

# BAB 1

## PENDAHULUAN

---

### 1.1 Latar Belakang

Kebutuhan koneksi internet yang semakin tinggi diikuti dengan berkembangnya teknologi *web* yang semakin hari semakin canggih, banyak *web framework* baru bermunculan. Proses membuat *web* pun semakin hari semakin mudah dengan banyaknya pilihan bahasa pemrograman yang bisa dipakai, bahkan terdapat proyek yang memungkinkan pengembang untuk menggantikan bahasa pemrograman javascript pada sisi *frontend / client* dengan bahasa pemrograman python dan masih banyak lagi lainnya. Bahkan perangkat lunak yang digunakan untuk mengakses (*web browser*) juga sudah dianggap sebagai OS (*Operating System*) karena kecanggihannya yang dimiliki, apalagi sekarang sudah memasuki era *web 3.0* [1], dimana semua aplikasi menjadi serba *cloud* dengan performa yang mirip dengan *native apps*. [2]

Hal ini tentu menimbulkan masalah baru, dikarenakan resource yang digunakan pada aplikasi web menjadi semakin bertambah berat karena fitur yang sangat kompleks dan teknologi yang digunakan semakin *advanced*. Maka dari itu, dibutuhkan sistem *load balancing* sebagai salah satu solusi dalam mengoptimalkan kinerja server yang digunakan supaya tidak banyak *resource* yang terbuang secara percuma dan aplikasi mendapatkan performa yang maksimal. Dengan begitu, akan dicapai performa yang diinginkan.

Maka dari itu, Central AI sebagai perusahaan layanan *chatbot* berbasis *Artificial Intelligence* melakukan implementasi *load balancing* pada server yang digunakan, dengan tujuan memberikan layanan dan pengalaman pengguna yang lebih baik dan cepat pada saat diakses.

### 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, penulis mendapatkan rumusan masalah yaitu bagaimana sebuah aplikasi dapat berjalan secara optimal pada *server* dan melakukan distribusi sumber daya (*resource*) secara maksimal. Karena semakin optimal kinerja sebuah *server*, maka jumlah *concurrent user* yang bisa di tampung juga semakin besar dan aplikasi akan semakin cepat untuk di akses *user*.

### 1.3 Tujuan

Tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut :

1. Mengimplementasikan *Application Load Balancer* berbasis Nginx dengan algoritma *Round Robin* pada *server* Central AI
2. Mengoptimalkan kinerja aplikasi yang berjalan pada server.
3. Membandingkan jumlah *concurrent user* yang dapat ditampung server sebelum dan sesudah menggunakan sistem *load balancer*.
4. Melakukan analisis performa sistem *load balancing* dalam melakukan *request* dengan parameter *response time* dan *request per second*

### 1.4 Batasan Masalah

Adapun batasan masalah dari sistem yang dirancang adalah sebagai berikut :

1. Algoritma yang digunakan adalah Round Robin
2. Aplikasi *web* menggunakan *framework* Python Flask
3. Sistem menggunakan 3 *server* dengan sistem operasi ubuntu 18.04.6 LTS
4. Menggunakan Nginx sebagai perangkat lunak *web server* sekaligus *Application load balancer*
5. Penelitian berfokus pada jumlah *concurrent user* dan kecepatan respon *server* terhadap *request* aplikasi