

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Cloud computing telah menjadi salah satu teknologi utama yang memungkinkan penyediaan layanan komputasi secara fleksibel, efisien, dan skalabel. Teknologi ini telah merevolusi cara organisasi mengelola dan memanfaatkan sumber daya komputasi, memungkinkan mereka untuk memenuhi kebutuhan yang terus berkembang tanpa investasi besar pada infrastruktur fisik. Namun, salah satu tantangan utama dalam implementasi cloud computing adalah manajemen beban yang efisien, yang secara langsung memengaruhi kinerja, keandalan, dan kualitas layanan yang diberikan. Load balancing, atau teknik pembagian beban, menjadi kunci untuk mengatasi tantangan ini. Proses ini memastikan bahwa beban kerja didistribusikan secara merata di antara server yang tersedia, sehingga mencegah terjadinya kondisi overloading pada satu server sekaligus mengoptimalkan pemanfaatan sumber daya lainnya (Fadida Zanetti Junaedy, 2024). Tanpa manajemen beban yang efektif, beberapa server dapat mengalami tekanan kerja berlebih, sementara server lain tetap tidak aktif atau kurang dimanfaatkan. Ketidakseimbangan ini dapat menyebabkan berbagai masalah, seperti bottleneck pada sistem, peningkatan latensi, penurunan throughput, hingga gangguan operasional yang dapat berdampak pada pengalaman pengguna.

Pada banyak implementasi sistem cloud, pengalokasian sumber daya yang tidak optimal sering kali menjadi penyebab inefisiensi. Ketika server tertentu terbebani secara berlebihan, potensi masalah seperti pemanasan perangkat keras, kegagalan sistem, atau perlambatan proses menjadi lebih tinggi. Selain itu, inefisiensi ini dapat berdampak pada pengeluaran energi yang lebih besar dan biaya operasional yang meningkat, yang bertolak belakang dengan tujuan utama cloud computing, yaitu efisiensi dan fleksibilitas. Oleh karena itu,

pengembangan strategi dan algoritma load balancing yang lebih cerdas dan adaptif, seperti yang berbasis kecerdasan buatan atau pembelajaran mesin, menjadi sangat penting untuk memastikan distribusi beban yang optimal dan mendukung performa sistem cloud secara keseluruhan (Erkamim et al., 2024). Salah satu algoritma yang dapat digunakan untuk mengatasi masalah distribusi beban pada server cloud computing adalah Ant Colony Optimization (ACO). Algoritma ini terinspirasi dari perilaku koloni semut dalam mencari jalur terpendek menuju sumber makanan. Dalam prosesnya, semut-semut saling berkomunikasi melalui feromon yang ditinggalkan di jalur yang mereka lalui. Jalur dengan tingkat feromon yang lebih tinggi menunjukkan efisiensi yang lebih baik, sehingga semut-semut lain cenderung mengikuti jalur tersebut (William et al., 2024). Pendekatan ini menjadikan ACO efektif dalam menemukan solusi optimal pada ruang pencarian yang luas dan kompleks, yang sangat sesuai dengan kebutuhan sistem load balancing dalam cloud computing. Keunggulan ACO terletak pada sifatnya yang adaptif, paralel, dan kemampuannya untuk menghindari solusi lokal optimum, menjadikannya pilihan yang menarik untuk mengatasi tantangan dalam manajemen beban. Dengan mengimplementasikan ACO, beban kerja dapat didistribusikan secara lebih merata di antara server, mengurangi kemungkinan bottleneck, dan meningkatkan efisiensi pemanfaatan sumber daya. Selain itu, ACO mampu beradaptasi dengan perubahan dinamika beban secara real-time, yang menjadi nilai tambah dalam sistem cloud yang bersifat dinamis dan terus berubah (Chandrashekhar et al., 2023).

Namun, meskipun algoritma ACO telah banyak diterapkan dalam berbagai domain seperti routing jaringan, penjadwalan, dan optimisasi transportasi, penelitian yang secara khusus mengintegrasikan ACO untuk load balancing dalam konteks cloud computing masih tergolong terbatas. Sebagian besar penelitian terdahulu lebih terfokus pada penggunaan algoritma heuristik lainnya atau pendekatan berbasis rule-based. Hal ini menciptakan celah penelitian yang signifikan, terutama dalam mengevaluasi bagaimana ACO

dapat dioptimalkan untuk menangani kompleksitas load balancing server yang melibatkan banyak variabel, seperti jumlah pengguna, jenis aplikasi, dan kebutuhan sumber daya yang dinamis. Untuk itu, penelitian ini bertujuan untuk menjawab tantangan tersebut dengan mengimplementasikan dan mengoptimalkan algoritma ACO dalam proses load balancing server pada cloud computing (Mulya et al., 2019). Penelitian ini tidak hanya bertujuan untuk meningkatkan efisiensi distribusi beban, tetapi juga untuk mengurangi latensi, meningkatkan throughput, dan mengoptimalkan kinerja sistem secara keseluruhan. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan solusi berbasis algoritma optimisasi untuk menghadapi tantangan di era cloud computing yang semakin kompleks (Mishra, 2012).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah yang akan dipecahkan dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengimplementasikan algoritma Ant Colony Optimization (ACO) untuk optimisasi load balancing pada server dalam sistem cloud computing?
2. Sejauh mana algoritma ACO dapat meningkatkan efisiensi penggunaan sumber daya dan kinerja sistem load balancing dibandingkan dengan algoritma lain yang umum digunakan?
3. Bagaimana pengaruh optimisasi load balancing terhadap kinerja server dan kualitas layanan pada cloud computing?

## 1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan implementasi algoritma Ant Colony Optimization (ACO) untuk optimisasi load balancing pada server dalam sistem cloud computing.

2. Menganalisis efisiensi penggunaan sumber daya server dan performa sistem setelah penerapan algoritma ACO.
3. Membandingkan kinerja algoritma ACO dengan algoritma load balancing lainnya dalam konteks cloud computing.
4. Menilai dampak optimisasi load balancing terhadap peningkatan kualitas layanan cloud computing.

#### **1.4 Hipotesis**

Berdasarkan rumusan masalah dan tujuan penelitian, hipotesis yang diajukan dalam penelitian ini adalah:

1. Algoritma Ant Colony Optimization (ACO) dapat meningkatkan efisiensi load balancing pada server dalam cloud computing dengan memaksimalkan pemanfaatan sumber daya secara merata.
2. Penerapan algoritma ACO dalam load balancing server dapat menghasilkan kinerja yang lebih baik dibandingkan dengan algoritma load balancing tradisional dalam cloud computing.
3. Optimisasi load balancing yang dilakukan dengan algoritma ACO dapat meningkatkan kualitas layanan, seperti menurunkan latensi dan meningkatkan throughput pada sistem cloud computing.

#### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat yang signifikan bagi berbagai pihak, antara lain:

1. Penelitian ini dapat memperkaya kajian ilmu komputer dan teknologi informasi di Universitas Narotama, khususnya di bidang cloud computing dan algoritma optimisasi. Penelitian ini juga diharapkan dapat memperkuat reputasi universitas dalam pengembangan teknologi informasi.
2. Penelitian ini dapat memberikan kontribusi ilmiah bagi mitra lembaga riset atau perguruan tinggi dalam pengembangan algoritma optimisasi

yang lebih efisien untuk cloud computing, serta dapat membuka peluang untuk kolaborasi riset lebih lanjut.

3. Bagi mahasiswa, penelitian ini memberikan pemahaman yang lebih dalam mengenai penerapan algoritma optimisasi dalam cloud computing. Penelitian ini juga memberikan pengalaman langsung dalam mengembangkan solusi berbasis algoritma cerdas yang dapat diterapkan dalam masalah praktis di dunia industri.

