

Code Running Assignment 5

Nick Miller

```
[nam99@had6110 NewMaxTemp]$ hadoop jar wc.jar MaxTemperature /user/chatree/CS1699/ncdc/
TestOut
```

```
16/10/25 16:01:25 INFO client.RMProxy: Connecting to ResourceManager at
had6110.cs.pitt.edu/136.142.167.12:8032
```

```
16/10/25 16:01:26 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not
performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
```

```
16/10/25 16:01:26 INFO input.FileInputFormat: Total input paths to process : 2
```

```
16/10/25 16:01:26 INFO mapreduce.JobSubmitter: number of splits:2
```

```
16/10/25 16:01:26 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1475179034150_0393
```

```
16/10/25 16:01:27 INFO impl.YarnClientImpl: Submitted application application_1475179034150_0393
```

```
16/10/25 16:01:27 INFO mapreduce.Job: The url to track the job:
http://had6110.cs.pitt.edu:8088/proxy/application_1475179034150_0393/
```

```
16/10/25 16:01:27 INFO mapreduce.Job: Running job: job_1475179034150_0393
```

```
16/10/25 16:01:35 INFO mapreduce.Job: Job job_1475179034150_0393 running in uber mode : false
```

```
16/10/25 16:01:35 INFO mapreduce.Job: map 0% reduce 0%
```

```
16/10/25 16:01:44 INFO mapreduce.Job: map 100% reduce 0%
```

```
16/10/25 16:01:52 INFO mapreduce.Job: map 100% reduce 100%
```

```
16/10/25 16:01:52 INFO mapreduce.Job: Job job_1475179034150_0393 completed successfully
```

```
16/10/25 16:01:52 INFO mapreduce.Job: Counters: 49
```

File System Counters

```
FILE: Number of bytes read=341386
```

```
FILE: Number of bytes written=1045402
```

```
FILE: Number of read operations=0
```

```
FILE: Number of large read operations=0
```

```
FILE: Number of write operations=0
```

```
HDFS: Number of bytes read=1777420
```

```
HDFS: Number of bytes written=103646
```

```
HDFS: Number of read operations=9
```

```
HDFS: Number of large read operations=0
```

Code Running Assignment 5

Nick Miller

HDFS: Number of write operations=2

Job Counters

Launched map tasks=2

Launched reduce tasks=1

Data-local map tasks=2

Total time spent by all maps in occupied slots (ms)=14005

Total time spent by all reduces in occupied slots (ms)=4841

Total time spent by all map tasks (ms)=14005

Total time spent by all reduce tasks (ms)=4841

Total vcore-milliseconds taken by all map tasks=14005

Total vcore-milliseconds taken by all reduce tasks=4841

Total megabyte-milliseconds taken by all map tasks=14341120

Total megabyte-milliseconds taken by all reduce tasks=4957184

Map-Reduce Framework

Map input records=13130

Map output records=13130

Map output bytes=315120

Map output materialized bytes=341392

Input split bytes=252

Combine input records=0

Combine output records=0

Reduce input groups=4380

Reduce shuffle bytes=341392

Reduce input records=13130

Reduce output records=4380

Spilled Records=26260

Shuffled Maps =2

Failed Shuffles=0

Merged Map outputs=2

Code Running Assignment 5

Nick Miller

GC time elapsed (ms)=179

CPU time spent (ms)=7290

Physical memory (bytes) snapshot=1594560512

Virtual memory (bytes) snapshot=15849668608

Total committed heap usage (bytes)=1463287808

Shuffle Errors

BAD_ID=0

CONNECTION=0

IO_ERROR=0

WRONG_LENGTH=0

WRONG_MAP=0

WRONG_REDUCE=0

File Input Format Counters

Bytes Read=1777168

File Output Format Counters

Bytes Written=103646

16/10/25 16:01:52 INFO client.RMProxy: Connecting to ResourceManager at had6110.cs.pitt.edu/136.142.167.12:8032

16/10/25 16:01:52 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

16/10/25 16:01:52 INFO input.FileInputFormat: Total input paths to process : 1

16/10/25 16:01:52 INFO mapreduce.JobSubmitter: number of splits:1

16/10/25 16:01:53 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1475179034150_0394

16/10/25 16:01:53 INFO impl.YarnClientImpl: Submitted application application_1475179034150_0394

16/10/25 16:01:53 INFO mapreduce.Job: The url to track the job:
http://had6110.cs.pitt.edu:8088/proxy/application_1475179034150_0394/

16/10/25 16:01:53 INFO mapreduce.Job: Running job: job_1475179034150_0394

16/10/25 16:02:06 INFO mapreduce.Job: Job job_1475179034150_0394 running in uber mode : false

16/10/25 16:02:06 INFO mapreduce.Job: map 0% reduce 0%

Code Running Assignment 5

Nick Miller

16/10/25 16:02:13 INFO mapreduce.Job: map 100% reduce 0%

16/10/25 16:02:20 INFO mapreduce.Job: map 100% reduce 100%

16/10/25 16:02:21 INFO mapreduce.Job: Job job_1475179034150_0394 completed successfully

16/10/25 16:02:21 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=96366

FILE: Number of bytes written=434471

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=103779

HDFS: Number of bytes written=43842

HDFS: Number of read operations=6

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

Job Counters

Launched map tasks=1

Launched reduce tasks=1

Data-local map tasks=1

Total time spent by all maps in occupied slots (ms)=4918

Total time spent by all reduces in occupied slots (ms)=4564

Total time spent by all map tasks (ms)=4918

Total time spent by all reduce tasks (ms)=4564

Total vcore-milliseconds taken by all map tasks=4918

Total vcore-milliseconds taken by all reduce tasks=4564

Total megabyte-milliseconds taken by all map tasks=5036032

Total megabyte-milliseconds taken by all reduce tasks=4673536

Map-Reduce Framework

Map input records=4380

Code Running Assignment 5

Nick Miller

Map output records=4380

Map output bytes=87600

Map output materialized bytes=96366

Input split bytes=133

Combine input records=0

Combine output records=0

Reduce input groups=2190

Reduce shuffle bytes=96366

Reduce input records=4380

Reduce output records=2190

Spilled Records=8760

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=119

CPU time spent (ms)=3810

Physical memory (bytes) snapshot=909004800

Virtual memory (bytes) snapshot=10587672576

Total committed heap usage (bytes)=820510720

Shuffle Errors

BAD_ID=0

CONNECTION=0

IO_ERROR=0

WRONG_LENGTH=0

WRONG_MAP=0

WRONG_REDUCE=0

File Input Format Counters

Bytes Read=103646

File Output Format Counters

Code Running Assignment 5
Nick Miller

Bytes Written=43842