CS 677 S21 Lab 3 - Bookstore
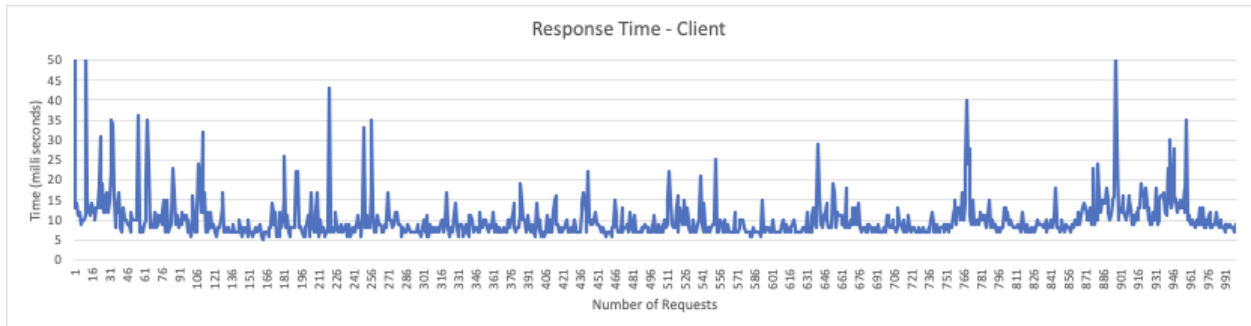
Performance Evaluation Document

Submitted By

Jagriti Singhal, Bhumika Kalavadia

[1]. Compute the average response time (query/buy) of your new systems as before. What is the response time with and without caching? How much does caching help? What is the latency of a subsequent request if it sees a cache miss?

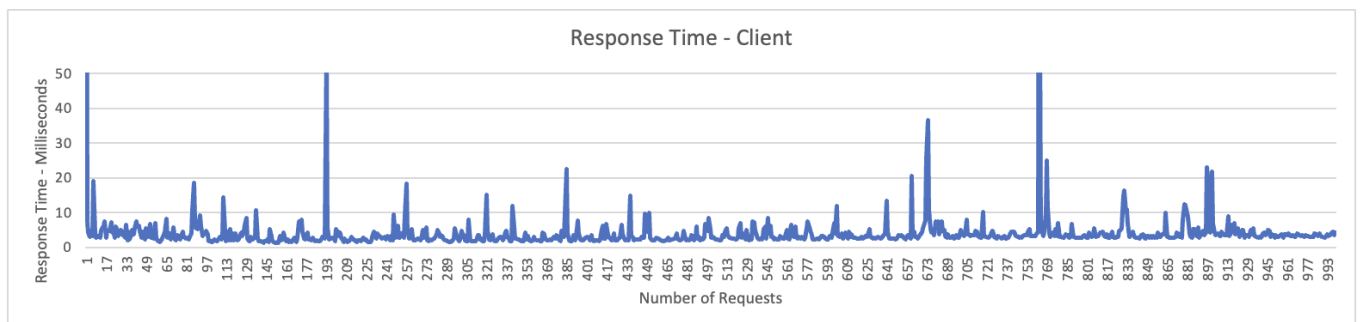**Response time without Caching for Lookup request:**

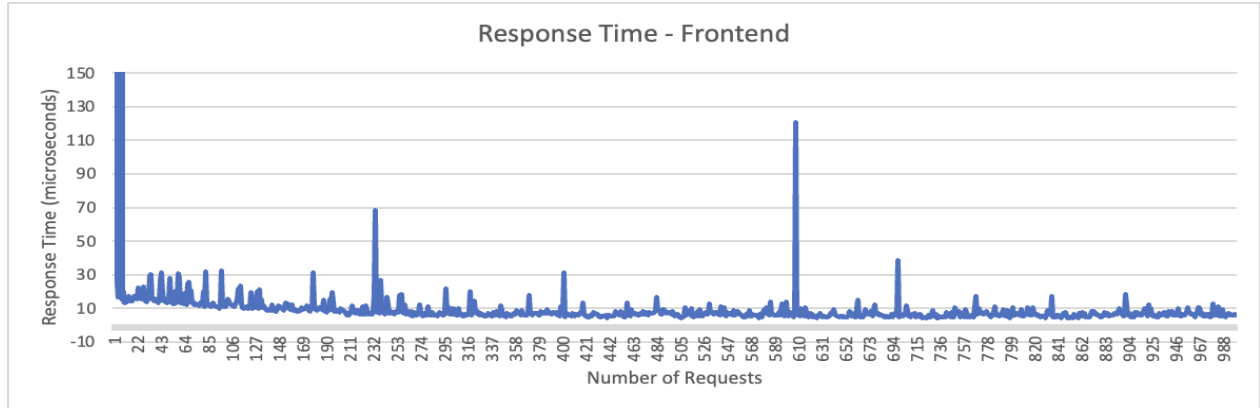

Average response time = 10.67 milliseconds
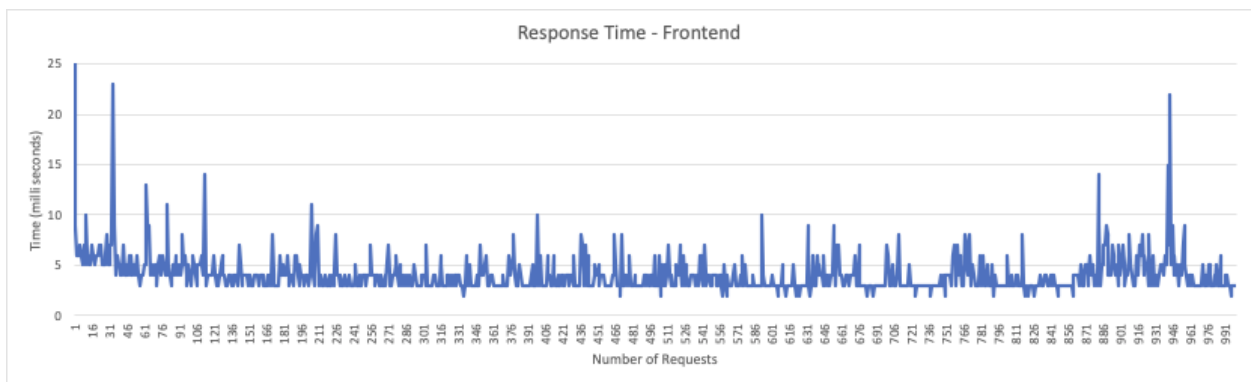
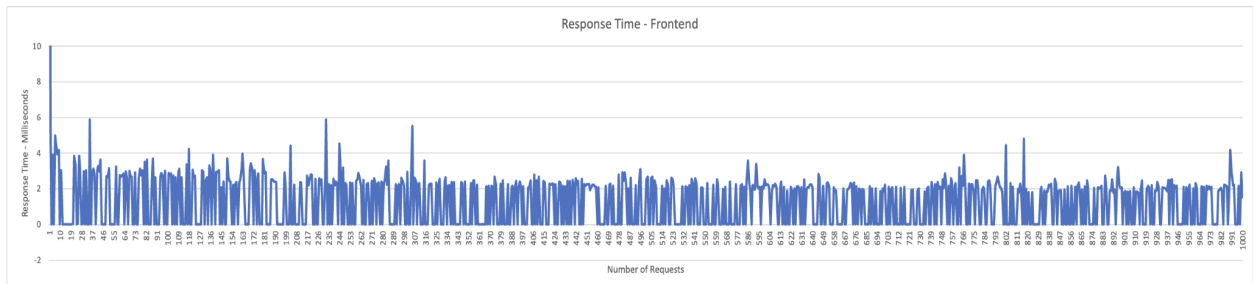Average response time = 4.24 milliseconds

Response time with Caching for Lookup request:

Average Response Time - 4.086 milliseconds

Average response time = 0.182 milliseconds (all requests served from cache)

Average response time = 1.53 milliseconds (with cache misses)

For this experiment, we triggered 1000 requests to the frontend server for lookup. There are 4 books for which we generated the lookup requests. We set the CACHE_SIZE = 2. We see clearly in the graph after every few requests, we see a spike in response time. This is a cache miss where the frontend server has to invoke catalog microservice to serve the request. In absence of such cache misses if request is always served from frontend cache the response time reduces further to just 0.182 milliseconds as seen above.

We observe that cache reduces the response time by more than 50%. This is because we overcome the network overhead involved while calling catalog microservice in without caching approach.

[2]. Construct a simple experiment that issues orders or catalog updates (i.e., database writes) to invalidate the cache and maintain cache consistency. What is the overhead of cache consistency operations?
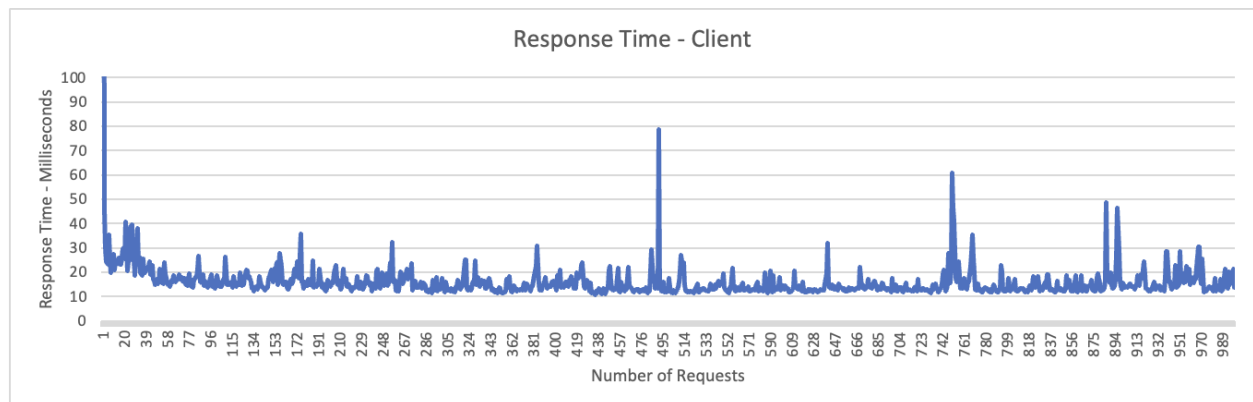
Response Time without Cache for Buy request i.e. no cache consistency overhead.

Average Response Time = 12.23 milliseconds
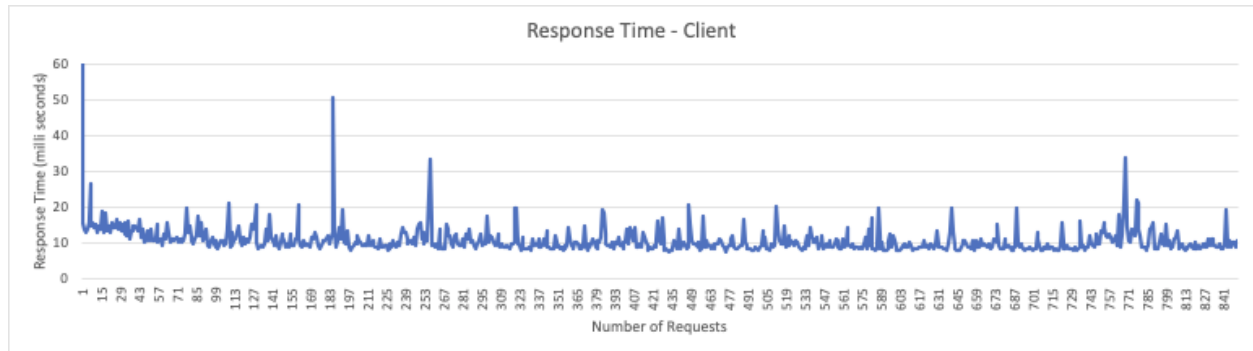
Response Time with Cache for Buy request i.e. involves cache consistency overhead.

Average Response Time = 17.98 milliseconds

For cache consistency, catalog microservice needs to invoke invalidate API on client to remove the dirty cache entry on successful buy request. This causes an increase in response time to serve the buy requests. We see an overhead of ~5.75 milliseconds.