CS425 MP1 Report

Team members: Taipeng Liu(taipeng2) & Xiang Li(xiangl14)

Design:

Our Project is implemented with the socket interface provided by Golang. Our Distributed Log Querier simply works this way, as long as any server is listening to the specific port(i.e. 8888), we can log in another VM to work as a client, then query the log file using grep with a certain pattern. Once connected, the servers are ready to read and check input from client, parse the input with an option and a pattern, "grep" on the log file locally. The client will receive the result and print the number of line count of each connecting server, and save the records into "MP1.log". If some servers are offline or crash during connection, the client will detect the error and close r/w connection. It won't interrupt the connections with other servers. Latency is counted as the time difference between the time client sends the 'grep' request and the time it receives messages. Besides the basic requirements, we also accomplish input format check on server-side, logger redirection and unit test on client-side.

Unit test:

We write a function "logCnt" to test the total line numbers of records from our client-server connection function. If the line counts match all the expected results, which is obtained by servers running corresponding 'grep' command locally, the unit test passes. As a test case, we choose a frequent pattern ("^[7-9]{1}"), normal pattern ("google") and infrequent pattern ("abc").

Visualization:

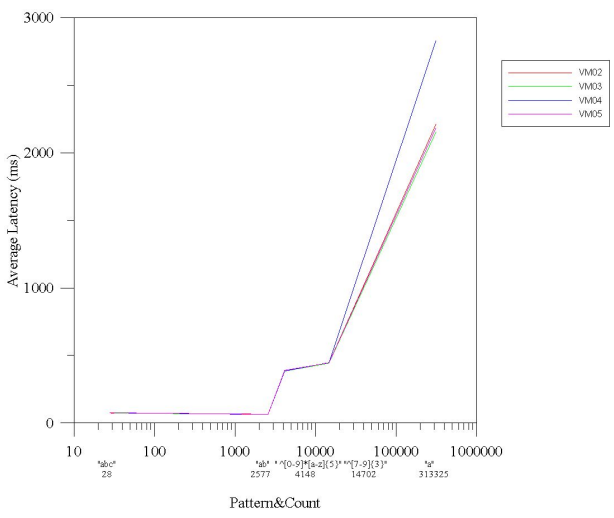

Fig1. Expectation of Latency

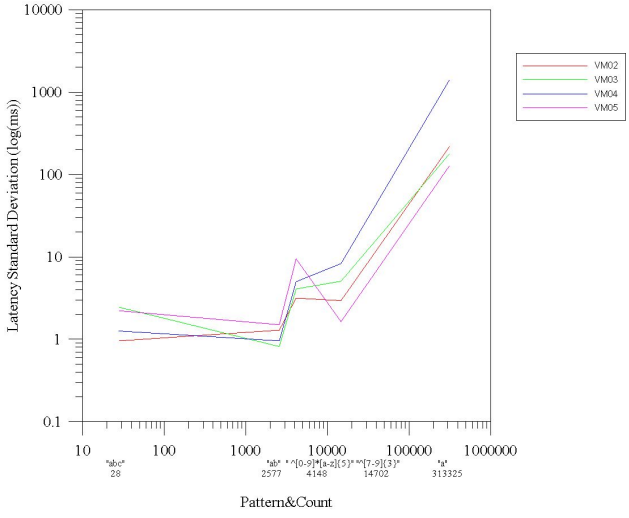

Fig2. Standard Deviation of Latency

As the result shows, to query a more frequent pattern(i.e. a larger line count) will have longer average latency time. Besides, since each VM has a unique connection, one single server might have longer latency due to unstable connection which will lead to a larger standard deviation at some data points. Our result also shows that the threshold for a distributed log query is about 50 ~ 60 ms even if there is no matched pattern. We will further test on our own-generated log files.