MP1 Design

Our distributed grep system utilizes a client-server architecture to perform parallel grep operations across multiple virtual machines. The algorithm works as follows:

Design

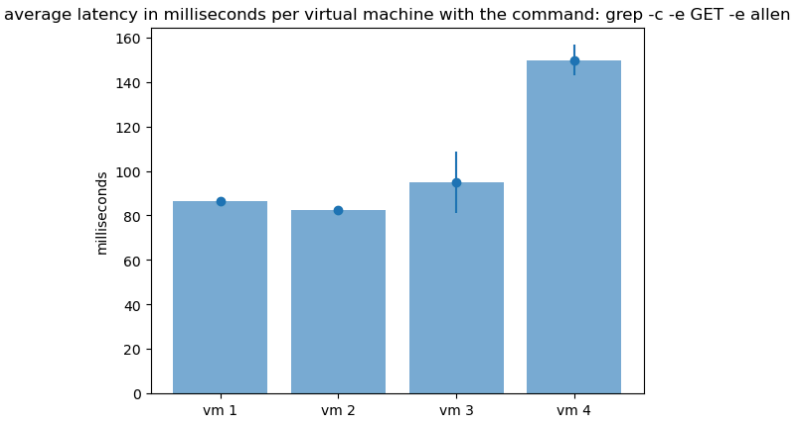
A server is started on every VM that listens to incoming requests from the client.
The client broadcasts grep requests to all VMs simultaneously.
Each VM executes the grep command locally on its log files.
Results are collected asynchronously and displayed as they arrive.
Even to run a grep on the VM where the client is running, we go through the server.
The client can be started on any VM.
There can be multiple clients running at the same time
This approach allows for efficient parallel processing, reducing overall query time.

Tests

TestGrepOnVMs

The function defines three test cases, each with a different pattern to search for:
A common word ("the") with an expected count of 6.
A rare word ("xylophone") with an expected count of 1.
A non-existent word ("qwerty123") with an expected count of 0.

RegexPatternTest



Averaging across five trials, this data indicates that a vm that has more latency on average, has more variance between the trials as with vm 3 compared to vm 1.

Also, this data shows that vm4 is consistently slower than the other three vms

VM1 Average Latency: 86.23896340000002
VM2 Average Latency: 82.2388342
VM3 Average Latency: 94.9741152
VM4 Average Latency: 149.719568