CprE 419 Lab 4: Sorting Using Hadoop MapReduce

Yu Zhang

1. How to run my mapreduce program

I have attached a readme in my submission. The command that I used is as follows:
First clean the output folder and any temp folder:

```
hadoop fs -rmr /scr/yuz1988/lab4
```

Then run my program:

```
hadoop jar /home/yuz1988/lab4/MySort.jar MySort /class/s14419x/lab4/gensort-out-250M
/scr/yuz1988/lab4/output    /scr/yuz1988/lab4/temp1    /scr/yuz1988/lab4/temp2/partlist
```

There are four arguments passed to my program. First is the input directory, second is the final output directory, third is the directory for the temp which contains the copy of the input files but the format is SequnceFile consisting of binary key/value pairs, forth argument is a path for a file (not a directory), so the name of my partition list file is partlist. The partition list should contain 14 keys (one less than the number of reducers).

Notice: Between my two map-reduce rounds is the sampling stage. As the sampling number is large, so it lasts a little bit long. After the first map-reduce round, it takes about 7~8 minutes (less than 10 minutes) before the second map-reduce to start.

2. The algorithm I used in my program

First I would thank Cody Hoover who gave me a lot of help and guided me to read the chapter 8 of the text book (definitive guide). Basically, I just did the very similar thing as the book does. First I reformat the input files to sequence files which are organized by key/value pairs. We know TextInputFormat reads each line and sets the line number as the key, content of the whole line as the value. But when we do the sampling, we need to read the first ten characters as the key, not the line number. So that is the reason we need to reformat the file that change each line to key/value pairs.

My program contains two map-reduce rounds. The first round is to generate the sequence files. It only contains map process and no reduce process. So in map process I set the first ten characters as the key and the remaining as the value.

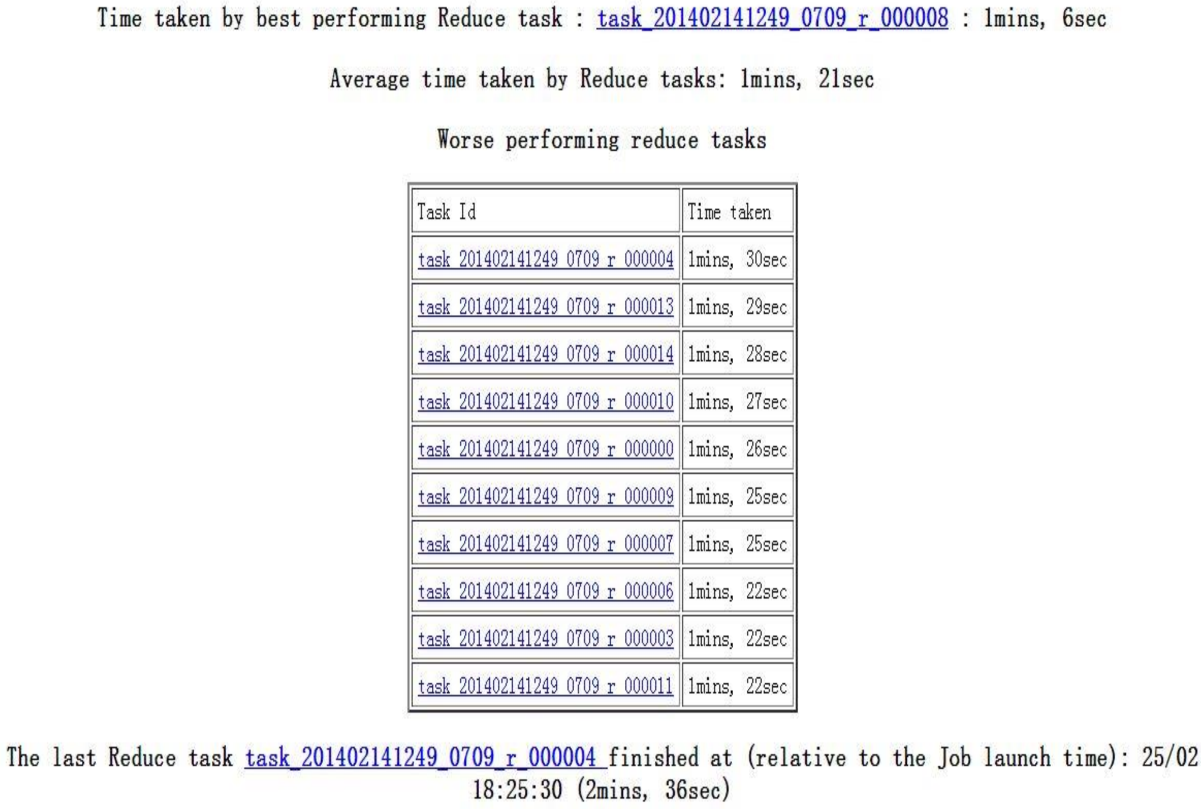
Before the second map-reduce round, I sample the dataset using RandomSampler and set the number of samples as 5% of the data. As 5% of 250M is 12.5M which is so large and may run out of the memory, so the maximum number of samples I set is 100,0000. Then the sampler generates the partition list and writes it to my partlist.

The second map-reduce round only contains reduce process. As we already parse each line to key/value pairs so we don't need map process. For each record, it reads the partition list and sends it to the corresponding reducer. And in total the output is ordered.

3. Screenshot of total reducer time

As stated in the second part, my first round map-reduce doesn't have a reduce round. So the time is 0sec.

The second map-reduce has 15 reducers, the time table is as follows:



Time taken by best performing Reduce task : task_201402141249_0709_r_000008 : 1mins, 6sec

Average time taken by Reduce tasks: 1mins, 21sec

Worse performing reduce tasks

Task Id	Time taken
task_201402141249_0709_r_000004	1mins, 30sec
task_201402141249_0709_r_000013	1mins, 29sec
task_201402141249_0709_r_000014	1mins, 28sec
task_201402141249_0709_r_000010	1mins, 27sec
task_201402141249_0709_r_000000	1mins, 26sec
task_201402141249_0709_r_000009	1mins, 25sec
task_201402141249_0709_r_000007	1mins, 25sec
task_201402141249_0709_r_000006	1mins, 22sec
task_201402141249_0709_r_000003	1mins, 22sec
task_201402141249_0709_r_000011	1mins, 22sec

The last Reduce task task_201402141249_0709_r_000004 finished at (relative to the Job launch time): 25/02 18:25:30 (2mins, 36sec)

So the total reducer time is **1min and 30sec**.