ABSTRACT

The increasing use of internet-based services in recent years requires server systems to have optimal load distribution capabilities. The main problem that arises is decreased performance due to server overload. To overcome this, this study compares two load balancing algorithms, namely Least Connection and Weighted Round Robin (WRR), which are implemented using HAProxy on the Google Cloud Platform infrastructure. Testing was carried out using the apache benchmark, with variations in the number of requests from 1000 to 5000 and the number of concurrent connections from 20 to 100. The test results show that the Weighted Round Robin algorithm is able to reduce delays by up to 0.317 ms, while the Least Connection algorithm is able to maintain a higher throughput of 3181.824 KB/s at maximum load. In addition, the jitter in the Weighted Round Robin algorithm is recorded as lower, reaching only 0.001 ms. This study contributes in the form of a performance analysis based on Quality of Service (QoS), which can be used as a reference in selecting a load balancing method for a system based on the Google Cloud Platform.

Keywords*: load balancing, least connection, HAProxy, weighted round robin, Google Cloud Platform*