**Hadoop Map Reduce**

To run the hadoop map reduce program initially we need to create a Cloudera setup in our machine. This is achieved through the usage of VMPlayer and linking the cloudera image through VM player.

The following steps are to be followed to install a VM image:

* Download the VMPlayer from the following link and install
  + <http://www.vmware.com/products/player/>
* Download the cloudera image and through using create a new VM machine option in VM player, link the cloudera image so as to load the cloudera onto VMPlayer.

Once the cloudera installation is completed now we need to run the hadoop programs i.e. word count example using hadoop.

Since the hadoop is installed on my local machine, there was no further re-installation done.

**Steps:**

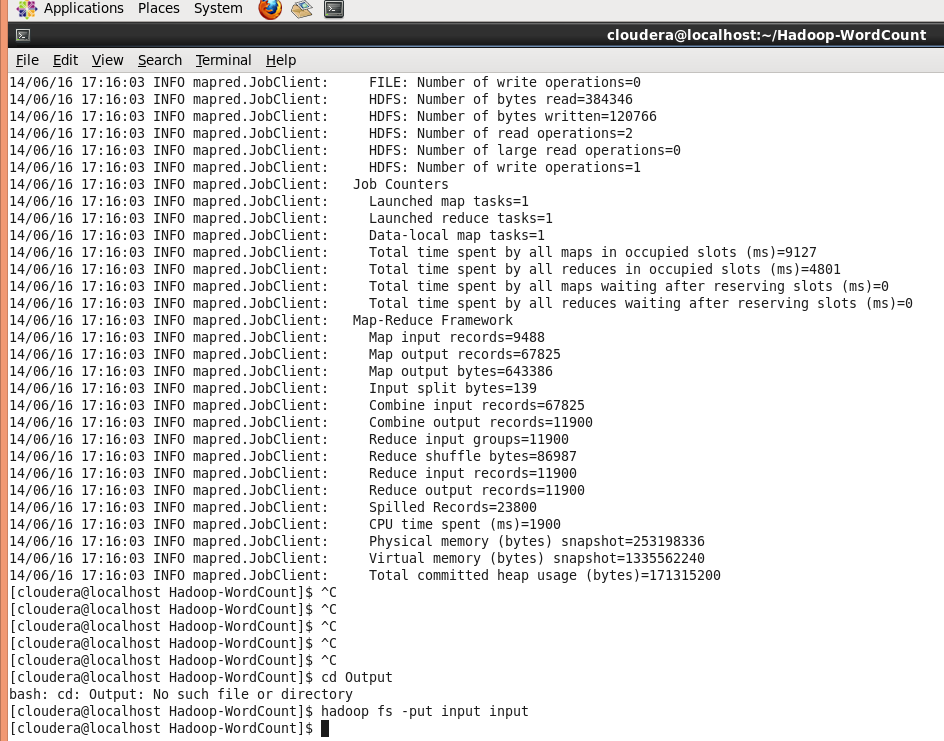
* Login to the cloudera image as cloudera user and go to the Word Count folder where the jars, input files etc. are located.
* Make sure you are in Word Count Directory, to validate it you can use the “pwd” command.
* Now run the following command in order to input the data onto the hadoop file system:
  + Hadoop fs –put input input
* The above command stores our input with the name input in our hadoop file system.
* 

Figure input for hdfs

* Once you have given the input run the following hadoop command for executing the word count example and retrieve the output for respective input files:
  + Hadoop jar wordcount.jar Wordcount input output
* The following figure gives the running of map reduce program for word count example:



Figure running of map reduce

* The output of the following map reduce program can be viewed using the following command:
  + Hadoop fs –cat output/\*
* The following figure shows the output of the word count example that is being run:

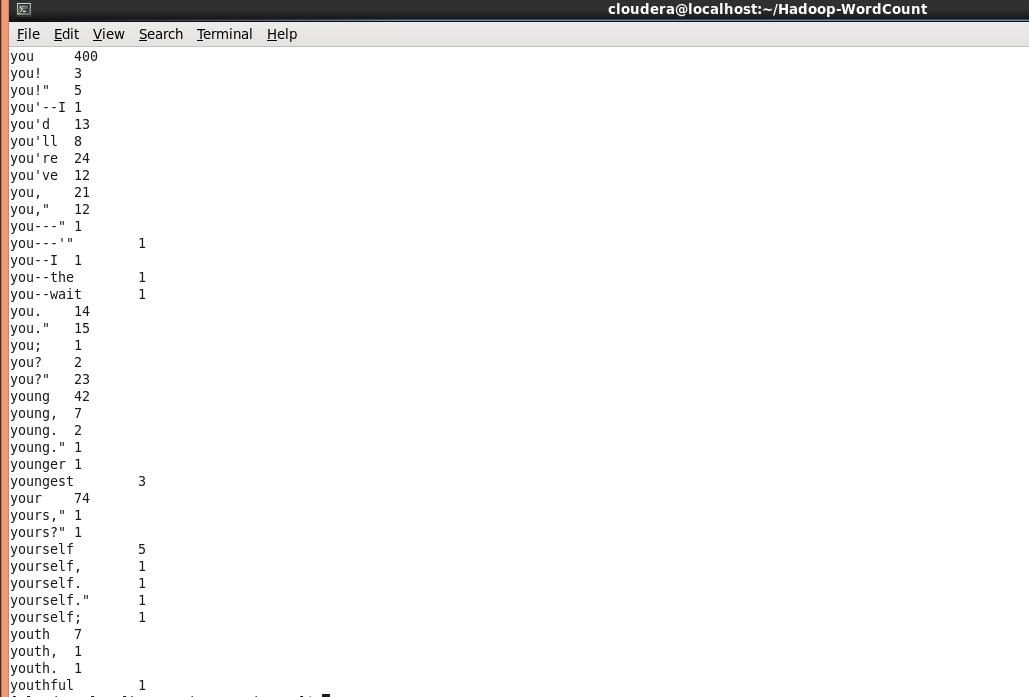


Figure output file