CS-425 MP1 Report Group No: 44

Lin Lyu(linlyu2), Ching-Hua Yu(cyu17)

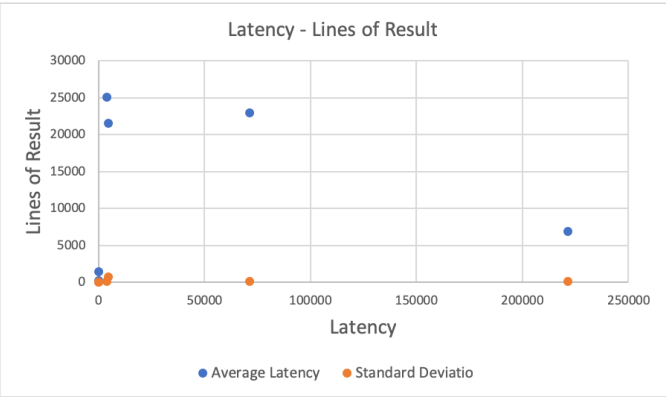
I. Design

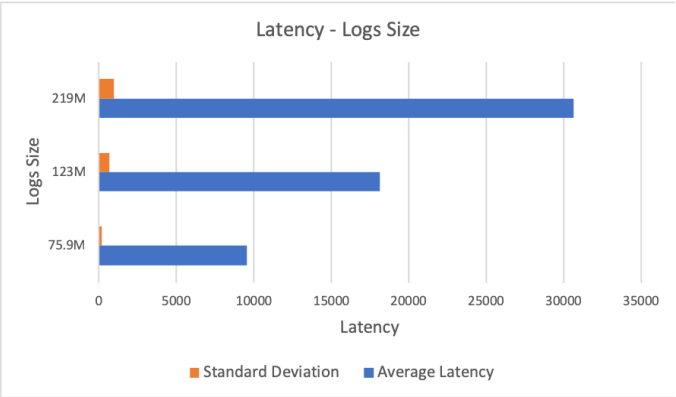
Following the master-slave framework, we build communication between one master (client) and many slaves (servers). First, we have each VM run a server daemon in the background. The server is designed to listen to a pre-coordinated port (8080 by default), receive the (permitted) requests from a client, execute the request and return the result. Second, a client, which can be run from everywhere in the network, connects to the pre-coordinated port, sends a request command (according to the arguments), to any combination of servers (according to the arguments), and receives and saves the result from the servers. In this MP, we only permit “grep” queries that are implemented by shell command on the server, with the framework potential to add more functions later. The client connects and issues requests to the servers in parallel using threads.

II. Unit Test

The test program first generates random patterns according to several parameters involving fixed patterns and frequencies and then store the frequency of some patterns to file. The test is then conducted on checking the consistencies between the answers and our distributed queries with some statistics report.

III. Some Results





Based on the chart above, we can see that query time does not go with the increasing of result lines, which may be caused by the fact that ‘grep’ needs to check all contents anyway. But it rises significantly when we increase the size of logs probably due to more contents to be checked. The difference of latency may result from the grep function itself.