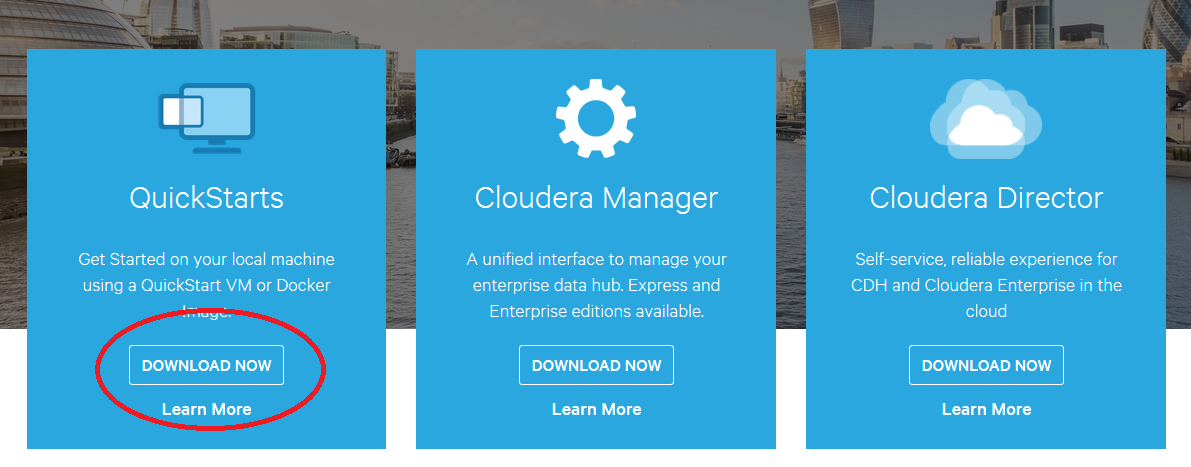
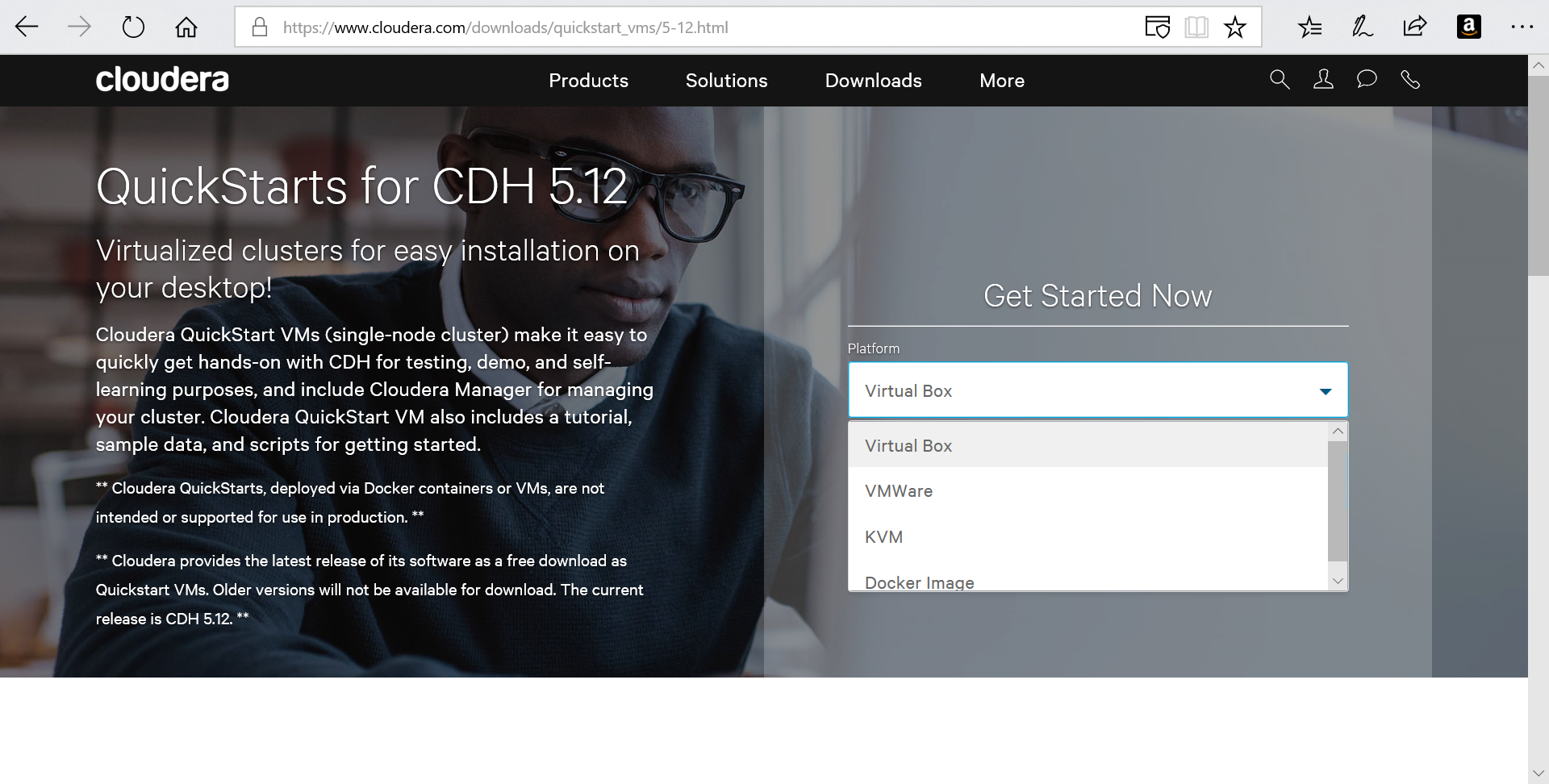
1. **Download Cloudera**

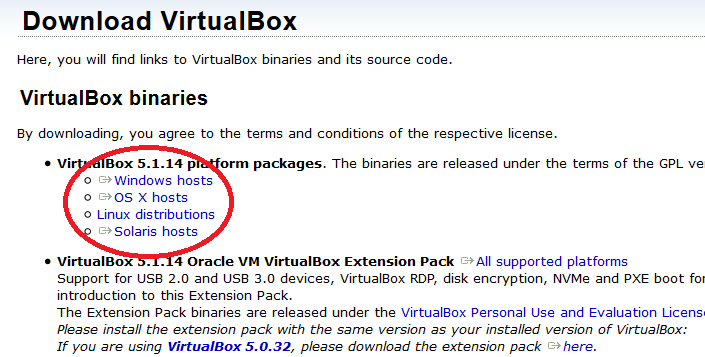
* Go to <http://www.cloudera.com/downloads.html>

****

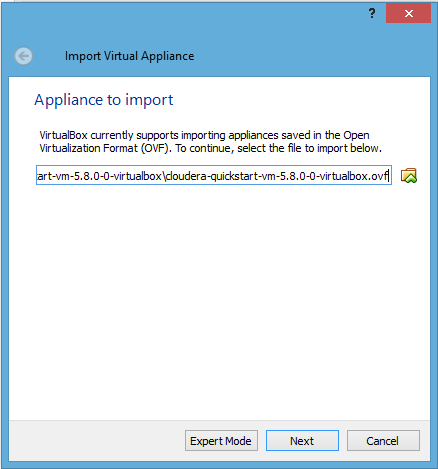
* Choose version and platform as image below

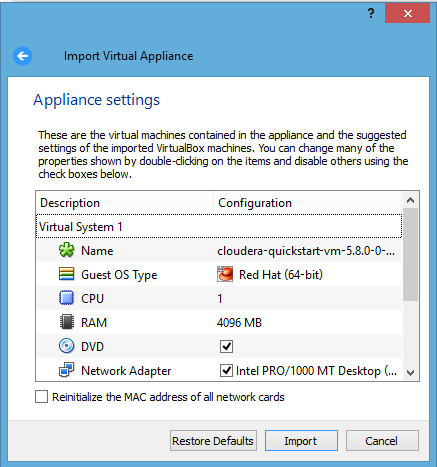


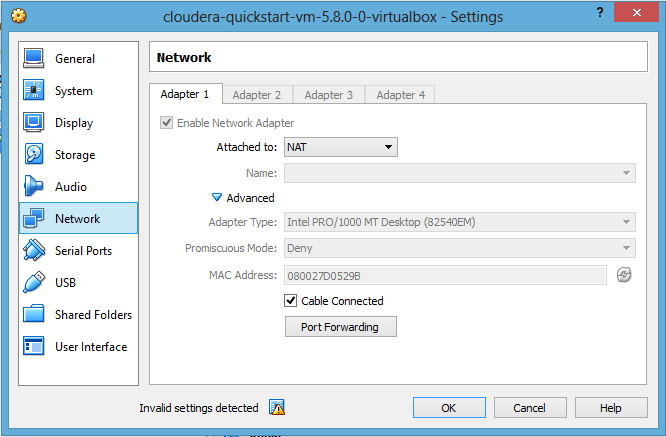
* **Download Virtual Box from** [**https://www.virtualbox.org/wiki/Downloads**](https://www.virtualbox.org/wiki/Downloads)

****

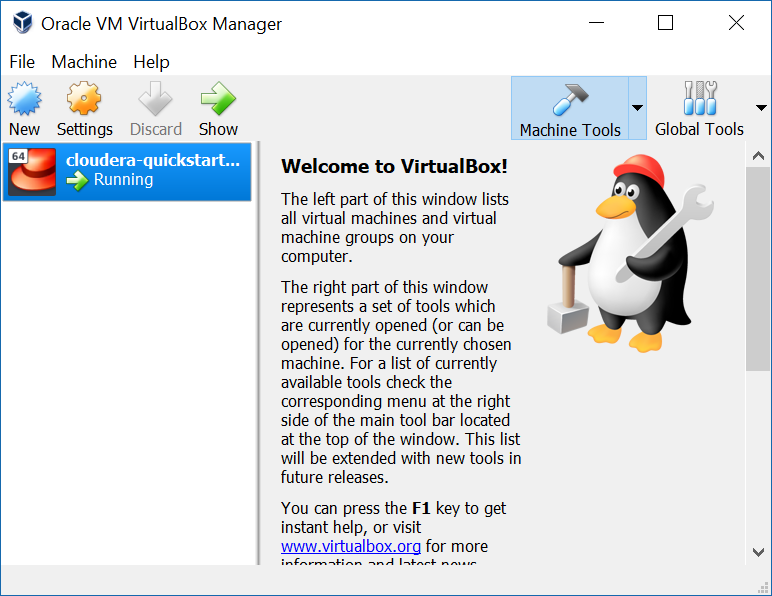
* **Click on File > Import Appliance**

****

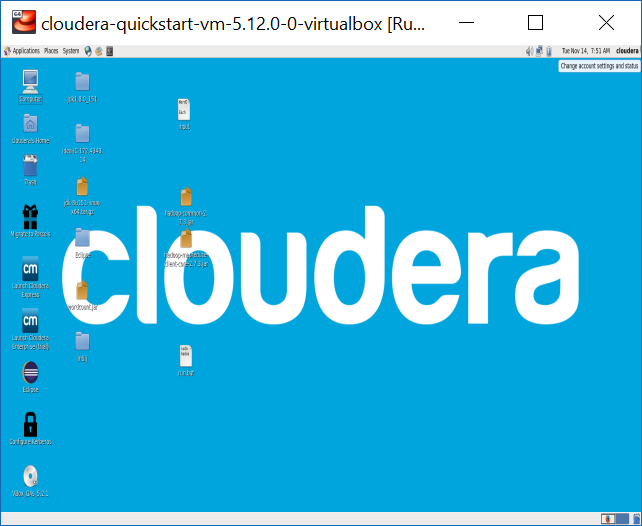
****

* **To enable network access go to Machine > Settings > Network and choose NAT in Attached to option as below**

**Now start cloudera by clicking start button:**



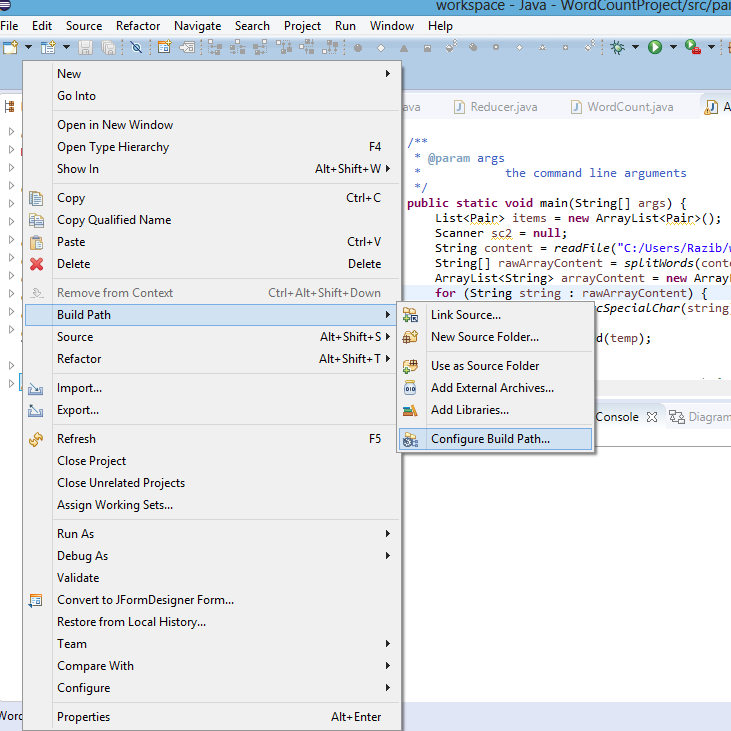
The desktop look like this:

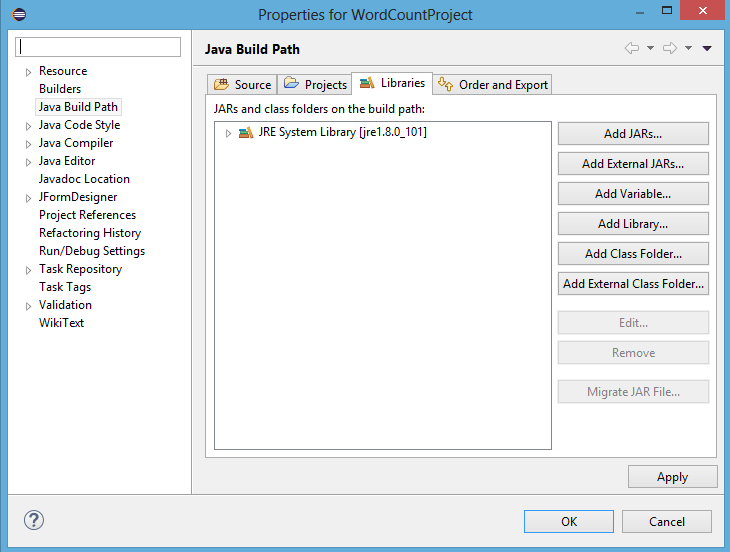


* **Open eclipse and start new project**
* **Download hadoop common and hadoop mapreduce client core from below link**

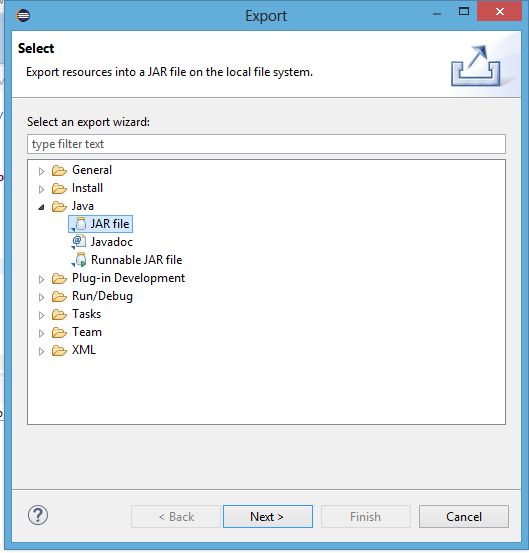
<http://mvnrepository.com/artifact/org.apache.hadoop/hadoop-common/2.7.3>  
<http://mvnrepository.com/artifact/org.apache.hadoop/hadoop-mapreduce-client-core/2.7.3>

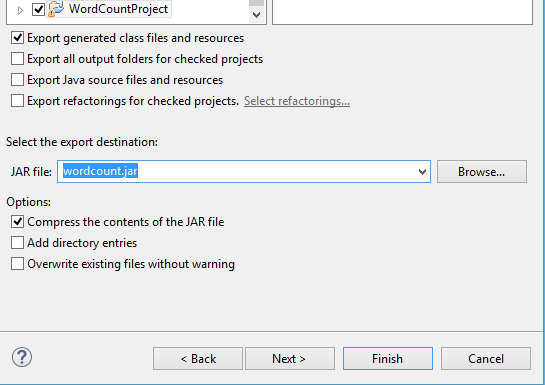
**Add jar file to project**

****

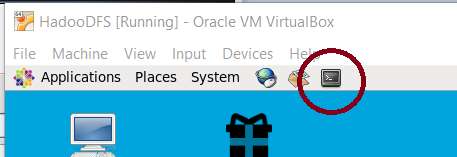
****

* **Get WordCount example from** [**https://wiki.apache.org/hadoop/WordCount**](https://wiki.apache.org/hadoop/WordCount)
* **Export to WordCount.jar to test the WordCount example. Right click on Project, choose “Export”. The Export window will appear, choose Java -> Runable JAR file. Click Next.**

****

****

* **Open terminal program**



* **Open Terminal and first create input and output locations in HDFS. Using the following commands:**
* hadoop fs -mkdir /user/cloudera

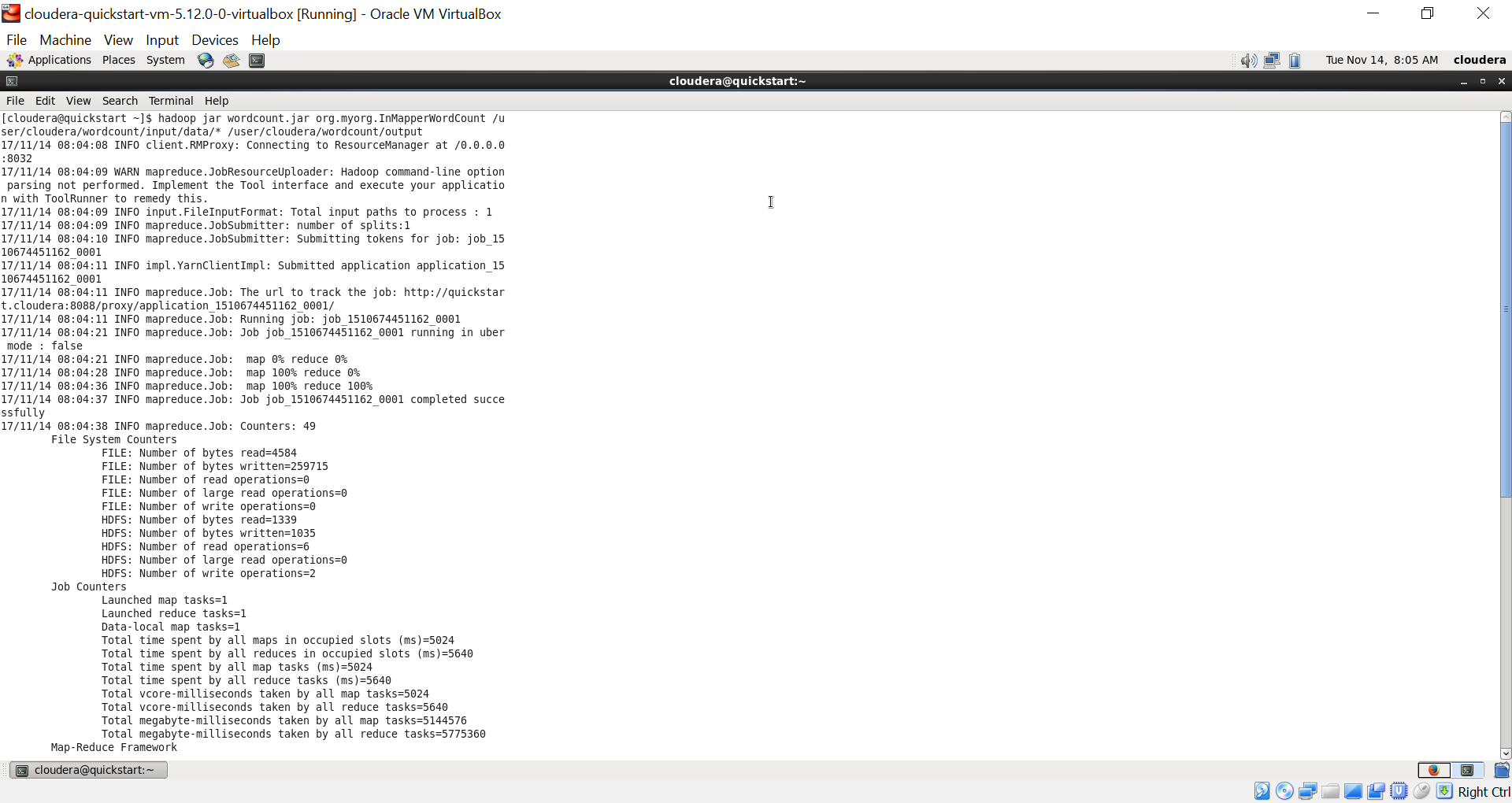
hadoop fs -mkdir /user/cloudera/input

Copy input files from local folder into HDFS input folder

hadoop fs -copyFromLocal /home/cloudera/data /user/cloudera/input

Run the WordCount application from the JAR file, given the paths to the input and output directories in HDFS.

hadoop jar wordcount.jar org.myorg.WordCount /user/cloudera/wordcount/input /user/cloudera/wordcount/output



**In order to see the output use the following command which will be as follows or you can see from the browser by downloading the output file.**

**hadoop fs -cat /user/cloudera/wordcount/output/\***

