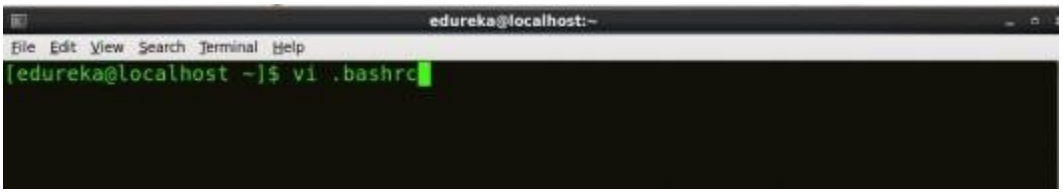


Fig: Hadoop Installation – Extracting Hadoop

*Files***Step 5:** Add the Hadoop and Java paths in the bash file

(.bashrc).Open**.bashrc**file.Now,addHadoopandJavaPathasshownbelow.

***Command*:**vi .bashrc



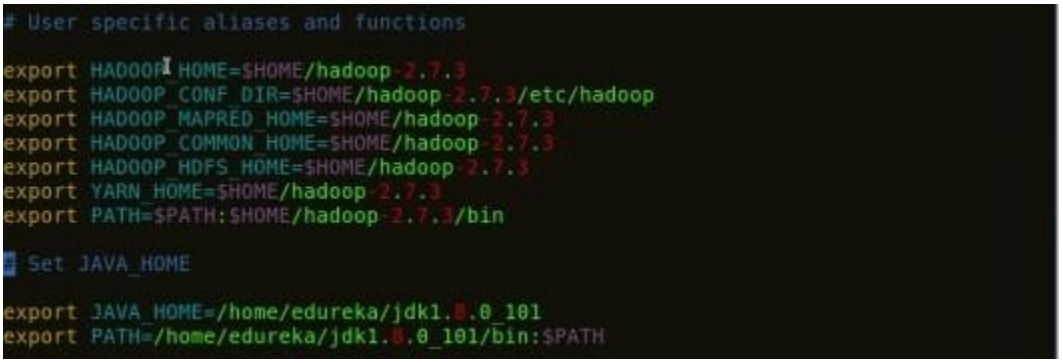


Fig:HadoopInstallation –SettingEnvironmentVariable

Then,savethebashfileandcloseit.

ForapplyingallthesechangestothecurrentTerminal,executethesourcecommand.

***Command*:**source.bashrc

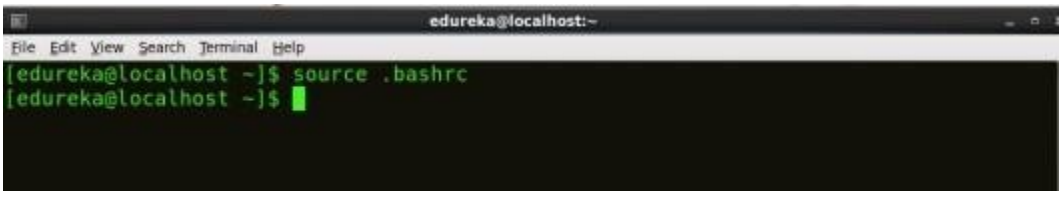


Fig:HadoopInstallation –Refreshingenvironmentvariables

TomakesurethatJavaandHadoophavebeenproperlyinstalledonyoursystemandcanbeaccessedthroughtheTerminal,executethejava-versionandhadoop

versioncommands.

Command:java-version

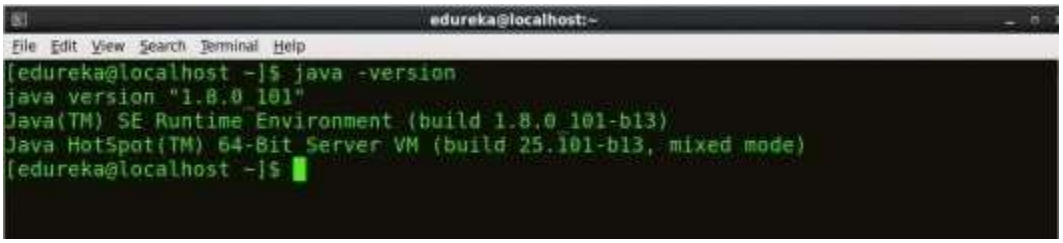


Fig:HadoopInstallation–CheckingJavaVersion

Command:hadoopversion

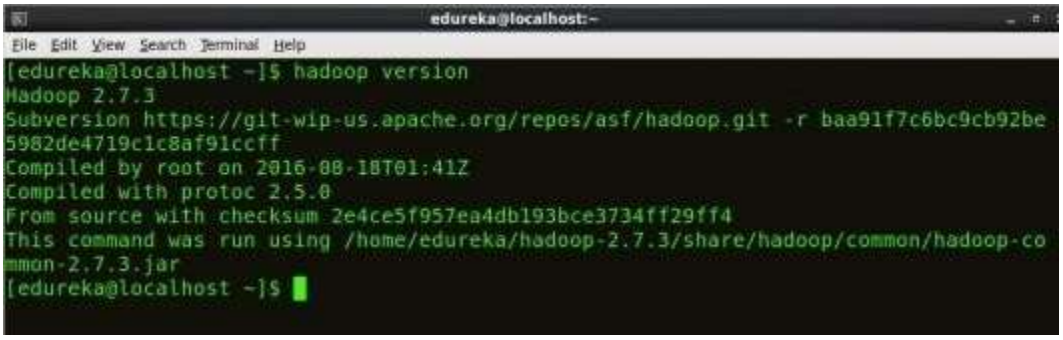


Fig:HadoopInstallation–CheckingHadoopVersion

Step 6: Editthe**HadoopConfigurationfiles**.

Command:cdhadoop-2.7.3/etc/hadoop/

Command:ls

All the Hadoop configuration files are located in **hadoop-2.7.3/etc/hadoop** directory asyoucan see inthe snapshotbelow:

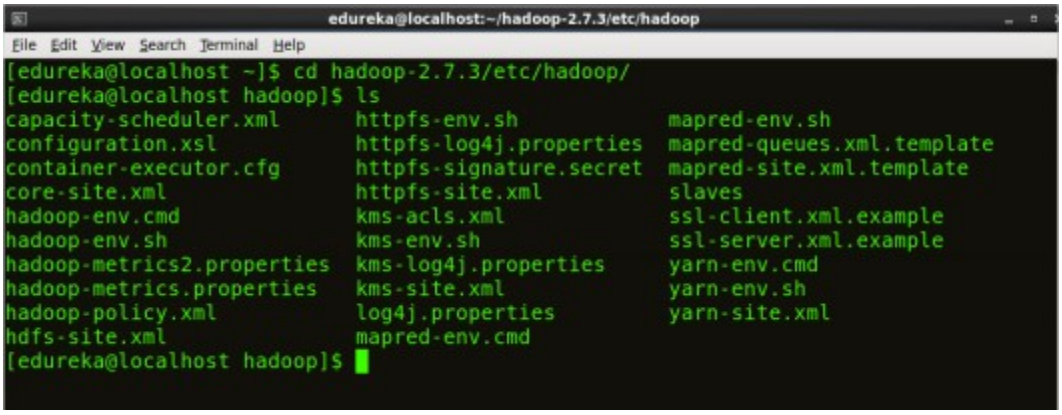


Fig:HadoopInstallation–HadoopConfigurationFiles

Step 7: Open *core-site.xml* and edit the property mentioned below insideconfigurationtag:

core-site.xml informs Hadoop daemon where NameNode runs in the cluster. It containsconfiguration settings of Hadoop core such as I/O settings that are common to HDFS &MapReduce.

***Command*:**vi core-site.xml

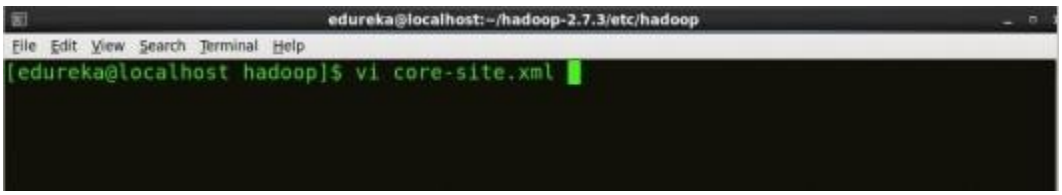


Fig:HadoopInstallation– Configuringcore-site.xml

```
<?xmlversion="1.0"encoding="UTF-8"?>
<?xml-stylesheettype="text/xsl"href="configuration.xsl"?>
<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>
```

Step8: Edit*hdfs-site.xml*andeditthepropertymentionedbelowinside **configurationtag:**

*hdfs-site.xml*containsconfigurationsettingsofHDFSdaemons(i.e.NameNode,DataNode, Secondary NameNode). It also includes the replication factor and block sizeof HDFS.

***Command*:**vihdfs-site.xml

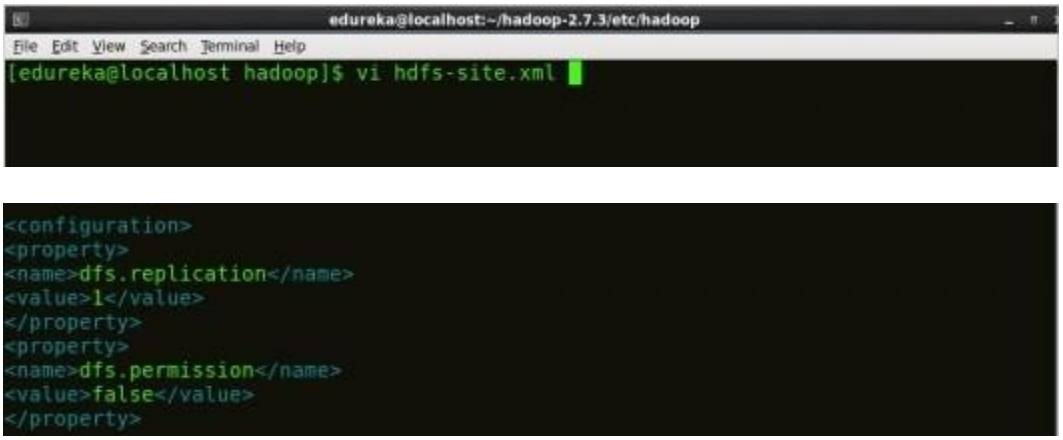


Fig:HadoopInstallation–Configuringhdfs-site.xml

```
1
2                <?xmlversion="1.0"encoding="UTF-8"?>
3    <?xml-stylesheettype="text/xsl"href="configuration.xsl"?>
4                          <configuration>
5                            <property>
6                     <name>dfs.replication</name>
7                           <value>1</value>
8                            </property>
9                            <property>
10                    <name>dfs.permission</name>
11                       <value>false</value>
                             </property>
                           </configuration>
```

Step9: Editthe*mapred-site.xml*fileandedittheproperymentionedbelow

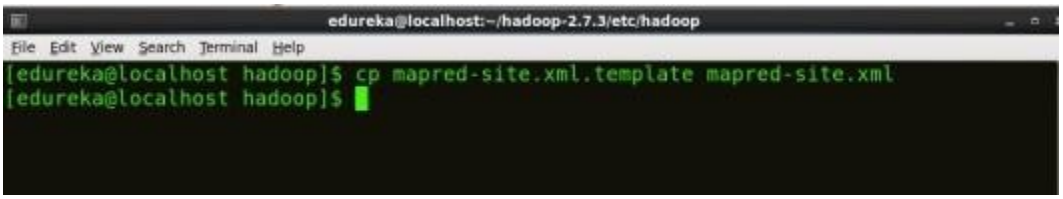
insideconfigurationtag:

*mapred-site.xml*containsconfigurationsettingsofMapReduceapplicationlikenumberof JVM that can run in parallel, the size of the mapper and the reducer process, CPUcoresavailable for aprocess,etc.

In some cases, mapred-site.xml file is not available. So, we have to create the mapred-site.xmlfileusing mapred-site.xml template.

***Command*:**cpmapred-site.xml.templatemapred-site.xml

***Command*:**vimapred-site.xml.



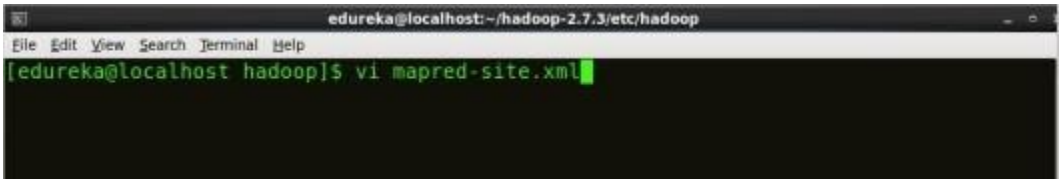


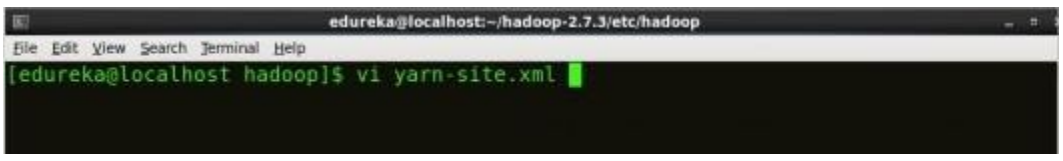
Fig:HadoopInstallation–Configuringmapred-site.xml

```
<?xmlversion="1.0"encoding="UTF-8"?>
<?xml-stylesheettype="text/xsl"href="configuration.xsl"?>
<configuration>
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
</configuration>
```

Step10: Edit *yarn-site.xml*andedit thepropertymentionedbelowinsideconfigurationtag:

yarn-site.xml contains configuration settings of ResourceManager and NodeManagerlikeapplicationmemorymanagementsize,theoperationneededonprogram& algorithm,etc.

***Command*:**vi yarn-site.xml



```
<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>
<value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
</configuration>
```

Fig:HadoopInstallation–Configuringyarn-site.xml

```
<?xmlversion="1.0">
<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</
name>
<value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
</configuration>
```

Step11: Edit*hadoop-env.sh*and addtheJavaPathas mentionedbelow:

hadoop-env.sh contains the environment variables that are used in the script to runHadooplike Java home path,etc.

***Command*:**vihadoop–env.sh

```
edureka@localhost:~/hadoop-2.7.3/etc/hadoop
File Edit View Search Terminal Help
[edureka@localhost hadoop]$ vi hadoop-env.sh
```

```
# The java implementation to use.
export JAVA_HOME=/home/edureka/jdk1.8.0_101
```

*Fig: Hadoop Installation – Configuring hadoop-env.sh***Step 12:** Go to Hadoop home directory and format the NameNode.***Command*:**cd

***Command*:**cdhadoop-2.7.3

***Command*:**bin/hadoopnamenode-format

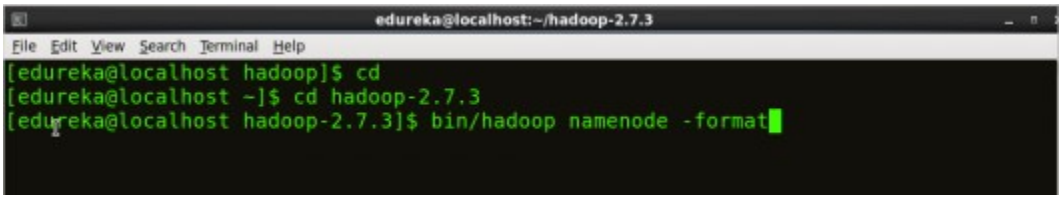


Fig:HadoopInstallation–FormattingNameNode

ThisformatstheHDFSviaNameNode.Thiscommandisonlyexecutedforthefirsttime. Formatting the file system means initializing the directory specified by thedfs.name.dirvariable.

Never format, up and running Hadoop filesystem. You will lose all your data stored intheHDFS.

Step 13: **Once the NameNode is formatted, go to hadoop-2.7.3/sbin directory and start all thedaemons.**

Command:cdhadoop-2.7.3/sbin

Eitheryoucan startalldaemonswithasinglecommandordoitindividually.

Command:./start-all.sh

The above command is a combination of ***start-dfs.sh, start-yarn.sh*** &***mr-jobhistory-daemon.sh***

Oryoucanrunalltheservicesindividuallyasbelow:

StartNameNode:

The NameNode is the centerpiece of an HDFS file system. It keeps the directory tree ofallfiles stored inthe HDFS andtracks allthefile storedacrossthecluster.

Command:./hadoop-daemon.shstartnamenode

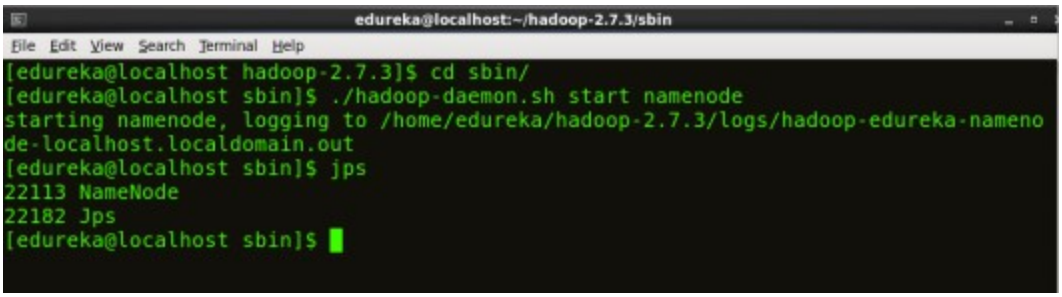


Fig:HadoopInstallation–StartingNameNode

StartDataNode:

Onstartup,aDataNodeconnectstotheNamenodeanditrespondstotherequestsfrom the Namenodefor differentoperations.

Command:./hadoop-daemon.shstartdatanode

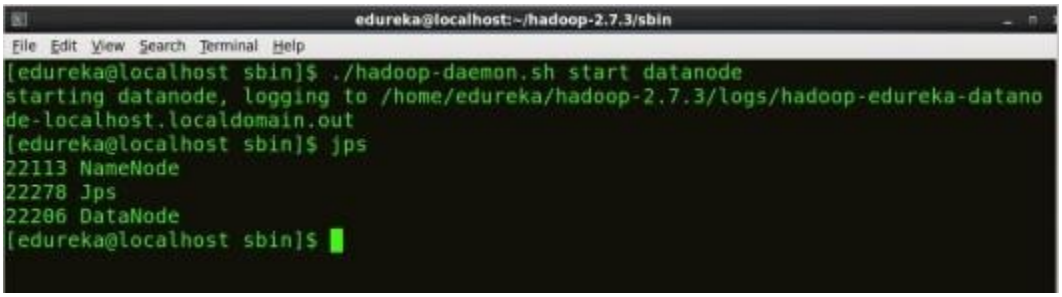


Fig:HadoopInstallation–StartingDataNode

StartResourceManager:

ResourceManager is the master that arbitrates all the available cluster resources andthushelpsinmanagingthedistributedapplicationsrunningontheYARNsystem.Itsworkist omanageeachNodeManagersandtheeachapplication'sApplicationMaster.

Command:./yarn-daemon.shstartresourcemanager

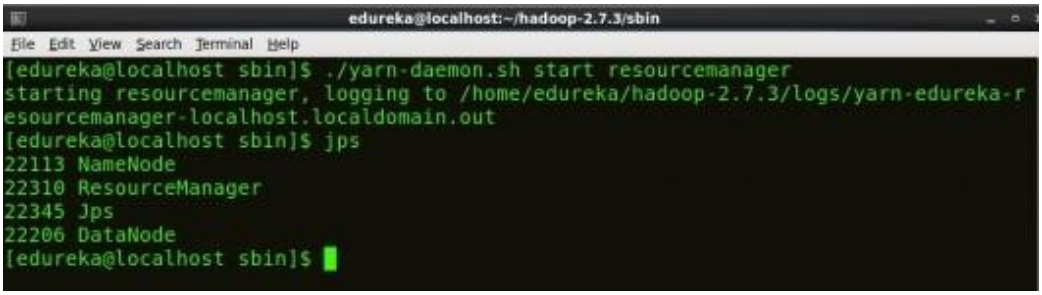


Fig:HadoopInstallation–StartingResourceManager

StartNodeManager:

The NodeManager in each machine framework is the agent which is responsible formanaging containers, monitoring their resource usage and reporting the same to theResourceManager.

Command:./yarn-daemon.shstartnodemanager

SeeBatchDetails

Fig:HadoopInstallation–StartingNodeManager

StartJobHistoryServer:

JobHistoryServerisresponsiblefor servicingalljobhistoryrelatedrequestsfromclient.

Command:./mr-jobhistory-daemon.shstarthistoryserver

Step14:**Tocheckthatallthe Hadoopservicesareup andrunning,runthebelowcommand.**

Command:jps

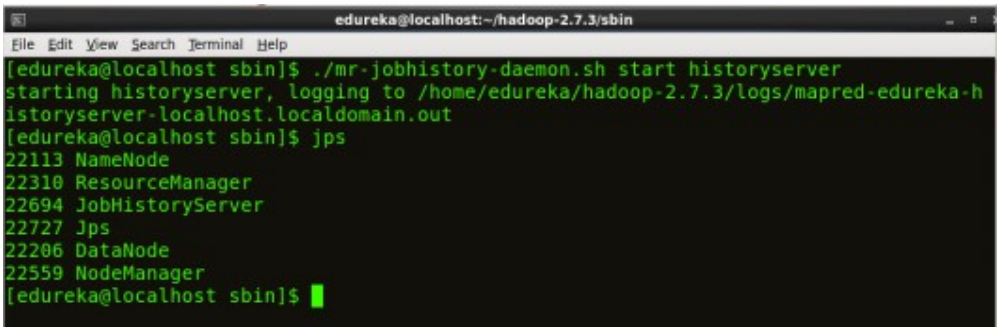


Fig:HadoopInstallation–Checking Daemons

Step15: Now opentheMozillabrowserandgo to**localhost**:**50070/dfshealth.html**tocheckttheNameNodeinterface.

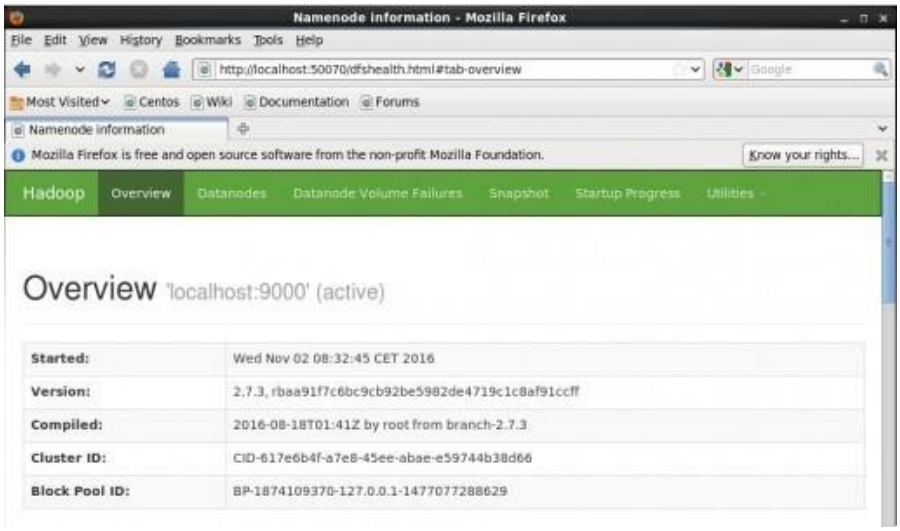


Fig:HadoopInstallation–StartingWebUI

Congratulations,youhavesuccessfullyinstalledasinglenodeHadoopcluster

Result:

Thus the Hadoop one cluster was installed and simple applications executedsuccessfully.

VIVAQUESTIONSANDANSWERS

1. DefineCloudComputingwithexample.

Cloudcomputingis amodel forenabling convenient,on-demand network access toa shared pool of configurable computing resources (e.g., networks, servers, storage, applications,and services) that can be rapidly provisioned and released with minimal management effort orserviceprovider interaction.

2. Whatis the working principleofCloud Computing?

The cloud is a collection of computers and servers that are publicly accessible via theInternet. This hardware is typically owned and operated by a third party on a consolidatedbasis in one or more data center locations. The machines can run any combination ofoperatingsystems.

3. Whatarethe advantagesanddisadvantages ofCloudComputing?

Advantages

Lower-Cost Computers for UsersImprovedPerformance
Lower IT Infrastructure CostsFewer Maintenance IssuesLowerSoftwareCosts
Instant Software UpdatesIncreased
Computing PowerUnlimited
Storage CapacityIncreasedData
Safety
Improved Compatibility Between Operating SystemsImprovedDocument Format Compatibility
Easier Group CollaborationUniversal Access to DocumentsLatestVersionAvailability
Removes the Tether to Specific DevicesDisadvantages

RequiresaConstant InternetConnection
Doesn't Work Well with Low-Speed ConnectionsCanBeSlow
FeaturesMightBeLimitedStoredDataMightNotBeSecure
IftheCloudLosesYour Data,You'reScrewed

4. Whatisdistributedsystem?

A*distributedsystem*isasoftwaresysteminwhichcomponentslocatedonnetworkedcomputerscommunicate and coordinate their actions by passing messages. The components interact with each other inorderto achieve acommon goal.

Threesignificantcharacteristicsofdistributedsystemsare:

- ✓ Concurrencyofcomponents
- ✓ Lackofaglobalclock
- ✓ Independentfailureofcomponents
- ✓ Whatiscluster?
- ✓ Acomputingclusterconsistsofinterconnectedstand-alonecomputerswhichworkcooperativelyasasingleintegratedcomputingresource.Inthepast,clustere

dcomputersystemshavedemonstrated

5. Whatisgridcomputing?

Grid Computing enables virtuals organizations to share geographically distributed resources as theypursue common goals, assuming the absence of central location, central control, omniscience, and anexistingtrust relationship.

(or)

- ✓ Gridtechnologydemandsnewdistributedcomputingmodels,software/middlewaresupport,networkprotocols,andhardwareinfrastructures.
- ✓ NationalgridprojectsarefollowedbyindustrialgridplatformdevelopmentbyIBM,Microsoft,Sun,HP,Dell,Cisco,EMC,PlatformComputing,andothers.Newgridserviceproviders(GSPs)andnewgridapplicationshaveemergedrapidly,similartothegrowthofInternetandwebservicesinthepasttwodecades.
- ✓ gridsystemsareclassifiedinessentiallytwocategories:computationalordatagridsandP2Pgrids.6.

Whatarethe business areas needs in Grid computing?

- ✓ LifeSciences
- ✓ Financialservices
- ✓ HigherEducation
- ✓ EngineeringServices
- ✓ Government
- ✓ Collaborativegames

7.ListouttheGridApplications:

- ✓ Applicationpartitioningthatinvolves breakingthe problem intodiscretepieces
- ✓ Discoveryandscheduling oftasks andworkflow
- ✓ Datacommunicationsdistributingtheproblemdatawhereand whenitisrequired
- ✓ Provisioninganddistributingapplicationcodestospecificsystemnodes
- ✓ Autonomic features such as self-configuration, self-optimization, self-recovery and self-management

8.Listsomegridcomputing toolkitsandframeworks?

- ✓ GlobusToolkitGlobusResourceAllocationManager(GRAM)
- ✓ GridSecurityInfrastructure(GSI)
- ✓ InformationServices
- ✓ Legion,CondorandCondor-G
- ✓ NIMROD,UNICORE,NMI.9.

WhatareDesktop Grids?

Thesearegridsthatleveragethecompute resourcesofdesktopcomputers.
Because of the true (but unfortunate) ubiquity of Microsoft® Windows® operatingsystem in corporations, desktop grids are assumed to apply to the Windows environment.TheMacOS™environmentissupported byalimited numberofvendors.

10. WhatareServerGrids?

- ✓ Somecorporations,whileadoptingGridComputing,keepitlimitedtoserverresourcesthatarewithinthe purviewof the IT department.
- ✓ Specialservers,insomecases,areboughtsolelyforthepurposeofcreatinganinternal“utilitygrid”withresources madeavailable to variousdepartments.
- ✓ Nodesktopsareincludedinservergrids.TheseusuallyrunsomeflavoroftheUnix/Linuxoperatingsystem.

11. DefineOpennebula.

OpenNebula is an open source management tool that helps virtualized data centers oversee private clouds,publicclouds andhybridclouds OpenNebulaisvendorneutral,aswellasplatform-andAPI-agnostic.It
can use KVM, Xen or VMware
hypervisors.12.DefineEclipse.

Eclipse is an integrated development environment (IDE) used in computer programming, and is the mostwidely used Java IDE. It contains a base workspace and an extensible plug-in system for customizing theenvironment.

13. DefineNetbeans.

NetBeans is an open-source integrated development environment (IDE) for developing with Java, PHP,C++, and other programming languages. NetBeans is also referred to as a platform of modular componentsusedfordeveloping Javadesktop applications.

14. DefineApache Tomcat.

Apache Tomcat (or Jakarta Tomcat or simply Tomcat) is an open source servlet container developed bytheApacheSoftwareFoundation(ASF).TomcatimplementstheJavaServletandtheJavaServerPages (JSP) specifications from Sun Microsystems, and provides a "pure Java" HTTP web server environment forJavacodetorun."

15. Whatisprivatecloud?

The *privatecloud* isbuiltwithinthedomainofanintranetownedbyasingleorganization.Therefore, they are client owned and managed. Their access is limited to the owning clients and theirpartners. Their deployment was not meant to sell capacity over the Internet through publicly accessibleinterfaces.Privatecloudsgivelocalusersaflexibleandagileprivateinfrastructuretorunserviceworkloadswithin their administrativedomains.

16. Whatispubliccloud?

A *public cloud* is built over the Internet, which can be accessed by any user who has paid for theservice. Public clouds are owned by service providers. They are accessed by subscription. Many companieshave built public clouds, namely Google App Engine, Amazon AWS, Microsoft Azure, IBM Blue Cloud,and Salesforce Force.com. These are commercial providers that offer a publicly accessible remote interfaceforcreating and managing VM instanceswithin theirproprietary infrastructure.

17. Whatis hybrid cloud?

A *hybridcloud* isbuiltwithbothpublicandprivateclouds, Privatecloudscanalsosupporta *hybrid cloud* model by supplementing local infrastructure with computing capacity from an externalpubliccloud. Forexample, the*researchcomputecloud* (RC2) is aprivatecloudbuilt by IBM.

18. Whatis a CommunityCloud?

A community cloud in computingis a collaborative effort in which infrastructure is shared betweenseveralorganizationsfromaspecificcommunitywithcommonconcerns(security,compliance,jurisdiction,etc.),whethermanagedinternallyorbyathird-partyandhostedinternallyorexternally.Thisiscontrolledandusedbyagroupoforganizationsthathavesharedinterest.Thecostsarespreadoverfewerusers thana publiccloud (but morethan aprivatecloud

19. DefineIaaS?

The IaaS layer offers storage and infrastructure resources that is needed to deliver the Cloudservices.Itonlycomprisesoftheinfrastructureorphysicalresource.TopIaaSCloudComputingCompanies: Amazon (EC2), Rackspace,GoGrid,Microsoft, TerremarkandGoogle.

20. DefinePaaS?

PaaS provides the combination of both, infrastructure and application. Hence, organisationsusing PaaSdon't have to worry for infrastructure nor for services. Top PaaS Cloud ComputingCompanies:Salesforce.com, Google, Concur Technologies, Ariba,Unisys and Cisco..

21. DefineSaaS?

In the SaaS layer, the Cloud service provider hosts the software upon their servers. It can be definedas a in model in which applications and softwares are hosted upon the server and made available tocustomersoveranetwork.TopSaaSCloudComputingCompanies:AmazonWebServices,AppScale,CA Technologies, Engine Yard, SalesforceandWindows Azure.

22. Whatismeantbyvirtualization?

Virtualizationisacomputerarchitecturetechnologybywhichmultiplevirtualmachines

(VMs)aremultiplexedinthesamehardwaremachine.TheideaofVMscanbedatedbacktothe1960s.Thepurposeofa VMistoenhanceresourcesharingbymanyusersandimprovecomputerperformanceintermsof

resourceutilizationandapplicationflexibility.

23. Whataretheimplementationlevelsofvirtualization?

Thevirtualizationtypes arefollowing

1. OS-levelvirtualization
2. ISAlevelvirtualization
3. User-ApplicationLevelvirtualization
4. hardwarelevelvirtualization
5. library level

virtualization24.Listtherequiremen
tsofVMM?

TherearethreerequirementsforaVMM.

First,aVMMshouldprovideanenvironmentforprogramswhichisessentiallyidenticaltotheoriginalmachine.

Second, programs run in this environment should show, at worst, only minor decreases in speed.Third,aVMM should bein completecontrol of thesystem resources.

25. ExplainHostOSand GuestOS?

Acomparisonofthedifferencesbetweenahostsystem,aguestsystem,andavirtualmachinewithinavirtual infrastructure.

Ahostsystem(hostoperatingsystem)wouldbetheprimary&firstinstalledoperatingsystem.Ifyouareusingabaremetal Virtualization platformlikeHyper-VorESX,thereareallyisn't ahostoperating system besides the Hypervisor. If you are using a Type-2 Hypervisor like VMware Server orVirtual Server, the host operating system is whatever operating system those applications are installedinto.

A guest system (guest operating system) is a virtual guest or virtual machine (VM) that is installedunderthehost operating system.Theguests aretheVMsthat you run inyour virtualization platform.

26. Writethe stepsforliveVMmigration?

Thefivesteps for liveVM migration is

Stage0: *Pre-Migration*
ActiveVMon HostA
Alternate physical host may be preselected for migrationBlock devices mirrored and free resources maintainedStage1:*Reservation*
Initialize a container on the target hostStage2:*Iterativepre-copy*
Enableshadowpaging
Copy dirty pages in successive rounds.Stage3:*Stop and copy*
SuspendVM on host A
Generate ARP to redirect traffic to Host
BSynchronize all remaining VM state to Host
BStage4:*Commitment*
VMstateonHost Ais released
Stage5: *Activation*
VM starts on Host
BConnects to local devicesResumesnormalop eration

27..DefineGlobusToolkit:GridComputingMiddleware

✓ Globusisopensourcegridsoftwarethataddressesthemostchallengingproblmesindistributedresourcessharing.

- ✓ TheGlobusToolkitincludessoftwareservicesandlibrariesfordistributedsecurity,resourcemanagement, monitoring and discovery,and datamanagement.

28. DefineBlocksinHDFS

- ✓ Adiskhasablocksize,whichistheminimumamountofdatathatitcanreadorwrite.Filesystemsfora single disk build on this by dealing with data in blocks, which are an integral multiple of the diskblock size. Filesystem blocks are typically a few kilobytes in size, while disk blocks are normally 512bytes. This is generally transparent to the filesystem user who is simply reading or writing a file—ofwhateverlength.

29. DefineNamenodes andDatanodes

- ✓ AnHDFS clusterhas twotypes ofnodeoperating ina master-worker pattern:
  - ➢ a*namenode* (themaster) and
  - ➢ anumberof*datanodes*(workers).
- ✓ The namenode manages the filesystem namespace. It maintains the filesystem tree and the metadatafor all the files and directories in the tree. This information is stored persistently on the local disk intheform oftwo files: thenamespaceimageand theeditlog.
- ✓ The namenode also knows the datanodes on which all the blocks for a given file are located,however, it does not store block locations persistently, since this information is reconstructed fromdatanodeswhen the system starts.

30. DefineHADOOP.

Hadoop is an open source, Java-based programming framework that supports the processing and storage ofextremely large data sets in a distributed computing environment. It is part of the Apache project sponsoredbythe ApacheSoftwareFoundation.

31. DefineHDFS.

Hadoop Distributed File System (HDFS) is a Java-based file system that provides scalable and reliable datastorage that is designed to span large clusters of commodity servers. HDFS, MapReduce, and YARN formthe coreof Apache™ Hadoop®.

32. Writeabout HADOOP.

Hadoop was created by Doug Cutting and Mike Cafarella in 2005. Cutting, who was working at Yahoo! atthe time, named it after his son's toy elephant. It was originally developed to support distribution for theNutchsearch engine project.

33. Definitionof *Grid Portal*:

A *GridPortal* providesanefficientinfrastructuretoputGrid-empoweredapplicationsoncorporateIntranet/Internet.

34. DefineGAE.

Google App Engine (often referred to as GAE or simply App Engine) is a Platform as a Serviceand cloudcomputingplatformfordevelopingand hostingwebapplicationsin Google-manageddata centers.
Applications are sandboxedand run across multiple servers. App Engine offers automatic scaling for webapplications—as the number of requests increases for an application, App Engine automatically allocatesmoreresources for theweb application to handletheadditional demand.

35. WhatisCloudsim?

CloudSim is a simulation toolkit that supports the modeling and simulation of the core functionality ofcloud, like job/task queue, processing of events, creation of cloud entities(datacenter, datacenter brokers,etc), communication between different entities, implementation of broker policies, etc. This toolkit allowsto:

- Testapplicationservices inarepeatableand controllableenvironment.
- Tunethesystem bottlenecksbeforedeployingapps inanactualcloud.
- Experiment with different workload mix and resource performance scenarios on simulatedinfrastructurefordeveloping andtesting adaptiveapplication provisioningtechniques

36. CorefeaturesofCloudSimare:

- The Support of modeling and simulation of large scale computing environment as federated clouddatacenters,virtualizedserverhosts,withcustomizablepoliciesforprovisioninghostresourcestovirtualmachines andenergy-aware computational resources
- It is a self-contained platform for modeling cloud's service brokers, provisioning, and allocationpolicies.
- Itsupportsthesimulationofnetworkconnectionsamongsimulatedsystemelements.
- Support for simulation of federated cloud environment, that inter-networks resources from bothprivateand publicdomains.
- Availability of a virtualization engine that aids in the creation and management of multipleindependentandco-hosted virtual services onadata centernode.
- Flexibility to switch between space shared and time shared allocation of processing cores tovirtualizedservices.

37. UsesofCloudsim.

- LoadBalancingofresourcesandtasks
- Taskschedulingandits migrations
- OptimizingtheVirtualmachineallocationandplacementpolicies
- Energy-awareConsolidationsorMigrationsofvirtualmachines
- Optimizingschemesfor Networklatenciesforvariouscloudscenarios

38. DefineOpenStack.

OpenStack is a cloud operating system that controls large pools of compute, storage, and networkingresources throughout a datacenter, all managed and provisioned through APIs with common authenticationmechanisms.

Adashboardisalsoavailable,givingadministratorscontrolwhileempoweringtheiruserstoprovisionresourcesthrough awebinterface.

39. DefineTrystack.

TryStackis a great way to take OpenStack for a spin without having to commit to a fulldeployment.

This free service lets you test what the cloud can do for you, offering networking, storage and computeinstances,without having to go all in with your own hardware.

It's a labor of love spearheaded by three Red Hat OpenStack experts Will Foster,KambizAghaiepourand Dan Radez.

TryStack's set-up must bear the load of anyone who wants to use it, but instead of an equallyboundless budget and paid staff, it was originally powered by donated equipment and volunteers fromCisco, Dell, Equinix, NetApp, Rackspace and Red Hat who pulled together for this OpenStack Foundationproject.

40. DefineHadoop.

Hadoop is an open-source software framework for storing data and running applications on clustersof commodity hardware. It provides massive storage for any kind of data, enormous processing power andtheability to handlevirtually limitless concurrent tasks orjobs.