

Activities

IntelliJ IDEA Community Edition

Apr 4 16:42

WordCountExample2

Version control

Project

WordCountExample2

src

main

java

org.example

WC_Mapper

WC_Reducer

WC_Runner

resources

test

java

target

classes

generated-sources

generated-test-sources

maven-archiver

maven-status

test-classes

WordCountExample2-1.0-SNAPSHOT.jar

gitignore

pom.xml

External Libraries

Scratches and Consoles

Terminal

Local

WordCountExample2 > src > main > java > org > example > WC_Reducer

WC_Reducer.java

WC_Mapper.java

WC_Runner.java

Current File

File

Edit

Find

Run

Settings

Tools

Help

20:2

LF

UTF-8

4 spaces

```
package org.example;

import java.io.IOException;
import java.util.Iterator;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.MapperContext;
import org.apache.hadoop.mapred.MapperRunner;
import org.apache.hadoop.mapred.Reducer;
import org.apache.hadoop.mapred.ReducerContext;
import org.apache.hadoop.mapred.Reporter;

2 usages

public class WC_Reducer extends MapReduceBase implements
    Reducer<Text, IntWritable, Text, IntWritable> {
    public void reduce(Text key, Iterator<IntWritable> values, OutputCollector<Text, IntWritable>
        reporter, Reporter reporter) throws IOException {

        int sum=0;
        while (values.hasNext()) {
            sum+=values.next().get();
        }
        output.collect(key, new IntWritable(sum));
    }
}
```

Activities

IntelliJ IDEA Community Edition

Apr 4 16:42

Current File

WC_Runner.java

WC_Reducer.java

WC.Mapper.java

WC_Runner.java

Project

WordCountExample2

src

main

java

org.example

WC.Mapper

WC.Reducer

WC.Runner

resources

test

java

target

classes

generated-sources

generated-test-sources

maven-archiver

maven-status

test-classes

WordCountExample2-1.0-SNAPSHOT.jar

.gitignore

pom.xml

External Libraries

Scratches and Consoles

Terminal

Local

WordCountExample2 > src > main > java > org > example > WC.Mapper

```
package org.example;

import java.io.IOException;
import java.util.StringTokenizer;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.OutputCollector;
import org.apache.hadoop.mapred.Reporter;

1 usage
public class WC_Mapper extends MapReduceBase implements
    Mapper<LongWritable,Text,Text,IntWritable>{
    1 usage
    private final static IntWritable one = new IntWritable( value: 1 );
    2 usages
    private Text word = new Text();public void map(LongWritable key, Text value,OutputCollector
        Reporter reporter) throws IOException{
        String line = value.toString();
        StringTokenizer tokenizer = new StringTokenizer(line);
        while (tokenizer.hasMoreTokens()){
            word.set(tokenizer.nextToken());
            output.collect(word, one);
        }
    }
}
```

24:2 LF UTF-8 4 spaces

IntelliJ IDEA Community Edition Apr 4 16:42

Project > WordCountExample2 ~\IdeaProjects\WordCountExample2

- > .idea
- > src
 - > main
 - > java
 - org.example
 - WC.Mapper
 - WC.Reducer
 - WC.Runner
 - resources
 - test
 - java
 - target
 - classes
 - generated-sources
 - generated-test-sources
 - maven-archiver
 - maven-status
 - test-classes
 - WordCountExample2-1.0-SNAPSHOT.jar
 - .gitignore
 - pom.xml
 - External Libraries
 - Scratches and Consoles

Terminal Local x

WordCountExample2 > src > main > java > org > example > WC.Mapper

WC.Reducer.java WC.Mapper.java WC.Runner.java

```
package org.example;

import java.io.IOException;
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.FileOutputFormat;
import org.apache.hadoop.mapred.JobClient;
import org.apache.hadoop.mapred.JobConf;
import org.apache.hadoop.mapred.TextInputFormat;
import org.apache.hadoop.mapred.TextOutputFormat;

public class WC_Runner {
    public static void main(String[] args) throws IOException {
        JobConf conf = new JobConf(WC_Runner.class);
        conf.setJobName("WordCount");
        conf.setOutputKeyClass(Text.class);
        conf.setOutputValueClass(IntWritable.class);
        conf.setMapperClass(WC_Mapper.class);
        conf.setCombinerClass(WC_Reducer.class);
        conf.setReducerClass(WC_Reducer.class);
        conf.setInputFormat(TextInputFormat.class);
        conf.setOutputFormat(TextOutputFormat.class);
        FileInputFormat.setInputPaths(conf, new Path(args[0]));
        FileOutputFormat.setOutputPath(conf, new Path(args[1]));
        JobClient.runJob(conf);
    }
}
```


IntelliJ IDEA Community Edition Apr 4 16:41

Project: WordCountExample2 Version control

Terminal: Local

```
at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
hduser@student-OptiPlex-3010: /home/student/IdeaProjects/WordCountExample2$ hadoop jar target/WordCountExample2-1.0-SNAPSHOT.jar org.example.WC_Runner
/INPUT/INPUT.txt /OUTPUT

WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
24/04/04 16:35:29 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
24/04/04 16:35:30 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session-id
24/04/04 16:35:30 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
24/04/04 16:35:30 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
24/04/04 16:35:30 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
24/04/04 16:35:30 INFO mapred.FileInputFormat: Total input paths to process : 1
24/04/04 16:35:30 INFO mapreduce.JobSubmitter: number of splits:1
24/04/04 16:35:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local36003458_0001
24/04/04 16:35:30 INFO mapreduce.Job: The url to track the job: http://localhost:8090/
24/04/04 16:35:30 INFO mapred.LocalJobRunner: OutputCommitter set in config null
24/04/04 16:35:30 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
24/04/04 16:35:30 INFO mapreduce.Job: Running job: job_local36003458_0001
24/04/04 16:35:31 INFO mapred.LocalJobRunner: Waiting for map tasks
24/04/04 16:35:31 INFO mapred.LocalJobRunner: Starting task: attempt_local36003458_0001_m_000000_0
24/04/04 16:35:31 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
24/04/04 16:35:31 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/INPUT/INPUT.txt:0+78
24/04/04 16:35:31 INFO mapred.MapTask: numReduceTasks: 1
```

WordCountExample2 > src > main > java > org > example > WC.Mapper

27:2 LF UTF-8 4 spaces

