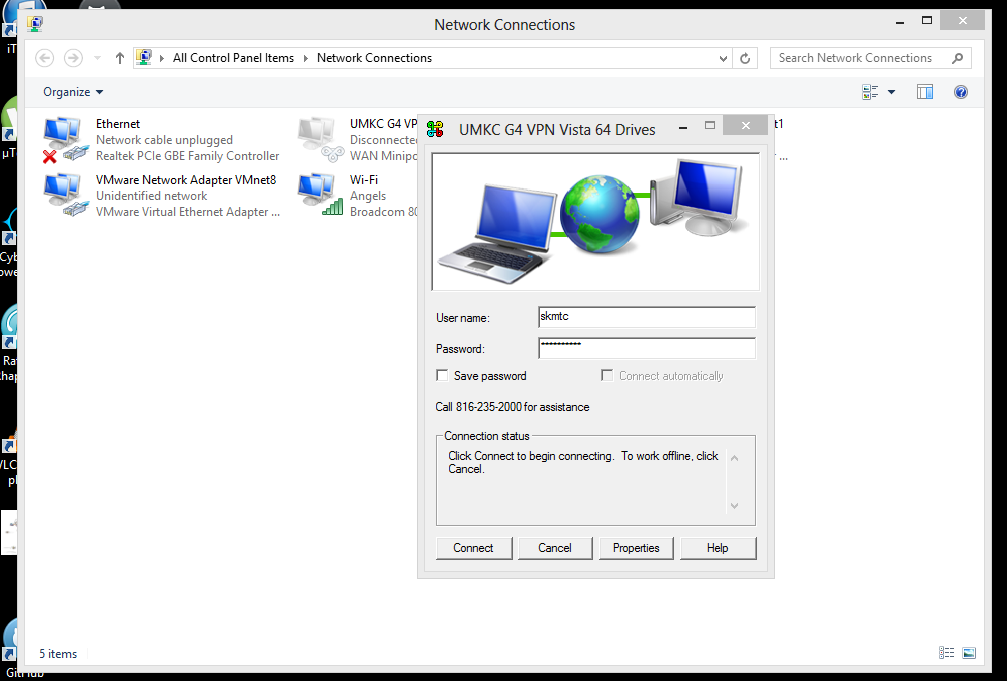
Task 2 :

SubTask1 :

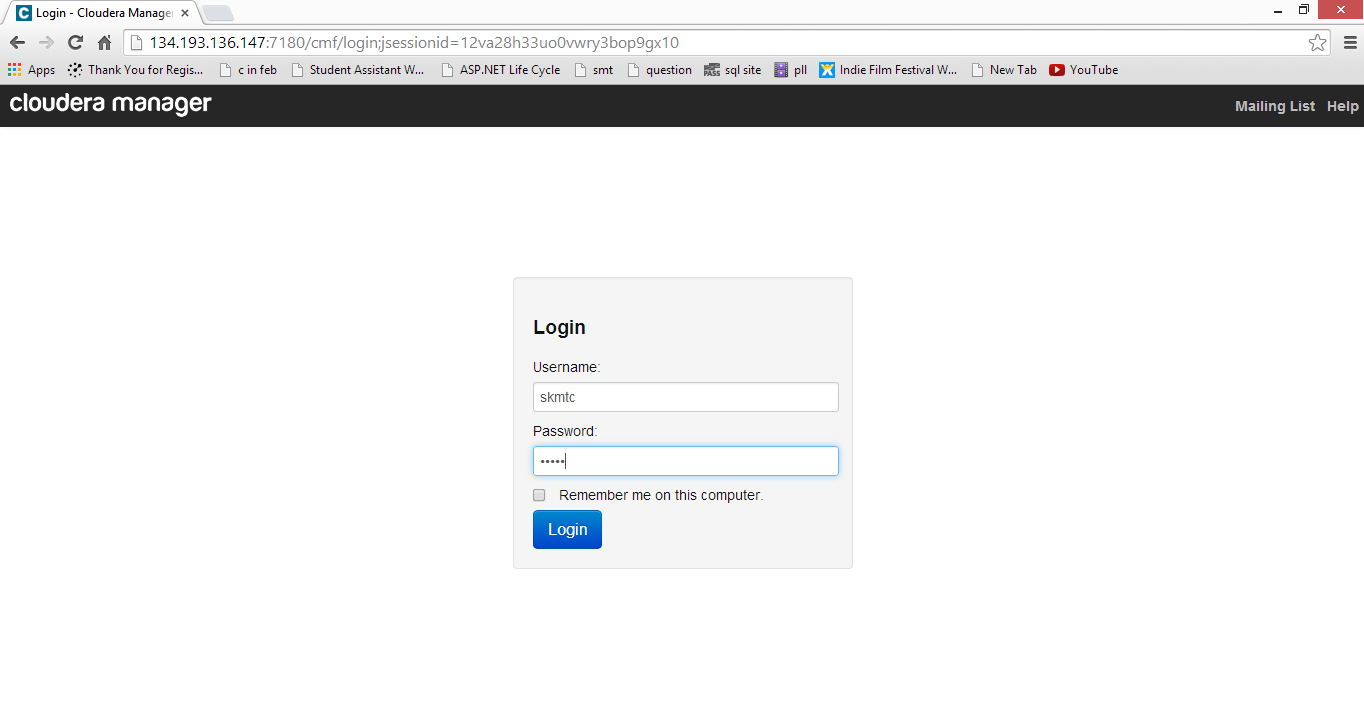
How to access UMKC cloudera Servers

First I connected to umkc VPN as I am outside of college.

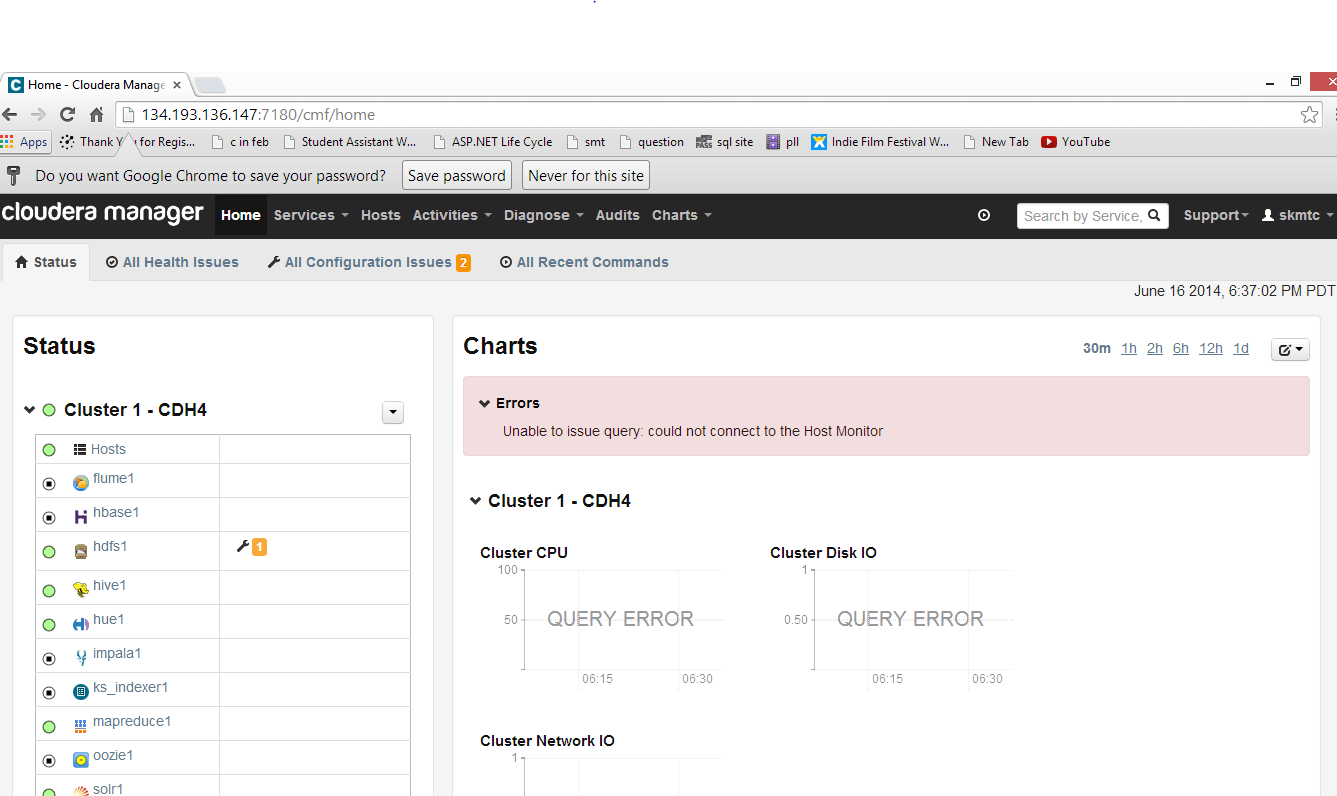


We should now login to the cloudera manager.For this we use the following <http://134.193.136.147:7180> link where 7180 is the port number to reach cloudera manager.

In this link we can enter our credentials in the following way.

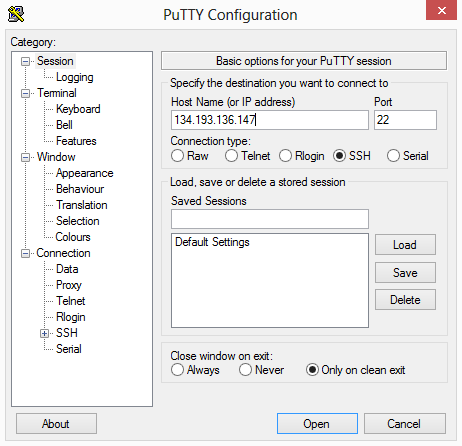


After logging in you can explore the page and use its various services .We can check whether it is having good health or not and proceed further according to it.

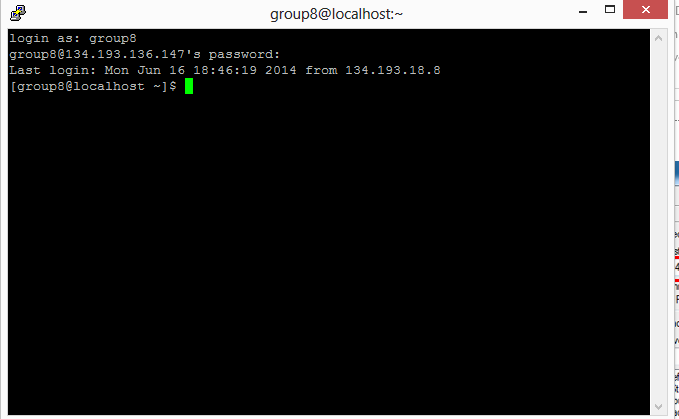


We have access to the remote machine terminal.

We can download putty and login to the system terminal with our credentials.



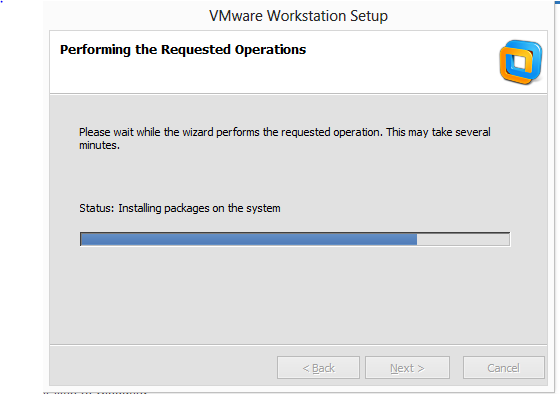
Then we can login using our credentials.



Subtask2 : Installing our own cloudera server

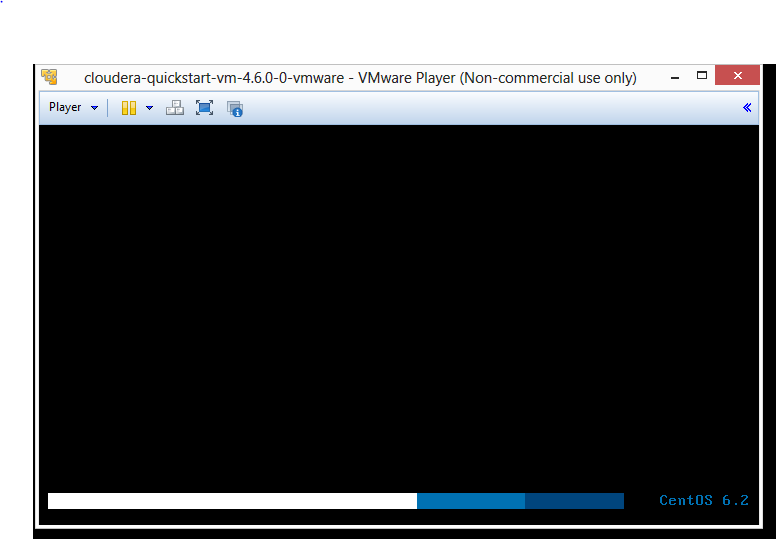
We should now install VM player from <http://www.vmware.com/products/player>





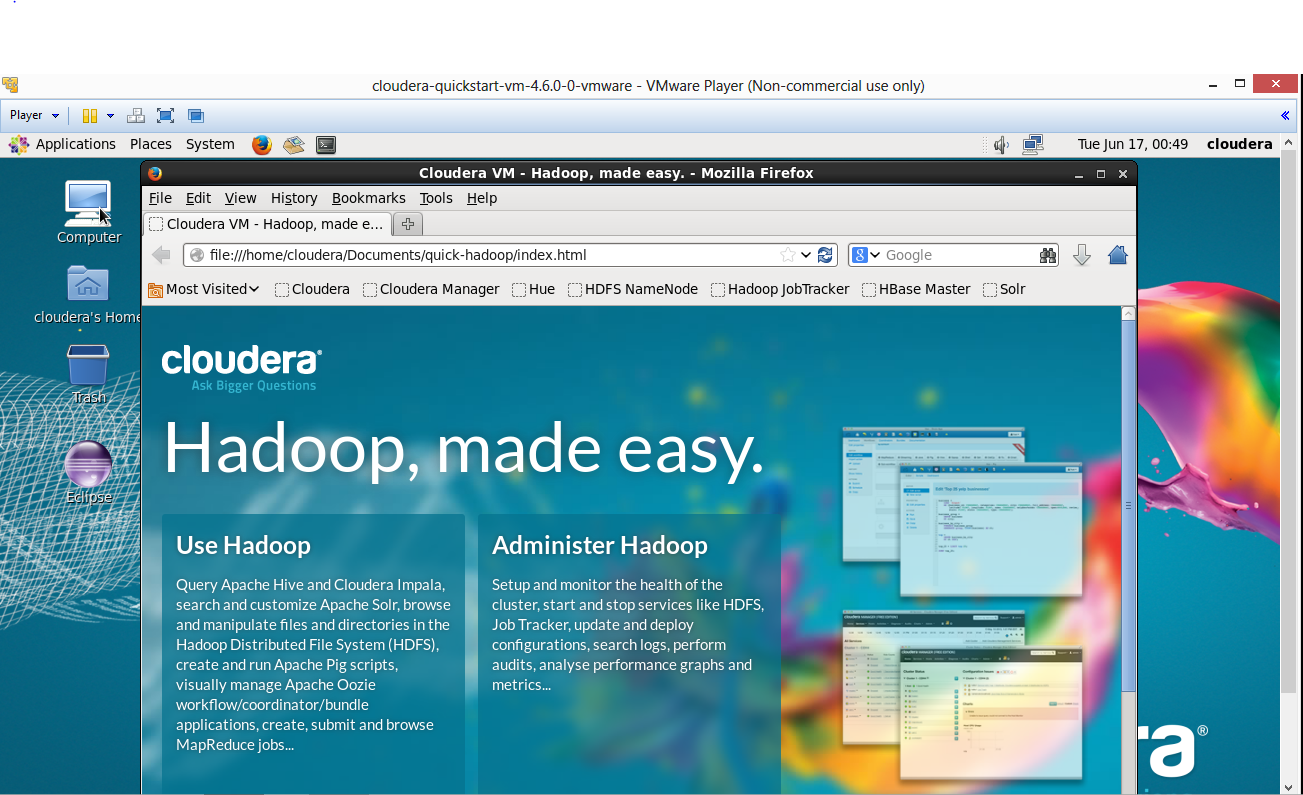
After installing we have to get the cloudera image from <http://www.cloudera.com/content/support/en/downloads/download-components/download-products.html?productID=F6mO278Rvo> .I installed cloudera 4.6 image .

Now let us open the image 4.6 in vmplayer

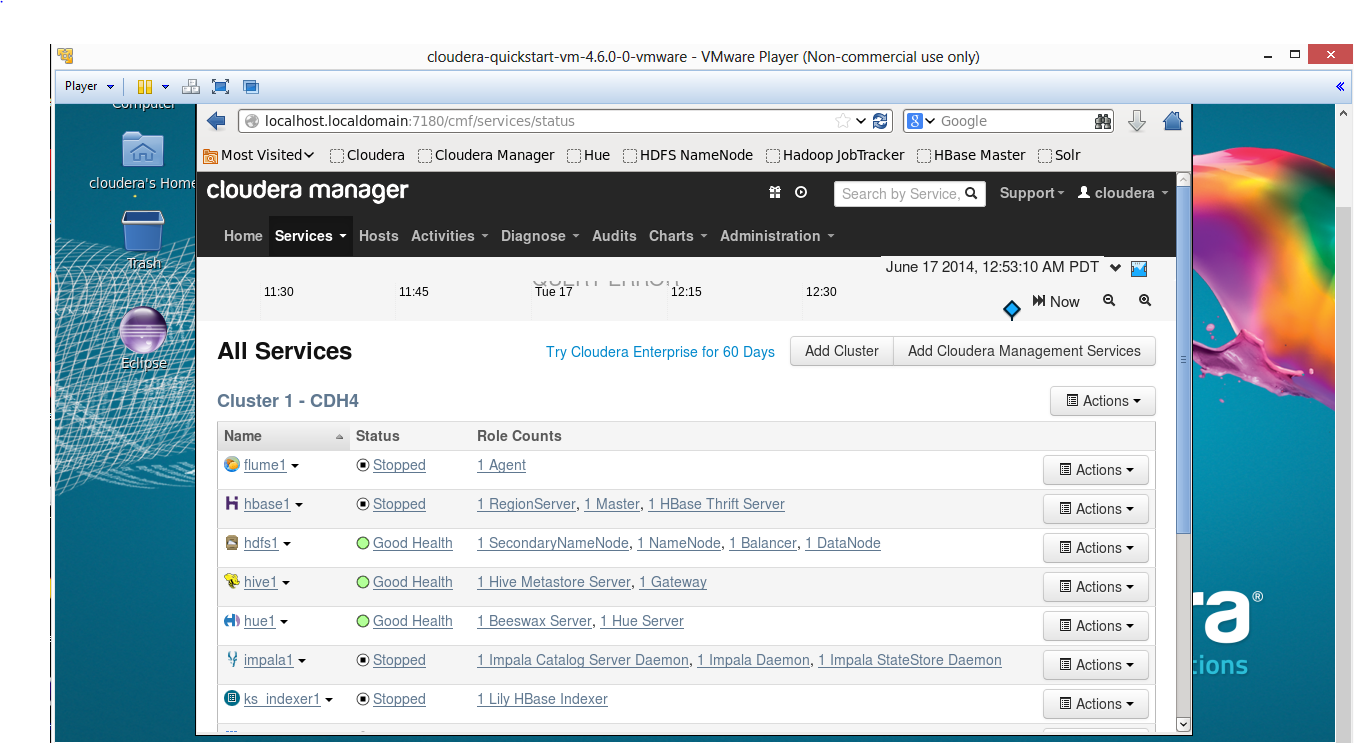


The VMplayer opens the image .

After that we willhave to see the desktop with firefox opening the following page.



We can select cloudera manager.And give the credentials as “cloudera” and login and we can access the cloudera manager.

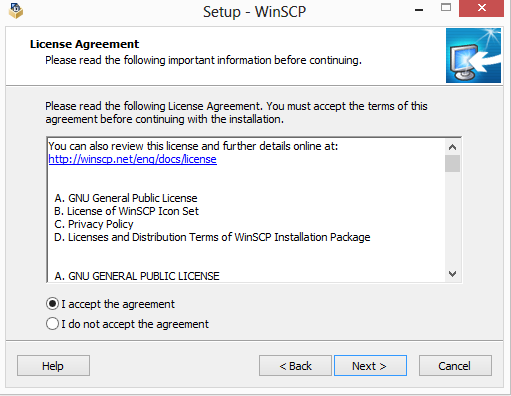


Here all the services are showed and the status of each service is also given.

We can also login with HUE credentials where we will be directed to HUE home page.

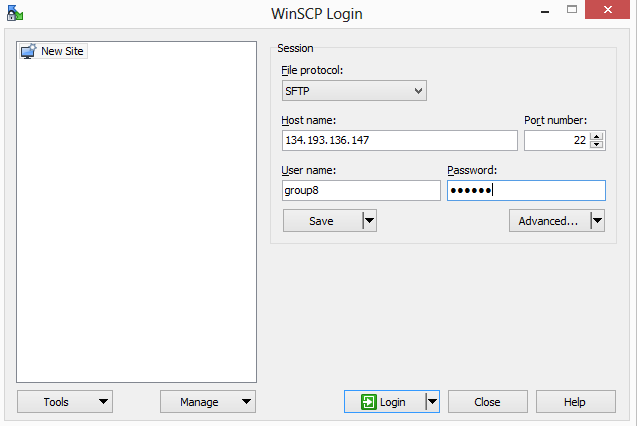
Subtask3 : Transferring files to Cloudera

WinSCP and FIlezilla can be used to transfer the files to cloudera. I downloaded WinSCP

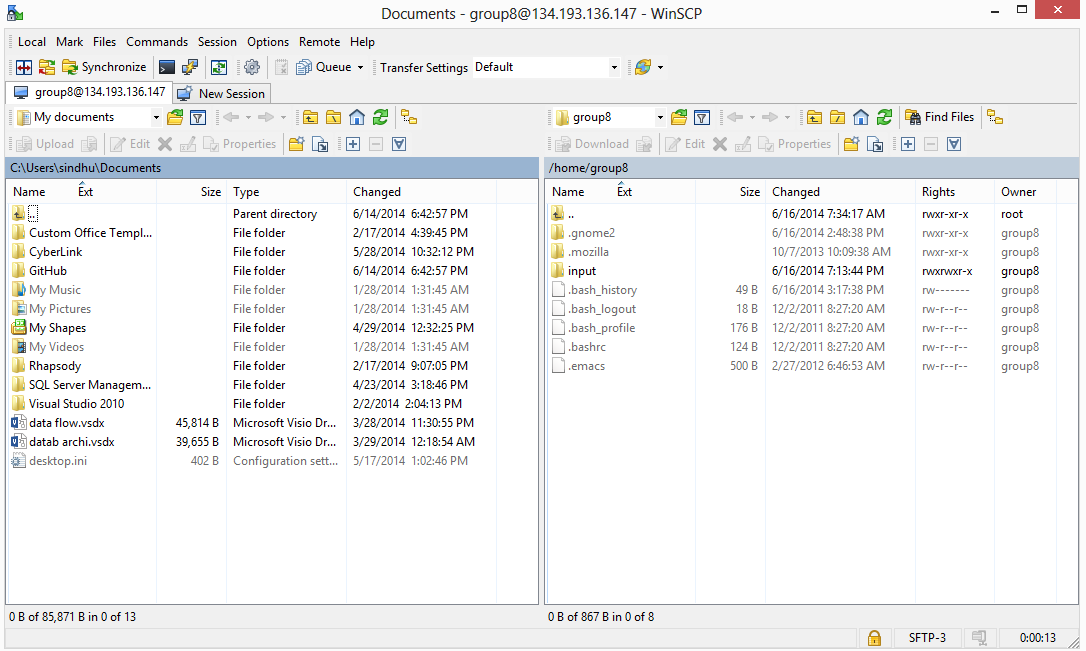


After downloading and accepting the conditions.

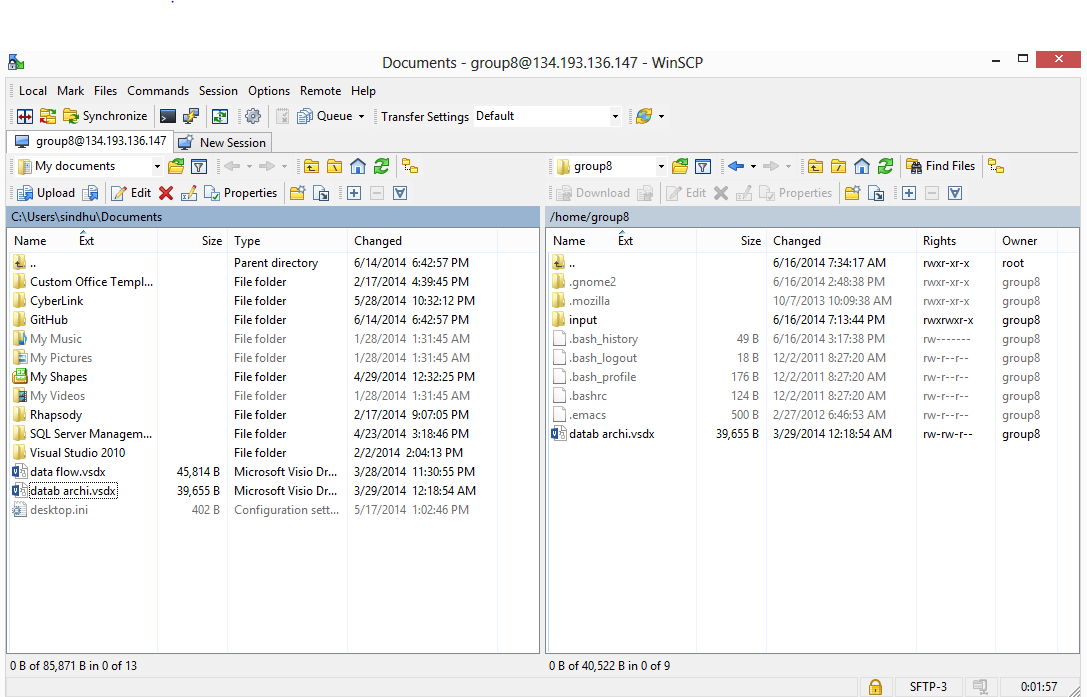
Now we need to connect by using the host name and username and also the password.



After logging in the screen looks like the following.



We will have the local files on one side and remote files on another side.We can just drag and drop the files to transfer them.I just transferred one datab archi file.



Subtask4:

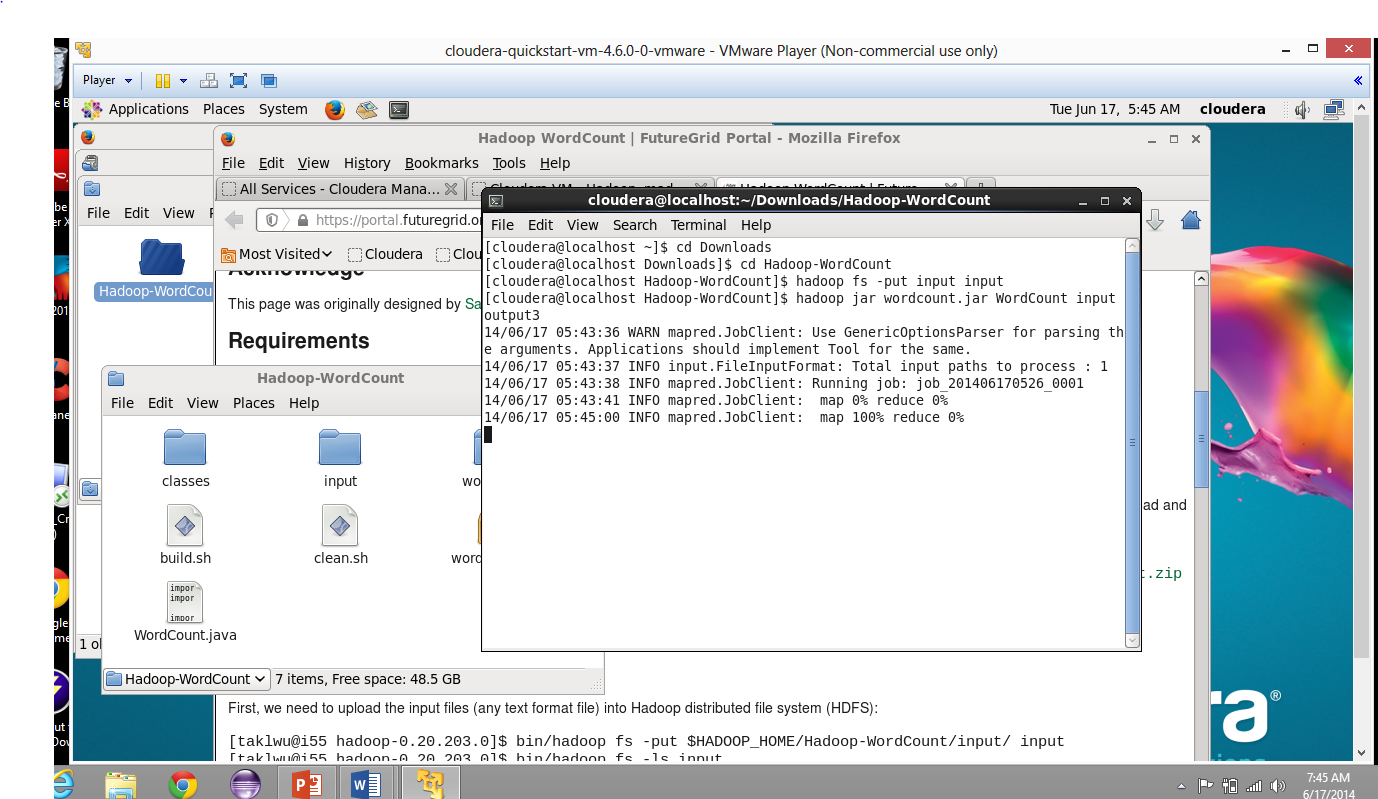
Running wordcount :

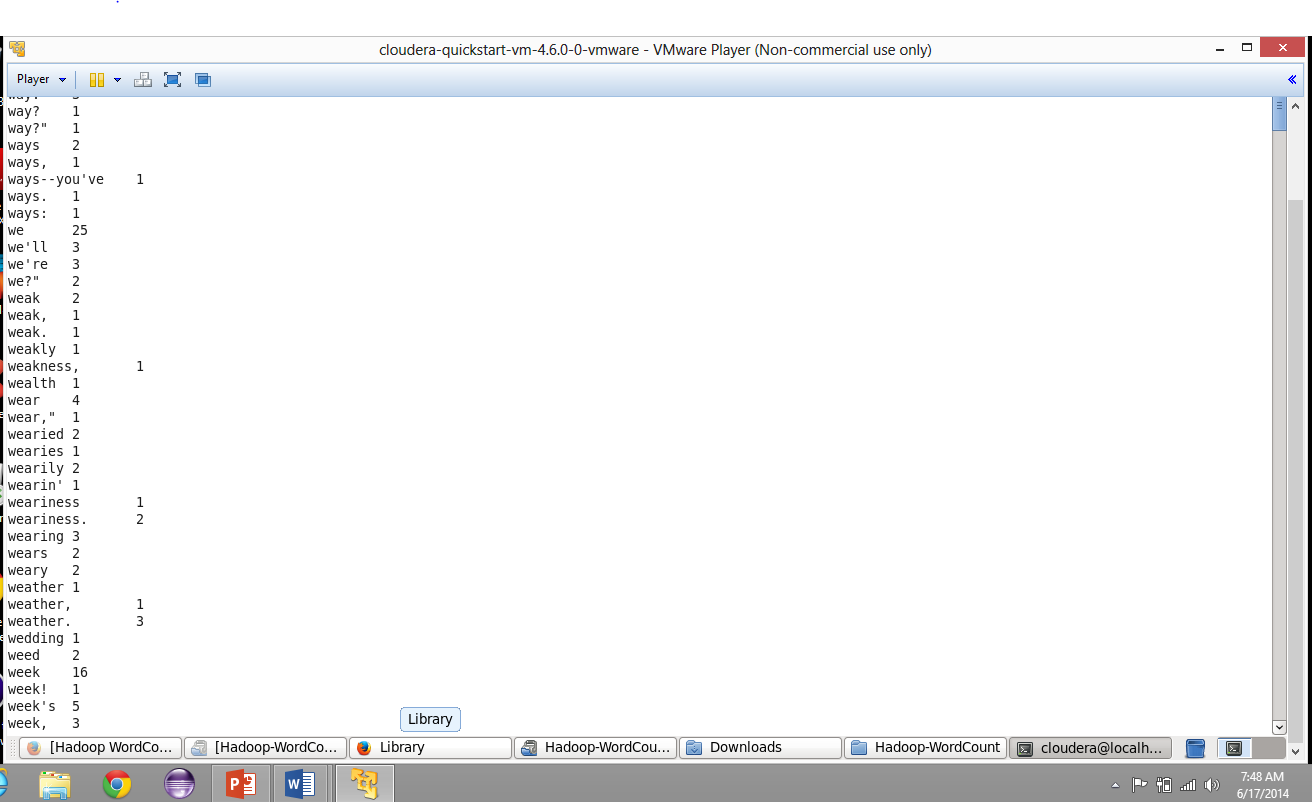
Now we can download the source code of word count example from <https://portal.futuregrid.org/manual/hadoop-wordcount>

This is a zipped file and we should unzip it.

We have to type some of the commands we have to keep the local file to Hadoop input directory Hadoop fs –put input input

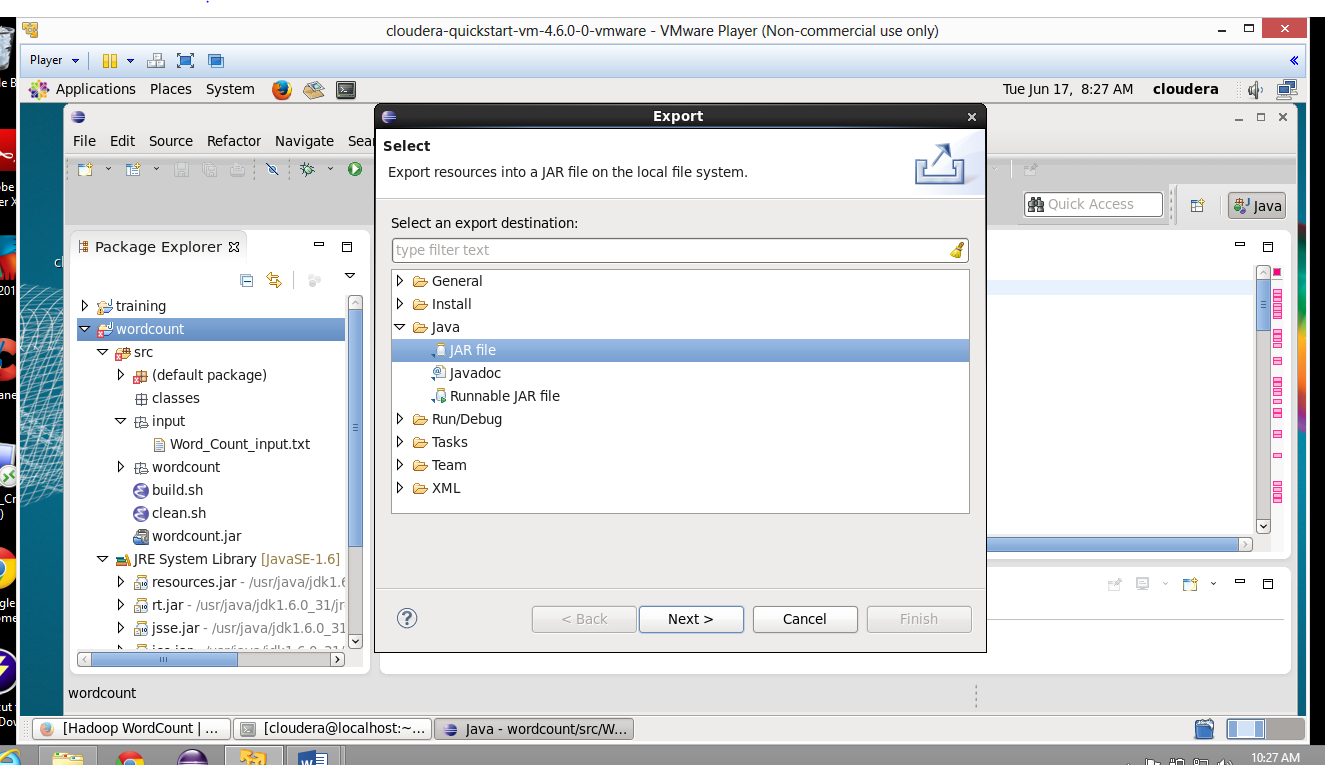
And then we should run Hadoop using hadoop jar wordcount.jar WordCount input output3

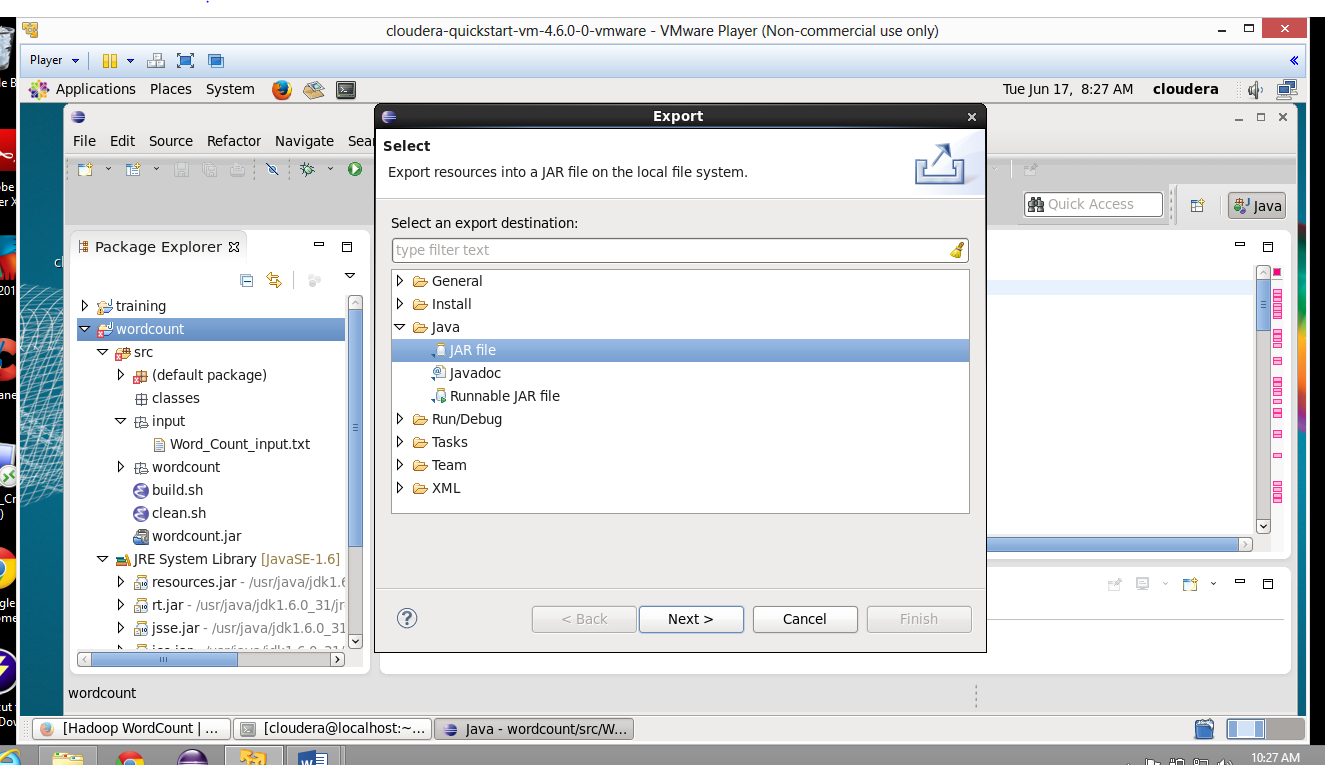


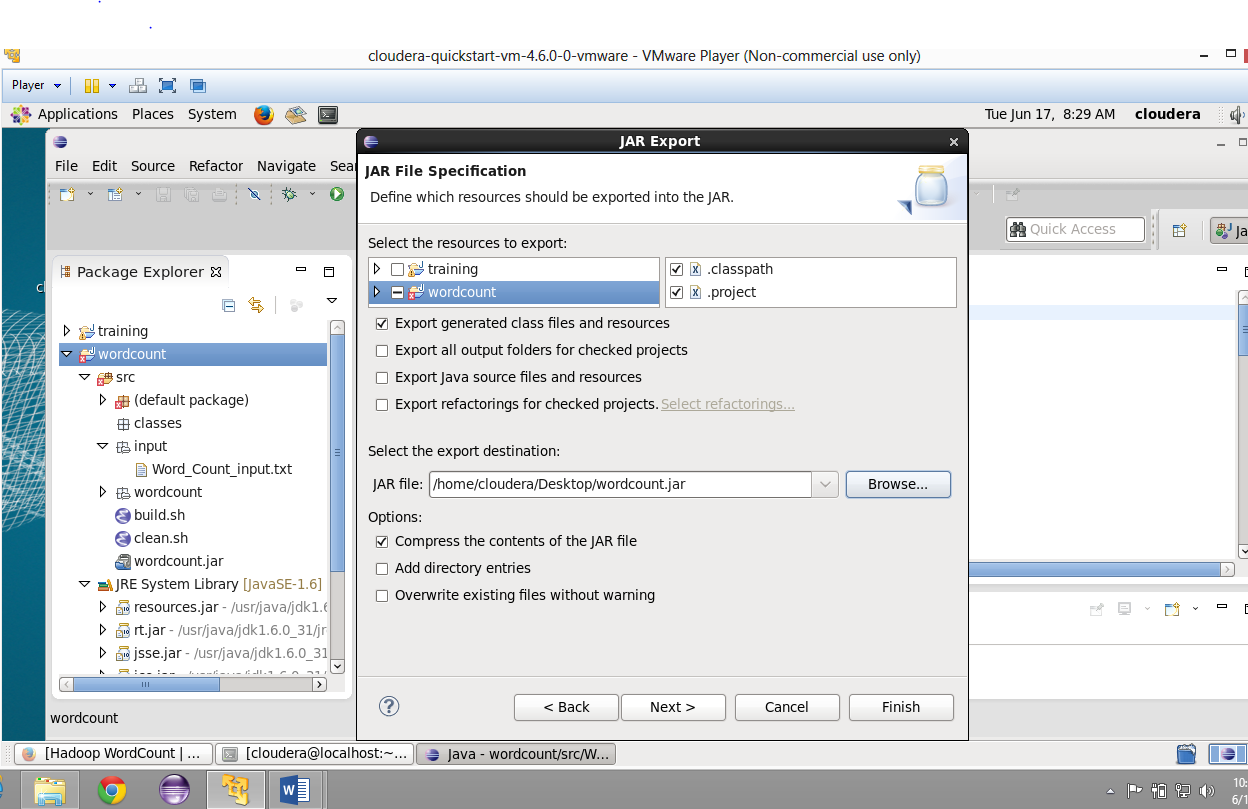


Subtask 5 : How to make hadoop jar and run it in cloudera

Import the wordcount project into the eclipse in linux machine. And now export it as jar file.

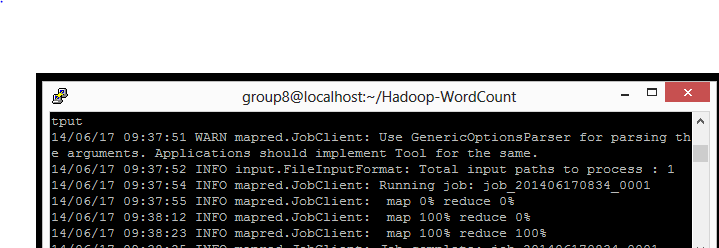






After we are done with exporting , we will go to the putty and connect to remote machine terminal.

Now we will go to current folder and then give the command hadoop jar wordcount.jar



We can check the result by giving appropriate command and it is like this.

