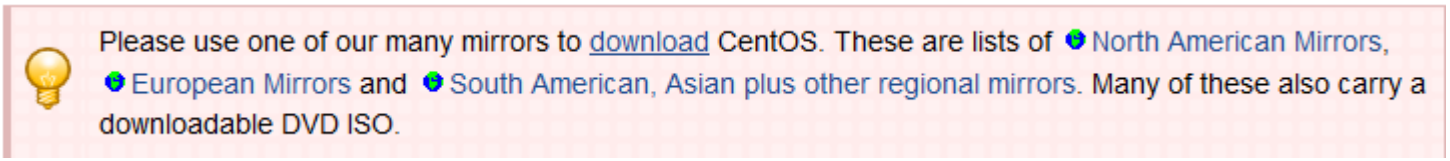
Creating
CentOS6.7 Virtual Machine

Csci E63 Big Data Analytics

Zoran B. Djordjević

Download CentOS 6.7, Select a mirror near you

- Please go to `http://wiki.centos.org/Download` and select your mirror.



- If you are in Massachusetts, use: `http://mirror.lug.udel.edu/pub/centos/`
- Click on HTTP and go to folder 6.4 There, please read the `readme` file. It will tell you that you should not be doing what you are doing.
- If you do know what you are doing and really want CentOS6.4, go to `http://vault.centos.org/` for your packages. Do as you are told.
- On `valut.centos.org`, you navigate to: `6.4 > isos > x86_64`
- I selected
  - `CentOS-6.4-x86_64-bin-DVD1.iso` and then
  - `CentOS-6.4-x86_64-bin-DVD2.iso`
- The second DVD is not used, at least not for the initial installation.
- With some downloads I got an impression that I have to burn the DVD and that CentOS would not install from an `iso` image. That might not be true on your machine.
- Normally, VMWare Workstation installs a Linux OS-s from an `iso` image.
- There is a bunch of other iso-s there, do not be tempted

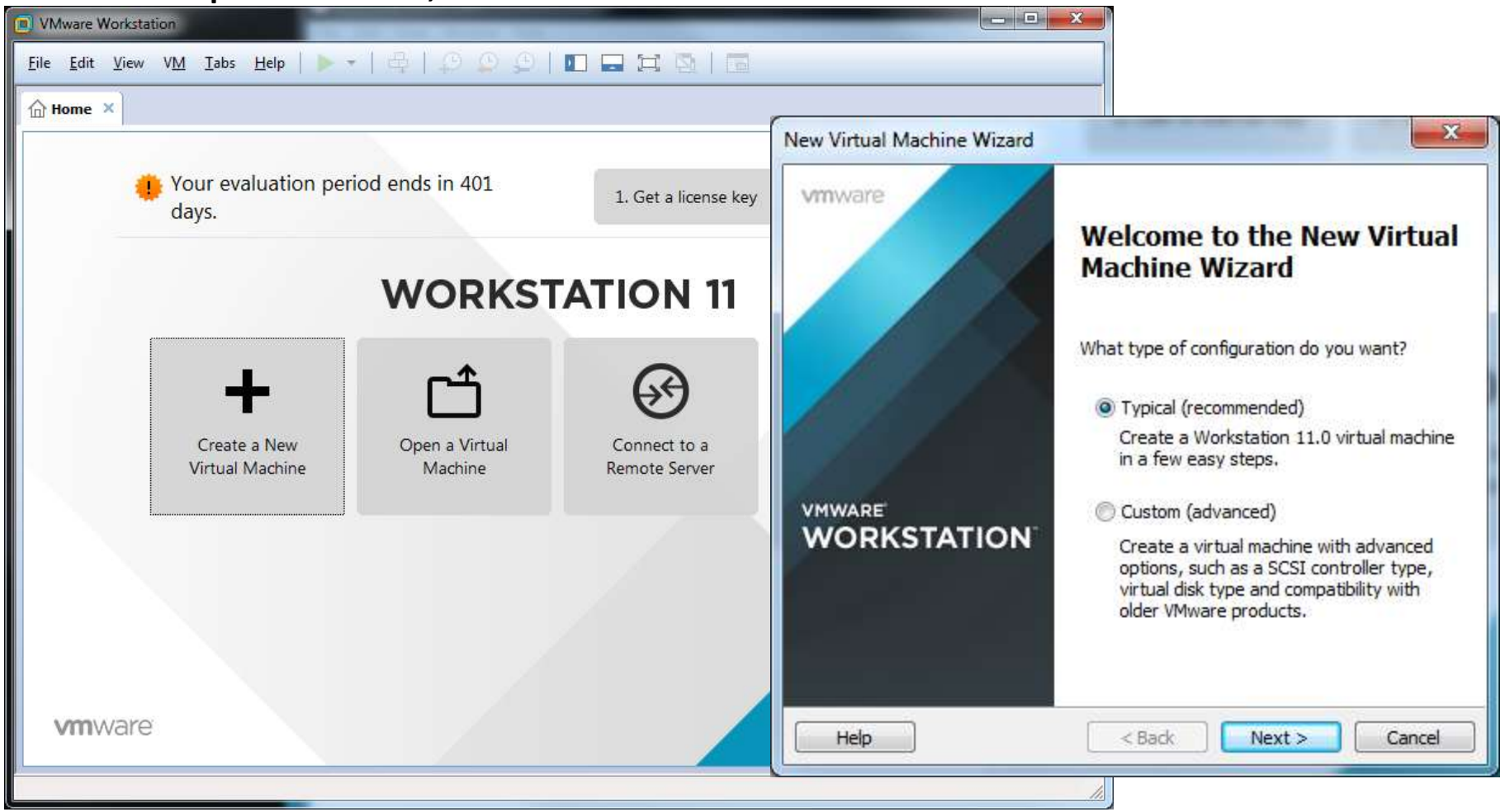
6.4 x86-64 folder distributions

Do not be tempted to use other iso-s, unless you know what you are doing

- CentOS-6.4-x86_64-netinstall.iso This is the network install and rescue image. This image is designed to be burned onto a CD. You then boot your computer off the CD CentOS-6.4-x86_64-minimal.iso The aim of this image is to install a very basic CentOS 6.7 system, with the minimum of packages needed to have a functional system. Please burn this image onto a CD and boot your computer off it. A preselected set of packages will be installed on your system Everything else needs to be installed using yum. Please read http://wiki.centos.org/Manuals/ReleaseNotes/CentOSMinimalCD6.4 for more details about this image. Beware that the set of packages installed by this image is NOT identical to the one installed when choosing the group named "Minimal" from the full DVD image.
- CentOS-6.4-x86_64-bin-DVD1.iso CentOS-6.4-x86_64-bin-DVD2.iso These two dvd images contain the entire base distribution. Please burn DVD1 onto a DVD and boot your computer off it. A basic install will not need DVD2. After the installation is complete, please run "yum update" in order to update your system.
- CentOS-6.4-x86_64-LiveCD.iso This is a CD live image of CentOS 6.7 designed to be burned onto a CD. You then boot your computer using that CD. Please read http://wiki.centos.org/Manuals/ReleaseNotes/CentOSLiveCD6.4 for more details about this image. The disk can also be used to install CentOS 6.7 onto your computer.
- CentOS-6.4-x86_64-LiveDVD.iso This is a DVD live image of CentOS 6.7 designed to be burned onto a DVD. You then boot your computer using that DVD. Please read http://wiki.centos.org/Manuals/ReleaseNotes/CentOSLiveDVD6.4 for more details about this image. The disk can also be used to install CentOS 6.7 onto your computer. Remember that in order to be able to partition your disk you will need to run the GUI installer which in turns needs enough RAM. The same is true for the network setup step. (http://wiki.centos.org/Manuals/ReleaseNotes/CentOS6.4) provide more details about these aspects.

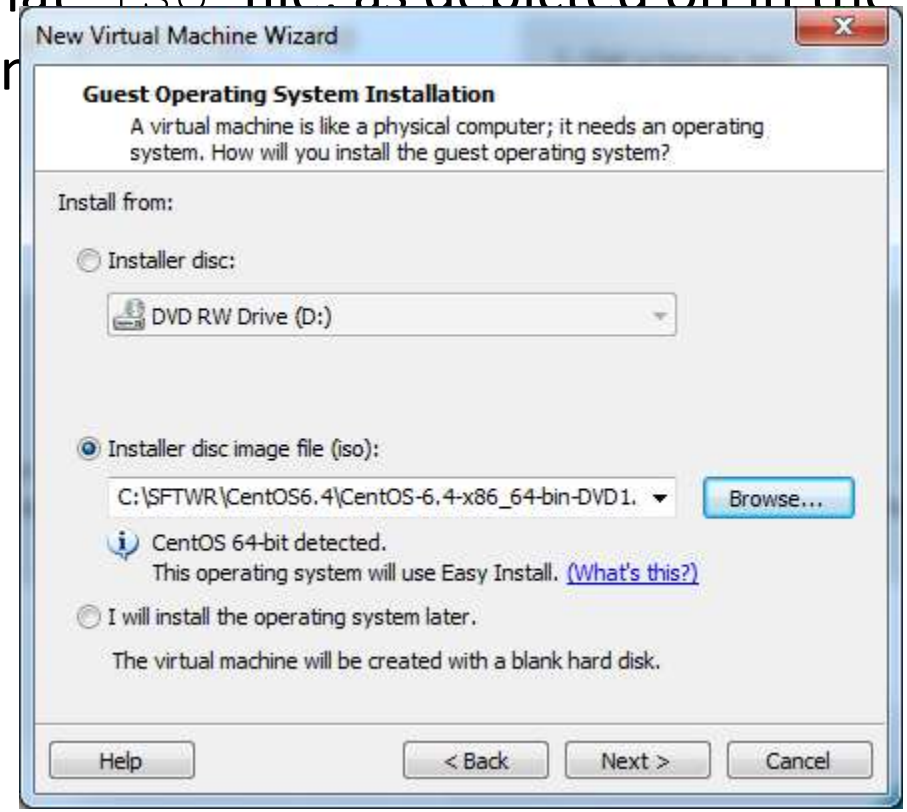
Start VMware Workstation

- Select "`Create a New Virtual Machine`".
- Accept Custom, Next >



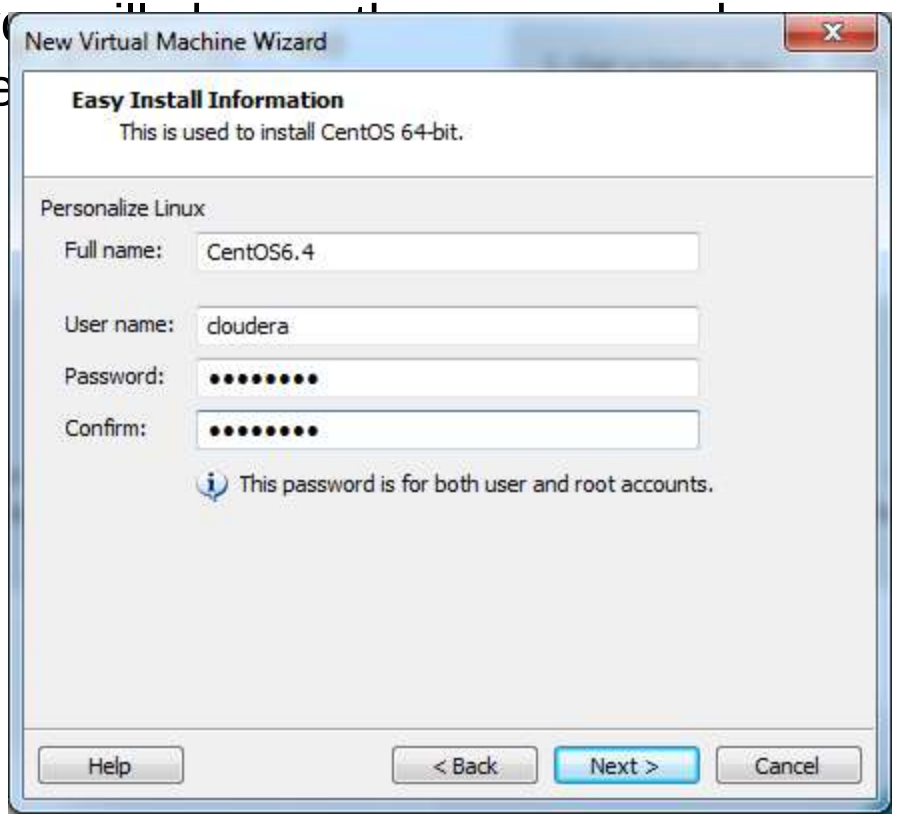
Guest Operating System Installation

- Check “Installer disc image file (iso):” and
- Select download CentOS 6.7 ISO file. Select Next >
- If VMWare Workstation accepts to read your `iso` file, it is faster and safer to use that `iso` file, as depicted on in the image on the below, thar



Create a Linux User

- Name you Linux instance and create new user.
- We will create user `cloudera`. If you do it here, user `cloudera` will belong to the Linux group `cloudera`. Typically, system administrators would create users and groups later on the command line.
- You are creating the initial password for user `root`, as well.
- In normal circumstances, y
  soon as you open the syste

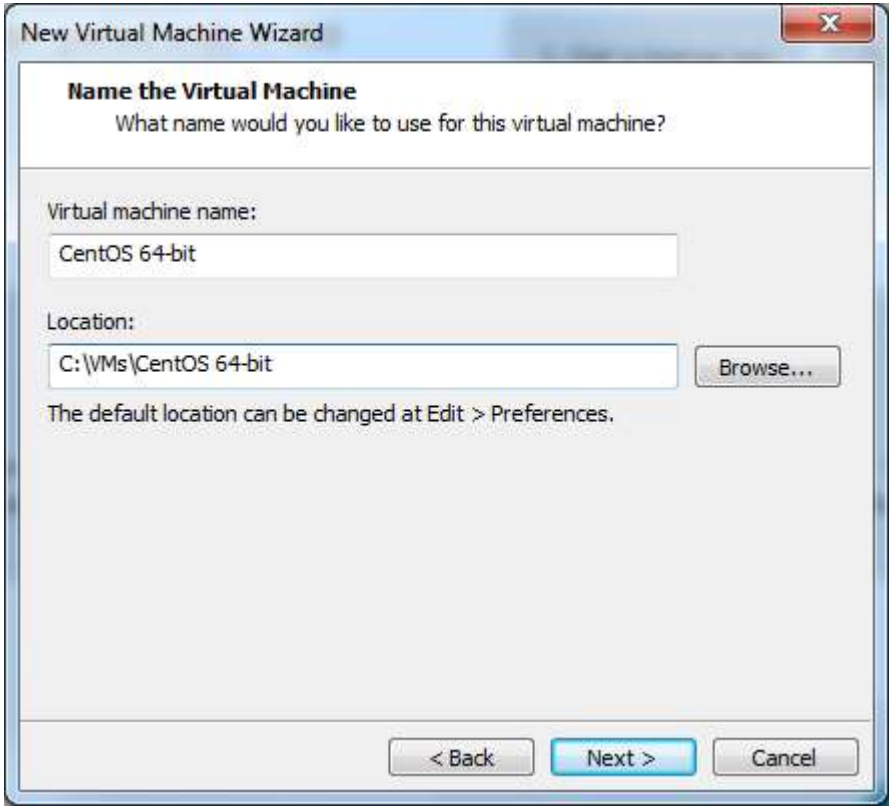


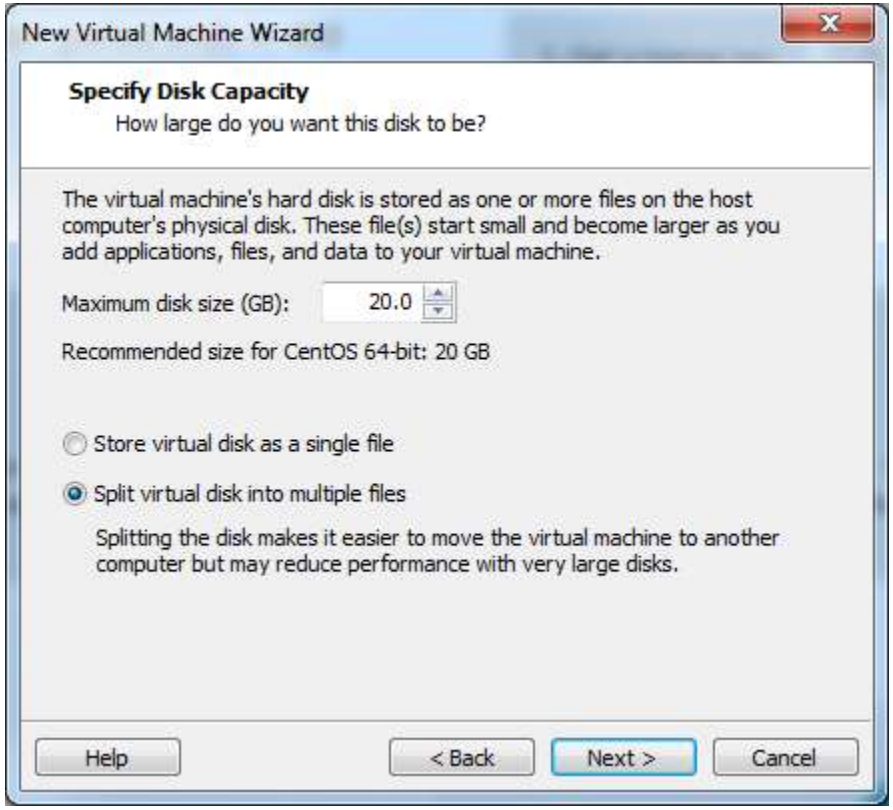
- If you do not plan to use this VM too often, stick with credentials:

  `cloudera/cloudera`

Name the VM, Select Directory

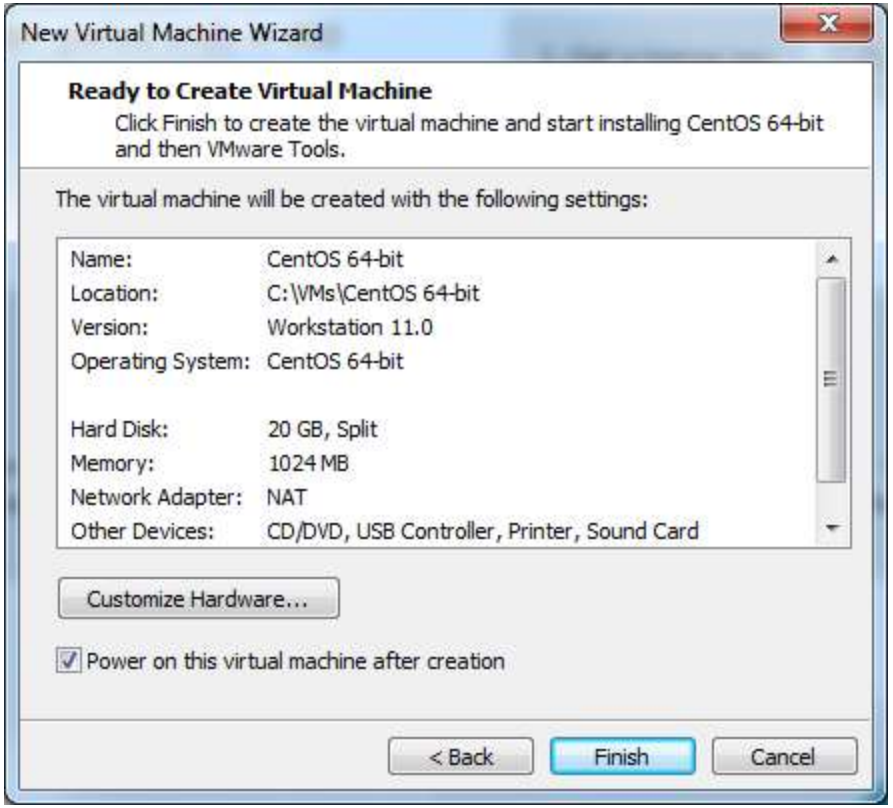
- You can place you VM anywhere, including a thumb drive or a USB external drive.
- Select Maximum disk size (20 GB) and most importantly “Split virtual disk into multiple files”. Click “Next >”





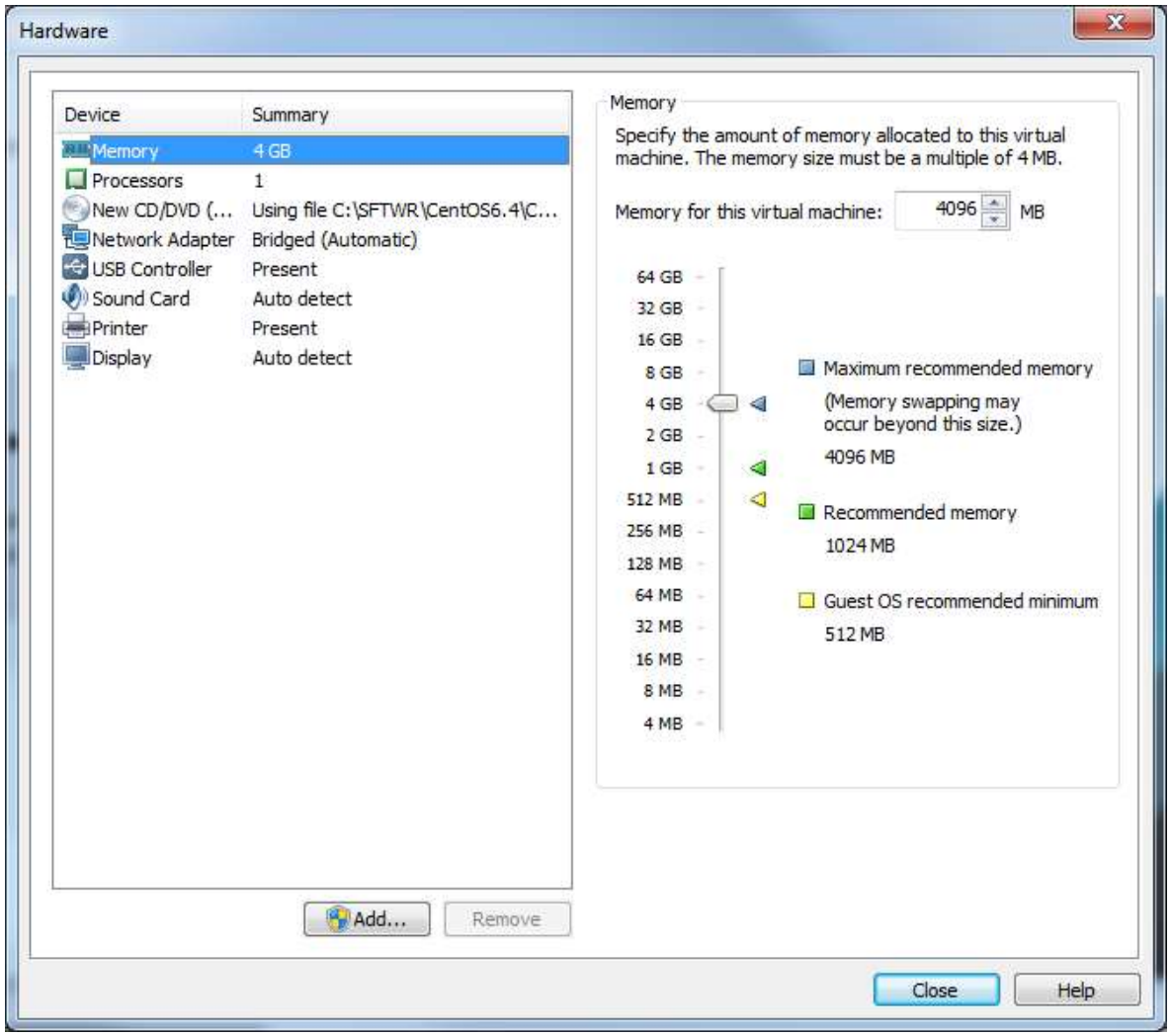
Customize Hardware, Select Memory, Network

- Select “Customize Hardware” in order to assign memory to your VM and select network arrangement.



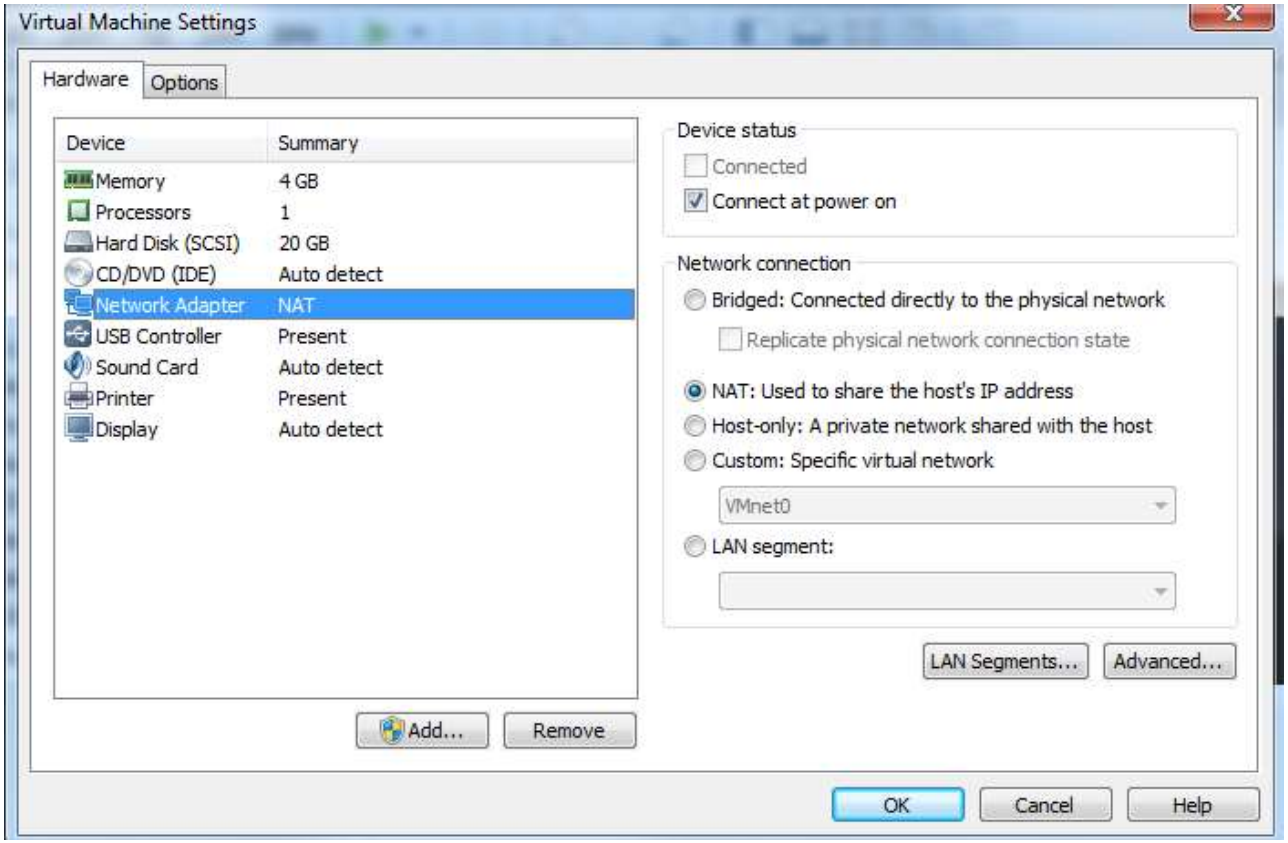
Select Memory

- More memory, faster the performance.
- Your VM will behave as a standalone machine with that much RAM.
-
- Some basic Hadoop processes could run on 2GB.
- Many processes require 4GB; some 8 or 10 GB.
- You could change (adjust) the size of memory later as you need it.
- I chose 4GB
- You might want to go buy new laptop, first. ☺



Network Setup

- Default network configuration contains one NAT (Natural Address Translation) adapter which has the same IP address as the host machine.
- That adapter is quite useful and we will leave it alone. We want another.
- On the Virtual Machine Settings select `Network Adapter` and then click `Add` button

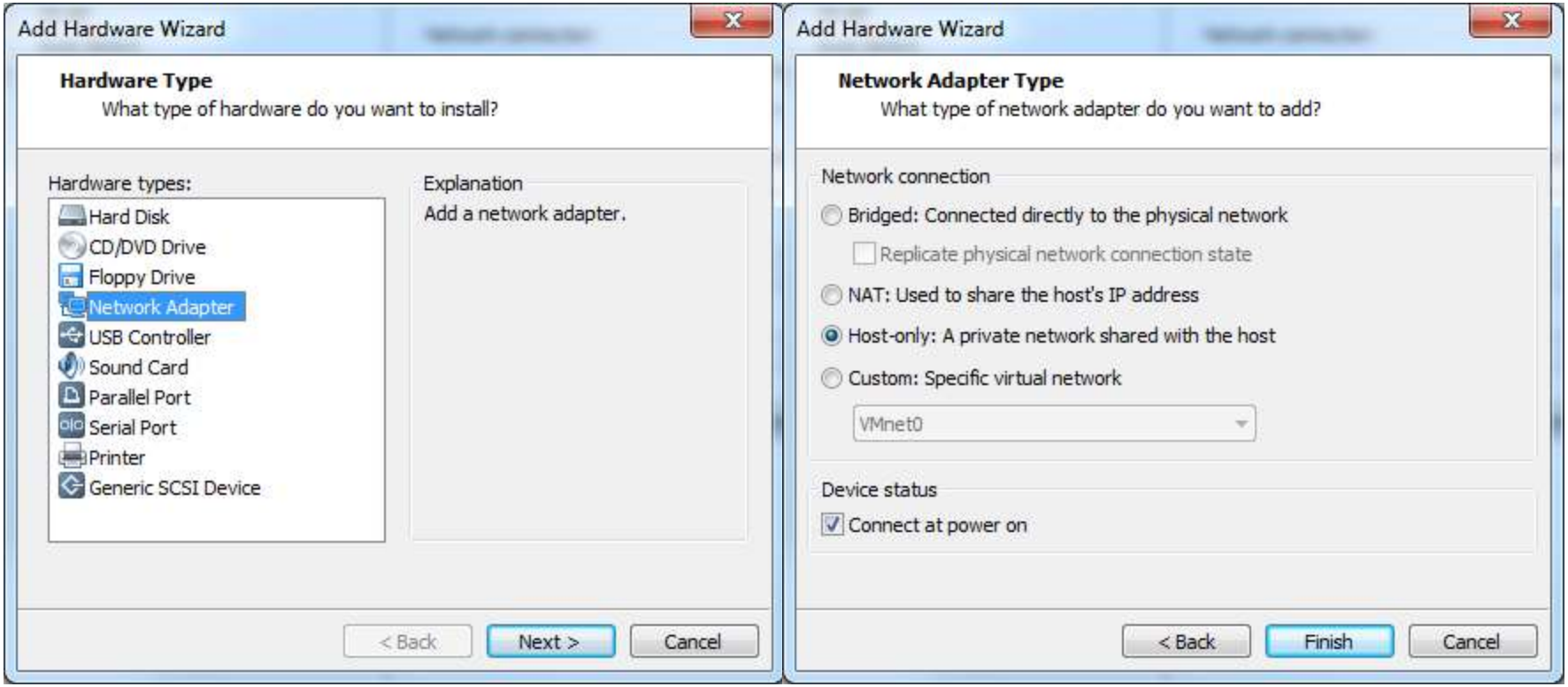


Adding Host-only Adapter

On the `Add Hardware Wizard` select `Network Adapter` again and hit `Next>`

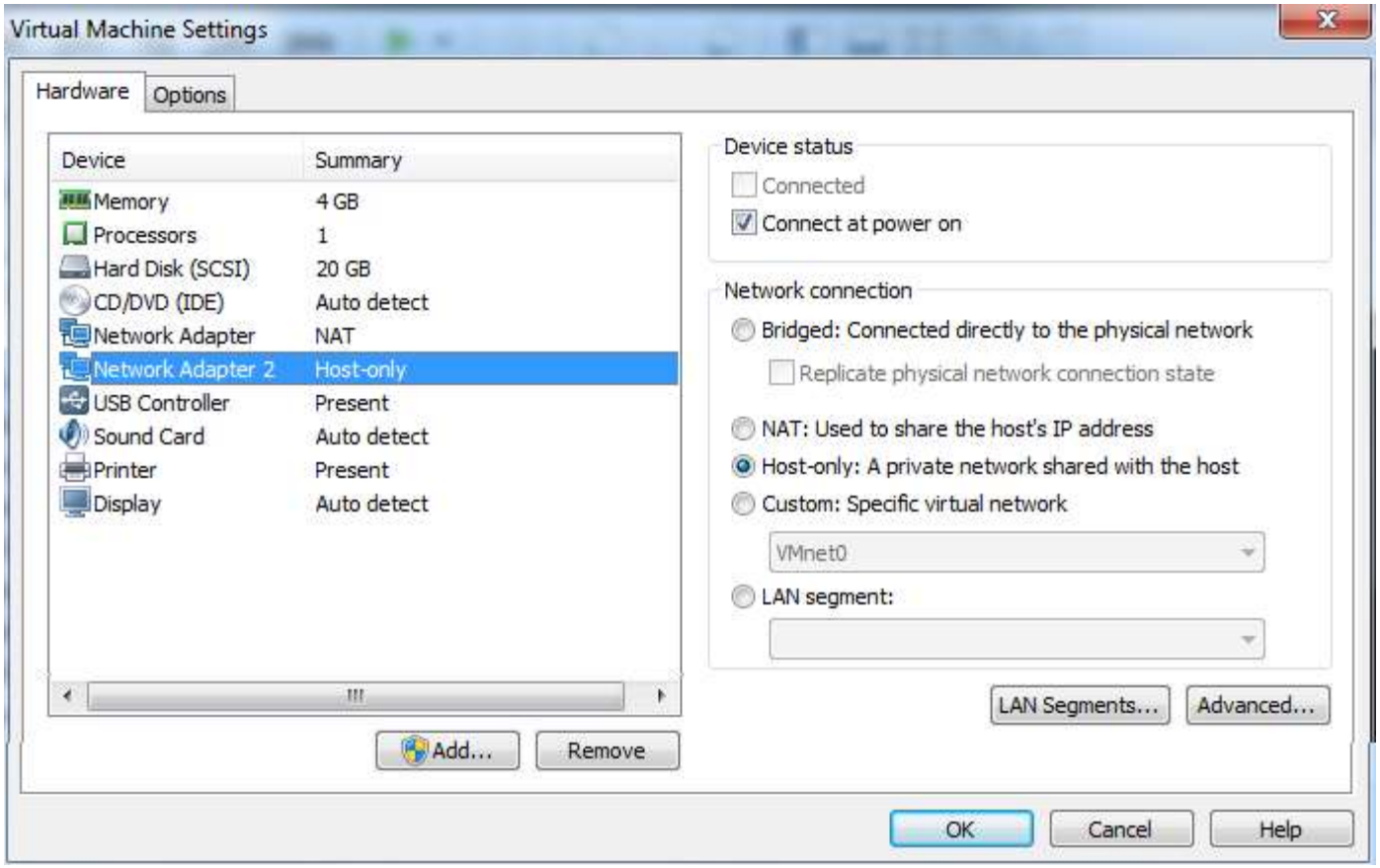
On the `Network Adapter Type` wizard select `Host-only: A private network shared with the host.`

Select `Finish`



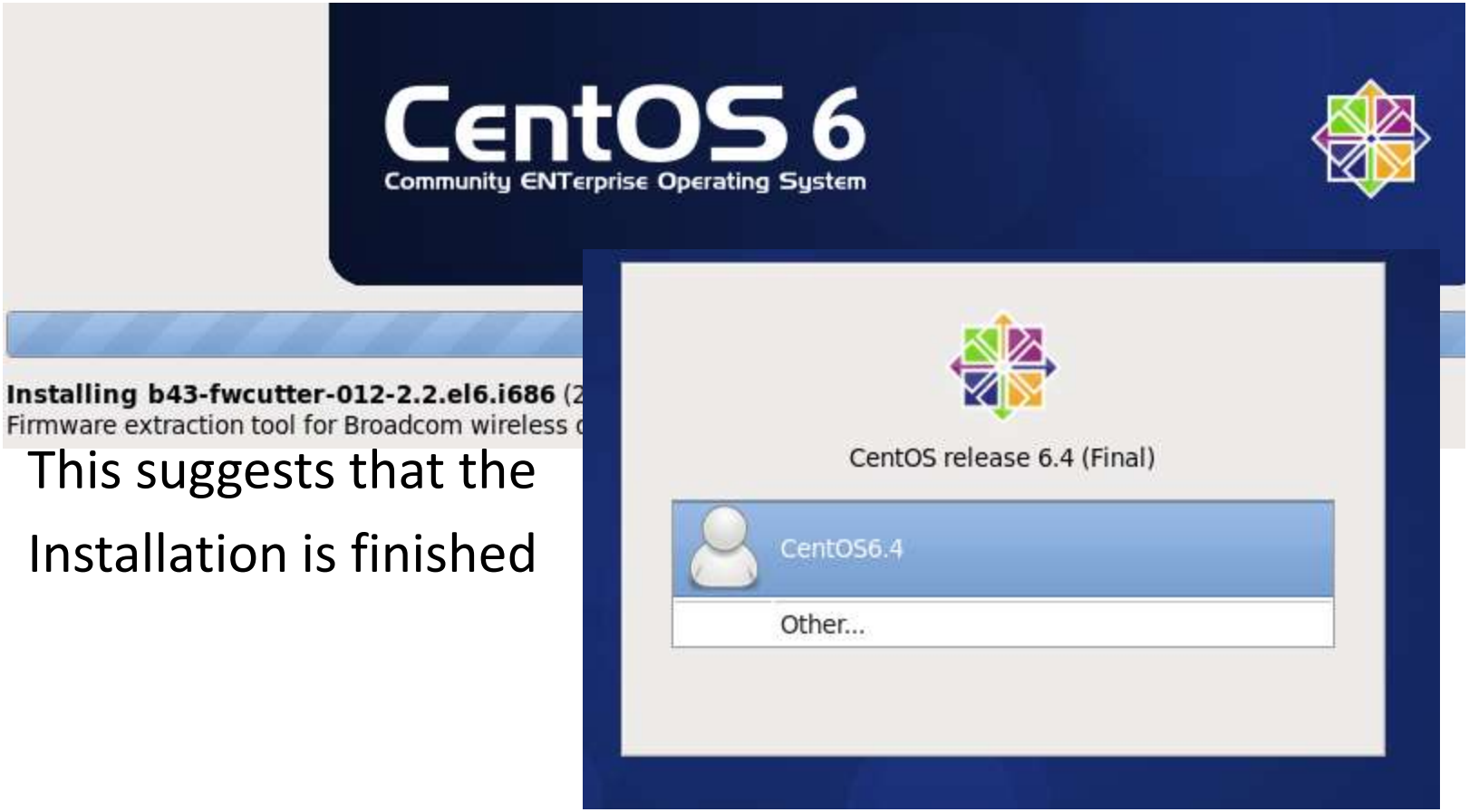
Two Network Adapters

- On the `Virtual Machine Settings` we now see two Network Adapters.
- One adapter is of type `NAT` and the other is of type `Host-only`.
- Hit `OK` and `Finish`



Eventually all packages are installed

- On a fast machine it might take 15 minutes. On a slow machine it might take 2 hours.



- This suggests that the
- Installation is finished

Login into the Linux VM

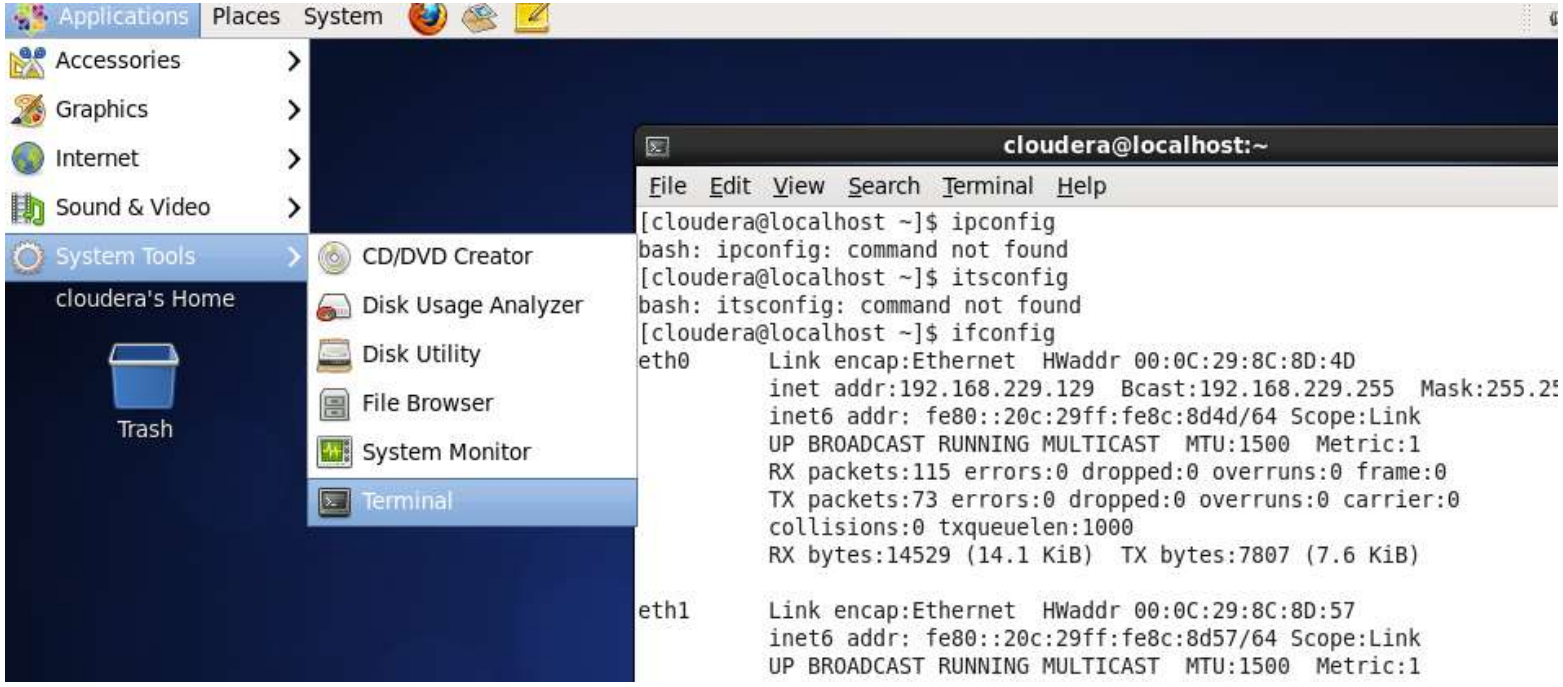
- If you remember `cloudera`'s password (`cloudera` ☺), you can log into your Linux box:

Open Terminal Window and find IP Address

- In the screen that opens we can go to

  `Applications > System Tools > Terminal`

- White (not Black) terminal window will open. If we type `ifconfig`, the system will produce the IP Address of the Host-only adapter we created. As we can see the IP address is `192.168.229.129.`
- We could use that IP address to connect to our VM as if it were a server.



Connect to VM as a Server, Transfer Files

- If we open a Cygwin window on our PC, we could issue `ssh` command:

```
$ ssh cloudera@192.168.229.129
cloudera@192.168.229.129's password:
[cloudera@localhost ~]$ pwd
/home/cloudera
[cloudera@localhost ~]$ ls
Desktop  Documents  Downloads  Music  Pictures  Public  Templates  Videos
[cloudera@localhost ~]$ exit
logout
Connection to 192.168.229.129 closed.
```

- Similarly, we could use scp command to transfer files to VM. For example:

```
$ scp JDBCSample.java cloudera@192.168.229.129:~
cloudera@192.168.229.129's password:
JDBCSample.java                                   100%  771
0.8KB/s   00:00
zdjordje@FDCE-ZDJORDJ-2 /cygdrive/c/CLASSES/code
$ ssh cloudera@192.168.229.129
cloudera@192.168.229.129's password:
Last login: Tue Feb 24 15:15:38 2015 from 192.168.229.1
[cloudera@localhost ~]$ ls
Desktop  Documents  Downloads  JDBCSample.java  Music  Pictures  Public
Templates  Videos
```

- We transferred file `JDBCSample.java` from my PC to the home directory (~) of user `cloudera` on the VM.

passphraseless `ssh`

- Various Hadoop processes have to navigate to different machines in the cluster and it would be a nuisance if they would have to submit a password on every entry. For example if you type

```
[cloudera@localhost Downloads]$ ssh localhost
```

- The system will ask you for cloudera's password

```
The authenticity of host 'localhost (::1)' can't be established.
RSA key fingerprint is
88:87:fc:e8:24:d0:c9:81:0e:f4:9c:9e:7a:24:b3:46.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'localhost' (RSA) to the list of known
hosts.
cloudera@localhost's password:xxxxxxxxx
Last login: Fri Feb 27 15:29:59 2015 from 192.168.72.1
[cloudera@localhost ~]$
```

- If you do `ls -la` in your home directory you will see directory `.ssh`
- Initially, the directory contains a file `known_hosts`

passphraseless `ssh`

- In the home directory of the user whom you want to equip with passphraseless `ssh` run the following commands:

```
[cloudera@localhost ~]$ ssh-keygen -t dsa -P '' -f ~/.ssh/id_dsa
Generating public/private dsa key pair.
Your identification has been saved in /home/cloudera/.ssh/id_dsa.
Your public key has been saved in /home/cloudera/.ssh/id_dsa.pub.
The key fingerprint is:
a6:f6:6b:1f:3b:77:0b:24:e1:b5:c0:c1:89:19:29:aa
cloudera@localhost.localdomain
The key's randomart image is:
+--[ DSA 1024]----+
|        .*..     |
|      . +.o.     |
|     . .  + .    |
|    .     . + .  |
|   .     S o o   |
|  E     o   o    |
|       o   . .   |
|      . .. .o... |
|       .oooo ... |
+-----------------+
[cloudera@localhost ~]$
```

passphraseless ssh

```
[cloudera@localhost ~]$ cat ~/.ssh/id_dsa.pub >> ~/.ssh/authorized_keys
[cloudera@localhost ~]$ cd .ssh
[cloudera@localhost .ssh]$ ls
authorized_keys  id_dsa  id_dsa.pub  known_hosts
```

- The above `cat` command has copied the public key to the `authorized_keys` file of the .ssh directory

```
[cloudera@localhost .ssh]$ cat authorized_keys
ssh-dss
AAAAB3NzaC1kc3MAAACBAI5Jsfv/wSHLNpC/KS8CPDR60zVUGzvc8K9L71igZeyd0xI8iNKGKM5
3+MMbgaIUHWxBWxdixFMkcOyIIee7ljZuBUPe6H6/AEY0MMnLetFLQt/DyYf6VpT0mpVUpjspOO
tbmqZrL+GRaQ0l3ApbPcucgYIavT0oHdW2ba5b07G9AAAAFQDb9BTZ81YbkQlUvxGaI+PbTMak/
QAAAIBb6mAcrdiHl+96/JwybKfMXaHhZZEzJjbki5S3UhUOiAoJXi4gLOtS2kr7NSKCh/J+TYVv
c5rxnyHiBZYGTARSFSw2nsQ+rJi9v70PoQ8Ij69+QhlPb0ugW/f4Pmn9fpFaMaSeQMRCfvYV11t
Fh1YjWrvYagnbzgiSk0YvE7BwAAAIATMX1giXJwglwnxNKalcH+krRAA7Xy38nnGYR4KJdGaQ/0
Z1zIObYb3T1EPFkQoWYlu2FR76eMFocCtK1f1sELE7afq/OVmkaF8Nu09ED086PiF24ZJp2p3VJ
IIwueLOz7EoSUItqumVoIJQVyMltpoPh+dyDtc5uLWOfVY2fISw==
cloudera@localhost.localdomain
```

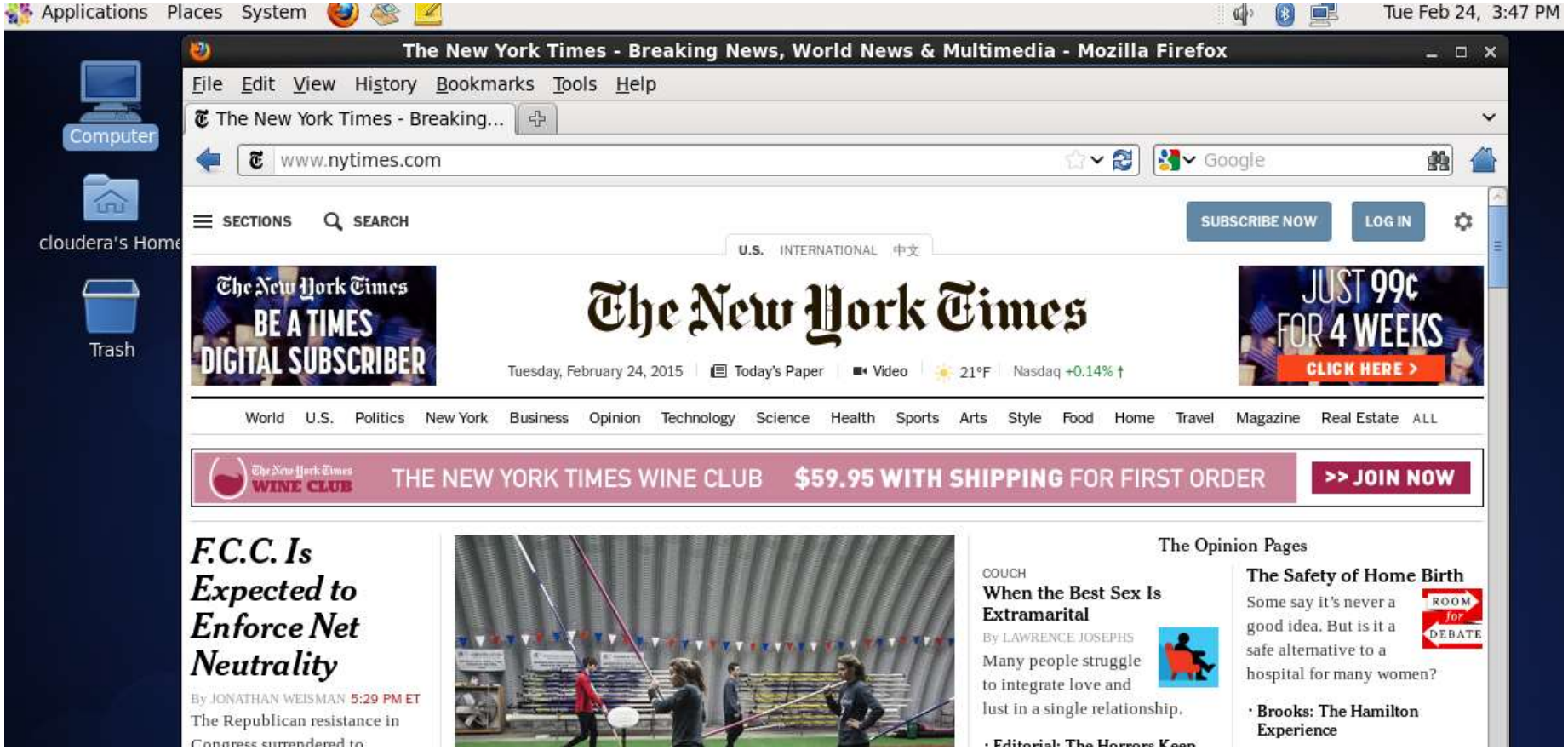
- You can transfer the private key (`.shh/id_dsa`) you just generated to any machine and then login to your VM using the `ssh` command

```
ssh -i id_dsa cloudera@192.168.255.129
```

- Without being asked for the password

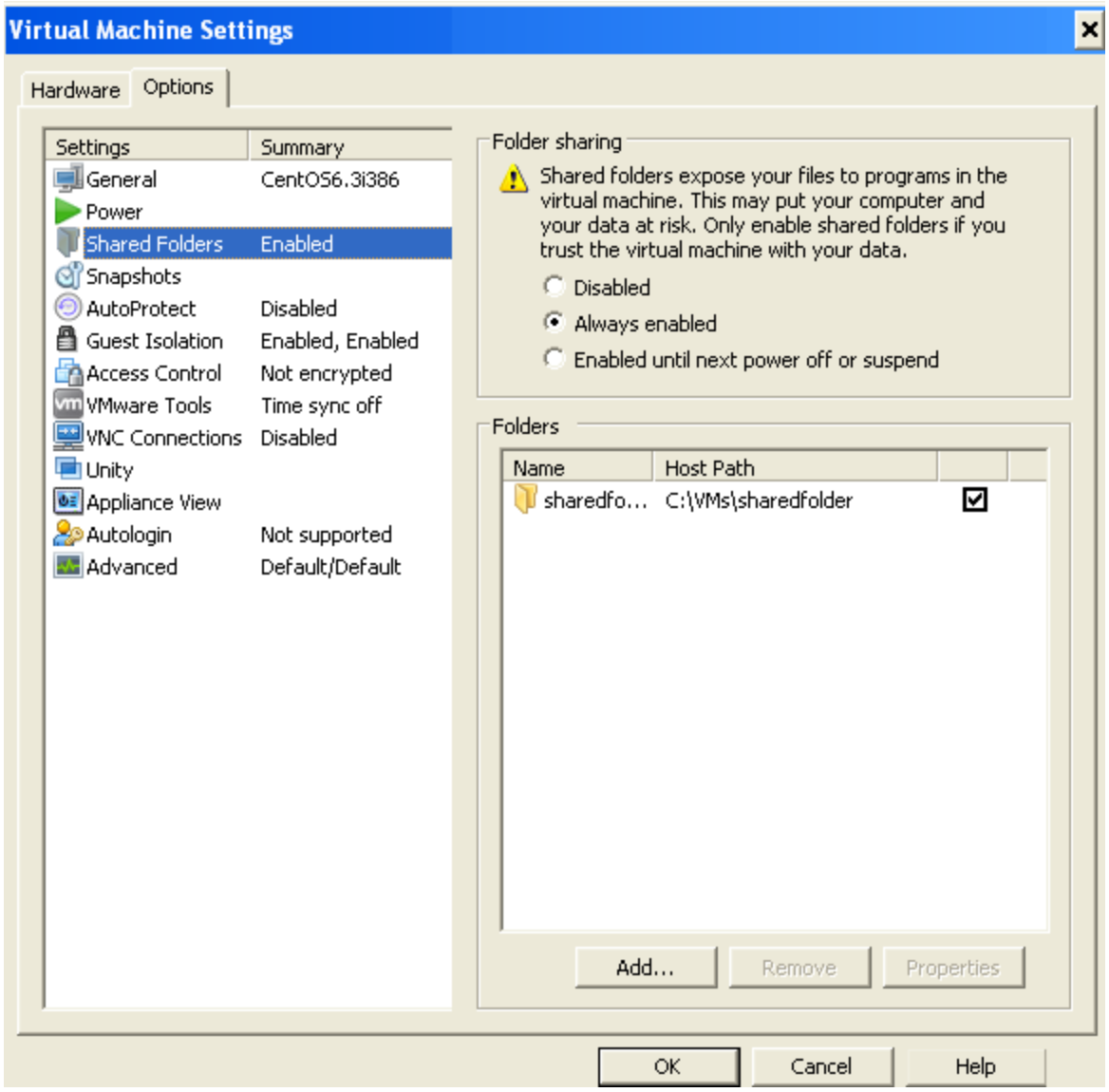
NET Adapter and connection to the Web

- We have another Network adapter (NAT) that uses IP address of the host machine and enables us to connect to the Internet from inside our VM.
- Open a browser and type the URL of the Ministry of Truth. You will see the truth. All you have to do is believe. All courtesy of NAT adapter.



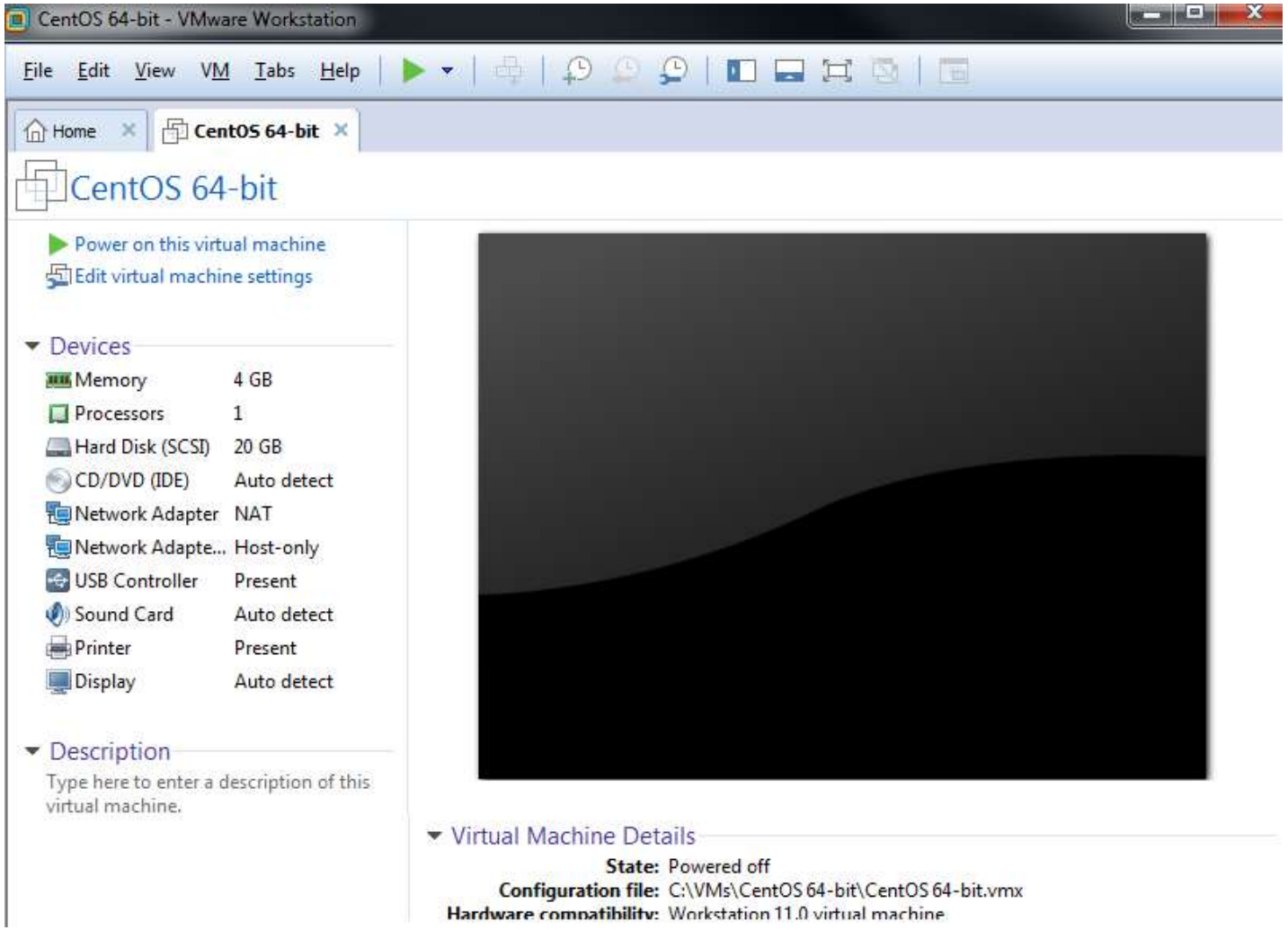
Enable Shared Folders

- Another way to share files with the host OS is to enable `Shared Folders`.
- Power down VM. Right click on the VM, select `Edit virtual machine settings` > `Options`
- Select `Shared Folders` > `Add`
- Add folder `c:\VMs\sharedfolder`
- Check `Always enable` > `Finish` > `OK`
- Power up VM
- Login as `cloudera`.
- Shared folder will shows as `/mnt/hgfs/sharefolder`



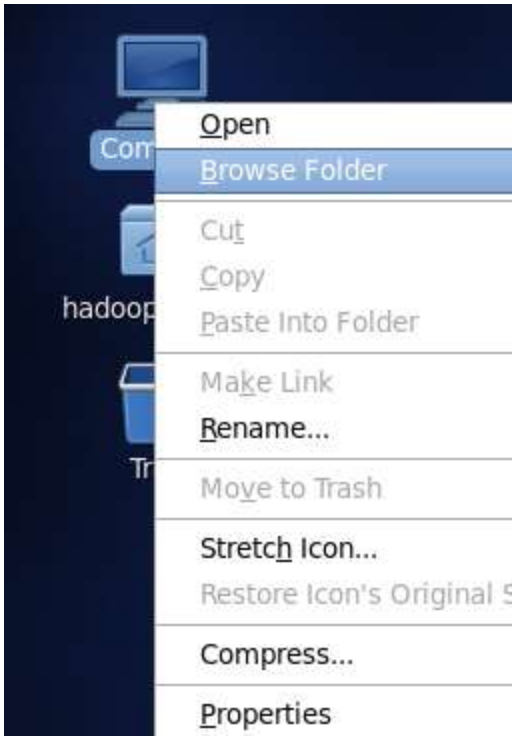
Open a Virtual Machine

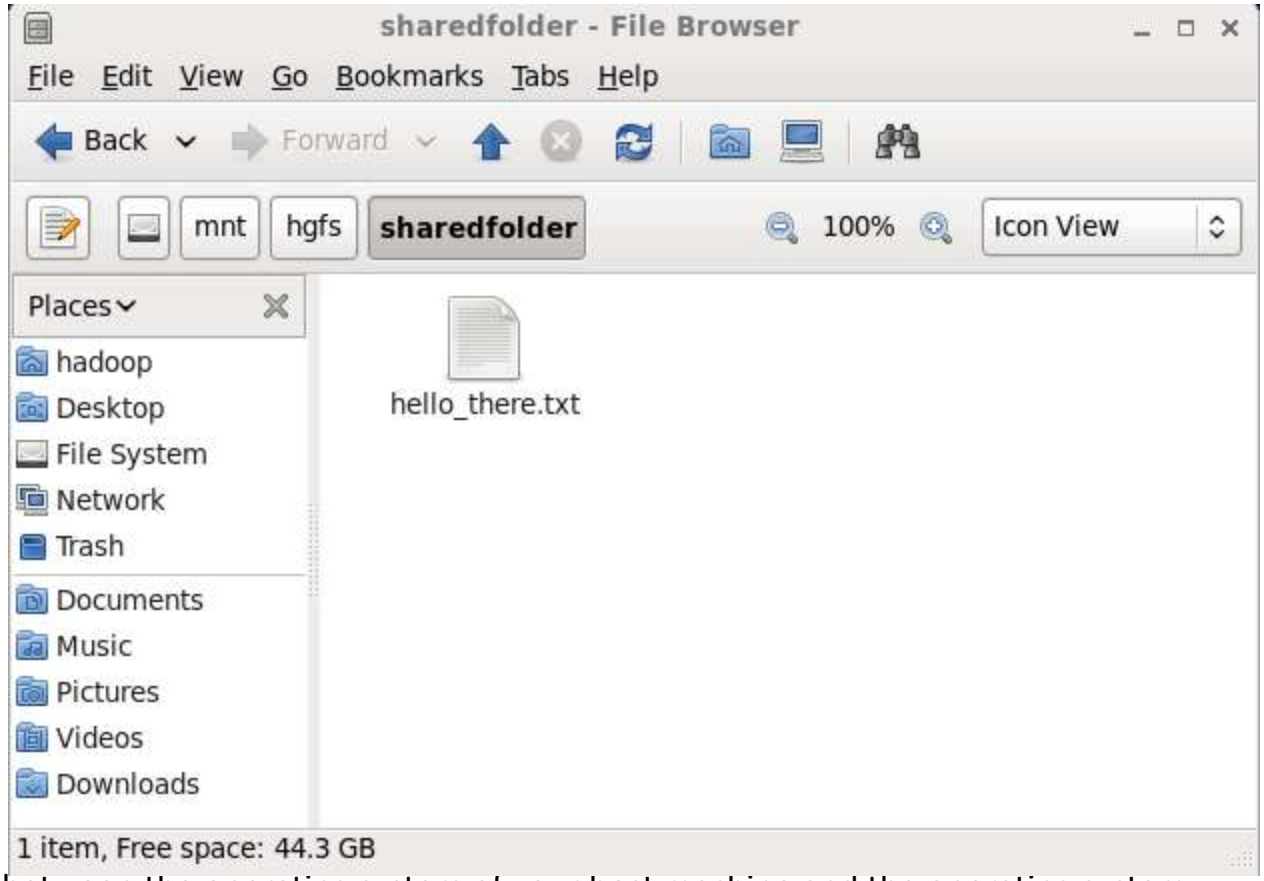
- Open the Workstation, select Open a Virtual Machine icon and select the VM file you want to run from your OS. To power the VM hit the green triangle. If you have enough memory, could run several VMs simultaneously.



Locate Shared Folder, `Browse Folder`

- Right click on `Computer` icon
  - `Browse Folder > File System > mnt > hgfs > sharefolder`





- We can use the `sharefolder`
- to share files back and forth between the operating system of your host machine and the operating system of your new VM.
- On some VMs sharefolder has to be mounted. Do not despair. You have seen it in recitation.

Select Terminal, `whoami`, Shell

- Select

`Applications > System Tools > Terminal.`

- Find out who you are,

```
$ whoami
```

- Examine `/etc/passwd` file
- Examine `/etc/group` file
- User `cloudera` belongs to group `cloudera` and has `bash` shell.

```
[cloudera@localhost ~]$ whoami
cloudera
[cloudera@localhost ~]$ cat /etc/passwd | grep cloudera
cloudera:x:500:500:CentOS6.4:/home/cloudera:/bin/bash
[cloudera@localhost ~]$ cat /etc/group | grep cloudera
cloudera:x:500
[cloudera@localhost ~]$ pwd
/home/cloudera
```

`mount`-ing sharedfolder

- Sometimes your will move (copy) your VM to another machine or your would download a VM and the above procedure for creating shared folder would simply not work. You will select the shard folder through VM Options but once you are in VM, the content of the shared folder will not be visible.
- At point you need to manually "`mount`" that folder. VM for whatever reasons failed to do it for you.
- Open Linux command prompt and issue the following command:

```
$ sudo mount -t vmhgfs .host:/ /mnt/hgfs
```

- The `sahredfolder` will appear under `/mnt/hgfs/sharedfolder`

Giving sudo privileges to user cloudera

- We need cloudera to be very powerful user. This is enabled by user root who grants "sudo" privilege to user cloudera.
- On the top menu bar, select System and "Log Out cloudera"
- On the next widget select Switch User.
- On the following widget select Other (user).
- Enter root as the username. > Log In
- root's password is still the same as the password of user cloudera.
- As user root open the terminal window and change permissions on file /etc/sudores

```
$chmod a+w /etc/sudoers
```

- As root, use tool visudo to add the following line to /etc/sudoers file:

```
root$ visudo /etc/sudoers
cloudera  ALL=(ALL) NOPASSWD: ALL
```

- User cloudera will not be asked for password after every sudo command.
- Exit and save modifications by typing Hold [shift]+press "X"
- Change permission on /etc/sudoers: root$ chmod –w /etc/sudoers

Giving `sudo` privileges to user `cloudera`

Allowing user `cloudera` to run commands without checking its password, is a security issue but is a great convenience. Interestingly, CDH installation does ask for a user with sudo privileges and no password.

- On some Linux systems, CentOS included, `sudo` command clears the environmental variables.
- In order to preserve some of those, you need to add lines to `/etc/sudoers` that read like **:**

```
Defaults env_keep+= "JAVA_HOME"
```

- Then change permissions of `/etc/sudoers` back to read only (440 mode). Note, if you do not do this your sudoers will not function properly.

```
$chmod 440 /etc/sudoers          # must do it as root
$ls –ls /etc/sudoers
-r—r----- . 1 root root 4035 Mar 8 06:56 /etc/sudoers
```

- Once you install Java JDK, you will be able to verify that `sudo` command does not remove `JAVA_HOME` environmental variable by typing:

```
$sudo env | grep JAVA_HOME
```