

ABSTRAK

Big Data dianggap sebagai solusi untuk pertumbuhan data yang sangat pesat, salah satu aplikasi yang menerapkan hal ini adalah Hadoop. Maka akan dilakukan pengujian terhadap Hadoop untuk membuktikan apakah Hadoop adalah solusi yang tepat, dengan melakukan percobaan pada 1 *server master* dan 2 *server slave* dengan jaringan *local* dan dilakukan pada *server* Linux Ubuntu yang berjalan pada Virtual Box. Lalu akan dilakukan percobaan stabilitas Hadoop dan waktu *transfer* data pada saat kondisi stabil maupun diberi gangguan yang berulang dengan menggunakan 3 *file* dengan ukuran 800MB, 1.2GB dan 2GB. Selanjutnya akan dilakukan pengukuran kecepatan penyebaran data berupa lamanya waktu yang dibutuhkan dan beban pada *processor* dan *ram*, sehingga akan didapatkan karakteristik dari Hadoop tersebut. Hasil yang didapat berupa data yang disebar merupakan *block data* yang memiliki ukuran *default* 128MB per block dan *META* data dan saat data gagal dikirim karena gangguan akan dikirim ulang beberapa saat kemudian, beban *processor* dan *ram* yang didapat pun beragam. Dengan adanya analisis terhadap Hadoop diharapkan dapat menjadi solusi di masa depan untuk menggantikan teknologi yang ada saat ini.

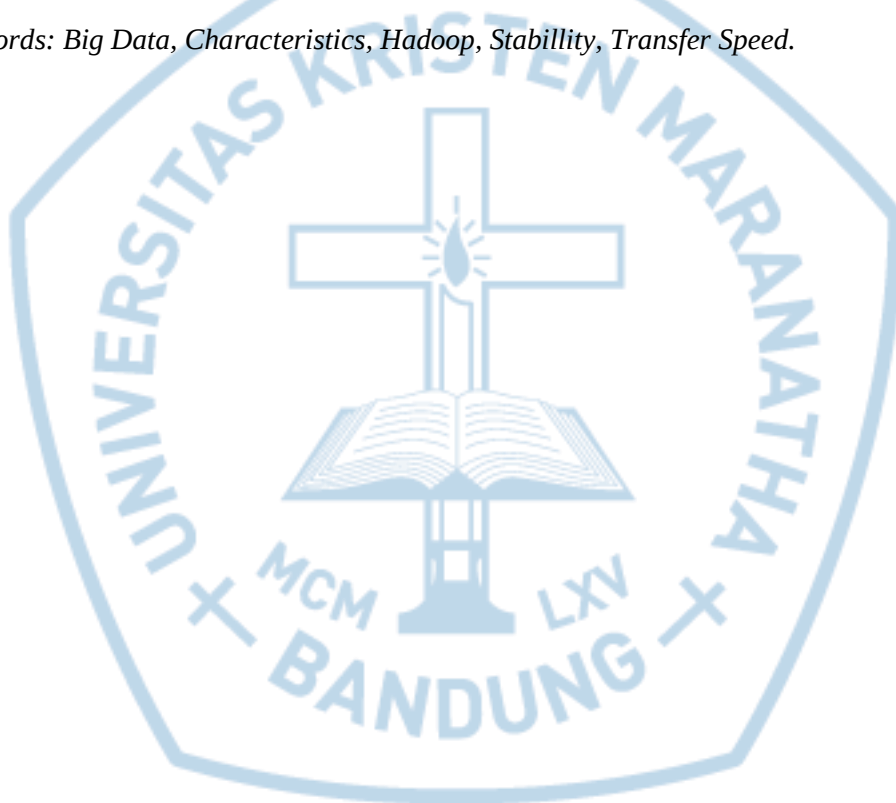
Kata kunci: *Big Data*, Hadoop, Karakteristik, Kecepatan Transfer, Stabilitas.



ABSTRACT

Big Data has been considered to be a solution for rapid data growth. One of Big Data applications is Hadoop. This final project is to test Hadoop is the right solution, the experiments is attempt on one master server and two slave server with local network and running on Ubuntu Linux server on a Virtual Box. Then the test conducted stability of Hadoop and data transfer time when in a stable condition and were given repeated interference by using 3 files with a size of 800MB, 1.2MB and 2GB. Next will be the measurement of the speed time of the data spread and the load on the processor and ram, so will obtain the characteristics of Hadoop. Results of distributed data is the default data block size of 128MB per block and META data and when the data failed to post because the interference will be reissued a few moments later, the load on processor and ram obtained also varied. With the analysis of Hadoop is expected to be a solution in the future to replace the technology that exists today.

Keywords: Big Data, Characteristics, Hadoop, Stabillity, Transfer Speed.



DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS LAPORAN PENELITIAN.....	ii
PERNYATAAN PUBLIKASI LAPORAN PENELITIAN	iii
PRAKATA.....	iv
ABSTRAK	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan Pembahasan	2
1.4 Ruang Lingkup.....	2
1.5 Metodologi Penelitian.....	2
1.6 Sistematika Laporan.....	3
BAB 2 DASAR TEORI.....	4
2.1 <i>Big Data</i>	4
2.2 Hadoop <i>Single-Node</i>	6
2.3 Hadoop <i>Multi-Node</i>	7
2.4 Map Reduce	8
2.5 Hadoop Distributed <i>File System</i> (HDFS)	8
2.6 OpenSSH.....	10
2.7 Cacti	11
2.8 Linux Ubuntu	11

2.9 Algoritama <i>FIFO (First In First Out)</i>	11
2.10 Penelitian Terkait	12
BAB 3 ANALISA DAN PEMODELAN.....	15
3.1 Perancangan	15
3.2 Skenario Percobaan.....	15
BAB 4 IMPLEMENTASI.....	18
4.1 <i>Add Lookup Library</i>	18
4.2 Konfigurasi SSH	18
4.3 Konfigurasi Hadoop.....	19
BAB 5 PENGUJIAN	26
5.1 Pengujian Skenario Pertama	26
5.2 Pengujian Skenario Kedua	27
5.3 Pengujian Skenario Ketiga.....	29
BAB 6 SIMPULAN DAN SARAN.....	32
6.1 Simpulan	32
6.2 Saran.....	32
DAFTAR PUSTAKA	33
LAMPIRAN A DATA PERCOBAAN.....	A-1
LAMPIRAN B INSTALASI	B-1

DAFTAR GAMBAR

Gambar 2.1 Definisi <i>Big Data</i> menurut IBM.....	4
Gambar 2.2 Gambaran 3V <i>Big Data</i>	5
Gambar 2.3 Gambaran 4 Dimensi <i>Big Data</i>	6
Gambar 2.4 Arsitektur Hadoop Single-Node [3]	7
Gambar 2.5 Arsitektur Hadoop <i>Multi-Node</i> [3].....	7
Gambar 2.6 Proses Penyerahan <i>Job</i> dari <i>Client</i> ke <i>MapReduce</i>	8
Gambar 2.7 Komponen <i>HDFS</i>	9
Gambar 2.8 <i>NameNode</i> pada <i>HDFS</i>	9
Gambar 2.9 Interaksi antara <i>Datanode</i> dan <i>NameNode</i>	10
Gambar 2.10 Ilustrasi <i>FIFO</i>	12
Gambar 2.11 Data dan Hasil Rata-Rata Penelitian UI.....	13
Gambar 2.12 Hasil Penelitian Universitas Kristen Petra	14
Gambar 3.1 Rancangan Percobaan	15
Gambar 3.2 Ilustrasi <i>Timing</i>	16
Gambar 4.1 Konfigurasi <i>nsswitch.conf</i>	18
Gambar 4.2 Konfigurasi <i>.bashrc</i>	19
Gambar 4.3 Konfigurasi <i>Hadoop-env.sh</i>	20
Gambar 4.4 Konfigurasi <i>core-site.xml</i>	21
Gambar 4.5 Konfigurasi <i>mapred-site.xml</i>	22
Gambar 4.6 Konfigurasi <i>hdfs-site.xml</i>	23
Gambar 4.7 Konfigurasi <i>yarn-site.xml</i>	24
Gambar 4.8 Service Pada Master	25
Gambar 5.1 Log pada Server Master	26
Gambar 5.2 Data dan Hasil Rata-Rata Pemrosesan Data Dengan <i>Block Size</i> 128 MB	27
Gambar 5.3 Data dan Hasil Rata-Rata Saat Diberi Gangguan.....	28
Gambar 5.4 Perintah Mengubah Ukuran <i>Block Size</i>	30
Gambar 5.5 Data Hasil Rata-Rata dengan <i>Block Size</i> Berbeda	30