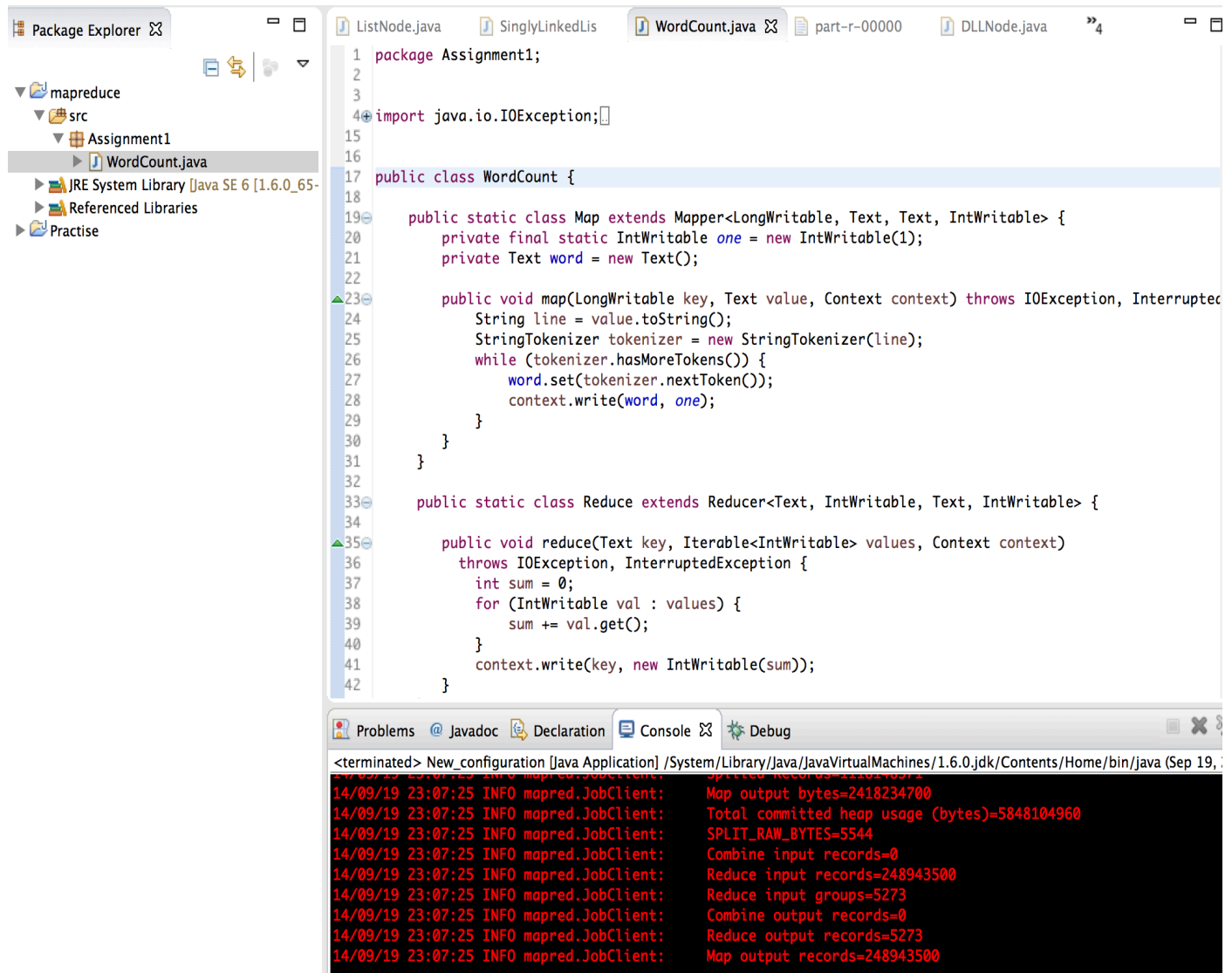


Class : CS 6240 | CRN 16804 | Tuesday 1:35 PM
Homework : HW 1
Name : Mahesh S.Bhandarkar

Local Execution



The screenshot displays an IDE with a Package Explorer on the left showing a project named 'mapreduce' with a source folder 'src' containing 'Assignment1' and 'WordCount.java'. The main editor shows the 'WordCount.java' file with the following code:

```
1 package Assignment1;
2
3
4 import java.io.IOException;
5
6
7
8
9
10
11
12
13
14
15
16
17 public class WordCount {
18
19     public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {
20         private final static IntWritable one = new IntWritable(1);
21         private Text word = new Text();
22
23         public void map(LongWritable key, Text value, Context context) throws IOException, InterruptedException {
24             String line = value.toString();
25             StringTokenizer tokenizer = new StringTokenizer(line);
26             while (tokenizer.hasMoreTokens()) {
27                 word.set(tokenizer.nextToken());
28                 context.write(word, one);
29             }
30         }
31     }
32
33     public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
34
35         public void reduce(Text key, Iterable<IntWritable> values, Context context)
36             throws IOException, InterruptedException {
37             int sum = 0;
38             for (IntWritable val : values) {
39                 sum += val.get();
40             }
41             context.write(key, new IntWritable(sum));
42         }
43     }
44 }
```

The bottom of the IDE shows the 'Console' tab with the following output:

```
<terminated> New_configuration [Java Application] /System/Library/Java/JavaVirtualMachines/1.6.0.jdk/Contents/Home/bin/java (Sep 19, 2014 23:07:25)
14/09/19 23:07:25 INFO mapred.JobClient: Map output bytes=2418234700
14/09/19 23:07:25 INFO mapred.JobClient: Total committed heap usage (bytes)=5848104960
14/09/19 23:07:25 INFO mapred.JobClient: SPLIT_RAW_BYTES=5544
14/09/19 23:07:25 INFO mapred.JobClient: Combine input records=0
14/09/19 23:07:25 INFO mapred.JobClient: Reduce input records=248943500
14/09/19 23:07:25 INFO mapred.JobClient: Reduce input groups=5273
14/09/19 23:07:25 INFO mapred.JobClient: Combine output records=0
14/09/19 23:07:25 INFO mapred.JobClient: Reduce output records=5273
14/09/19 23:07:25 INFO mapred.JobClient: Map output records=248943500
```

Last 22 lines of console output

```
14/09/19 23:07:25 INFO mapred.JobClient: map 100% reduce 100%
14/09/19 23:07:25 INFO mapred.JobClient: Job complete: job_local73949850_0001
14/09/19 23:07:25 INFO mapred.JobClient: Bytes Written=73395
14/09/19 23:07:25 INFO mapred.JobClient: FileSystemCounters
14/09/19 23:07:25 INFO mapred.JobClient: FILE_BYTES_READ=153668749092
14/09/19 23:07:25 INFO mapred.JobClient:
FILE_BYTES_WRITTEN=185590723000
14/09/19 23:07:25 INFO mapred.JobClient: File Input Format Counters
14/09/19 23:07:25 INFO mapred.JobClient: Bytes Read=1454183628
14/09/19 23:07:25 INFO mapred.JobClient: Map-Reduce Framework
14/09/19 23:07:25 INFO mapred.JobClient: Map output materialized
bytes=2916121964
14/09/19 23:07:25 INFO mapred.JobClient: Map input records=21907700
14/09/19 23:07:25 INFO mapred.JobClient: Reduce shuffle bytes=0
14/09/19 23:07:25 INFO mapred.JobClient: Spilled Records=1118148571
14/09/19 23:07:25 INFO mapred.JobClient: Map output bytes=2418234700
14/09/19 23:07:25 INFO mapred.JobClient: Total committed heap usage
(bytes)=5848104960
14/09/19 23:07:25 INFO mapred.JobClient: SPLIT_RAW_BYTES=5544
14/09/19 23:07:25 INFO mapred.JobClient: Combine input records=0
14/09/19 23:07:25 INFO mapred.JobClient: Reduce input records=248943500
14/09/19 23:07:25 INFO mapred.JobClient: Reduce input groups=5273
14/09/19 23:07:25 INFO mapred.JobClient: Combine output records=0
14/09/19 23:07:25 INFO mapred.JobClient: Reduce output records=5273
14/09/19 23:07:25 INFO mapred.JobClient: Map output records=248943500
```

AWS Execution

Cluster: ClusterWordCount_11:13PM Terminated Steps completed



Connections: --
Master public DNS: ec2-54-68-244-42.us-west-2.compute.amazonaws.com [SSH](#)
Tags: --

Summary	Configuration Details	Security/Network	Hardware
ID: j-2K15CTOMQTL4W	AMI version: 3.2.1	Availability zone: us-west-2a	Master: Terminated 1 m1.medium
Creation date: 2014-09-16 23:15 (UTC-4)	Hadoop distribution: Amazon 2.4.0	Subnet ID: subnet-f262bd97	Core: Terminated 2 m1.medium
End date: 2014-09-16 23:43 (UTC-4)	Applications: --	Key name: --	Task: --
Elapsed time: 27 minutes	Log URI: s3://hw1wordcount/hw1logs/	EC2 instance profile: --	
Auto-terminate: Yes	EMRFS consistent view: Disabled	EMR role: --	
Termination protection: Off		Visible to all users: None Change	

▼ Steps

<div>Add step</div>							
<div>Steps<div>View all interactive jobs View all jobs</div></div>							
<div>Filter: All steps <input type="text" value="Filter steps ..."/> 2 steps (all loaded)</div>							
	ID	Name	Status	Start time (UTC-4) ▼	Elapsed time	Log files	Actions
▶	s-3JAJB1WIGU1IB	WordCountJAR	Completed	2014-09-16 23:22 (UTC-4)	18 minutes	View logs	View jobs
▶	s-8JJACH26U6IN	Setup hadoop debugging	Completed	2014-09-16 23:21 (UTC-4)	27 seconds	View logs	View jobs