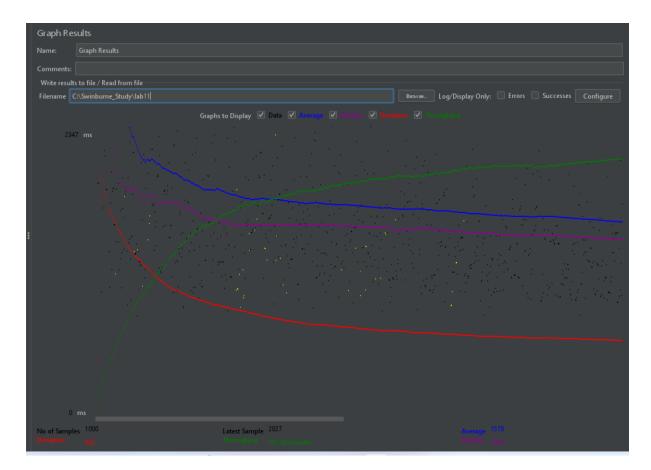


1. Graph Results



2. Test Results Analysis

Throughput represents the number of transactions/requests processed per unit time by the server.

Latency is the unit time taken to process a single transaction/request.

Here, the maximum throughput achieved was **503.88 /minute**, which means an average of **8.38** requests/transactions are processed by the server every second.

As sample load increases, the throughput increases and finally hits a plateau and hits a saturation point after which it cannot increase.

This is the maximum throughput of the server being load tested.

The latency is initially high at about **2347 milli seconds or 2.347 seconds**, but it gradually reduces as the throughput increases and the latency also reaches a saturation point, when the throughput reaches its saturation, latency reaches a minimal latency of **602 milli seconds or 0.6 seconds** when the system reaches maximum throughput.

Therefore, for most systems in general, throughput and latency are inversely proportional. When throughput increases, the latency decreases and vice versa.

But this is not true for all cases, for example, when the system is overloaded or encounters a bottleneck, then in that case, increasing the number of requests per second (throughput) would increase the individual request processing time as well (latency).