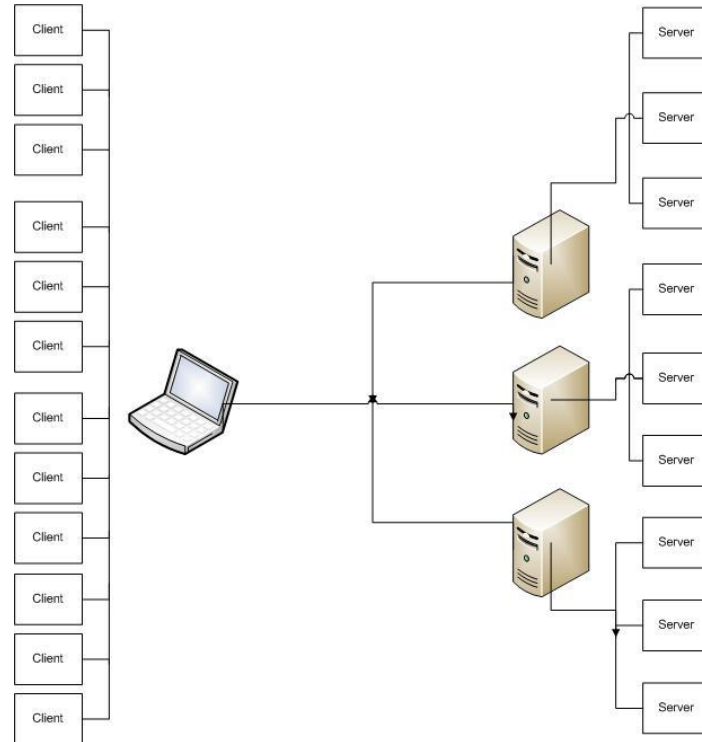# Distributed Systems Report

The Testing Environment

To mimic a real world scenario all tests were performed on the following setup:

Every VPS had been set up with an Ubuntu environment running on a single CPU and 1GB Ram. New servers were initiated on every VPS and would run as a separate process on a separate port. Client processes would be created on the main machine. This could run up to 20 clients in total. The entire environment was itself run on one machine and therefore network transfer speeds were minimal.

Methodology

In every round the Enron data files were recursively read in and a random sample set was divided among each client. Because the time it took in every round to read in all data files we had to limit the number of files to Kaminsky-V folder which consists of 6905 files. Our base random sample set consisted of 2000 record, which were distributed across multiple clients in every round. (i.e. 5 clients – every client has 400 requests). Upon initiating the servers every client went through their list of items to add to the network and the time for a PUT was measured. The same was done for the GET requests however these where chosen randomly off a list of available entries. At the end of the round the average duration for PUT and GET was determined and recorded.

Results

The following where our findings:
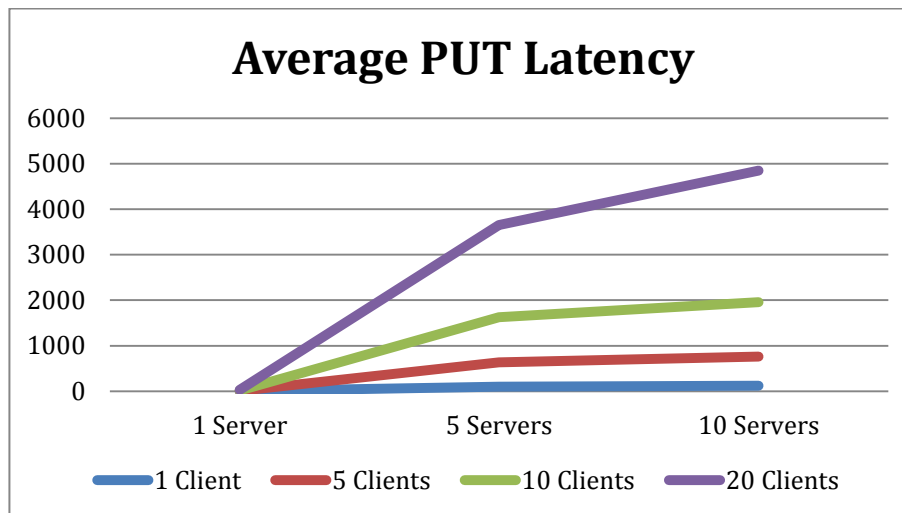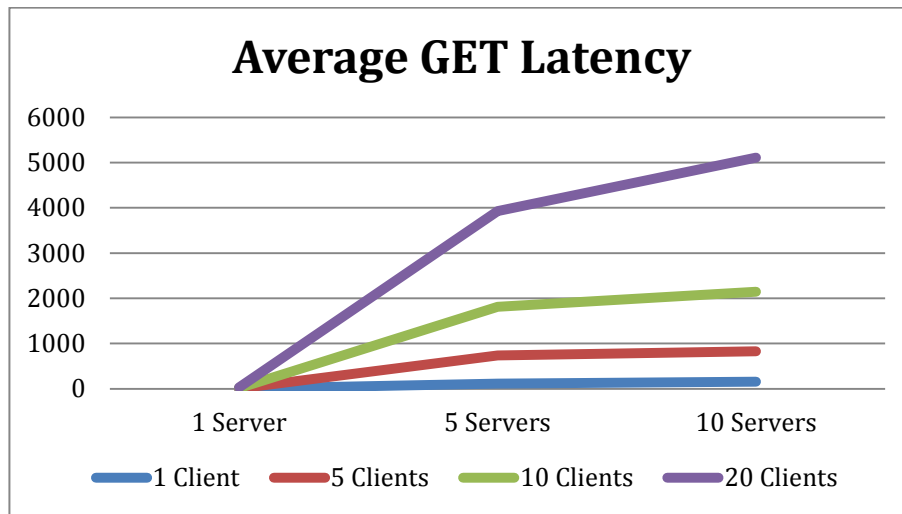
Average PUT & GET latencies increased for the increasing number of clients and servers. The rate of change however (slope) between 1 to 5 servers and 5 to 10 servers indicates a decrease. The throughput of the clients can also be shown as below: