

Bab 1

Pendahuluan

1.1. Latar Belakang

Semakin berkembangnya dunia teknologi, dimana pertumbuhan data yang sangat cepat dan semakin lama semakin meningkat. Hal ini memungkinkan suatu strategi khusus untuk menyasati cara pengolahan data yang besar. Tetapi infrastruktur dari sebuah media penyimpanan memiliki kendala terutama dalam kecepatan transfer, sehingga membutuhkan waktu yang cukup lama untuk membaca keseluruhan data.

Adakalanya aplikasi yang dibuat membutuhkan komputer dengan sumber daya yang tinggi sebagai lingkungan implementasi dan biasanya harga komputer dengan sumber daya yang tinggi tidaklah murah, sedangkan untuk komputer dengan spesifikasi yang tidak terlalu tinggi akan kurang dalam menangani data yang begitu besar . Dengan demikian, teknologi dalam skala besar yang berhubungan dengan peningkatan kinerja sistem benar-benar di butuhkan

Apache Hadoop menerapkan komputasi paralel dengan menggunakan aplikasi terdistribusi, yang dirancang secara efisien mendistribusikan sejumlah pekerjaan besar dalam mesin cluster. *Apache Hadoop* mengembangkan beberapa Project diantaranya Hadoop Distributed System (HDFS) dan Mapreduce framework yang berbasiskan *java* dan *opensource*. Hadoop Cluster mempunyai kelebihan dari segi ekonomi karena tidak berbayar dan dapat diimplementasikan pada perangkat keras dengan spesifikasi yang tidak terlalu tinggi.

Berdasarkan pemaparan diatas, Karya akhir ini akan dirancang sistem *hive* pada *hadoop* untuk melakukan pengolahan data dan analisis kinerja sistem. Diharapkan dengan bertambahnya jumlah *node* dapat meningkatkan kinerja *cluster* dalam menyelesaikan suatu proses.

1.2. Tujuan

Adapun tujuan dari pembuatan karya akhir ini adalah sebagai berikut.

- a. Menerapkan sistem yang telah berkembang seperti hive ke dalam hadoop

1.3. Batasan Masalah

Batasan masalah dari karya akhir ini adalah sebagai berikut.

- a. Tidak membahas masalah topologi jaringan, *switching*, *signalling*, *routing*, dan *network security*.
- b. Tidak membahas *web programming*.

1.4. Sistematika Penulisan

Bab 1 menjelaskan latar belakang, tujuan, batasan masalah dan sistematika penulisan.

Bab 2 menjelaskan arsitektur sistem.

Bab 3 menjelaskan pembuatan simulasi.

Bab 4 menjelaskan penggunaan simulasi.

Bab 5 berisi penutup yang terdiri dari hambatan yang dialami dan saran pengembangan.

