1NT19IS108
CHAITANYA P
C1 BATCH

Exercise-3: MAPREDUCE

Use the Hadoop framework to write a custom MapReduce program to perform word count operation on a custom data set.

first create a new project, package and class in eclipse to run a java code.

To install jar files:

Right click on project (Mapreduce)

Click on -> build a path -> add external archives -> Hadoop 3.2.1 -> share

In share 1. Click on common ->open hadoop-common-3.2.1.jar

2. Click on mapreduce -> open hadoop-mapreduce-client-core-3.2.1.jar

Right click on project -> export -> java -> jar file -> next

Browse the address of the java file and save it in desktop/document/downloads and name it

IN TERMINAL:

Run the commands:

```
cd $HADOOP_HOME

cd sbin

jps

start-all.sh
```

hdfs dfs -mkdir -p ~ /input

hdfs dfs -appendToFile - ~/input/text.txt

Create a file and add content to it. ->(ctrl D two times)

hdfs dfs -mkdir -p ~ /input

hdfs dfs -appendToFile - ~/input/text.txt

Create a file and add content to it. ->(ctrl D two times)

hdfs dfs -cat ~/out/part*

The word count of all the words in the file are 1.

Add more content to the file created.

Again run the above commands

hdfs dfs -appendToFile - ~/input/test.txt

hadoop jar /home/hdoop/Desktop/name.jar

hadoop jar /home/hdoop/Desktop/laisha.jar ~/input ~/output

// use a new output dir when u append content to existing file

// u can see map and reduce to be 100% in the picture

hdfs dfs -cat ~/output/part*

U can see the wordcount of the words in file