Ex.No.9

# DEMONSTRATE THE MAP REDUCE PROGRAMMING MODEL BY COUNTING THE NUMBER OF WORDS IN A FILE

**AIM:**

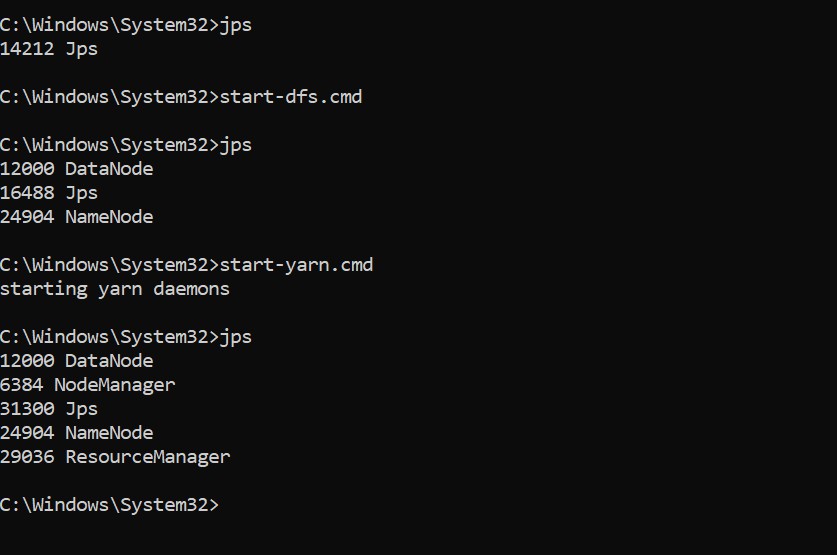
To demonstrate the MAP REDUCE programming model for counting the number of words in a file.

# PROCEDURE:

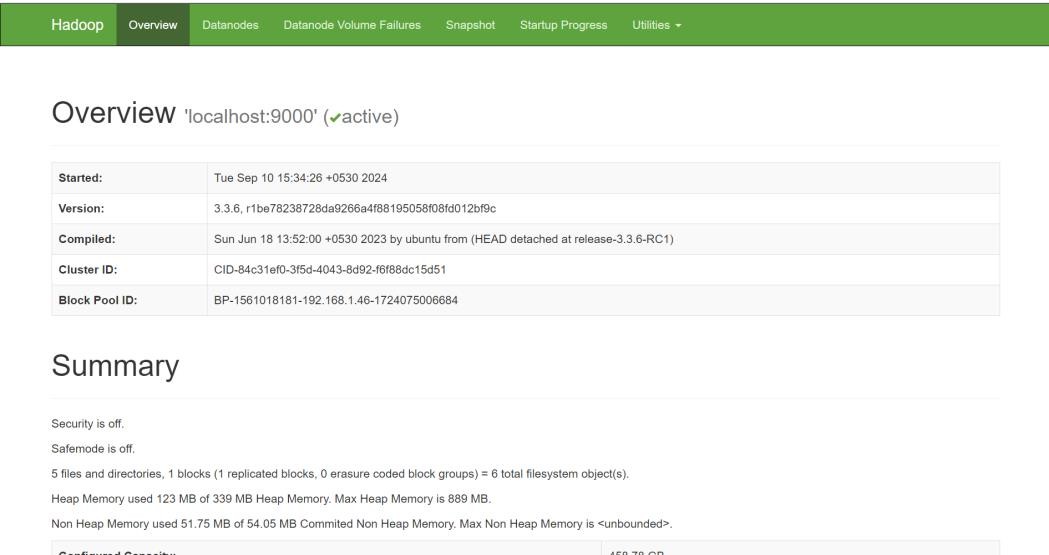
Open command prompt and run as administrator

Start Hadoop services by typing in the following commands:

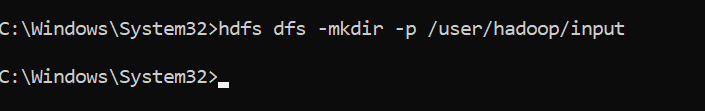
* start-dfs.cmd
* start-yarn.cmd

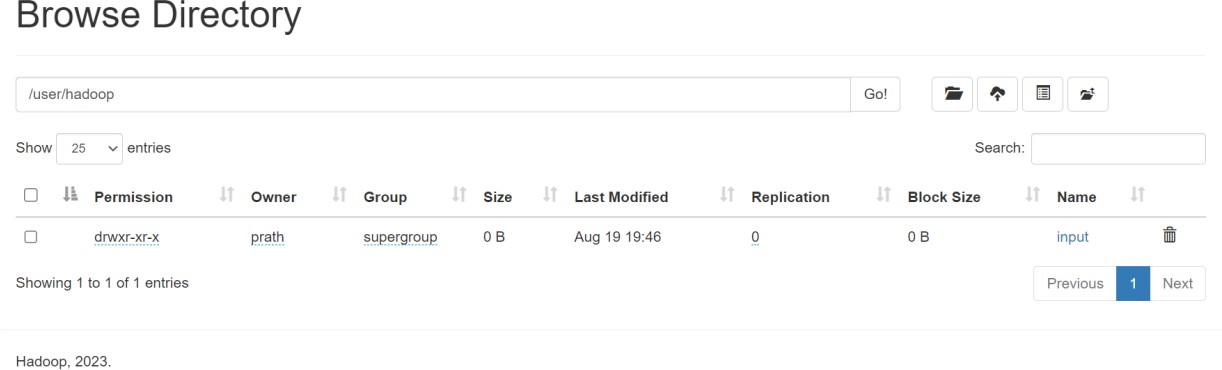


Open the browser and go to the URL localhost:9870



Create a directory in HDFS using the command: hdfs dfs -mkdir -p /user/hadoop/input



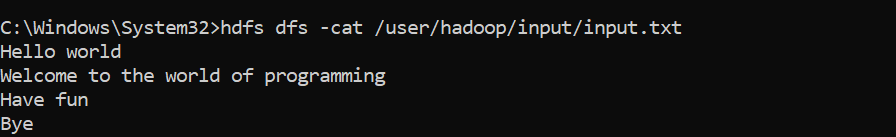


Copy the input file to HDFS using the command:

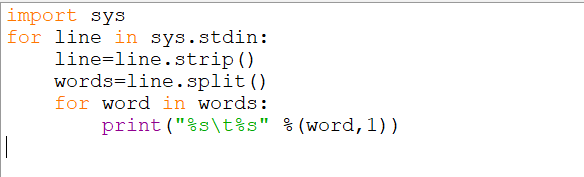
hdfs dfs -put C:/Semester7/DataAnalytics/Lab/input.txt /user/hadoop/input



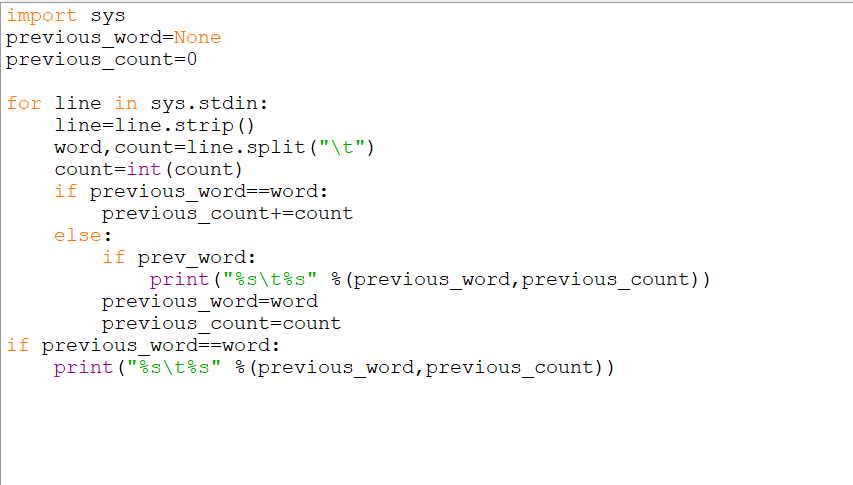
Display the contents of the file using this command: hdfs dfs -cat /user/hadoop/input/input.txt



Create mapper.py and reducer.py files mapper.py



reducer.py



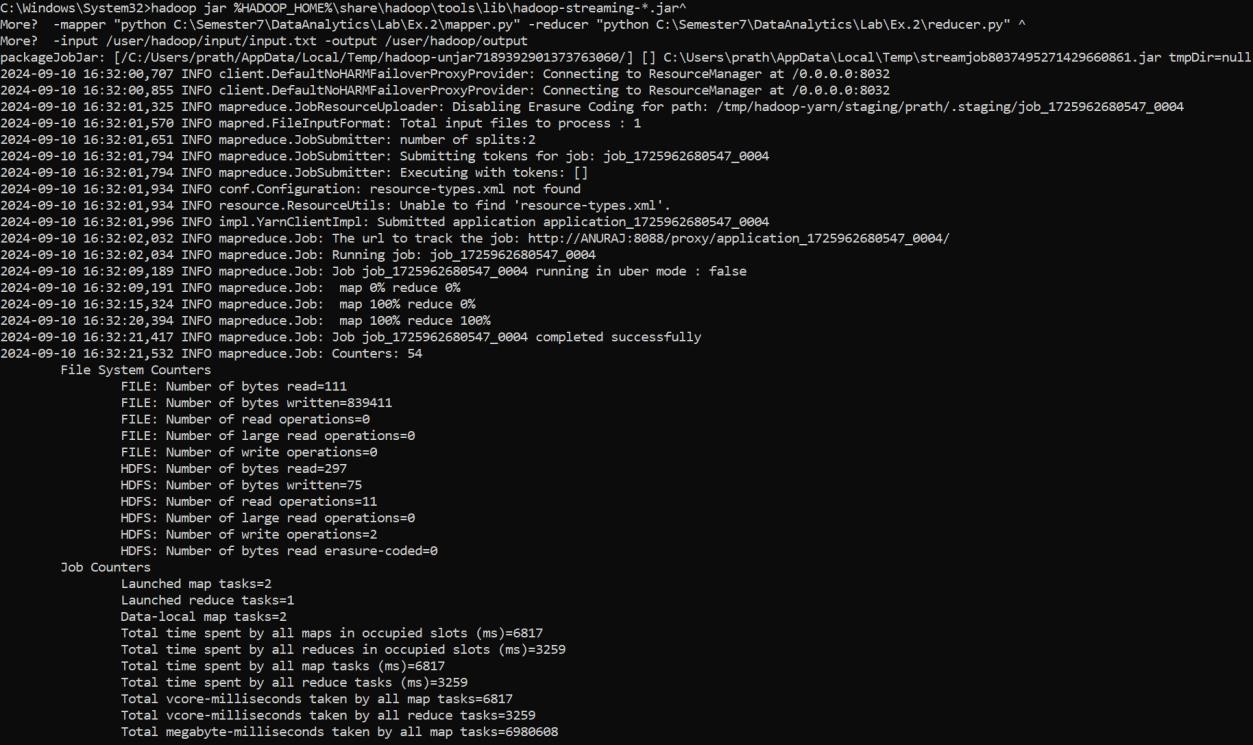
Run the Hadoop Streaming Job and give the file paths to the input, mapper and reducer using the following command:

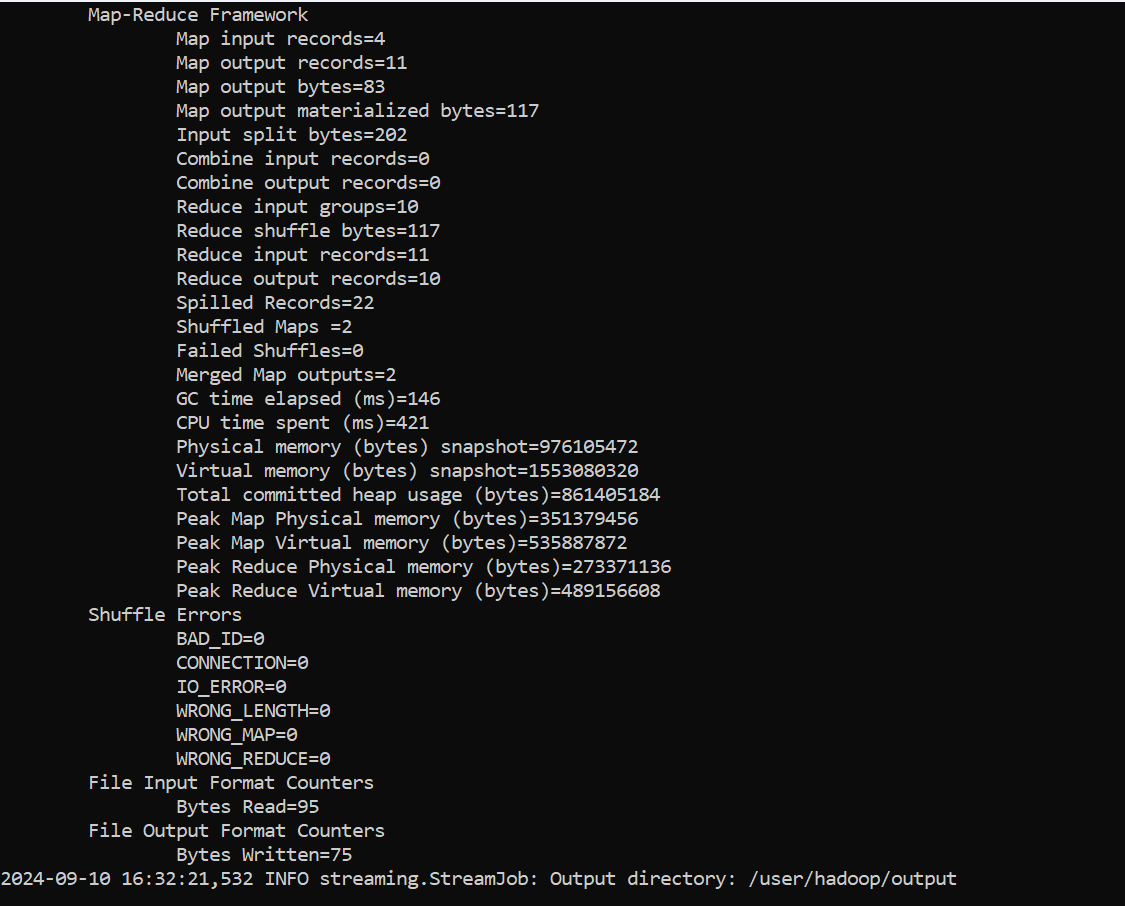
hadoop jar %HADOOP\_HOME%\share\hadoop\tools\lib\hadoop-streaming-

\*.jar^

-mapper "python C:\Semester7\DataAnalytics\Lab\Ex.2\mapper.py" -reducer "python C:\Semester7\DataAnalytics\Lab\Ex.2\reducer.py"^

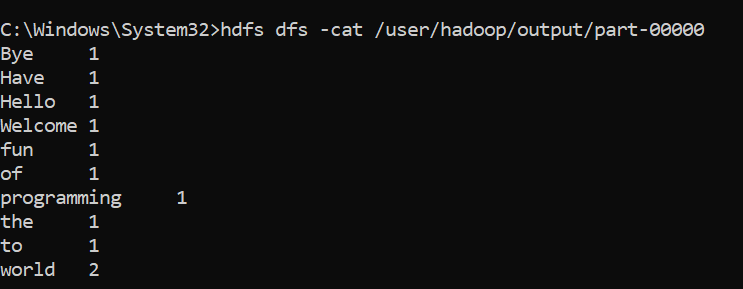
-input /user/hadoop/input/input.txt -output /user/hadoop/output



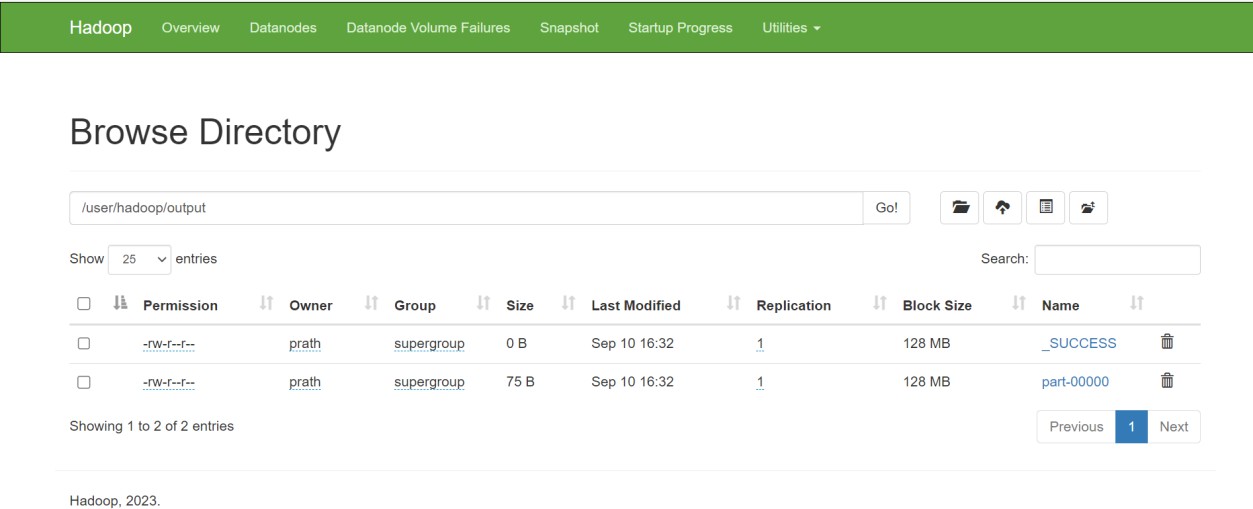


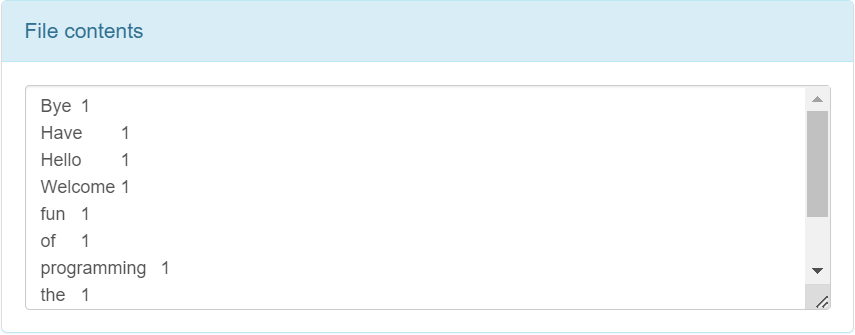
View the output using the command:

hdfs dfs -cat /user/hadoop/output/part-00000



Check the output on the file system in the browser





# RESULT:

Thus, to demonstrate the MAP REDUCE programming model for counting the number of words in a file was completed successfully.