1. How many extra days did you use for the project?

   10 days. Finished on 24th Apr. Actually, we finished the majority part before deadline, but the tail job goes slowly with preparing for exams and other things.

2. Given the same goals, how would you complete the project differently if you didn't have any restrictions imposed by the instructor? This could involve using a particular library, programming language, etc. Be sure to provide sufficient detail to justify your response.

   Unlike DFS in project 1, the MapReduce part actually doesn't involve much multithread job, so I guess other languages like Java or python might also work well.

   Among all the framework and tools we learnt in this class, zookeeper might be a good choice to help with manage nodes and scheduling jobs. As it is great at electing leaders and coming to a consensus.

3. Let's imagine that your next project was to improve and extend P2. What are the features/functionality you would add, use cases you would support, etc? Are there any weaknesses in your current implementation that you would like to improve upon? This should include at least three areas you would improve/extend.

   (1) Support chaining MapReduce job.

   (2) Improved fault tolerance: The current system's successful operation relies on users strictly adhering to the standards, and the entire system remaining stable, such as no worker failure during a MapReduce job, etc. If we have more time, we would like to provide more comprehensive error handling capabilities.

   (3) User defined key value type instead of byte array of the current version.

   (4) We have been thinking whether multi-thread in doing map tasks would help.

4. Give a rough estimate of how long you spent completing this assignment. Additionally, what part of the assignment took the most time?

   About 10 full days overall, i.e. (5 days for each person).

   Mapper and Reducer take the most time.

5. What did you learn from completing this project? Is there anything you would change about the project?

   We have gained a solid proficiency in Go! And a more detailed understanding of MapReduce even deeper than we learnt in the preparation of presentation. Also how to clusters work together.

   We can't think of anything need change. It's perfect, complicated enough to be put on resume, also gives a great sense of achievement when completing it.

6. If you worked as a group, how did you divide the workload? What went well with your team, and what did not go well?

   Unlike divided by functions in project 1, this time, we do the project in turns. Luna started the project from client side. She finished the communication between client and MapReduce manager and the job assignment part. I took over from there and finished the map, shuffle and reduce method. Then from here, Luna did two

MapReduce jobs, report progress and testing work. And we did the clean up work together.

It went pretty well with our team, I can’ t think of anything that goes wrong really.