

LAB - 5

Name	Keval D Gandevia
Roll Number	CE046
ID	19CEUEG017
Subject	Big Data and Analytics

Aim: Finding mutual friends for every two friends using MapReduce framework.

❖ **Java code:**

```
import java.io.IOException;
import java.util.ArrayList;
import java.util.StringTokenizer;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.*;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class MutualFriends
{
    public static class MutualFriendMapper extends Mapper <LongWritable, Text,
    Text, Text>
```

```

{
    public void map(LongWritable key, Text value, Context context) throws
IOException, InterruptedException
    {
        String[] Line = value.toString().split("=>");
        String commonFriend = Line[0];
        Text c = new Text();
        c.set(commonFriend);
        String[] friends = Line[1].split(",");

        for(int i = 0; i < friends.length; i++)
        {
            for(int j = i + 1; j < friends.length; j++)
            {
                Text t1 = new Text();
                t1.set((friends[i] + "," + friends[j]).toString());
                context.write(t1, c);
            }
        }
    }

    public static class MutualFriendReducer extends Reducer <Text, Text, Text,
Text>
    {
        public void reduce(Text key, Iterable <Text> values, Context context)
throws IOException,InterruptedException
        {
            String resultVal = "";
            Text t1 = new Text();
            for(Text t:values)
            {
                String res = t.toString();
                resultVal += (res + ",");
            }
            t1.set(resultVal);
            context.write(key, t1);
        }
    }

}

public static void main(String[] args) throws Exception
{
    Configuration conf = new Configuration();
    Job job = Job.getInstance(conf, "Mutual Friends");
    job.setJarByClass(MutualFriends.class);
    job.setMapperClass(MutualFriendMapper.class);
    job.setCombinerClass(MutualFriendReducer.class);
    job.setReducerClass(MutualFriendReducer.class);
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(Text.class);
}

```

```

FileInputFormat.addInputPath(job, new Path(args[0]));
FileOutputFormat.setOutputPath(job, new Path(args[1]));
System.exit(job.waitForCompletion(true) ? 0 : 1);
}
}

```

❖ Steps to get the output:

(1) Starting the Hadoop server, yarn, and history server.

```

hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~$ start-dfs.sh
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting namenodes on [localhost]
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting datanodes
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting secondary namenodes [localhost]
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 09:59:50,761 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~$ start-yarn.sh
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting resourcemanager
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting nodemanagers
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~$ jps
16176 SecondaryNameNode
15954 DataNode
15763 NameNode
16422 ResourceManager
16970 Jps
16764 NodeManager
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~$ mapred --daemon start historyserver
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~$ jps
16176 SecondaryNameNode
17057 JobHistoryServer
15954 DataNode
15763 NameNode
16422 ResourceManager
16764 NodeManager
17134 Jps

```

(2) Compiling the code and creating a jar file.

```

hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop com.sun.tools.javac.Main MutualFriends.java
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ jar cf mf.jar MutualFriends*.class

```

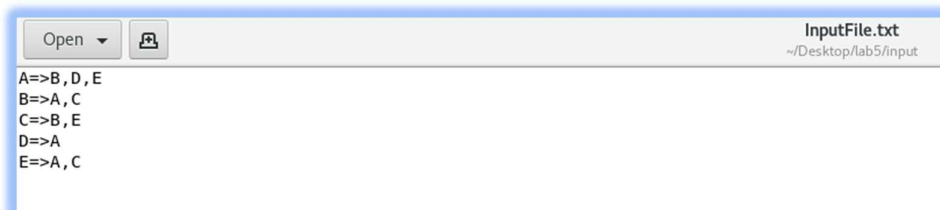
(3) Creating input directory.

```
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -ls
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:05:43,194 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 4 items
drwxr-xr-x - hadoop supergroup          0 2022-07-22 09:36 100
drwxr-xr-x - hadoop supergroup          0 2022-07-21 09:33 Test
drwxr-xr-x - hadoop supergroup          0 2022-07-28 09:11 input
drwxr-xr-x - hadoop supergroup          0 2022-07-28 09:16 wordcount
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -mkdir -p /user/hadoop/MutualFriends
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:06:09,621 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

(4) Creating a text file which contains the input.



(5) Uploading file to HDFS.

```
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -moveFromLocal ./sample.txt /user/hadoop/MutualFriends
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:09:00,280 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

(6) Removing the output directory.

```
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -ls /user/hadoop/MutualFriends
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:10:23,457 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxrwxr-x - hadoop supergroup          0 2022-08-24 10:07 /user/hadoop/MutualFriends/Output
-rwxrwxr-x 3 hadoop supergroup          61 2022-08-24 10:09 /user/hadoop/MutualFriends/sample.txt
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -rm -r /user/hadoop/MutualFriends/Output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:11:11,622 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Deleted /user/hadoop/MutualFriends/Output
```

(7) Executing the program.

```
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop jar mf.jar MutualFriends /user/hadoop/MutualFriends /user/hadoop/MutualFriends/output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:17:29,663 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2022-08-24 10:17:29,931 INFO client.DefaultNoHARMFalloverProxyProvider: Connecting to ResourceManager at localhost/127.0.0.1:8032
2022-08-24 10:17:30,065 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2022-08-24 10:17:30,075 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job_1661315399190_0002
2022-08-24 10:17:30,180 INFO input.FileInputFormat: Total input files to process : 1
2022-08-24 10:17:30,207 INFO mapreduce.JobSubmitter: number of splits:1
2022-08-24 10:17:30,311 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1661315399190_0002
2022-08-24 10:17:30,311 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-08-24 10:17:30,381 INFO conf.Configuration: resource-types.xml not found
2022-08-24 10:17:30,382 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-08-24 10:17:30,407 INFO impl.YarnClientImpl: Submitted application application_1661315399190_0002
2022-08-24 10:17:30,422 INFO mapreduce.Job: The url to track the job: http://celab2-ThinkCentre-neo-50s-Gen-3:8088/proxy/application_1661315399190_0002/
2022-08-24 10:17:30,423 INFO mapreduce.Job: Running job: job_1661315399190_0002
2022-08-24 10:17:34,470 INFO mapreduce.Job: Job job_1661315399190_0002 running in uber mode : false
2022-08-24 10:17:34,471 INFO mapreduce.Job: map 0% reduce 0%
2022-08-24 10:17:37,508 INFO mapreduce.Job: map 100% reduce 0%
2022-08-24 10:17:41,532 INFO mapreduce.Job: map 100% reduce 100%
2022-08-24 10:17:41,541 INFO mapreduce.Job: Job job_1661315399190_0002 completed successfully
2022-08-24 10:17:41,587 INFO mapreduce.Job: Counters: 54
File System Counters
FILE: Number of bytes read=177
```

(8) Final output.

```
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -ls /user/hadoop/MutualFriends/output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:18:15,708 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 3 hadoop supergroup 0 2022-08-24 10:17 /user/hadoop/MutualFriends/output/_SUCCESS
-rw-r--r-- 3 hadoop supergroup 156 2022-08-24 10:17 /user/hadoop/MutualFriends/output/part-r-000000
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$ hadoop dfs -cat /user/hadoop/MutualFriends/output/part-r-000000
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2022-08-24 10:18:34,486 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1,2 3,6,5,4,,
1,3 2,4,,
1,4 5,2,3,,
1,5 2,4,,
1,6 2,,
2,3 1,4,,
2,4 1,5,3,,
2,5 4,1,,
2,6 1,,
3,4 1,2,,
3,5 2,1,4,,
3,6 2,1,,
4,5 2,1,,
4,6 2,1,,
5,6 2,1,,
hadoop@celab2-ThinkCentre-neo-50s-Gen-3:~/Lab-5-BDA$
```