

EXP. 24: LAUNCH THE HADOOP 2.X AND PERFORM MAPREDUCE PROGRAM FOR A WORD COUNT PROBLEM

AIM: LAUNCH THE HADOOP 2.X AND PERFORM MAPREDUCE PROGRAM FOR 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 Click on Project tab and go to Properties.Under Libraries tab, click Add External JARs and select all the jars in the folder (click on 1st jar, and Press Shift and Click on last jar to select all jars in between and click ok)

```
/usr/local/hadoop/hadoop-2.7.2/share/hadoop/commonand  
/usr/local/hadoop/hadoop-2.7.2/share/hadoop/mapreduce folders.
```

Step -8 – WordCount Program Create 3 java files named

- WordCount.java
- WordCountMapper.java
- WordCountReducer.java

Step 9 - Creatr JAR file

Now Click on the Run tab and click Run-Configurations. Click on New Configuration button on the left-top side and Apply after filling the following properties.

Step 10 - Export JAR file

Now click on File tab and select Export. under Java, select Runnable Jar.

In Launch Config – select the config fie you created in **Step 9** (WordCountConfig).

Select an export destination (lets say desktop.)

Under Library handling, select Extract Required Libraries into generated JAR and click Finish.

Right-Click the jar file, go to Properties and under **Permissionstab**, Check Allow executing file as a program. and give Read and Write access to all the users

Step 11 – To view results in old Terminal

IMPLEMENTATION:

The screenshot shows a Mozilla Firefox window titled "Browsing HDFS - Mozilla Firefox". The address bar displays the URL "localhost:50070/explorer.html#/output". The main content area is titled "Browse Directory" and contains a search bar with the text "/output". Below the search bar is a table listing two files:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hduser	supergroup	0 B	8/11/2016, 9:54:38 PM	1	128 MB	_SUCCESS
-rw-r--r--	hduser	supergroup	44 B	8/11/2016, 9:54:38 PM	1	128 MB	part-00000

Browsing HDFS - Mozilla Firefox

Browsing HDFS

localhost:50070/explorer.html#/

Permission	Owner	Group	Size	Last Modified	Replication	Size	Name
drwxr-xr-x	hduser	supergroup	0 B	8/12/2016, 12:20:50 AM	0	0 B	cloud
drwxr-xr-x	hduser	supergroup	0 B	8/11/2016, 1:47:41 AM	0	0 B	cse
drwxr-xr-x	hduser	supergroup	0 B	8/4/2016, 11:37:37 PM	0	0 B	folder
drwxr-xr-x	hduser	supergroup	0 B	8/11/2016, 9:52:15 PM	0	0 B	grid
drwxr-xr-x	hduser	supergroup	0 B	8/11/2016, 9:54:38 PM	0	0 B	output
drwxr-xr-x	hduser	supergroup	0 B	8/11/2016, 11:54:23 PM	0	0 B	project
drwx-----	hduser	supergroup	0 B	8/4/2016, 11:40:37 PM	0	0 B	tmp

RESULT: To Launch The Hadoop 2.X And Perform Mapreduce Program For A Word Count Problem Have Successfully Implemented.