EXP. 24: LAUNCH THE HADOOP 2.X AND PERFORM MAPREDUCE PROGRAMFOR A WORD

COUNT PROBLEM

AIM: LAUNCH THE HADOOP 2.X AND PERFORM MAPREDUCE PROGRAMFOR A WORD COUNT

PROBLEM

PROCEDURE:

Step 1 - Open Terminal

$ su

hduser

Password:

Step 2 - Start dfs and mapreduce services

$ cd /usr/local/hadoop/hadoop-2.7.2/sbin

$ start-dfs.sh

$ start-yarn.sh

$ jps

Step 3 - Check Hadoop through web UI

// Go to browser type http://localhost:8088 – All Applications Hadoop Cluster

// Go to browser type http://localhost:50070 – Hadoop Namenode

Step 4 – Open New Terminal

$ cd Desktop/

$ mkdir inputdata

$ cd inputdata/

$ echo “Hai, Hello, How are you? How is your health?” >> hello.txt

$ cat >> hello.txt

Step 5 – Go back to old Terminal

$ hadoop fs –copyFromLocal/home/hduser/Desktop/inputdata/hello.txt/folder/hduser

// Check in hello.txt in Namenode using Web UI

Step 6 – Download and open eclipse by creating workspace

Create a new java project.

Step 7 – Add jar to the project

You need to remove dependencies by adding jar files in the hadoop source folder. Now Clickon

Project tab and go to Properties.Under Libraries tab, click Add External JARs and select all the



