

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada awalnya Big Data adalah sebuah teknologi yang diperkenalkan untuk menanggulangi perkembangan data dan informasi yang semakin pesat, makin bertambahnya pengguna perangkat mobile dan internet di seluruh dunia sangat mempengaruhi perkembangan volume dan jenis data yang terus bertambah.

Kemunculan Big Data tersebut dipandang sebagai solusi dari pertumbuhan data yang sangat signifikan dari waktu ke waktu yang terus melampaui batas kemampuan media penyimpanan maupun sistem database yang ada saat ini. Big Data dapat diasumsikan sebagai sebuah media penyimpanan data yang menawarkan ruang tak terbatas, serta kemampuan untuk mengakodasi dan memproses berbagai jenis data dengan sangat cepat.

Salah satu aplikasi yang menerapkan Big Data tersebut adalah Hadoop, eksistensi Hadoop sudah diakui oleh perusahaan-perusahaan besar di dunia bahkan menggunakan Hadoop sebagai fondasi dasar aplikasi mereka, seperti Microsoft Azure, Google, Facebook, Twitter dan lainnya.

Sumber referensi yang masih sangat minim dan didukung dengan fakta perusahaan besar sudah mulai menggunakan Big Data tersebut, maka akan dilakukan pengujian terhadap pernyataan pembuat aplikasi Hadoop untuk membuktikan kehandalan aplikasi tersebut dalam menangani data. Diperlukan percobaan untuk mengetahui apakah data dapat dikirim melalui perangkat jaringan lalu mengukur kecepatan dalam pengiriman dan penerimaan data, kemudian ketika hal-hal buruk terjadi beberapa kemungkinan dapat terjadi seperti data tiba-tiba corrupt ataupun hilang.

1.2 Rumusan Masalah

Rumusan masalah yang didapatkan adalah sebagai berikut :

1. Bagaimana pengaruh ukuran besar *file* terhadap waktu saat perpindahan *data*?
2. Bagaimana karakteristik penyebaran data dari *master* menuju *slave*?
3. Bagaimana keterkaitan konfigurasi terhadap *block data* yang dikirim ke *slave*?

1.3 Tujuan Pembahasan

Tujuan pembahasan yang didapat dari rumusan masalah adalah sebagai berikut :

1. Mengamati pengaruh ukuran besar file terhadap waktu pengiriman ke *slave*.
2. Mengamati karakteristik pengaturan penyebaran file saat diberi gangguan.
3. Mengamati dan mengukur besar *block data* yang dikirim ke *slave* berdasar konfigurasi pada *server*.

1.4 Ruang Lingkup

Agar tidak menyimpang dari tujuan yang semula direncanakan, maka ditetapkan batasan-batasan sebagai berikut :

1. Perancangan Hadoop dari instalasi, *setting* hingga percobaan data hanya dilakukan pada *server* dengan metode *multi-node*.
2. Percobaan kepada *slave* akan dilakukan saat *slave* sudah siap untuk digunakan.
3. Jaringan menggunakan local *LAN* yang hanya terkoneksi pada lab internet 1 Universitas Kristen Maranatha.
4. PC menggunakan *processor* Intel Core i5-2320 3.00GHz, dan RAM 4,00GB di lab internet 1 dan dilakukan pada jaringan *local* pada Universitas Kristen Maranatha.
5. Aplikasi Hadoop akan digunakan pada infrastruktur Ubuntu versi 14.04 LTS(Trusty Tahr) dan berjalan pada Virtual Box Versi 5.0.0 dengan spesifikasi *processor one core*, *ram* 768MB.
6. Data yang digunakan untuk percobaan terbatas pada data yang berukuran antara 800MB(*MegaByte*) hingga 2GB(*GigaByte*) dan dapat berbentuk text, foto maupun video.
7. Konfigurasi block size dilakukan pada tiga size berbeda 64MB, 128MB, dan 256MB.

1.5 Metodologi Penelitian

Penelitian yang dilakukan dimulai dari pengumpulan informasi tentang Hadoop maupun *Big Data*, lalu akan dilakukan percobaan aplikasi tersebut

dengan minimal 3 PC(*Personal Computer*) dengan spesifikasi yang sama, sehingga pada akhirnya dapat ditarik kesimpulan seberapa cepat dan beban terhadap *processor* maupun *ram*.

1.6 Sistematika Laporan

Dalam proses penyusunan laporan, sistematika penulisan yang akan digunakan adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini dijelaskan latar belakang, rumusan masalah, tujuan pembahasan, batasan masalah, metodologi penelitian dan sistematika laporan pada penelitian ini.

BAB II DASAR TEORI

Bab ini akan dijelaskan teori-teori yang menunjang percobaan yang dilakukan.

BAB III ANALISA DAN PEMODELAN

Bab ini akan dijelaskan mengenai rancangan topologi dari sistem Hadoop yang dibuat beserta detail skenario yang hendak dilakukan.

BAB IV HASIL IMPLEMENTASI

Bab ini menjelaskan konfigurasi yang dibutuhkan sistem Hadoop.

BAB V PENGUJIAN

Bab ini menjelaskan hasil pengujian dan analisis untuk seluruh skenario yang terdapat pada bab 3.

BAB VI SIMPULAN DAN SARAN

Bab ini menjelaskan simpulan dan saran penelitian sistem Hadoop, yang dapat digunakan untuk penelitian lebih lanjut.