Final Report
CS265, Spring 2020
James Bond

Formatting
Your report should be in the order of 10 single column pages. Use 10 point font. Spend roughly two pages describing your design and the rest of the report should be about the experimental analysis (roughly one page per experiment).

Guidelines
The design description (Section 1) describes the overall design in detail. Do not include any code. You may include pseudocode if you want to describe an algorithm that goes beyond what we discussed in class or if it contains one or more of optimizations that you’ve proposed. Do include a lot of structure (Sections, subsections and figures) so that it is easy to read. Reports should be polished to be typo free and easy to read. There should be at least one experiment for each of the operators (Section 2) and the three optimizations (Sections 3). We expect at least 3 additional experiments with three different workloads (read-intensive, mixed, and write-intensive) in Section 4. You can add more experiments in sections of additional experiments in the end. These additional experiments are the perfect opportunity to show-case any design features and/or performance tuning you have implemented that are beyond the basic scope of the project or to go deeper to analyze the existing work.

Tips for Running Experiments
1) Prepare scripts for running your experiments so you can easily redo experiments, and update the setup.
2) Isolate performance as best as possible (close all apps, recreate the same environment every time).
3) Test by changing one parameter at a time to find out what affects performance

1. Design Description (2 pages)

1.1. Introduction

Summarise in one paragraph what you developed in the project. In terms of technical details, this section should be easily understandable to someone with familiarity with data systems. For example, don’t assume that the reader knows we are building a write-optimized NoSQL key-value store, but you can assume that they know what that is.

1.2 Technical Description

This should be a detailed overview of the design considerations made throughout the project. For each design decision, you should specify (a) the intuition behind decision in the first place, (b) detailed description of the design elements involved, and c) how it was implemented at a high level in your code. Additionally, you can include deeper details with supporting materials such as,
architecture diagrams, flowcharts, and/or pseudocode for clarification. Some examples of design considerations are as follows:

*Examples:*
*How did you design the on-disk data layout?*
*How do you store bloom filters and fence pointers in-memory?*
*How did you decide on the data structure design of the memtable?*
*How do you compact data across levels?*

...

1.3 Additional Optimization

You should list the three optimizations (or more) that you’ve additionally implemented on top of the basic design requirements. The format should be the same as subsection 1.2.

1.4. Challenges

You should list all challenges that you’ve faced during the implementation of this project and how you resolved them.

2. Performance Analysis of Basic Operators (3 page)

2.1 Get

In this subsection, you’ll describe the details of the experiments to benchmark the performance of get operations. For each operator, you should include the following paragraphs:

*Paragraph 1* **Process of Benchmarking.** Explain how you benchmarked the performance of gets -- number of queries used for the experiment, size of the dataset, distribution used to generate the dataset, the hardware used, and number of parallel clients supported. Ideally, you should repeat each experiment for n times and plot the data points with a confidence interval to eliminate any kind of bias and error that may have been introduced during query or data generation processes.

*Paragraph 2* **Performance Analysis.** Same as Section 3.1 (discussed below), you should add at least 2 graphs explaining the performance of the query. You should include a short text on what one can learn from the observed results.

2.2 Put

Same as 5.1

2.3 Range

Same as 5.1
2.4 Delete
Same as 5.1

2.5 Load
Same as 5.1

3. Experimental Analysis for Optimizations

This section is to evaluate the performance of each of the optimizations that you've implemented on top of the basic design.

3.1 Optimization 1 (1 page)

For each experiment, write the following three paragraphs.

(Paragraph 1) Motivation. Explain your goal and the motivation for doing the specific experiment. What are you trying to investigate and why?

(Paragraph 2) Set-up. Explain your data, queries and hardware for the experiment. Someone else should be able to read that description and fully reproduce the experiment. Also the setup should connect with the original goals set in the previous paragraph.

![Graph](image1.png)  
Figure 1. Caption of performance graph

![Graph](image2.png)  
Figure 2. Caption of additional graph

(Paragraph 3) Result discussions. Show at least one performance graph and at least one additional graph which explain the results (similar to Figure 1 and Figure 2). The performance graph should illustrate the improvement achieved over the basic design. For example the result graph can report latency (2 plots -- with and without the optimization) while the additional graph can report cache misses which helps explain the observed latency. Then provide a paragraph of text to discuss the results and explain the reason why we see the resulting behavior. Your captions should make a convincing point about a performance trend or result, not merely restating/describing what is observed.

3.2 Optimization 2 (1 page)
Same as Sec 2.

3.3 Optimization 3 (1 page)

4. Performance Analysis of Workloads (1 page)

In this section, you should show the overall performance measures of your system on different workloads. For this purpose, you should use a read-intensive, a mixed, and a write-intensive workload. The idea is to show the performance tradeoffs for different workloads as you vary the proportion of different queries.

5. Additional Experiment 1 (1 page)

...

6. Additional Experiment 2 (1 page)

...

7. Additional Experiment 3 (1 page)

...