MAP-REDUCE ON SAMPLE SQL FILE GIVEN

1. Start the Cluster using hadoop_single_node_cluster.pdf
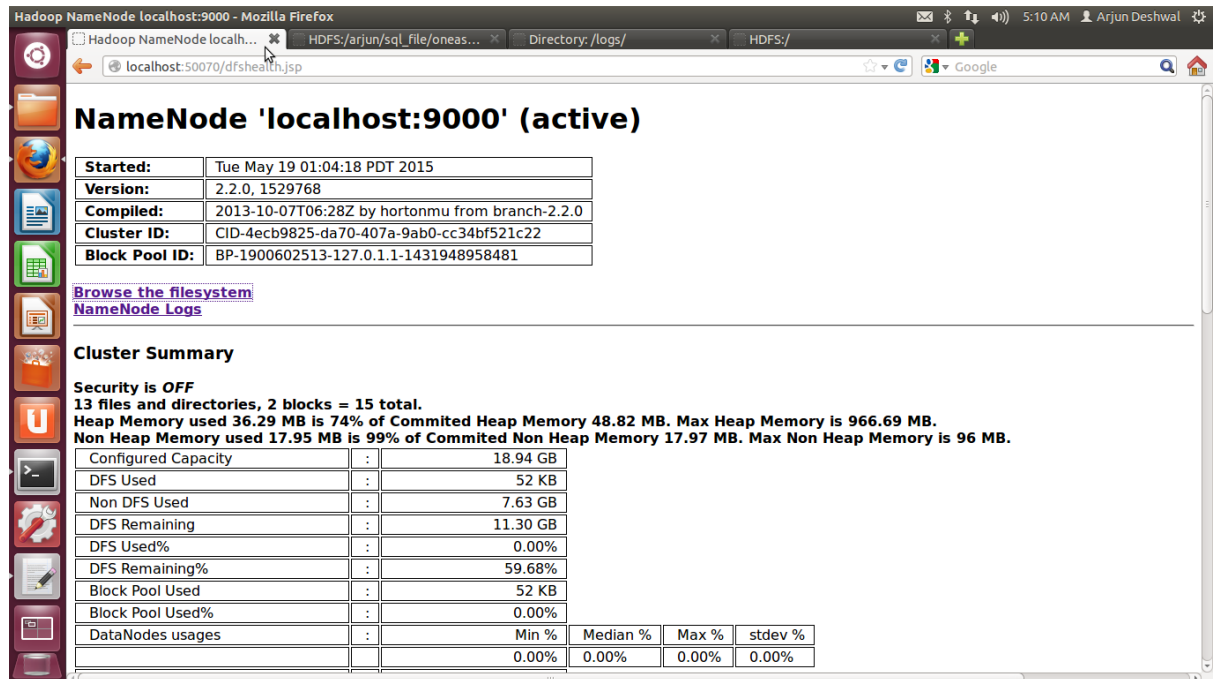
   link :- http://localhost:50070

Fig 1

2. Create a folder in hdfs where the input file(oneassist_060515.sql) can be placed

   2.1)Create directory data_any

   command :-bin/hdfs dfs -mkdir /arjun/data_any

   2.2)Create another directory input in data_any

   command :- bin/hdfs dfs -mkdir /arjun/data_any/input

   2.3)copy the given file into the input folder

   command:- bin/hdfs -put <local_file_place> <hdfs_folder>

   <local_file_place> :-/home/arjun/oneassist_060515.sql

   <hdfs_folder> :-/arjun/data_any/input

All the folder can be visible by clicking the Browse the filesystem on the link:-
http://localhost:50070

Hadoop NameNode localh... | HDFS:/arjun/sql_file/oneas... | Directory: /logs/ | HDFS:/arjun/data_any

localhost:50075/browseDirectory.jsp?dir=%2Farjun%2Fdata_any&namenodeInfoPort=50070&nnaddr=127.0.0.1:9000

Google

Contents of directory /arjun/data_any

Goto : /arjun/data_any [go]

Go to parent directory

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
input	dir				2015-05-19 04:18	rwxr-xr-x	arjun	supergroup

Go back to DFS home

Local logs

Log directory

Hadoop, 2015.

Hadoop NameNode localh... | HDFS:/arjun/sql_file/oneas... | Directory: /logs/ | HDFS:/arjun

localhost:50075/browseDirectory.jsp?dir=%2Farjun&namenodeInfoPort=50070&nnaddr=127.0.0.1:9000

Google

Contents of directory /arjun

Goto : /arjun [go]

Go to parent directory

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
data_any	dir				2015-05-19 05:13	rwxr-xr-x	arjun	supergroup
wordcount	dir				2015-05-18 04:50	rwxr-xr-x	arjun	supergroup

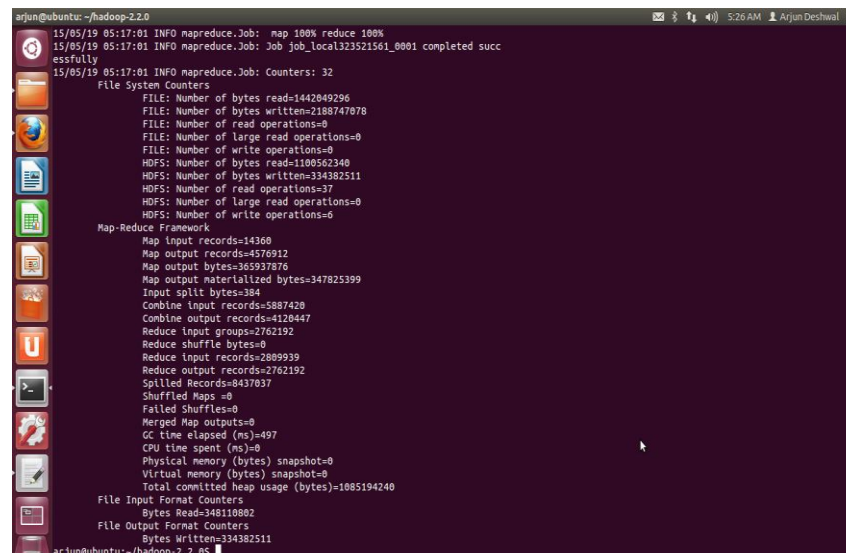Go back to DFS home

Local logs

Log directory

Hadoop, 2015.

Data can be visible here

3. Run the wc.jar file on the given file

        3.1) bin/hdfs jar wc.jar wordcount /arjun/data_any/input

        /arjun/data_any/output

4. Output can be visible on web_console