CS425 Distribute Systems

MP1-Distributed Log Querier

System Design& Implementation

The design concept of Distributed Log Querier is based on the Client -Server Model, and we implemented the grep functionality to query distributed log files on multiple machines.

We implemented the Distributed Log Querier in java and used Sockets for the communications between the Client and Server. The Server component has two classes—GrepServer and GrepServerThread. The GrepServer starts with specified port and spawns threads to keep listening to the request from each connecting Client. The GrepServerThread handles the grep query request locally and returns the result to the client over the client socket. The Client also consists of two class—GrepClient and GrepClientThread. The GrepClient gets the user input and starts the GrepClientThread class. The GrepClientThread take charge of spawning thread to communicate with Server, and storing query results into local files.

Unit Test

The unit test is implemented with the GrepUnitTest. First, we generated the log file by LogfileTest and distribute them on Servers.Then, we did local grep on frequent patterns, infrequent patterns and regular expressions to get query results and stored them into files.At last, we run the unit test with same commands to get the query results from the Distributed Log Querier and compared them with former results.

Query Latency

The graph below shows the Average Query Latency of 4 VMs in the Distributed Log Querier with various query sizes and patterns.