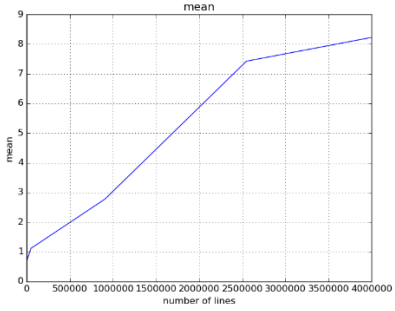
CS 425 MP1 Report

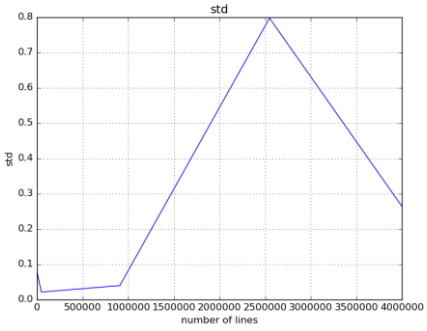
Author: Carl Zhang(hz5), Ruian Pan(ruianp2)

Design: We based the systems on the socket functions in C libraries. Each machine has got two threads (created by pthread), one for server (listen and respond) and one for client (wait for user input and connect). The server thread is an infinite loop which keeps listening for incoming connections, grep locally and send the messages back to the client. The client thread is an infinite loop which, once received a keyboard input, connect to the targets and receive messages sent by the server. For grep functionality, we simply used popen() and system() which takes in the linux command directly and execute locally.

Testing: For our unit test, we generated our own separate log files and grep from them. For the infrequent pattern, we generated logs with 6 lines on each vms. For the frequent pattern, we generated random logs of size 10 MB. We saved the grepped results in the client machine, and compare it with the grepped results using diff. If everything goes fine, the 2 results should match.

Plot: The following is the plot of four machines, with patterns of different frequency: grep 0(4000000), z(2549425), 12(908808),123(49609), zhy(974), and zhzl(28). We plotted the mean time and std deviation versus frequency. We skipped the printing on terminal phase to ignore the time spent on printing onto the screen and focus only on the transmission and grep phase.





We can see that the mean time it takes changes drastically from grep z to grep zhy, since they have much different frequency, which is expected.

There is a big lump in the std plot, partly because of the 5 data points, there is one odd point that influences the std value. Otherwise, we can see that when the frequency gets lower, the standard deviation tends to go down in the big scale.