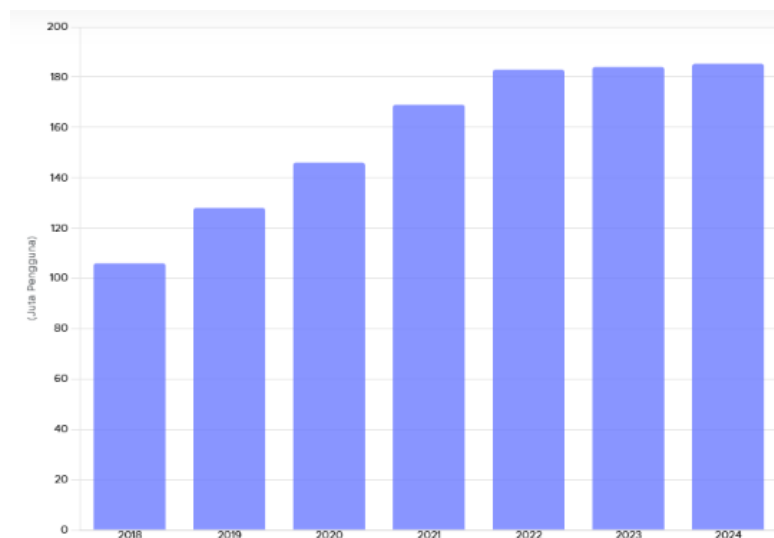


BAB I

PENDAHULUAN

1.1 Latar Belakang

Dewasa ini jumlah pengguna internet mengalami peningkatan yang cukup pesat. Berdasarkan data dari Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), pada tahun 2024 terdapat 221.563.479 jiwa pengguna internet di Indonesia dari total populasi 278.696.200 jiwa penduduk Indonesia. Dari hasil survei yang dirilis APJII tersebut, maka cakupan penggunaan internet Indonesia mencapai angka 79,5%. Hal ini menunjukkan peningkatan 1,4%, dibandingkan dengan periode sebelumnya [1]. Selain itu, menurut goodstat.id, di Indonesia terjadi peningkatan pengguna internet tiap tahunnya dari tahun 2018 hingga tahun 2024. Peningkatan paling signifikan, yakni 24,6%, terjadi pada tahun 2018. [2].



Gambar 1. 1 Pengguna Internet di Indonesia [2]

Berdasarkan Gambar 1. 1 Pengguna Internet di Indonesia tercatat per Januari 2024, sebanyak 185,3 juta orang Indonesia menggunakan internet. Angka ini mengalami peningkatan sebesar 0,8% daripada tahun sebelumnya. Pertumbuhan pengguna internet ini juga terlihat pada beberapa tahun sebelum 2024. Sejak 2018, terjadi lonjakan signifikan sebesar 24,6%, menjadikan total pengguna saat itu sebanyak 106 juta. Peningkatan berlanjut pada tahun 2019, di mana jumlah pengguna internet di Indonesia mencapai 128 juta jiwa. Jumlah tersebut

menunjukkan kenaikan sebesar 20,7% dari tahun sebelumnya. Terdapat 146 juta jiwa papengguna internet Indonesia, tercatat pada tahun 2020. Selain itu, pada tahun 2021 juga mengalamikenaikan yaitu sebesar 16,5%, atau sebanyak 169 juta pengguna internet.

Besarnya jumlah pengguna internet di Indonesia tentu berpengaruh terhadap volume lalu lintas data pada server[3]. Permasalahan yang kerap terjadi adalah kurang optimalnya manajemen server web oleh pengelolanya. Akibatnya, sering terjadi server *downtime* dan kelebihan beban (*overload*) ketika server menerima akses secara bersamaan dari banyak pengguna, terutama jika hanya menggunakan satu server untuk menangani seluruh permintaan. Oleh karena itu, dibutuhkan sebuah sistem server yang dapat menangani banyak akses sekaligus dan dapat di tingkatkan skalabilitasnya sesuai dengan kebutuhan. Sistem yang dapat memenuhi kebutuhan tersebut adalah server *load balancing*.

Load balancing merupakan proses membagi beban lalu lintas jaringan yang masuk ke beberapa server, instant, atau sumber daya untuk memastikan tidak ada satu pun sumber daya yang kelebihan beban (*overload*) atau kewalahan [4]. Dengan menggunakan *load balancing*, situs web dapat tetap online dan stabil walaupun mengalami peningkatan lalu lintas data karena banyaknya pengunjung. Tujuan utama dari *load balancing* adalah untuk mengurangi beban pada server tertentu, meningkatkan ketersediaan layanan, dan memastikan sumber daya server digunakan secara optimal. Dengan pembagian beban secara merata, *load balancing* dapat mencegah server menjadi terlalu sibuk, sehingga kinerja keseluruhan sistem dapat dipertahankan. Jika salah satu server *down* atau tidak tersedia, *load balancer* dapat secara otomatis memindahkan lalu lintas ke server lain. Dengan *load balancing*, sistem dapat dengan mudah menambah atau mengurangi server sesuai kebutuhan, yang memungkinkan skalabilitas yang lebih baik.

Salah satu perangkat lunak *load balancer* yang sering digunakan dan *open-source* adalah HAProxy (*High Availability Proxy*). HAProxy mampu membagi permintaan dari pengguna ke server web menggunakan berbagai strategi algoritma, dua diantaranya yaitu *Least Connection* dan *Weighted Round Robin* (WRR). Algoritma *Least Connection* akan mengarahkan permintaan dari pengguna ke

server dengan jumlah koneksi aktif terendah, dengan harapan mengoptimalkan penggunaan sumber daya yang ada, sedangkan WRR mendistribusikan permintaan berdasarkan bobot (*weight*) yang ditetapkan pada masing-masing server. Kedua algoritma ini memiliki karakteristik berbeda dan dapat mempengaruhi performa layanan secara signifikan. Namun, implementasi algoritma ini memerlukan analisis lebih lanjut untuk memahami bagaimana performansi sistem dapat dipengaruhi oleh penggunaan algoritma ini.

Dalam pengaplikasiannya, *load balancing* dapat dilakukan pada beberapa platform, salah satunya yaitu pada layanan *cloud computing* [5]. Cloud computing merupakan suatu model layanan teknologi informasi yang memberikan kemudahan bagi pengguna dalam mengakses dan memanfaatkan sumber daya seperti penyimpanan data, server, basis data, jaringan, dan perangkat lunak dengan sistem pembayaran berdasarkan penggunaan [6]. *Cloud computing* memungkinkan pengguna untuk mengatur sumber daya IT mereka dengan lebih efisien, elastis, dan hemat biaya. Beberapa penyedia *cloud load balancing* yang terkenal antara lain AWS, Google Cloud Platform, Alibaba Cloud dan Microsoft Azure. Masing-masing memiliki fitur dan opsi yang berbeda untuk menunjang kebutuhan *load balancing* yang berbeda.

Pada penelitian ini menggunakan Google Cloud Platform sebagai media implementasi karena mendukung penyediaan infrastruktur virtual secara cepat dan efisien. Dengan memanfaatkan platform *cloud*, pengujian dapat dilakukan secara lebih realistis dan mendekati skenario produksi. Berdasarkan hal tersebut, dalam penelitian ini diangkat judul “Analisis Performansi Algoritma *Load Balancing Least Connection* dan *Weighted Round Robin* Berbasis Google Cloud Platform”.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan masalah yang akan dijawab dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana kinerja algoritma *load balancing Least Connection* dan *Weighted Round Robin* dalam lingkungan Google Cloud Platform
- b. Algoritma manakah yang memberikan performansi lebih baik berdasarkan parameter *Quality of Service* (delay, throughput dan jitter)?

1.3 Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah menganalisis dan membandingkan performa algoritma *Least Connection* dan *Weighted Round Robin* dalam lingkungan Google Cloud Platform. Penelitian ini memberikan pengetahuan dalam perancangan server *load balancing*, dan mengetahui metode apa yang memiliki kinerja lebih bagus untuk diterapkan dalam *load balancing server*.

1.4 Batasan Masalah Penelitian

Adapun batasan pada penelitian ini antara lain :

- a. Penelitian ini dilakukan menggunakan HAProxy sebagai perangkat lunak *load balancer*.
- b. Algoritma yang dibandingkan terbatas pada *Least Connection* dan *Weighted Round Robin*
- c. Lingkungan pengujian dilakukan pada infrastruktur Google Cloud Platform (GCP)
- d. Parameter performa yang dianalisis hanya mencakup throughput, jitter, dan delay.
- e. Tidak membahas protokol routing
- f. Tidak menangani kelebihan beban kerja

1.5 Metode Penelitian

Penelitian ini dilakukan dengan penelitian eksperimen, metode penelitian yang bertujuan untuk menjelaskan hubungan sebab-akibat (kausalitas) antara satu variabel dengan lainnya (variabel X dan variabel Y) [7]. Pada penelitian ini dilakukan dengan cara menerapkan, menguji, dan menganalisis dari dua algoritma *load balancing*, yakni *least connection* dan *weighted round robin* (WRR). Tahap akhir dari penelitian ini adalah melakukan analisis terhadap hasil evaluasi dan menarik kesimpulan berdasarkan *performa* model, serta memberikan saran untuk pengembangan sistem segmentasi lebih lanjut pada masa mendatang.