

Lab-07

20-11-21

To perform map Reduce program for word count using Eclipse by creating jar file

Open Eclipse IDE

> Open Eclipse and write programs for WCRReducer.java, WMapper.java & WCDriver.java

> Export the file to jar or jar file

> Open terminal

> su hcluser

> star-ll-sh

> jps

> hadoop fs -copyfromlocal /home/ll/Desktop/hadoop/Desktop/file.txt /tmp/text

output

are 1
brother 1
family 1
he 1
how 5
is 4
job 1
Sister 1
Gyan 1
Gyan 1

2.11.21

21-11-2021

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer x

WordCount

- JRE System Library [JavaSE-21]
- src
 - WordCount
 - WCDriver.java
 - WCDriver
 - WCMapper.java**
 - WCReducer.java
 - Referenced Libraries

WCDriver.java WCMapper.java x WCReducer.java

```
1 package WordCount;
2 import java.io.IOException;
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.LongWritable;
5 import org.apache.hadoop.io.Text;
6 import org.apache.hadoop.mapred.MapReduceBase;
7 import org.apache.hadoop.mapred.Mapper;
8 import org.apache.hadoop.mapred.OutputCollector;
9
10 import org.apache.hadoop.mapred.Reporter;
11 public class WCMapper extends MapReduceBase implements
12 Mapper<LongWritable,Text, Text, IntWritable> {
13 // Map function
14 public void map(LongWritable key, Text value, OutputCollector<Text,
15 IntWritable> output, Reporter rep) throws IOException
16 {
17 {
18 String line = value.toString();
19 // Splitting the line on spaces
20 for (String word : line.split(" "))
21 {
22 if (word.length() > 0)
23 {
24 output.collect(new Text(word), new IntWritable(1));
25 }
26 }
27 }
28 }
```

Task List x

Find

All Activate...

Connect Mylyn

Connect to your task and ALM tools or [create](#) a local task.

Outline x

WordCount

- WCMapper

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer x

WordCount

- JRE System Library [JavaSE-21]
- src
 - WordCount
 - WCDriver.java
 - WCMapper.java
 - WCReducer.java**
 - Referenced Libraries

WCDriver.java WCMapper.java WCReducer.java x

```
1 package WordCount;
2 import java.io.IOException;
3 import java.util.Iterator;
4 import org.apache.hadoop.io.IntWritable;
5 import org.apache.hadoop.io.Text;
6 import org.apache.hadoop.mapred.MapReduceBase;
7 import org.apache.hadoop.mapred.OutputCollector;
8 import org.apache.hadoop.mapred.Reducer;
9 import org.apache.hadoop.mapred.Reporter;
10
11 public class WCReducer extends MapReduceBase implements
12     Reducer<Text,IntWritable, Text, IntWritable> {
13     // Reduce function
14     public void reduce(Text key, Iterator<IntWritable> value,
15         OutputCollector<Text, IntWritable> output,
16         Reporter rep) throws IOException
17     {
18
19         int count = 0;
20         // Counting the frequency of each words
21         while (value.hasNext())
22         {
23             IntWritable i = value.next();
24             count += i.get();
25         }
26         output.collect(key, new IntWritable(count));
27     }
28 }
29 }
```

Task List x

Find

All Activate...

Connect Mylyn

[Connect](#) to your task and ALM tools or [create](#) a local task.

Outline x

WordCount

- WCReducer

File Edit Source Refactor Navigate Search Project Run Window Help

Package Explorer x

WordCount

- JRE System Library [JavaSE-21]
- src
 - WordCount
 - WCDriver.java
 - WCMapper.java
 - WCReducer.java
 - Referenced Libraries

WCDriver.java x WCMapper.java WCReducer.java

```
1 package WordCount;
2 import java.io.IOException;
3 import org.apache.hadoop.conf.Configured;
4 import org.apache.hadoop.fs.Path;
5 import org.apache.hadoop.io.IntWritable;
6 import org.apache.hadoop.io.Text;
7 import org.apache.hadoop.mapred.FileInputFormat;
8 import org.apache.hadoop.mapred.FileOutputFormat;
9 import org.apache.hadoop.mapred.JobClient;
10 import org.apache.hadoop.mapred.JobConf;
11 import org.apache.hadoop.util.Tool;
12 import org.apache.hadoop.util.ToolRunner;
13 public class WCDriver extends Configured implements Tool {
14     public int run(String args[]) throws IOException
15     {
16         if (args.length < 2)
17         {
18             System.out.println("Please give valid inputs");
19             return -1;
20         }
21         JobConf conf = new JobConf(WCDriver.class);
22         FileInputFormat.setInputPaths(conf, new Path(args[0]));
23         FileOutputFormat.setOutputPath(conf, new Path(args[1]));
24         conf.setMapperClass(WCMapper.class);
25         conf.setReducerClass(WCReducer.class);
26         conf.setMapOutputKeyClass(Text.class);
27         conf.setMapOutputValueClass(IntWritable.class);
28         conf.setOutputKeyClass(Text.class);
29         conf.setOutputValueClass(IntWritable.class);
30         JobClient.runJob(conf);
31         return 0;
32     }
33     // Main Method
34     public static void main(String args[]) throws Exception
35     {
36         int exitCode = ToolRunner.run(new WCDriver(), args);
37         System.out.println(exitCode);
38     }
39 }
```

Task List x

Find

All Activate...

Connect Mylyn

Connect to your task and ALM tools or [create](#) a local task.

Outline x

WordCount

- WCDriver
 - run(String[]): int
 - main(String[]): void

Open ▾



***sample.txt**
~/Desktop

mapred-site.xml



```
1 hi how are you  
2 how is your job  
3 how is your family  
4 how is your brother  
5 how is your sister  
6
```



```
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -copyFromLocal -f /home/hadoop/Desktop/file1.txt /rgs/test.txt
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop jar /home/hadoop/Desktop/WordCount.jar wordcount.WordCount /rgs/test.txt /output
JAR does not exist or is not a normal file: /home/hadoop/Desktop/WordCount.jar
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop jar /home/hadoop/Desktop/Word_Count.jar wordcount.WordCount /rgs/test.txt /output
Exception in thread "main" java.lang.ClassNotFoundException: wordcount.WordCount
    at java.base/java.net.URLClassLoader.findClass(URLClassLoader.java:476)
    at java.base/java.lang.ClassLoader.loadClass(ClassLoader.java:594)
    at java.base/java.lang.ClassLoader.loadClass(ClassLoader.java:527)
    at java.base/java.lang.Class.forName0(Native Method)
    at java.base/java.lang.Class.forName(Class.java:398)
    at org.apache.hadoop.util.RunJar.run(RunJar.java:321)
    at org.apache.hadoop.util.RunJar.main(RunJar.java:241)
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -cat /output/part-00000
are      1
brother  1
family   1
hi       1
how      5
is       4
job      1
sister   1
you      1
your     4
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -ls /output
Found 2 items
-rw-r--r--  1 hadoop supergroup      0 2024-05-21 15:21 /output/_SUCCESS
-rw-r--r--  1 hadoop supergroup    69 2024-05-21 15:21 /output/part-00000
```