```java
import java.io.IOException;
import java.net.URI;
import java.net.URISyntaxException;

import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
//import org.apache.hadoop.mapred.FileInputFormat;
//import org.apache.hadoop.mapred.FileInputFormat;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class WordCountDriver {
	public static void main(String[] args) throws IOException,
ClassNotFoundException, InterruptedException, URISyntaxException
	{
		Configuration conf = new Configuration();
		//FileSystem fs = FileSystem.get(conf);
		//System.out.println(fs.getUri());
		Job j = new Job();// getConf(), "Max temperature");
		j.setJobName("My First Job");
		j.setJarByClass(WordCountDriver.class );
		j.setMapperClass(WordCountMapper.class );
		//j.setCombinerClass(WordCountReducer.class);
		//j.setPartitionerClass(WordCountPartitioner.class);
		j.setReducerClass(WordCountReducer.class);
		//j.setNumReduceTasks(0);
		//j.setCombinerClass(WordCountReducer.class);
		//j.setCombinerClass(WordCountReducer.class);
		//
		//j.setNumReduceTasks(0);

		j.setOutputKeyClass(Text.class);
		j.setOutputValueClass(IntWritable.class);

		FileInputFormat.addInputPath(j, new Path(args[0]));
		FileOutputFormat.setOutputPath(j, new Path(args[1]));

		URI uri = new URI(args[1].toString());

		FileSystem fs =  FileSystem.get(uri, conf);

		boolean x = fs.delete(new Path(uri),true);


		int xxx =  j.waitForCompletion(true) ? 0 : 1;

	}

}
```

```java
import java.io.IOException;
import java.util.StringTokenizer;

import org.apache.hadoop.io.DoubleWritable;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Mapper;

public class WordCountMapper extends Mapper<LongWritable,  Text, Text,
IntWritable>

{

      @Override
      protected void map(LongWritable key, Text value,
                  org.apache.hadoop.mapreduce.Mapper.Context context)
                  throws IOException, InterruptedException {
String inputstring = value.toString();
//StringTokenizer stk = new StringTokenizer(inputstring);
//while(stk.hasMoreTokens())
//{
//    context.write(new Text(stk.nextToken()),new IntWritable(1));
//}

//for(int i =0; i<inputstring.split(" ").length;i++)
//{
//    context.write(new Text(inputstring.split(" ")[i]),new
IntWritable(1));

for(String x : inputstring.split(" "))
      //how are you
      //[0] how
      //[1] are
      //[2] you
{
      context.write(new Text(x),new IntWritable(1));
      // how 1
      //are 1
      //you 1
}
      }

}
```

```
import java.io.IOException;

public class WordCountReducer extends Reducer<Text, IntWritable, Text, IntWritable>{

    @Override
    public void reduce(Text key, Iterable<IntWritable> values, Context context)
    throws IOException, InterruptedException {
        int y = 0;
        for(IntWritable x : values)
        {
            y ++;

        }
        context.write(key, new IntWritable(y) );

    }
}
```

HUE   Query   Search data and saved documents...   Jobs   cloudera

File Browser

cloudera

Even.txt
EvenOdd.txt
Extern_table
OPEven
OPEvenOP
OPEvenS
OPofEven
OPword
StudentData.txt
WordCountH.txt
evenodd.txt
intern_table
oOPEven
poem.txt

View as binary   Home / user / cloudera / OPHadoop / part-r-00000   Page 1 to 1 of 1

Edit file

Download

View file location

Refresh

Last modified
06/19/2021 4:48 AM

User
cloudera

Group
cloudera

Size
1.25 KB

Mode
100644

```
(            1
(HDFS),  1
/həˈduːp/)      1
All      1
Apache  2
Distributed      1
File     1
Hadoop  6
It       3
MapReduce  2
System  1
The      1
This     2
a        10
access  1
across  1
advantage        1
allows  1
also     1
amounts 1
and      7
approach         1
architecture     1
are      3
as       1
assumption       1
```

Hue - File Browser - Mozilla Firefox

[Java - WordCount/src/... cloudera@quickstart:... [Cloudera Live : Welco... Hue - File Browser - M...

he mouse pointer inside or press Ctrl+G.

30°C Light rain   10:19

Applications Places System 🌐 🖨 🖵   Fri Jun 18, 9:50 PM 🔊 🖥 cloudera

cloudera@quickstart:~/workspace

File Edit View Search Terminal Help

```
[cloudera@quickstart Desktop]$ ls
Eclipse.desktop       Kerberos.desktop    StudentData.txt~   WordCountH.txt
Enterprise.desktop  Parcels.desktop     untitled folder
Express.desktop       StudentData.txt     Word Count
[cloudera@quickstart Desktop]$ cd ..
[cloudera@quickstart ~]$ ls
cloudera-manager  Downloads                 kerberos  Pictures    workspace
cm_api.py          eclipse                   lib       Public
Desktop            enterprise-deployment.json  Music     Templates
Documents          express-deployment.json   parcels   Videos
[cloudera@quickstart ~]$ cd workspace/
[cloudera@quickstart workspace]$ ls
EvenOdd         OddEven.jar     SumOfEven     word
evenodd.txt  StudentData.txt  SumOfEven.jar  WordCount
Hadoop.jar  SumEven            Sumprime.jar  WordCount.jar
OddEven         SumEven.jar     training      word.jar
[cloudera@quickstart workspace]$ hadoop jar Hadoop.jar WordCountDriver WordCount
H.txt OPHadoop
21/06/18 21:48:31 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
21/06/18 21:48:31 WARN mapreduce.JobResourceUploader: Hadoop command-line option
 parsing not performed. Implement the Tool interface and execute your applicatio
n with ToolRunner to remedy this.
21/06/18 21:48:31 INFO input.FileInputFormat: Total input paths to process : 1
21/06/18 21:48:31 WARN hdfs.DFSClient: Caught exception
java.lang.InterruptedException
        at java.lang.Object.wait(Native Method)
        at java.lang.Thread.join(Thread.java:1281)
        at java.lang.Thread.join(Thread.java:1355)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.closeResponder(DF
SOutputStream.java:967)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.endBlock(DFSOutpu
tStream.java:705)
        at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.run(DFSOutputStre
am.java:894)
21/06/18 21:48:31 INFO mapreduce.JobSubmitter: number of splits:1
21/06/18 21:48:32 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_16
24071776421_0006
21/06/18 21:48:32 INFO impl.YarnClientImpl: Submitted application application_16
24071776421_0006
21/06/18 21:48:32 INFO mapreduce.Job: The url to track the job: http://quickstar
t.cloudera:8088/proxy/application_1624071776421_0006/
21/06/18 21:48:32 INFO mapreduce.Job: Running job: job_1624071776421_0006
21/06/18 21:48:39 INFO mapreduce.Job: Job job_1624071776421_0006 running in uber mode : false
21/06/18 21:48:39 INFO mapreduce.Job:  map 0% reduce 0%
21/06/18 21:48:46 INFO mapreduce.Job:  map 100% reduce 0%
```