Distributed Shell Writeup

Ryan Baker - rdbaker

Experiment Design

I needed to figure out how long a network connection would take for a trivial command (on the CPU) and a small amount of data. I ended up sending the "pwd" command 10,000 times to the server, which returned a response of 8 bytes.

I used sysbench to figure out how fast the CPU could run on the client (CCC) and server (EC2) machines. Using "sysbench --test=cpu --cpu-max-prime=20000 run" seemed like a good way to benchmark CPU between the two servers.

I also used sysbench to benchmark how fast the file I/O is on the different servers. I ran the command "sysbench --text=fileio --tele-total-size=16G --file-test-mode=rndrw --init-rng=on --max-time=30 --max-requests=0 run".

Experiment Results

Simple network connection takes on average (over 10,000 trials sending 8 bytes): 0.187ms (1873.7895s for the full trial).

Here are the results of running the aforementioned commands on the server and client:

Command	Server time	Client time
`sysbench --test=cpu --cpu-max-prime=20000 run`	32.8384s	25.4975s
`sysbench --test=fileio --file-total-size=16G --file-test-mode=rndrw --init-rng=on --max-time=30 --max-requests=0 run`	52.052Mb/sec	18.28Mb/sec

If we let server CPU be x, we can show that the client CPU is 22.35% faster. So client CPU is 0.7765 if the server CPU is 1. Likewise, it can be shown that the server is 284.75% faster than the client. So client I/O is 2.8475 if the server CPU is 1.

Experiment Analysis

Local_CPU + Local_File_I/O = n * Network + Remote_CPU + Remote_File_I/O
Assuming we're allowing server (remote) values to be 1, we get the following equation:
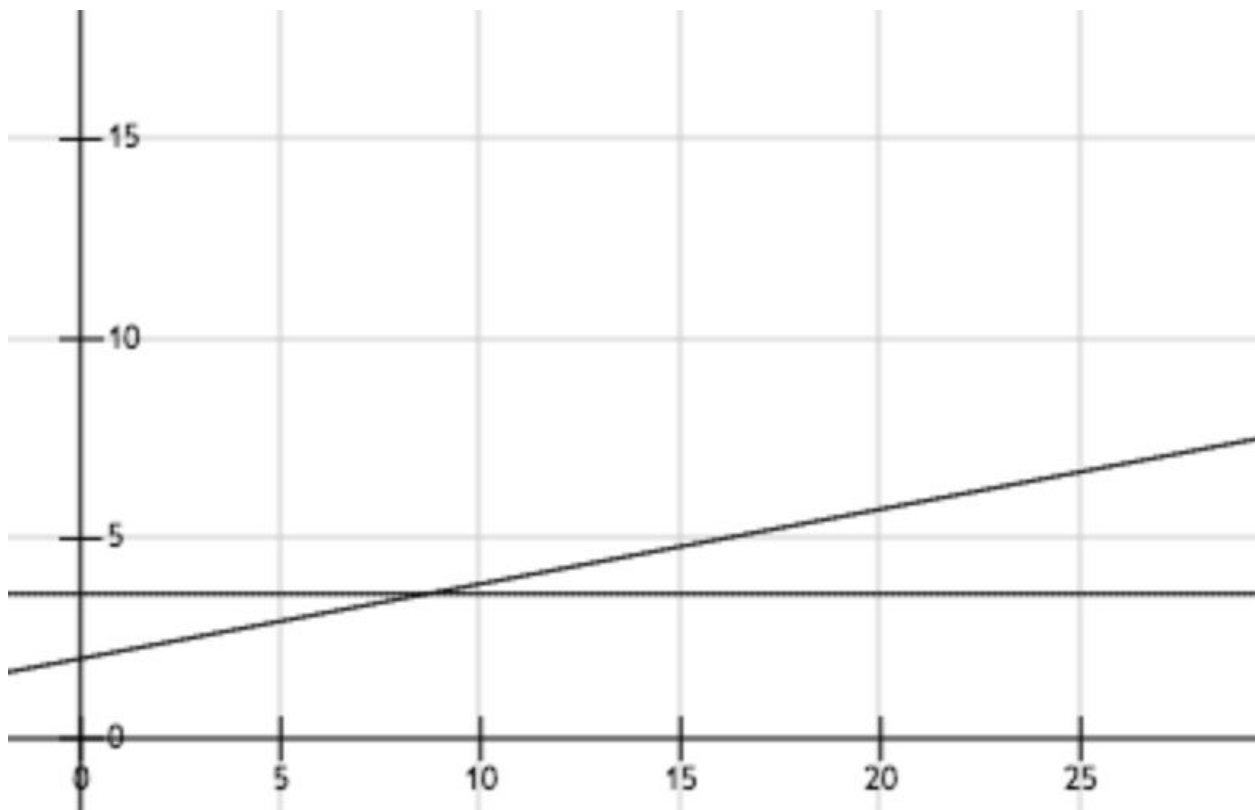0.7765 + 2.8475 = n * 0.187 + 1 + 1

we can simplify this to:
3.624 = 0.187n + 2

then the final result is:
n = 8.68

Giving us the chart:

Where the X-axis is the file size and the Y-axis is the total time. We find that it's less expensive to do the computation on the remote server and transfer the file for anything less than 8.68MB.