I have measured the performance of the system using a Java program that I wrote which included in the repository.

Below is the table of average response time of ten thousand “lookup” and “buy” request each, with and without cache

Average response time for request	With cache	Without cache
Lookup request	Average time: 13.0385	Average time: 19.5437
Buy request	Average time: 105.144	Average time: 93.742

As seen from the table the cache helps the “lookup” request by nearly 1.5 times.

Also, we can see that the cache has a significant performance overhead on the “buy” request

To test the cache consistency operation (invalidates) and cache miss latency I will simulate the following scenario with the following http requests

1- Lookup for book 1 – cache miss
2- Lookup for book 3 – cache miss
3- Lookup for book 1 – cache hit
4- Buy book 1
5- Lookup for book 1 - cache miss

We will compare these requests with the previous that we got in the first table. Since we don’t have many requests, I am going to use Postman to get the latency of each request.

Latencies:

1- 27 ms
2- 21 ms
3- 10 ms
4- 141 ms
5- 19 ms

From the latency list we can see that the “buy” causes an overhead when lookup comes after it.