PERFORMANCE METRICS DOCUMENT

COMPSCI-677

LAB 2 - PYGMY.COM

In order to conduct performance metrics evaluation of our system, we implemented the following steps:

- Record the starting time stamp for each new request that is generated by any client
- Generate a unique ID for that request, which is forwarded along with the request to each server that is involved with completing that request
- Record the time stamp when a response is received and calculate the difference in milliseconds
- Record the time taken for each request to get completed along with its unique request ID
- Calculate total time taken for all requests to be completed for a particular client as “end to end time of completion”
- Track number of requests generated in order to calculate average response time for each client across all its requests
- Save the list of request IDs, response times, avg response times and end to end completion times to file titled ‘client_<client number>_metrics.txt’ in the appropriate tests subdirectory
- In addition to this, each of the servers (front-end, catalog, order) also logs the request ID and time taken to service that request for each request they receive at each API end point - Thus, the path of each request can be traced across the client and all the servers by finding the appropriate request ID within the log files of each of the servers and seeing how much time was taken to process that specific request at that specific server.

Below, you can see the recorded metric values from one execution cycle of tests on EDLAB machines:

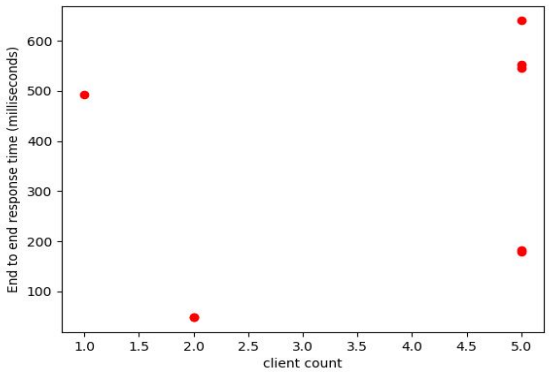
	Client	**Avg. response time** (milliseconds)	**End to end response time** (milliseconds)
Test 1	Client 1	163.81766666666667	491.453
Test 2	Client 1 Client 2	16.321666666666665 16.019333333333332	48.964999999999996 48.05799999999999
Test 3	Client 1 Client 2 Client 3 Client 4 Client 5	184.33666666666667 213.01399999999998 181.49833333333333 59.51433333333333 60.58533333333333	553.01 639.0419999999999 544.495 178.543 181.756
Test 4	Client 1	93.67237800000008	93672.37800000008
Test 5	Client 1 Client 2	154.738773 159.607368	154738.77300000002 159607.36800000002

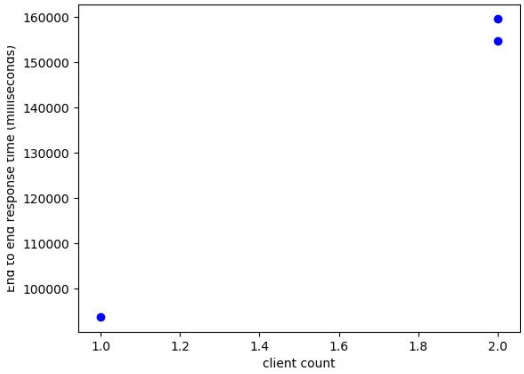
It may be noted that since requests are generated randomly, and update (buy) requests have a larger turn around time than query (search/lookup) requests. As a result of this the metrics may be skewed depending on the distribution of these requests across clients

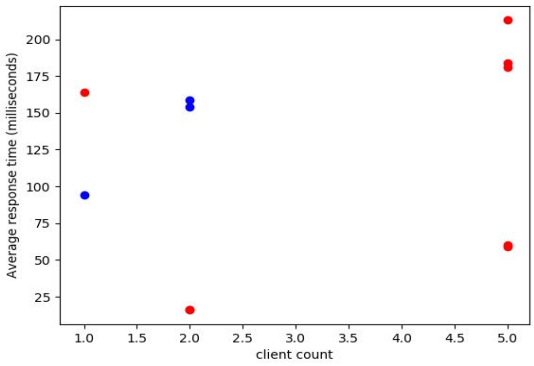
Sample file showing how the order server logs the unique request IDs and the request service time:

Request ID	Request completion time (milliseconds)
c17d9498-782f-11ea-9714-7304c17ec9ca	226.526
c1a93cd8-782f-11ea-9714-7304c17ec9ca	183.067
c20d6294-782f-11ea-8781-28f10e08bdf0	250.796
c21572f4-782f-11ea-a5ba-cbfbcd6061b9	450.153
c217453e-782f-11ea-a5ba-cbfbcd6061b9	571.889
c299be42-782f-11ea-b838-f7046ed58e93	197.609
c2bf8a82-782f-11ea-b838-f7046ed58e93	149.331
c2de5408-782f-11ea-b838-f7046ed58e93	181.373
c30222f2-782f-11ea-b838-f7046ed58e93	177.507
c323a710-782f-11ea-b838-f7046ed58e93	171.033

c3514148-782f-11ea-b838-f7046ed58e93 183.396







The above plots show clients running 1000 requests in blue, and clients running 3 requests in red. The trends suggest that as the number of clients increases, average and end to end request processing time also tends to increase