CS425 DS MP1 - Distributed Log Querier Report

Contributors: Shubham Singhal & Nirupam K N (ss77, nirupam2)@illinois.edu

1. Design & Implementation:

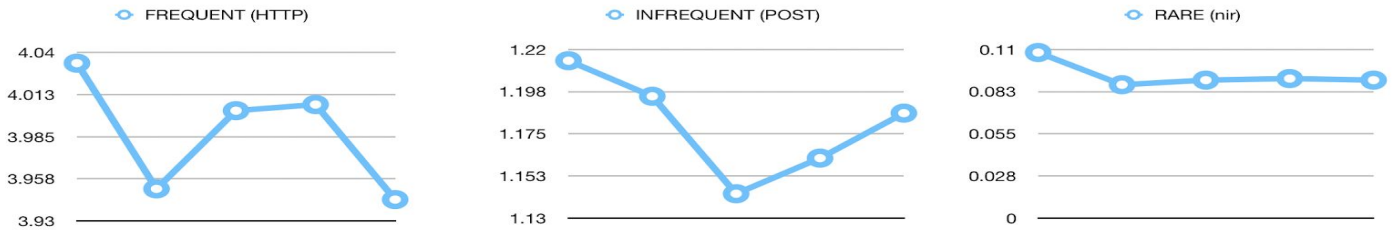
Client: The client takes the pattern from the user as a command-line argument and concurrently calls the servers (goroutines) to get the matched lines from their respective log files. The matched output from the server is sent to the channel in the client code which helps in synchronisation of the output from multiple servers. From this channel, the client reads the desired output and prints it on the terminal.

Server: The server code accepts the connection and spawns a goroutine to process the log file to match the given pattern from the client. The server runs a sub-process grep command to match the pattern and return the matched lines to the client.

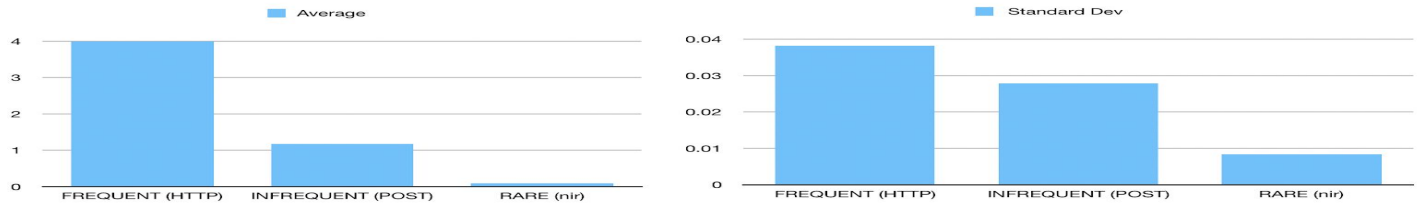
2. Unit Tests:

We have written comprehensive test cases to cover frequent, somewhat frequent and rare pattern on a single server. Along with that we have tested the presence of patterns on single, some and all the servers. To achieve this we have generated the necessary log files and pushed to the respective servers. The unit test is performed on 1 Client & 4 Servers topology.

3. Performance: (Performance is tested on 1 Client & 4 Servers topology. File size on each server = ~ 75M)



Line Graph: X-axis: Trial number Y-axis: Time(in seconds)



Bar Graph: X-axis: Pattern Frequency Y-axis: Time(in seconds)

Average time to execute the functionality is **3.987s** for Frequent pattern - "**HTTP**"
Average time to execute the functionality is **1.180s** for Infrequent pattern - "**POST**"
Average time to execute the functionality is **0.009s** for Rare pattern - "**nir**"

AvgTime(Frequent) > AvgTime(Infrequent) > AvgTime(Rare)

The above mentioned graphs are inline with the expected results with respect to execution time.