Lab 01 report

Individual Tasks

Name

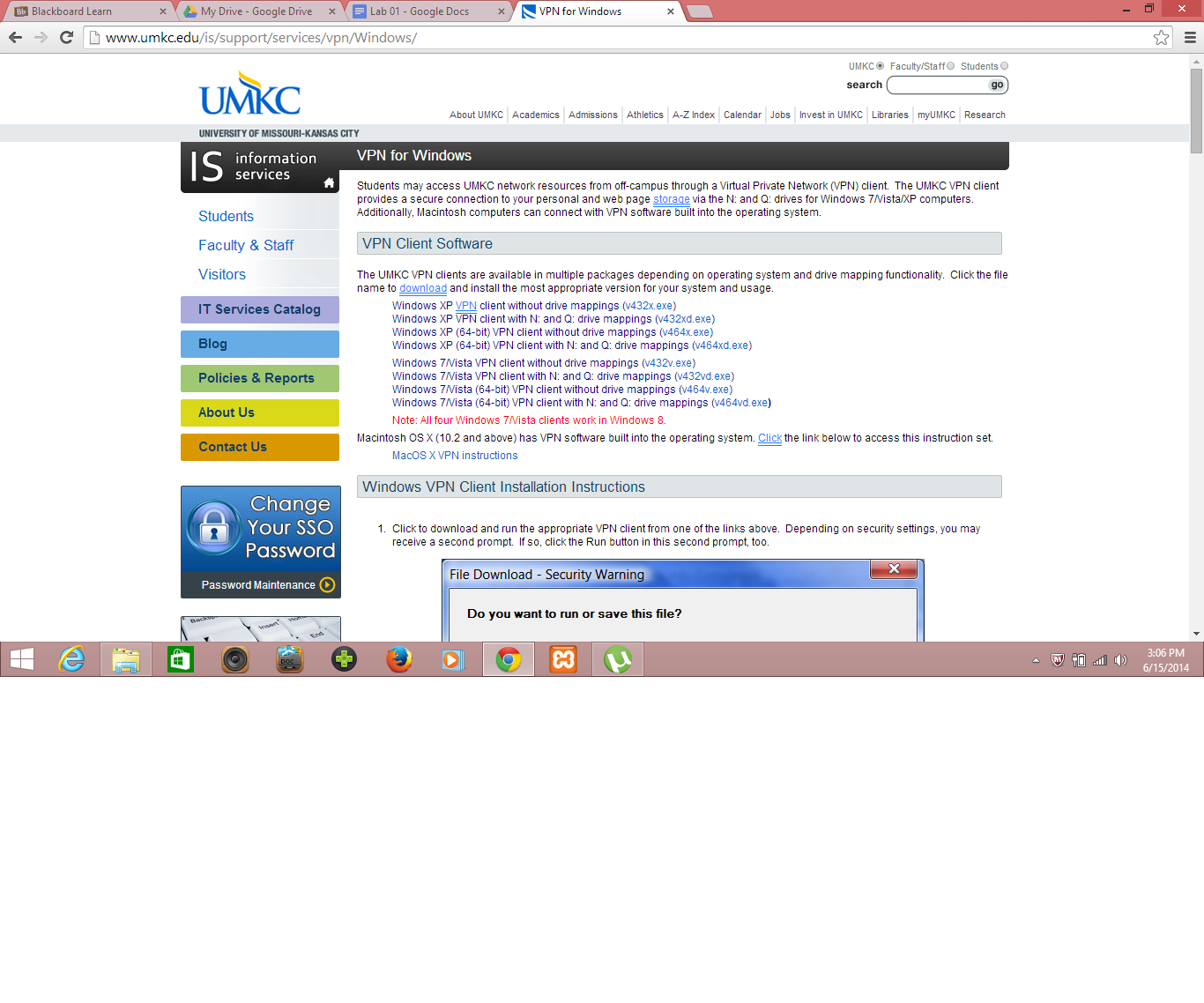
Tirumala Reddy Konireddy - 16174347

**Tasks 02**

**Subtask 01**

**Connecting to UMKC Cloudera Servers:**

First, I connected to UMKC VPN using v464.exe downloaded from UMKC website.

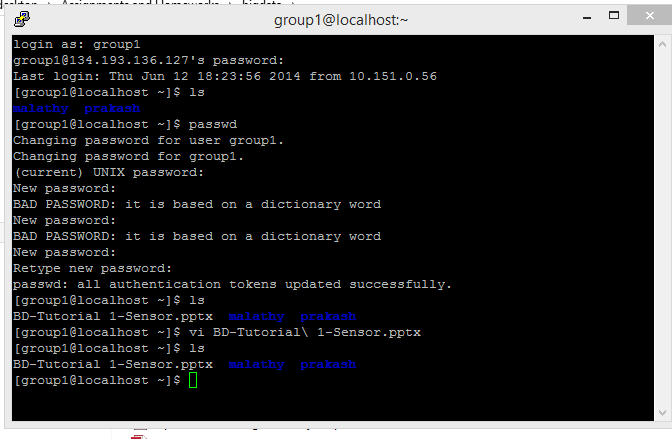


I installed v464.exe file and run UMKC VPN to get connect into UMKC VPN.

Later, I used our group credentials to connect to College Cloudera Servers as below as I got access denied with my SSO credentials.



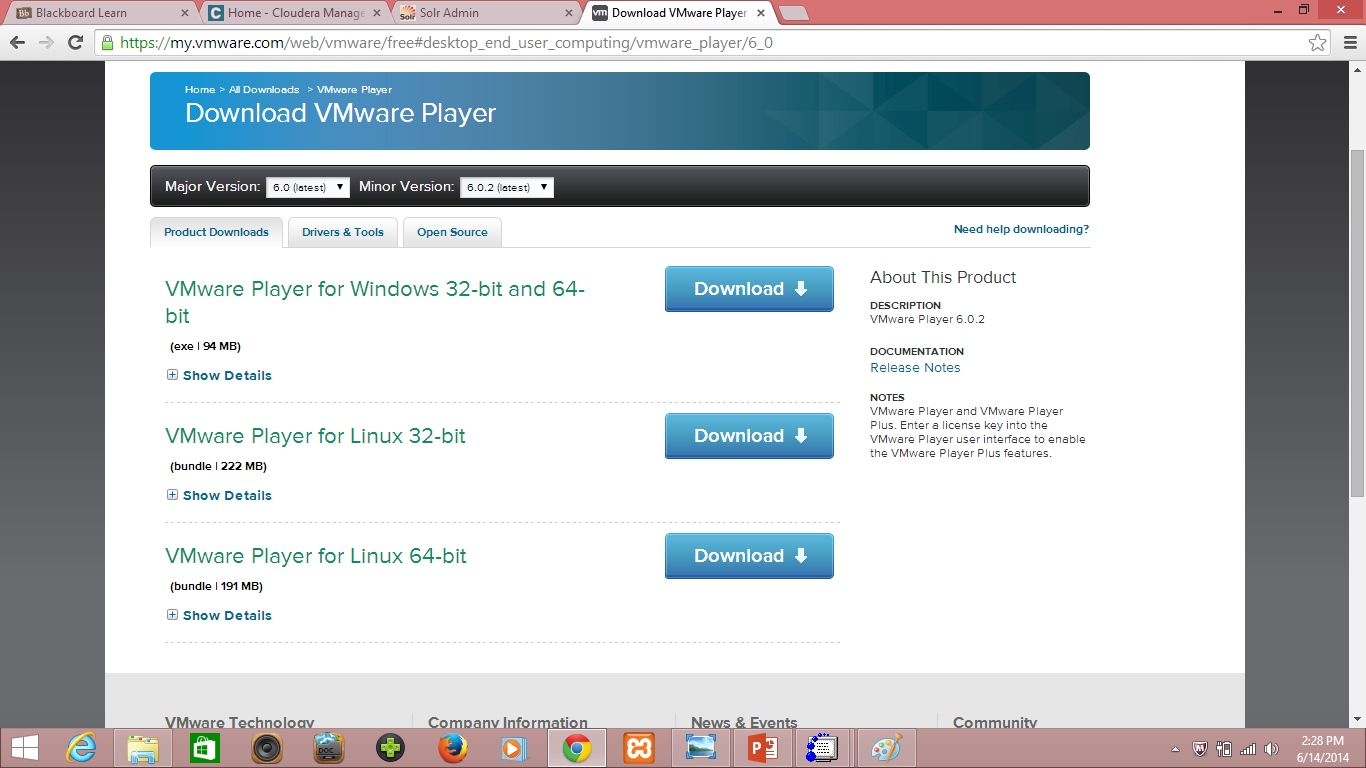
First, I used Putty to execute commands in College Cloudera server.



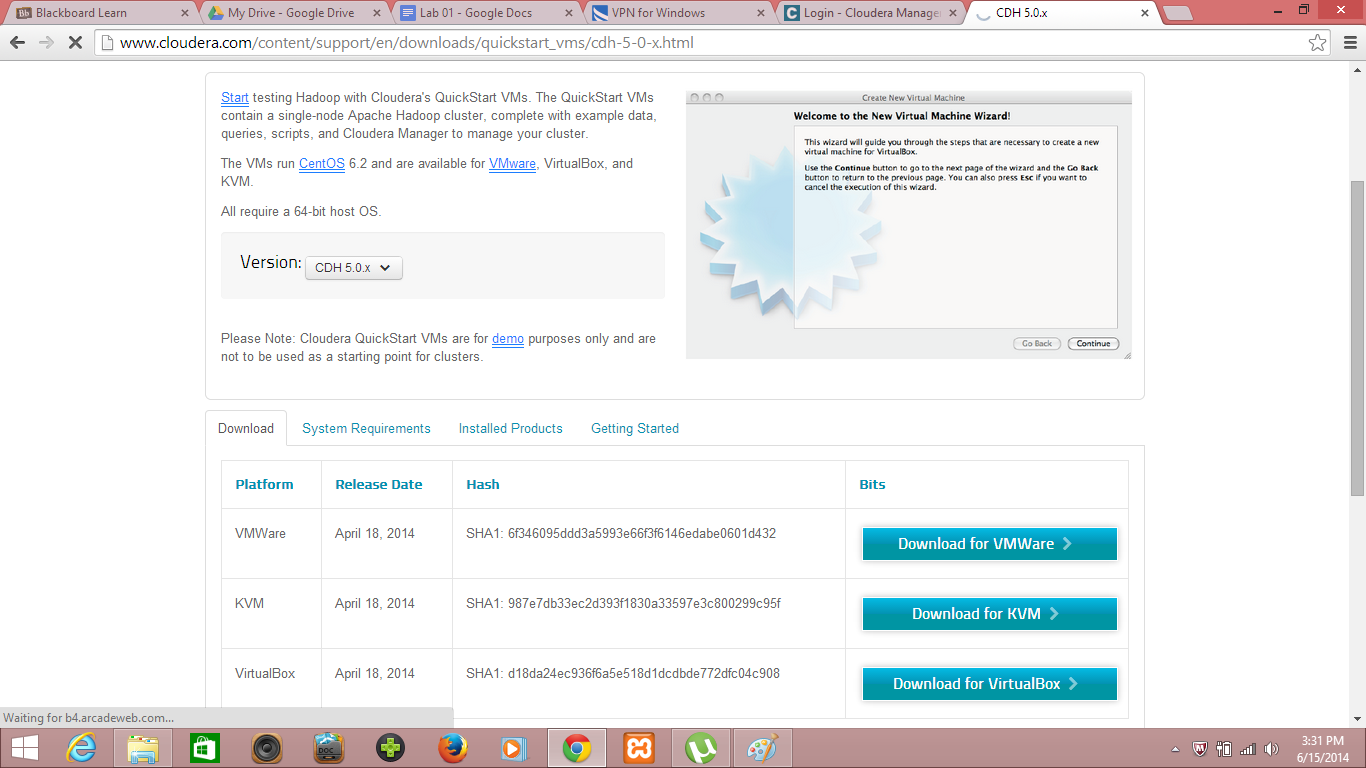
**Subtask 02**

**Installing a own Cloudera Server:**

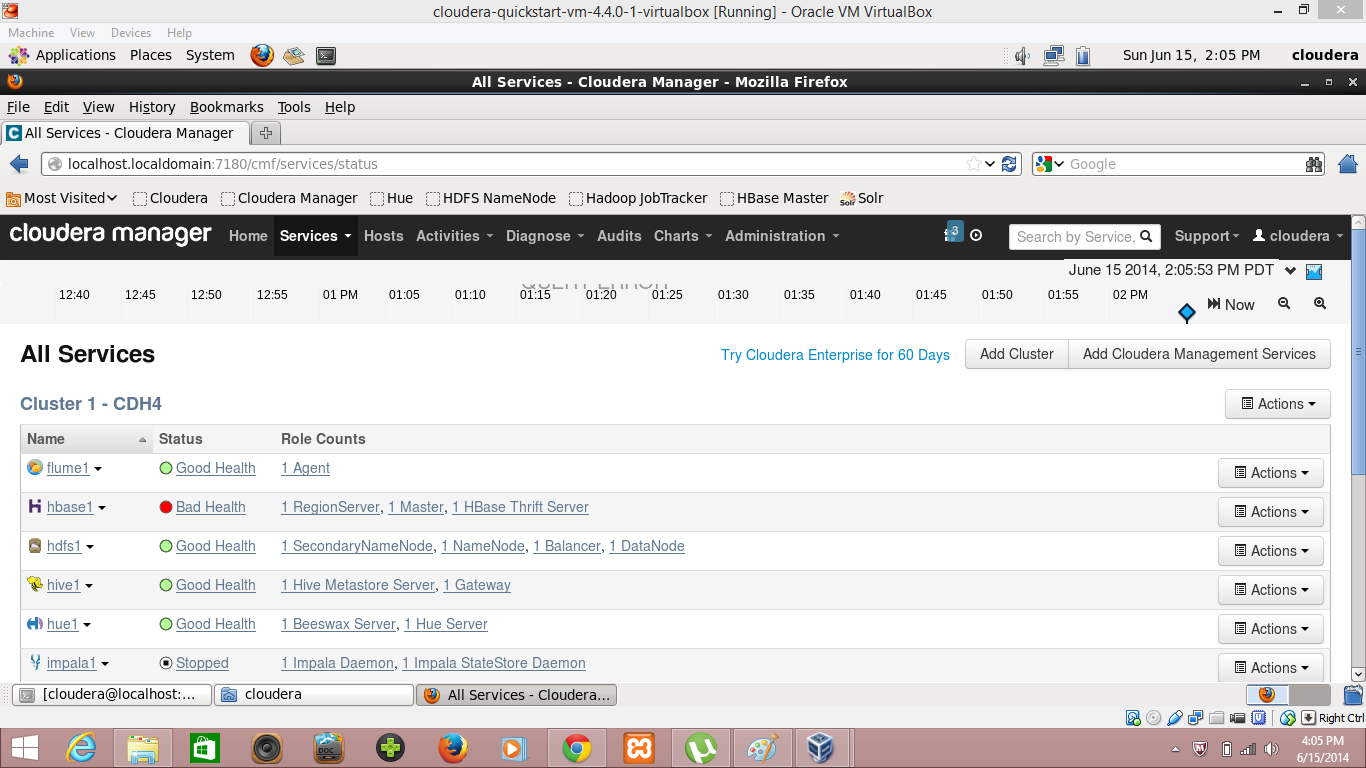
I had installed VMware player downloaded from the link given in the Tutorial.

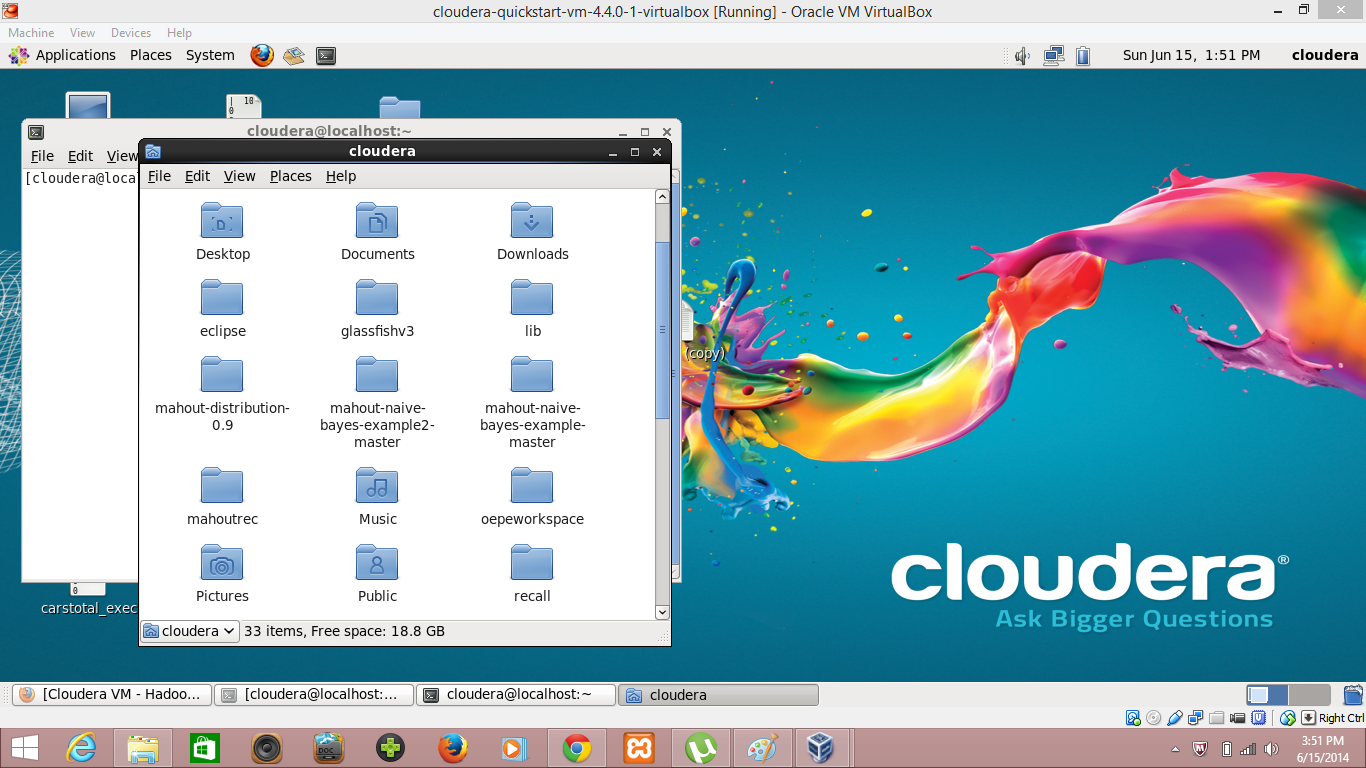


I had downloaded the Cloudera image for vmware player from the given link in tutorial.



Then, I had changed my machine BIOS settings to allow virtual machine. But, unfortunately my vmware player is unable to play the cloudera image. Then, I tried virtual box image and player to install own local cloudera server. I successfully installed my cloudera server using virtual box image as below.

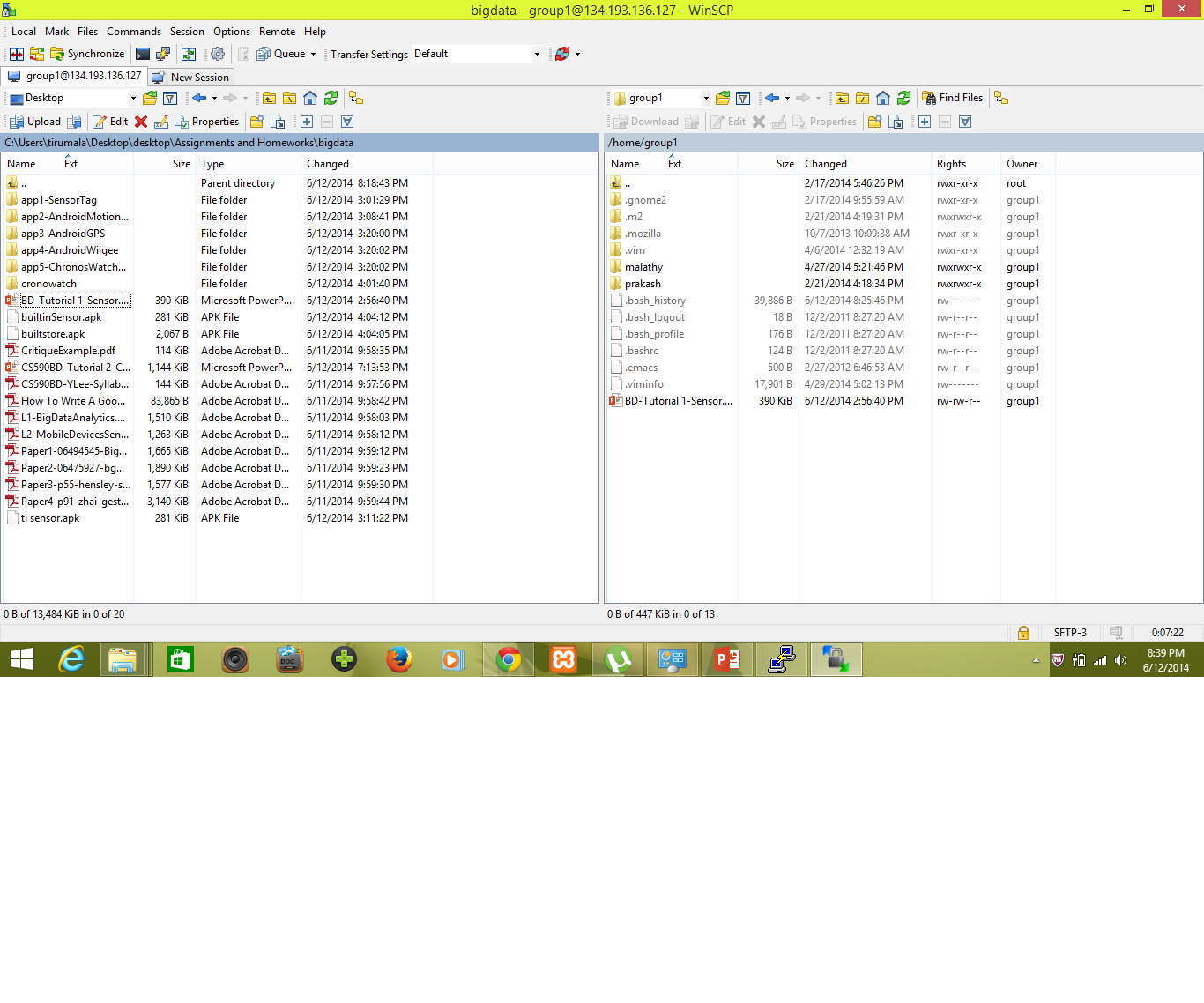




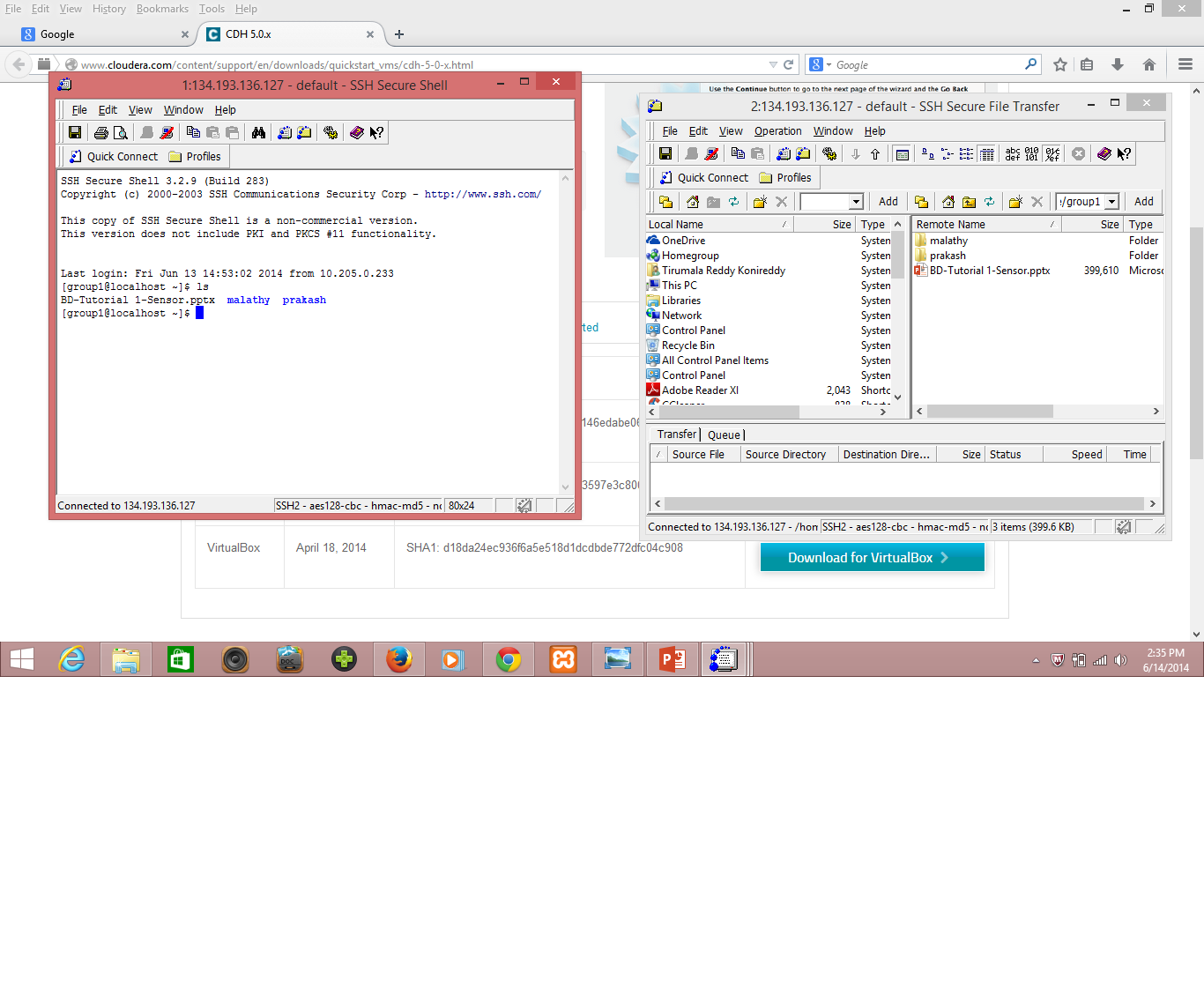
**Subtask 03**

**File transfer between Local drive and UMKC Cloudera Server:**

I used Winscp to do file transfer from my local drive to college Cloudera drive as below.



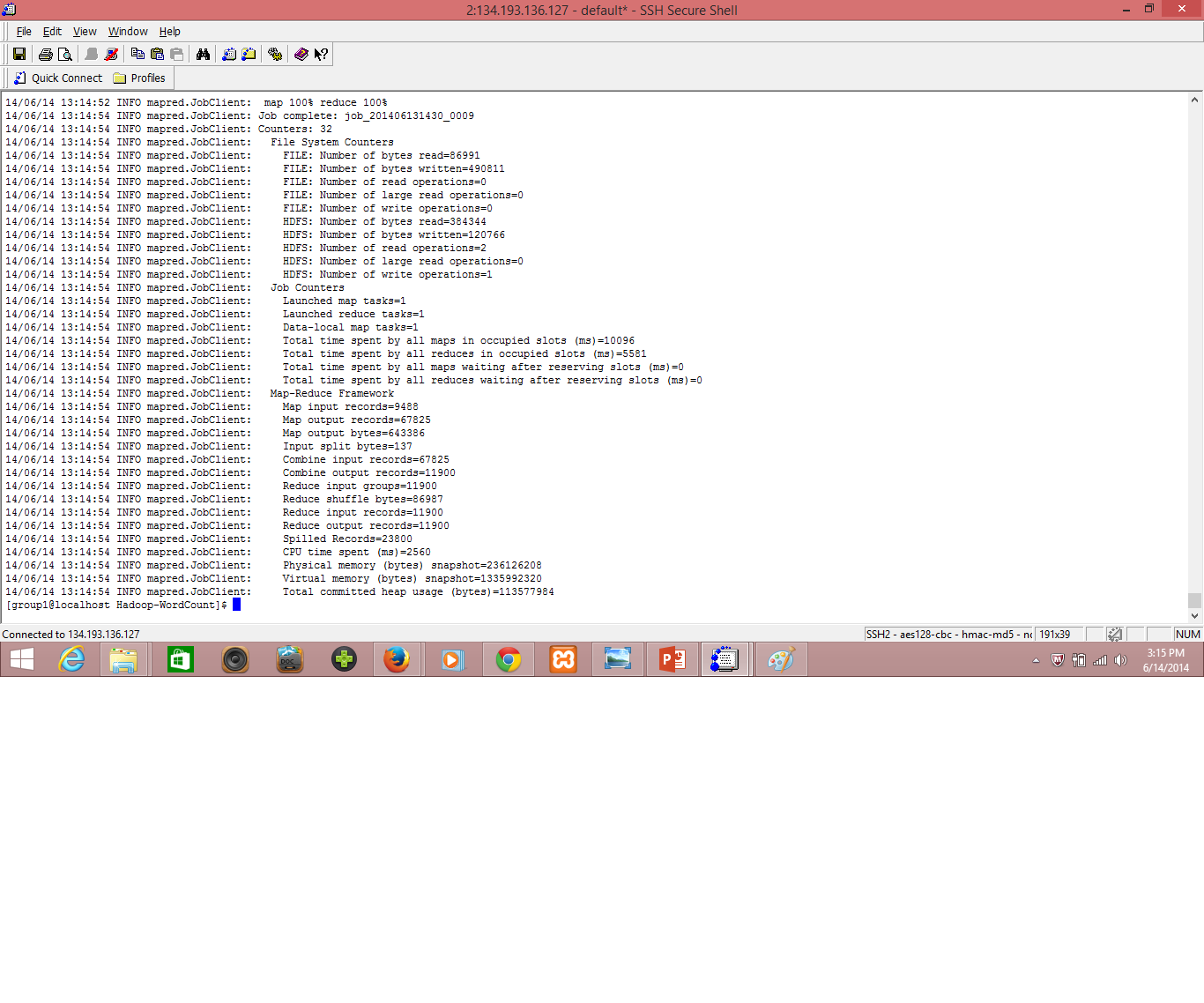
Later, I found Secure shell through which I can execute commands in cloudera server and file transfer between Local Drive and Cloudera Server.

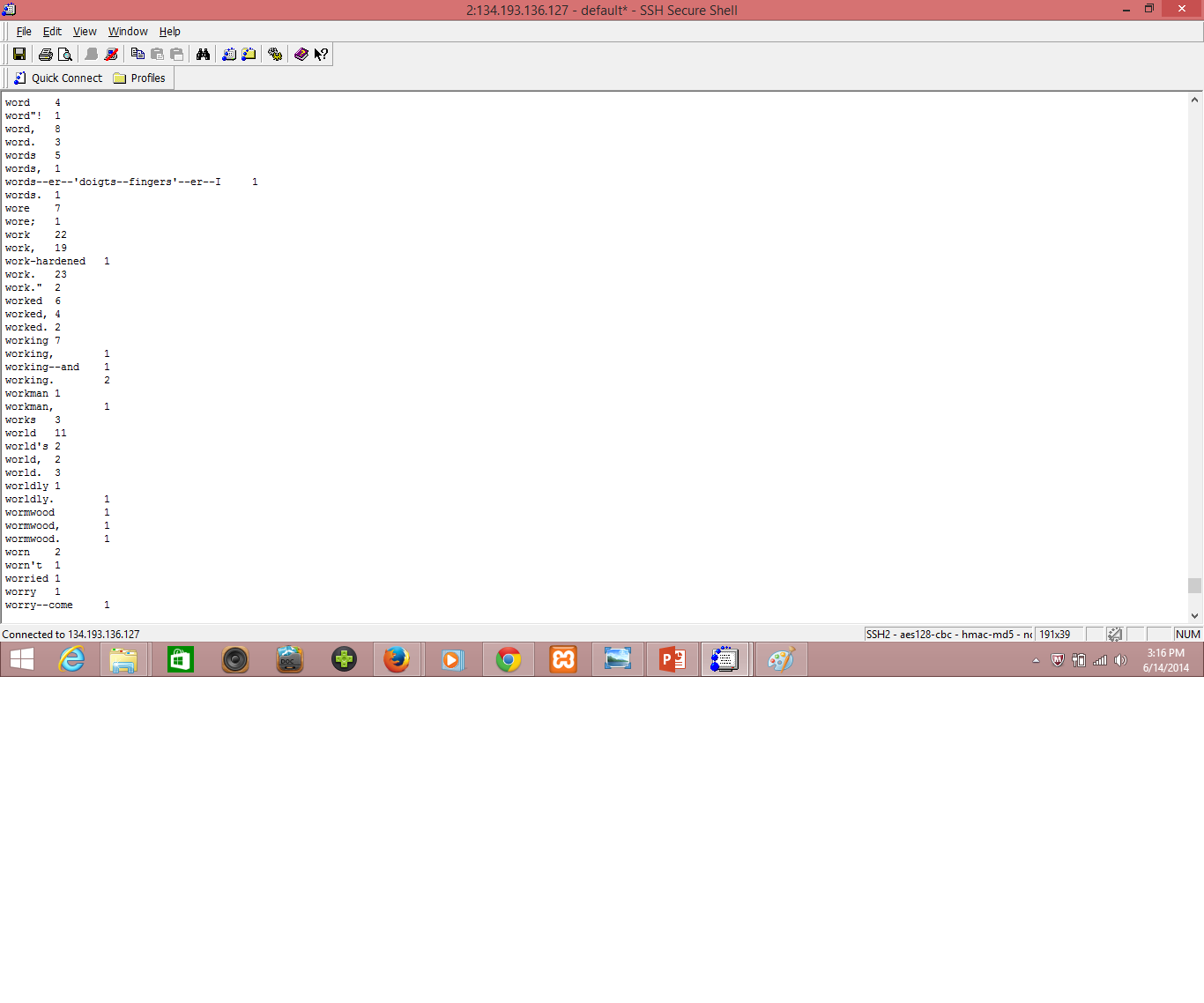


**Subtask 04**

**Running Word Count Example:**

I had well understood the Mapper and Reduce classes of the WordCount Class. I found that, we are using every word as a key and the number of its occurrences in the document as the value in the mapper and reduce class. I am familiar with this program as I had already worked on this type programs in Knowledge Discovery Management course.

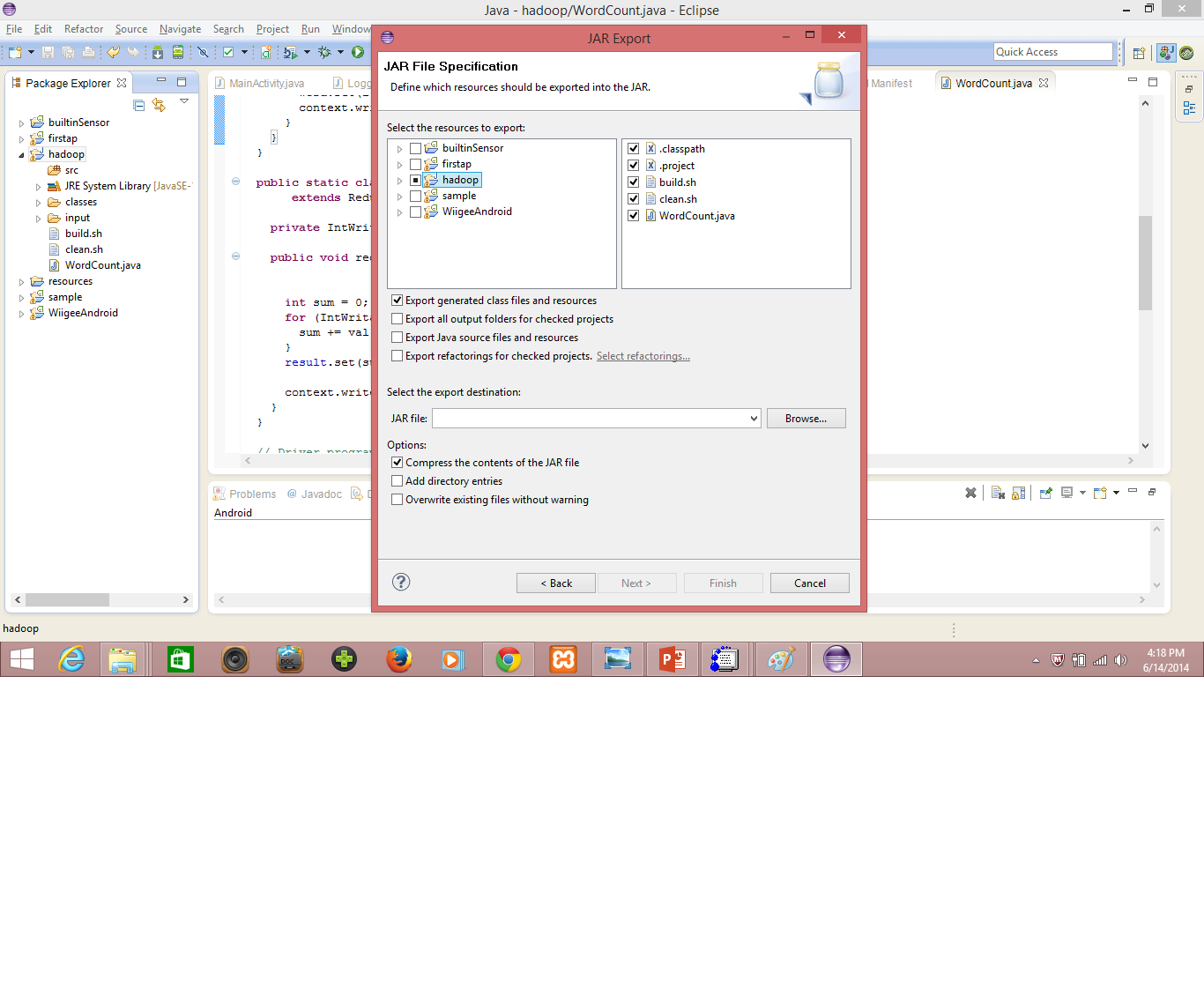




**Subtask 05**

**Creating Hadoop Jar File:**

I had downloaded the word count program source code from the link given in the tutorial. I imported the whole code into my eclipse under hadoop project. Then, I export that hadoop project as jar file as below.



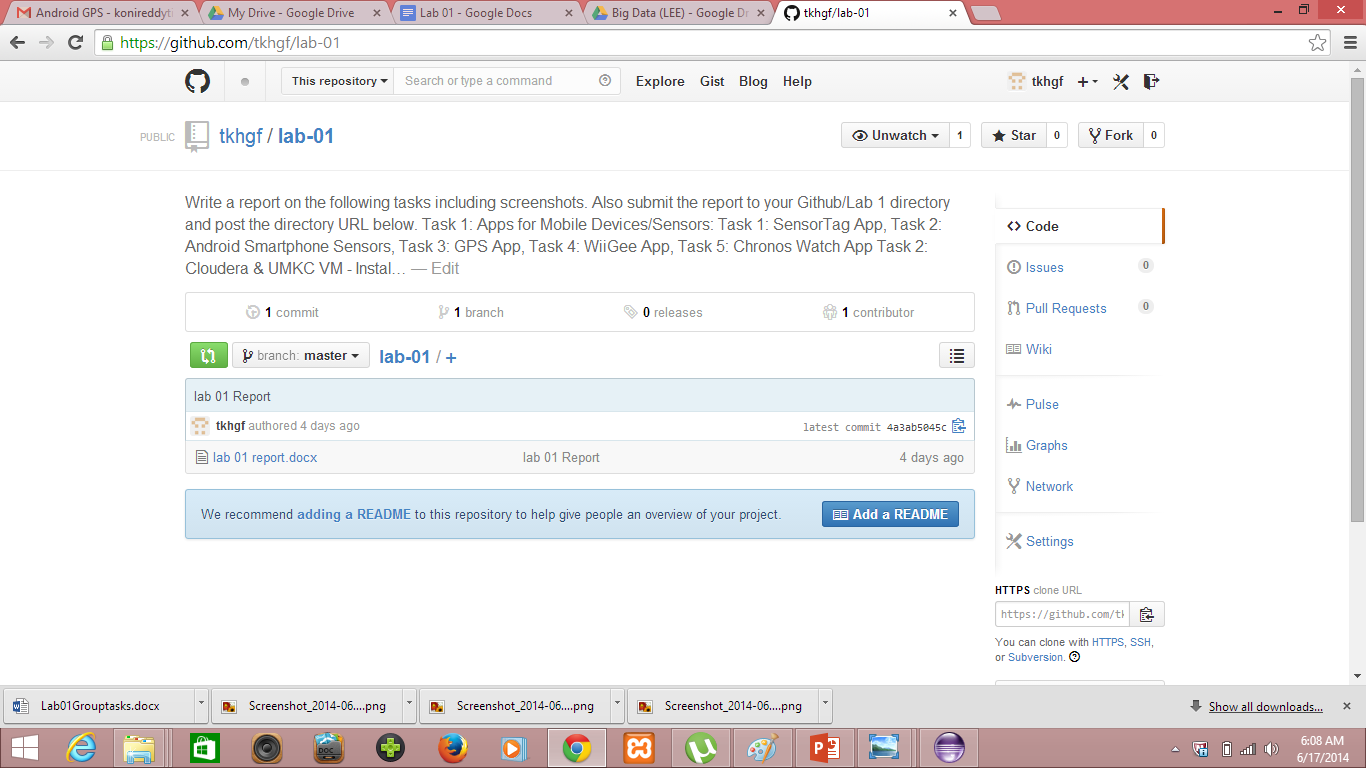
Then, I runned that jar file for word count exercise and got the same results as before in word count example.

**Tasks 03**

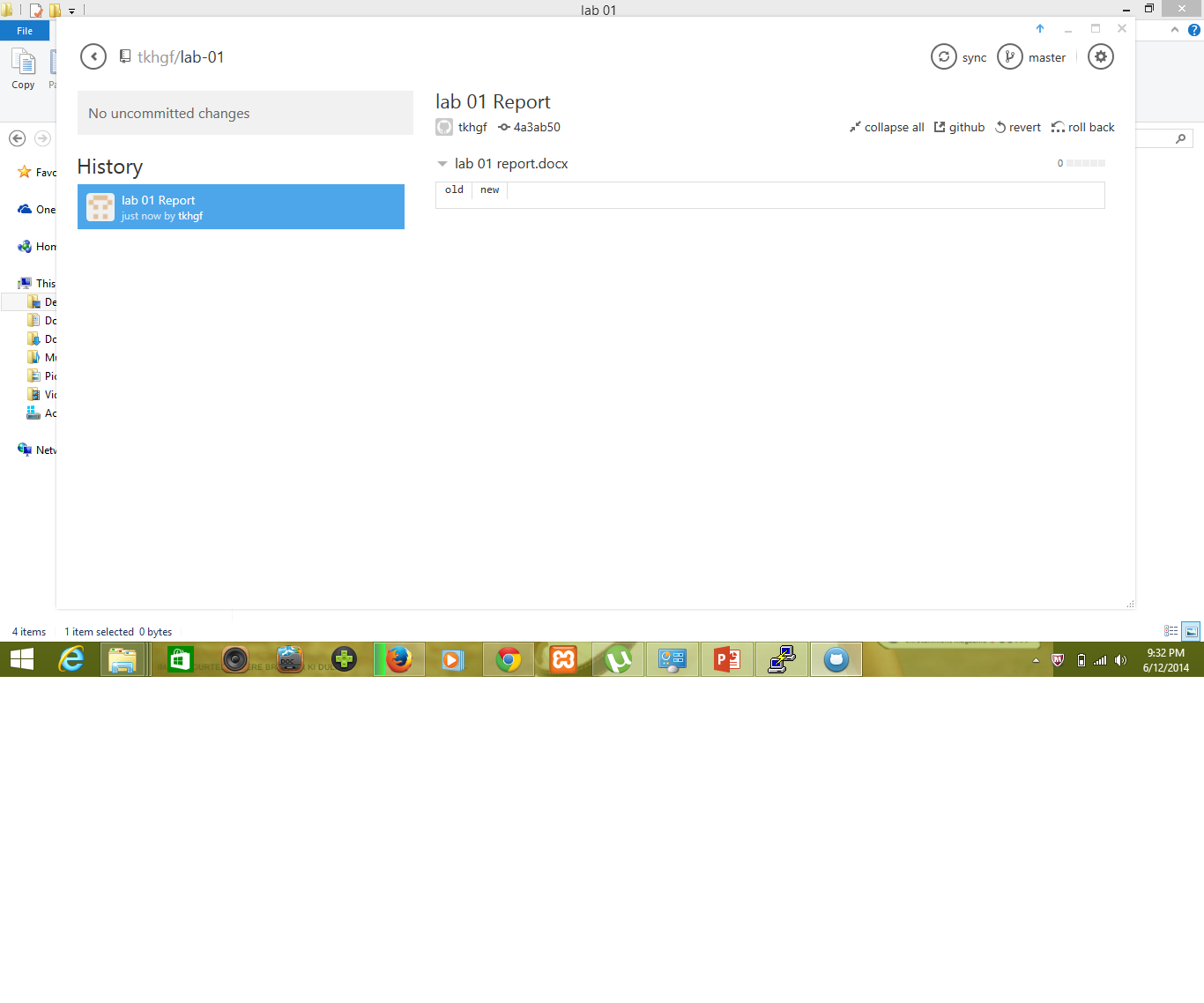
**Subtask 01**

**Github Account Creation:**

I had already a Github account with username tkhgf. I had created this account as part of my KDM course. I am familiar with uploading and deploying documents, programs to Github. I had created a new repository as part of my Lab 01 assignment.

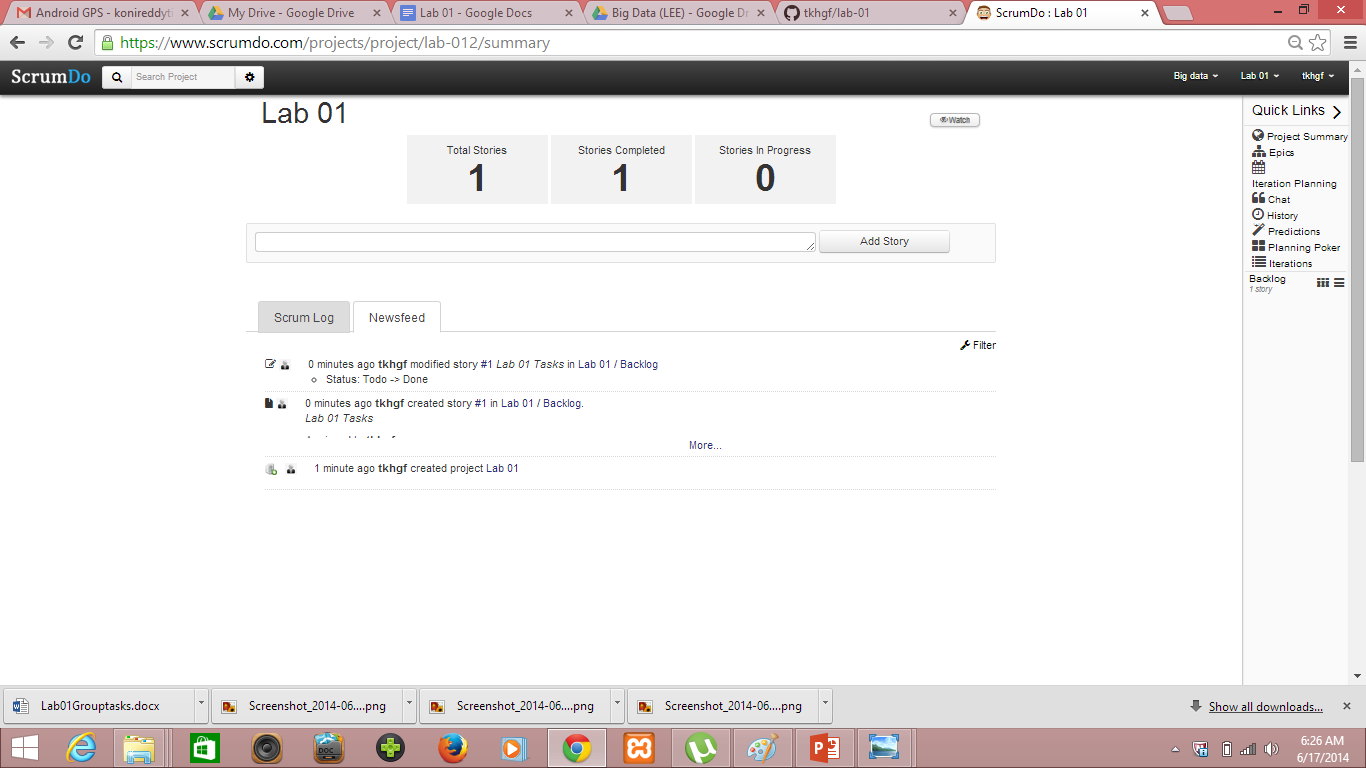


I uploaded my files and publish them to Github as below.



**Scrumdo Account Creation:**

I had already a Scrumdo account with username tkhgf. I had created this account as part of my KDM course. I am familiar with stories, iteration planning services provided by this Agile Tool Scrumdo as we used those services for our Recall Application Project as part of Knowledge Discovery Management Course. Here, I had defined my Lab 01 Assignment as one story and added Group tasks and Individual tasks under iteration1.



**Scrumdo project summary**

**The End**