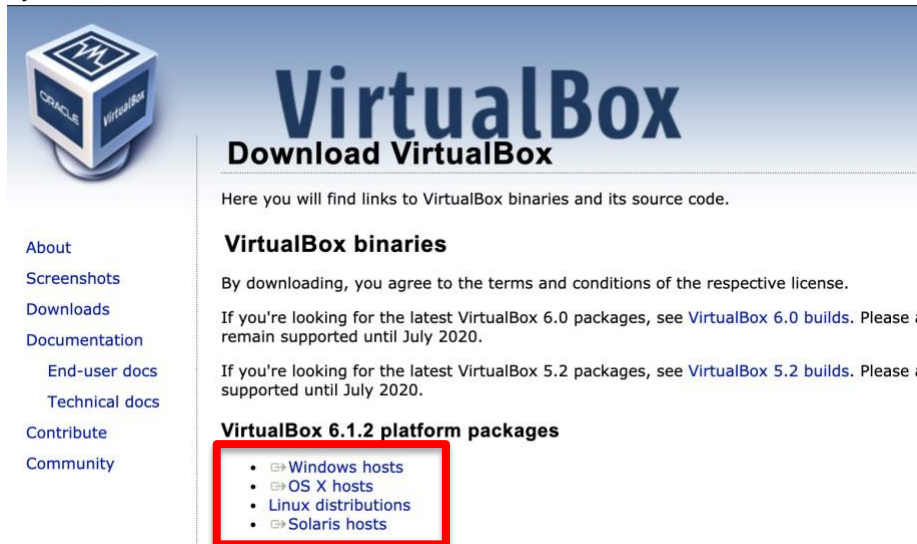In this lab, we will set up Cloudera on the VirtualBox platform.

Part 1: Download and install VirtualBox
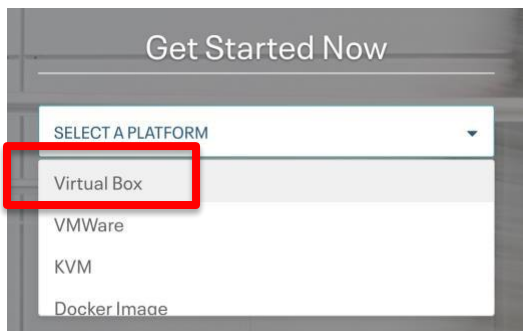
Go to https://www.virtualbox.org/wiki/Downloads, select the correct version for your operating system.

Follow the installation prompt and finish the installation process with **default** settings.

Part 2: Download Cloudera QuickStart VM

Go to https://www.cloudera.com/downloads/quickstart_vms/5-13.html, click **SELECT A PLATFORM** and choose VirtualBox.
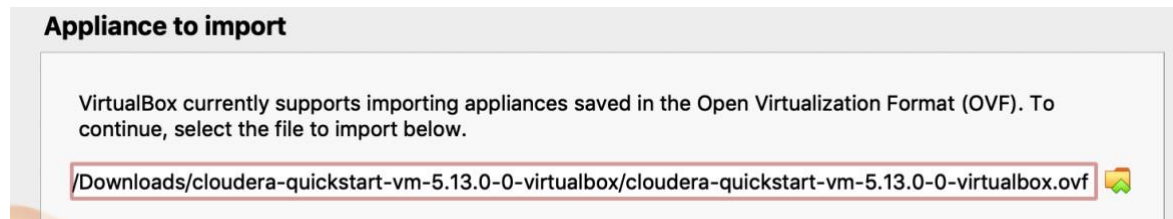

Click on the **GET IT NOW** and register with your student information.

Now you will get the **cloudera-quickstart-vm-5.13.0-0-virtualbox.zip**. Download and extract the zip file.

Part 3: Create a new virtual machine and setup Cloudera

Open **VirtualBox**, go to **File** and select **Import Appliance**.
Browse your folders, find the extracted file, and select the **.ovf** file to import.

Appliance to import

VirtualBox currently supports importing appliances saved in the Open Virtualization Format (OVF). To continue, select the file to import below.

/Downloads/cloudera-quickstart-vm-5.13.0-0-virtualbox/cloudera-quickstart-vm-5.13.0-0-virtualbox.ovf
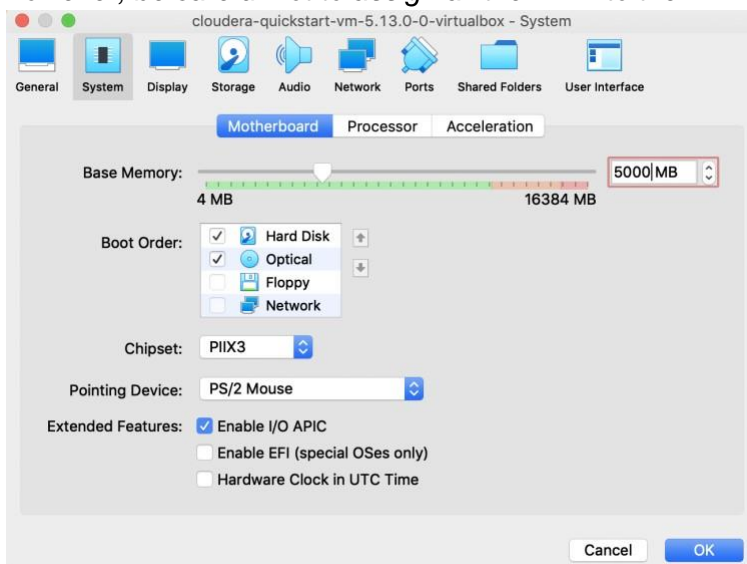
Click **Continue** and then **Import**. Keep the **default** settings for now.

Then, it will take a few minutes to set up the new VM. After the process is completed, right-click on the machine you just created and go to **settings**.

Go to **System**. 4096 MB is the minimum amount of RAM required. If you have more amount of RAM, you can assign more than 4096 MB. The more, the faster. Recommend 5000 MB.
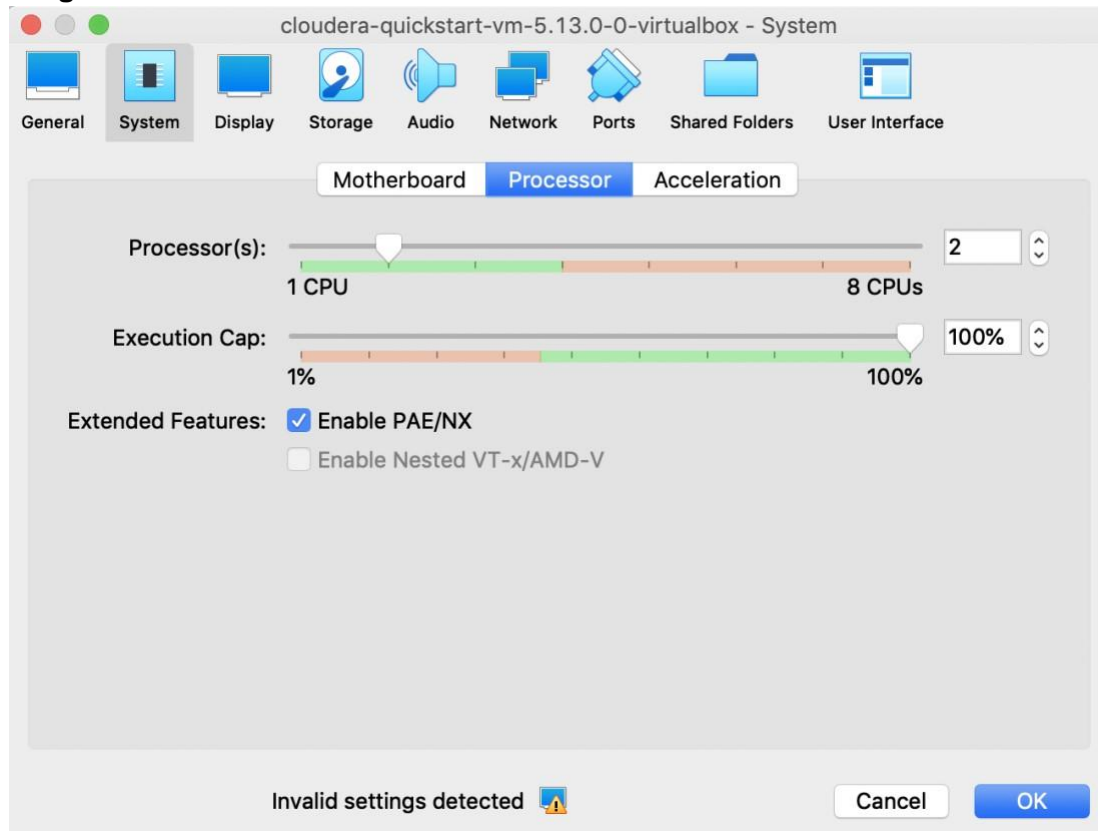*However, be careful not to assign all the RAM to the VM.*

Under the **System** tab, go to **Processor**. And assign the number of cores that you want your VM to run on.
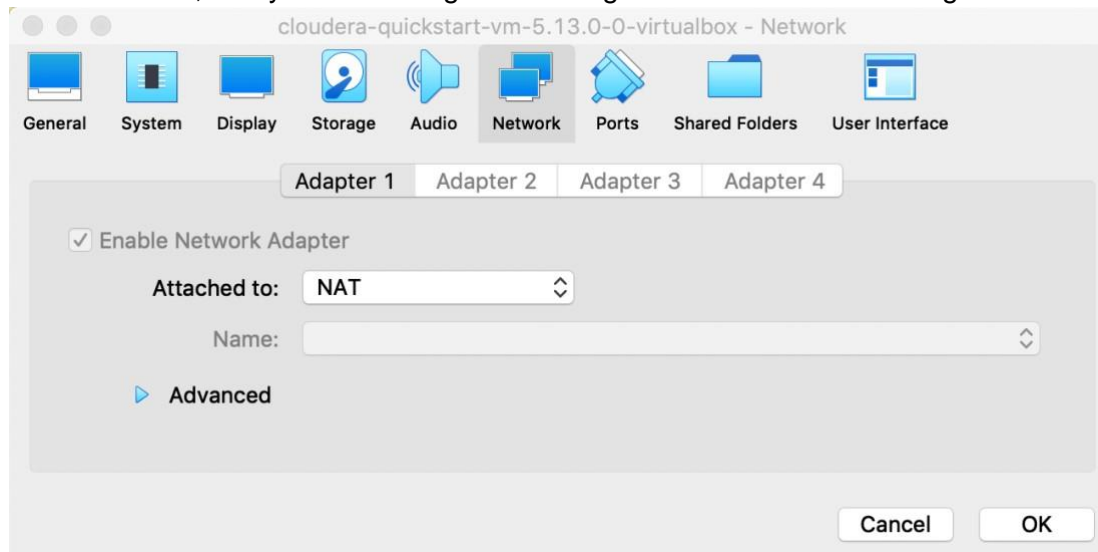*Do not Assign a greater number of cores than you physically have on your machine.*

If it is quad core, assign 4 cores maximum.
If it is dual core, assign 2 cores maximum. Here,
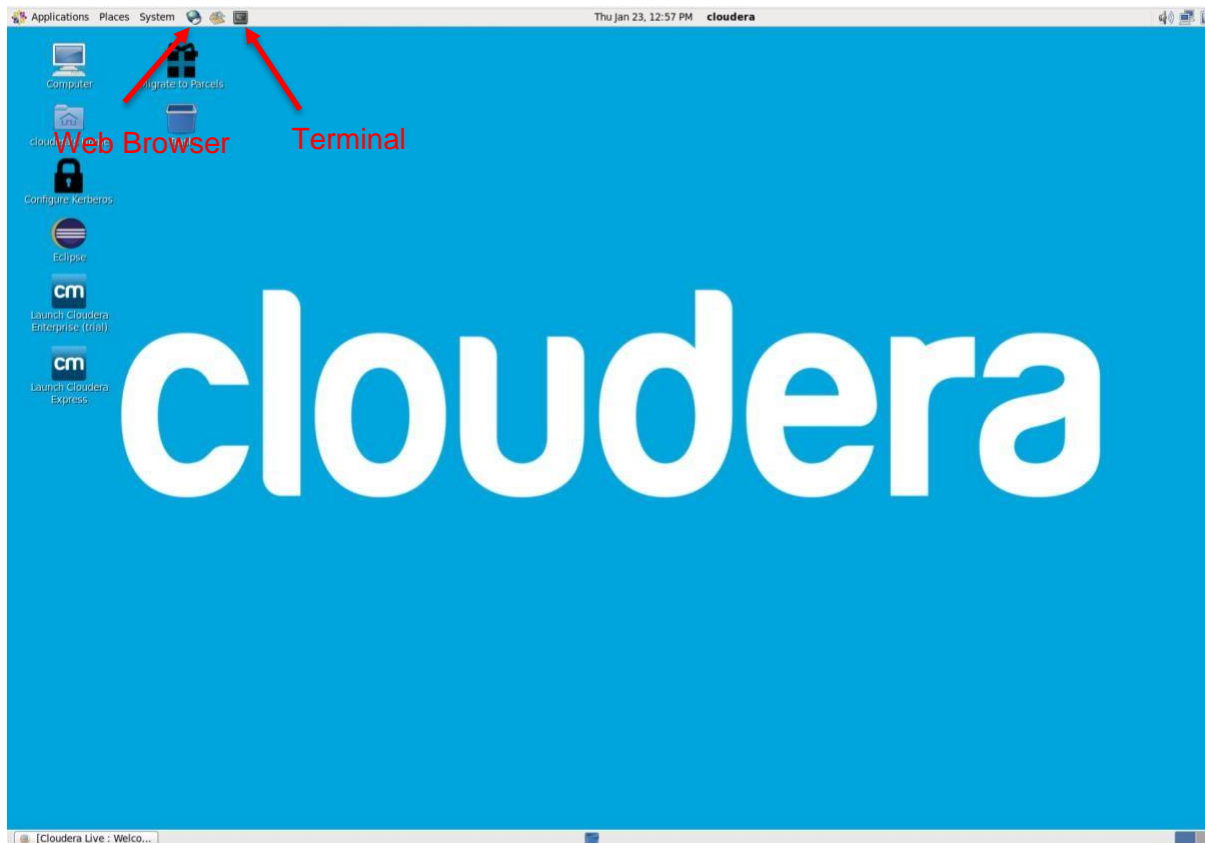**assign 2 cores** to the VM.

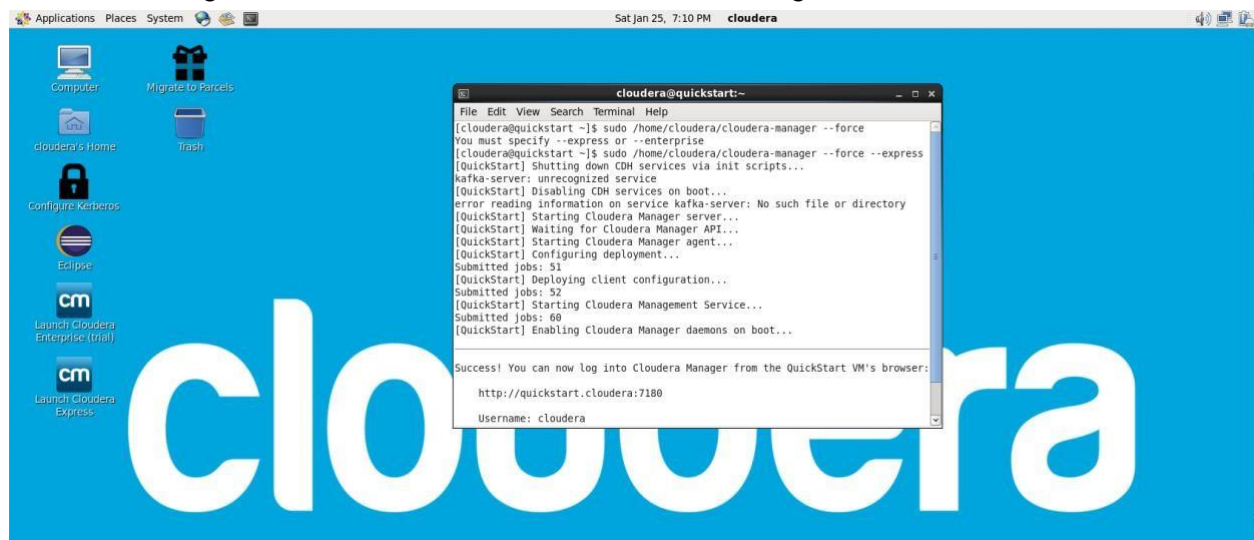Go to **Network**, and you can change the settings here. The default setting here is fine.

Click on **OK** and click the **Start** button to start the VM. It will take a few minutes for your VM to initialize.
Note: The terminal and web browser are located at the top left corner.
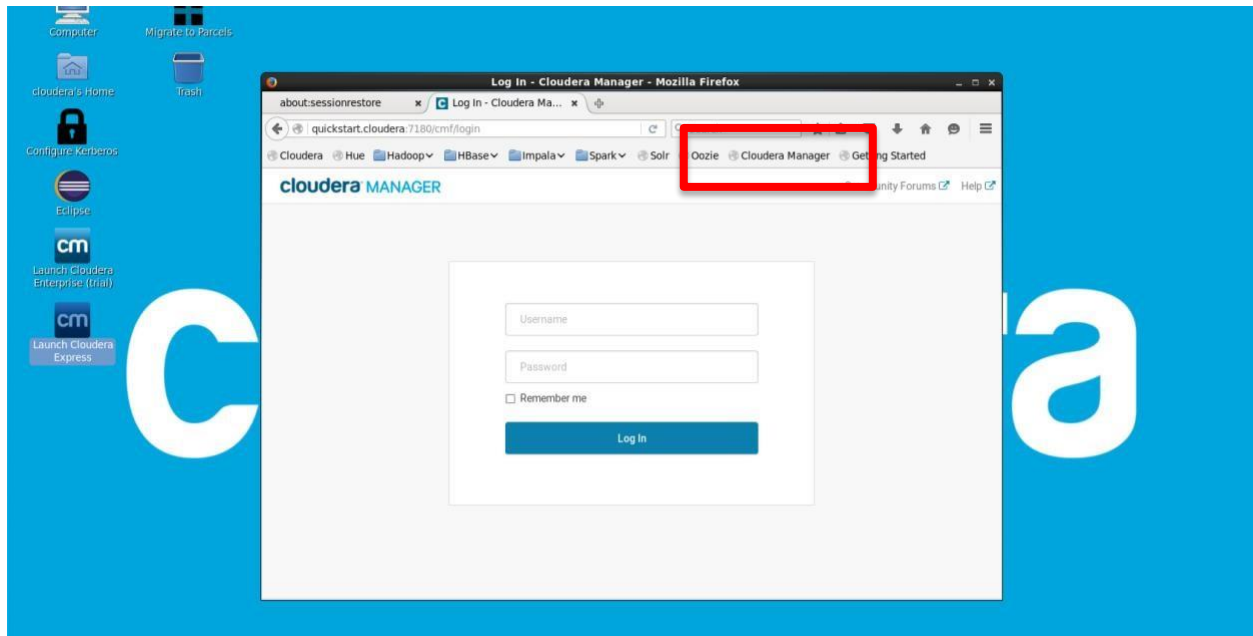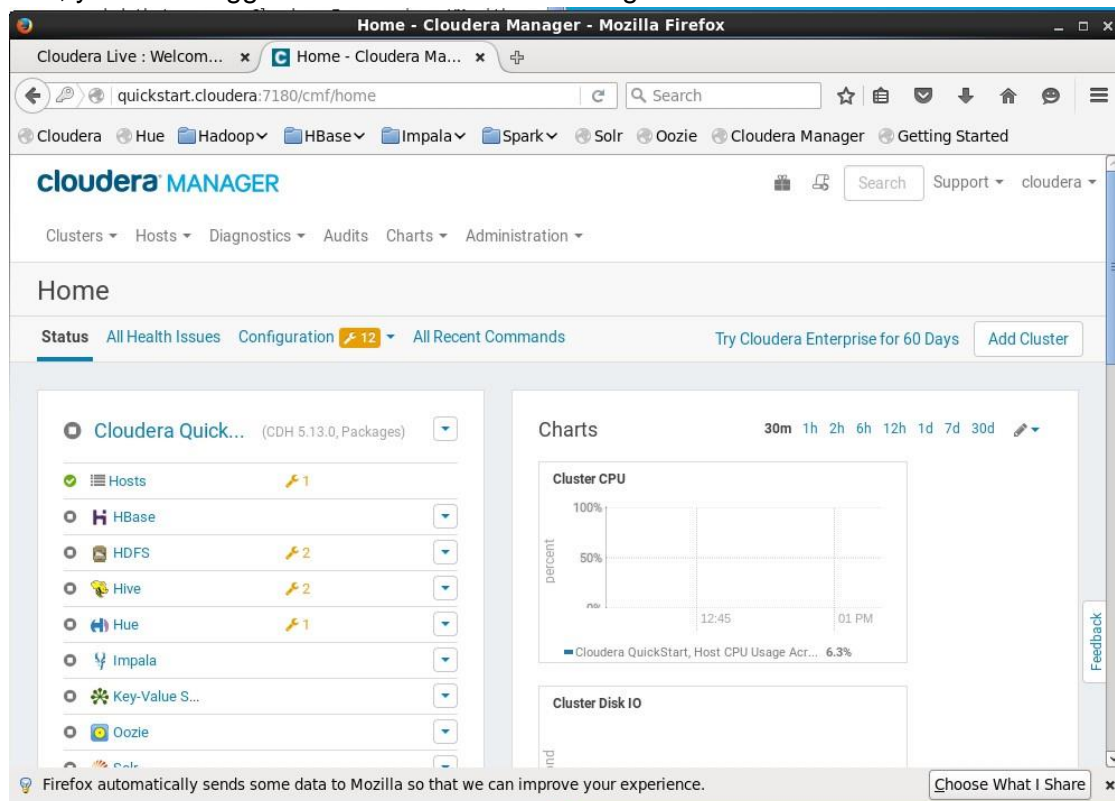
Part 4: Check out the Cloudera Manager

To make sure the Cloudera Manager is running on your VM, open the terminal and run `sudo /home/cloudera/cloudera-manager --force --express` command to reboot the Cloudera Manager server. It will take a few seconds for the agent to restart.
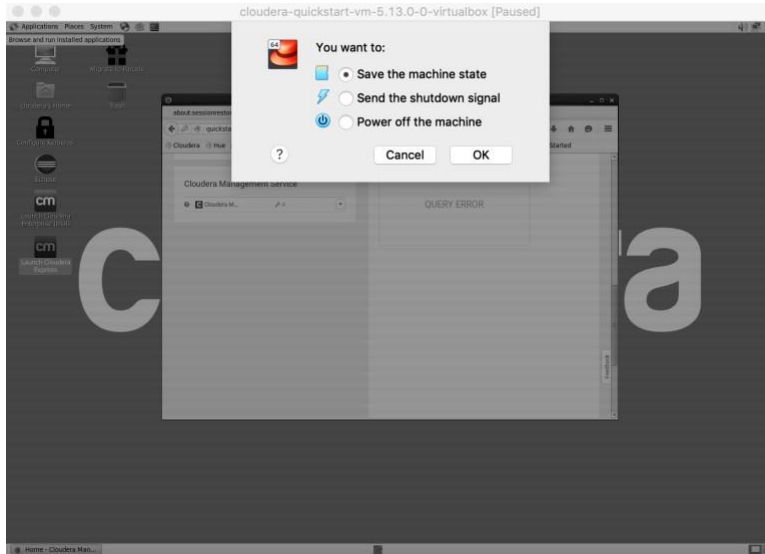

After the terminal finishing executing, open the **browser**, click, and go to the **Cloudera Manager**.

When you are at the login page, type in **cloudera** for both the username and the password. Now, you have logged on to the Cloudera Manager.

Part 5: Close the VM

Remember to **save the machine state** after you have done working on it. It will save the machine state exactly the same at the time you exit, so when you start the machine next time, you can continue working on what was left.

For more detailed set-up tutorial, check this video:
https://www.youtube.com/watch?v=HP4g2BU7-xU&t=578s

Connect to Hive:

```
to check status  sudo service hadoop-hdfs-

namenode status

to start

sudo service hadoop-hdfs-namenode start
```