DATABASE 1

Phase2

Junyong Lee

In this report, I am going to compare the performance of Single Insertion operation and Bulk Insertion operation of SQL. For implementing the task, I created a web interface by using NodeJS and basic web structure.

This program was written with NodeJS, so you need to install NodeJS before you run this program.

And please run the following two commands

npm install

npm start

The webpage would be at : localhost:3000/

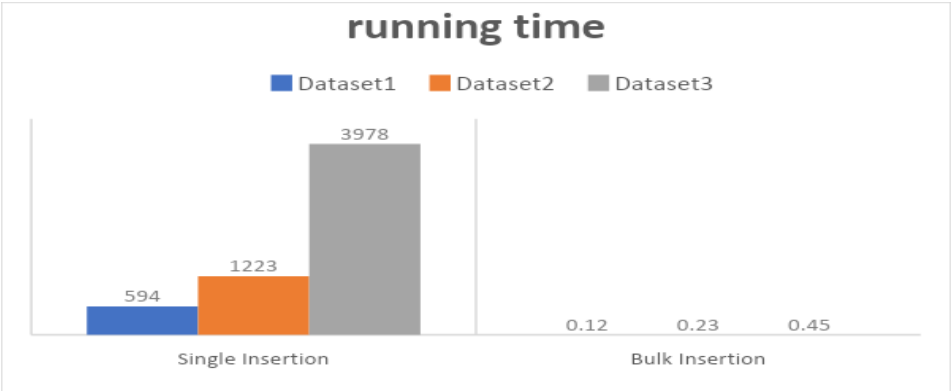
In NodeJS, there is a library called 'mysql2' for managing many different operations between NodeJS code and mySQL database including database connections and query delivery.

I defined several different functions in my code to read and parse the data in the text file, and also combine them with general queries that execute single insertion and bulk insertion respectively.

And also, since NodeJS uses async functions, at this time, I used workbench to define the schema of the table. However, the data insertions were done by code.

Following is the execution time result for each method on 3 datasets.

	Dataset1	Dataset2	Dataset3
Single insertion	594sec (about 10 min)	1223sec (about 20min)	3978sec (about 66min)
Bulk insertion	0.12sec	0.23sec	0.25sec



Takeaway and Conclusion

As we can see above, the bulk insertion took significantly less time than single insertion.

Generally, Single Insertion took much more time than bulk insertion. One of the interesting things that I noticed was that there was some delay of time for the database to actually display the inserted data. So, I was able to see the increasing number of data at the same time. (Details are explained in the video).

In addition, there was no big difference of execution time from bulk insertion for 3 datasets. However, Single insertion showed the longer execution time as the input data increase.

From the result, we can guess that the bulk insertion is much faster because in case of bulk insertion, only extent allocations are logged instead of the actual data being inserted. This will provide much better performance than just single insertion.