

ABSTRAK

Seiring dengan perkembangan data yang sangat pesat, cara – cara penyimpanan data dalam *server* telah banyak dikembangkan dan disediakan dalam media yang berbeda. Diantaranya ada Hadoop yang merupakan software framework untuk *server*. Demi menggali pengetahuan lebih dari *server* maka saya melakukan penelitian yang berjudul karakteristik Hadoop Multi node pada *master server* terhadap data Kecil. Hadoop sendiri merupakan project opensource dari apache yang sudah berjalan kurang lebih 9 tahun, yang dimana 2 tahun setelah Hadoop lahir dipakai google untuk prosesing data. Hasil yang didapat memuat karakteristik dari Hadoop, apakah *HDFS* menjadi tempat penyimpanan yang baik untuk data. Metode yang akan dipakai akan banyak mengambil dan menaruh data dalam 5 mesin yang sudah menggunakan Hadoop. Mesin yang terbagi dalam 1 *master server* dan 4 *slave server* akan diberikan data dan diukur dengan melihat waktu dari log masing – masing *server*. Data yang didapat diolah menjadi grafik agar bisa terlihat apakah terjadi anomali yang bisa dibilang unik untuk Hadoop?. Akhir dari pengukuran data diharapkan mendapat kesimpulan apa yang menjadi keunikan dari Hadoop dan bagaimana data diproses dalam *HDFS*?

Kata Kunci : Hadoop, data, *file*, *HDFS*



ABSTRACT

The growth of data over years increase very rapidly. Because of this growth, the data volume that stored in the server has been increase and in different media type. Hadoop is one of framework to handle big data on server. Hadoop itself is an open source project which developed by Apache and has been going for approximately nine years. This experiments have an objective to look at how Hadoop work to handle data less than 500MB. This experiment used five servers with Hadoop installed on each server (one as a master and the other as slave servers). Three experiment has been conducted, each experiment gave its own conclusion. Experiment one showed if a node on web UI gave 500 ms from last contact then the node declared to be dead. The second experiment showed the transfer rate 61.39 ms per MB. The last experiment showed if a new node is added as a slave, the Hadoop server's capacity increased size, but the new slave didn't get any data from other slave.

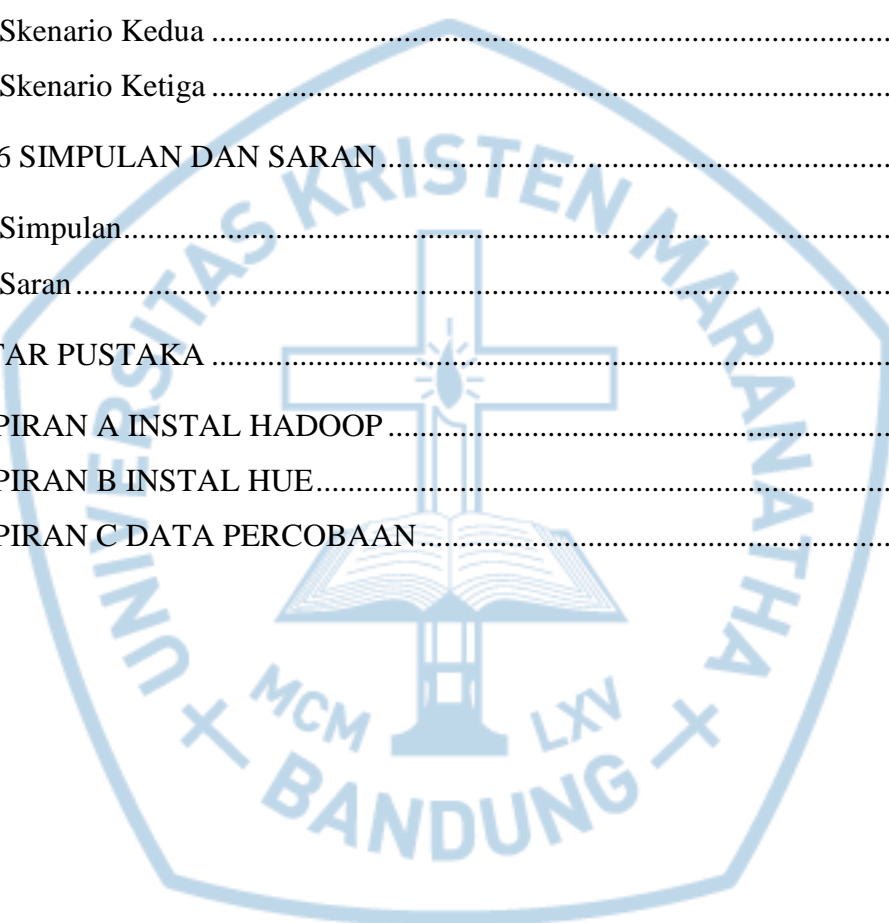
Keywords : Hadoop, data, *file*, *HDFS*



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	2
1.3 Tujuan.....	2
1.4 Ruang Lingkup Penelitian	2
1.5 Sumber Data	2
1.6 Sistematika Penyajian.....	3
BAB 2 KAJIAN TEORI	4
2.1 Hadoop	4
2.2 Hadoop Single–Node	7
2.3 Hadoop Multi–Node.....	8
BAB 3 ANALISIS DAN RANCANGAN SISTEM.....	9
3.1 Pemodelan	9
3.2 Skenario.....	10
3.2.1 Skenario Pertama	10
3.2.2 Skenario Kedua.....	10
3.2.3 Skenario Ketiga.....	11

BAB 4 KONFIGURASI PENGUJIAN	13
4.1 SSH.....	13
4.2 Host dan hostname	13
4.3 Hadoop	16
4.4 Hue	18
BAB 5 PENGUJIAN	19
5.1 Skenario Pertama.....	19
5.2 Skenario Kedua	20
5.3 Skenario Ketiga	26
BAB 6 SIMPULAN DAN SARAN.....	27
6.1 Simpulan.....	27
6.2 Saran	27
DAFTAR PUSTAKA	28
LAMPIRAN A INSTAL HADOOP	A-1
LAMPIRAN B INSTAL HUE.....	B-1
LAMPIRAN C DATA PERCOBAAN.....	C-1



DAFTAR GAMBAR

Gambar 2.1 : Hadoop Core / Hadoop Common.....	4
Gambar 2.2 : Cara Kerja <i>HDFS</i>	5
Gambar 2.3 : Data Penelitian Terkait.....	7
Gambar 2.4 : Contoh <i>HDFS</i> Architectue Single Node	8
Gambar 2.5 : Contoh <i>HDFS</i> Architectue Multi Node.....	8
Gambar 3.1 : Analysis Modeling Example	9
Gambar 3.2 : Skenario 1.....	10
Gambar 3.3 : Skenario 2.....	11
Gambar 3.4 : Skenario 3.....	12
Gambar 4.1 : Hostname.....	13
Gambar 4.2 : Hosts.....	14
Gambar 4.3 : Hadoop config folder	14
Gambar 4.4 : core-site.xml.....	14
Gambar 4.5 : yarn-site.xml	15
Gambar 4.6 : <i>HDFS</i> -site.xml	15
Gambar 4.7 : masters.....	16
Gambar 4.8 : slaves	16
Gambar 4.9 : Masuk Ke <i>File</i> bashrc	16
Gambar 4.10 : <i>File</i> .bashrc.....	17
Gambar 4.11 : <i>File</i> hadoop-env.sh.....	17
Gambar 4.12 : <i>File</i> hue.ini	18
Gambar 5.1 : Hadooslave2 <i>Last Contact</i>	19
Gambar 5.2 : Hadooslave2 Setelah <i>Last Contact</i> melebihi 500	19
Gambar 5.3 : Hasil <i>Summry</i> Menyatakan Jumlah <i>Live</i> dan <i>Dead</i> Node.....	19
Gambar 5.4 : Keadaan Node Setelah Mesin Kembali Di Nyalakan	20
Gambar 5.5 : <i>Summary</i> Node	20
Gambar 5.6 : <i>File</i> Information Dari Port 50070	21
Gambar 5.7 : Namenode Log	21
Gambar 5.8 : Penyimpanan Data Di <i>Master</i> Server.....	22
Gambar 5.9 : <i>File</i> Information Dan Folder Tempat <i>File</i> Disimpan.....	22
Gambar 5.10 : Grafik Durasi Percobaan 1	23

Gambar 5.11 : Hasil <i>Traffic</i> Percobaan 1 [127.00 MB].....	23
Gambar 5.12 : Grafik Durasi Percobaan 2	24
Gambar 5.13 : Hasil <i>Traffic</i> Percobaan 2 [235.16 MB].....	24
Gambar 5.14 : Grafik Durasi Percobaan 3	25
Gambar 5.15 : Hasil <i>Traffic</i> Percobaan 3 [314.00 MB].....	25
Gambar 5.16 : Edit <i>File Hosts</i>	26
Gambar 5.17 : Edit <i>File Slaves</i>	26

