

ABSTRAK

Seiring dengan berkembangnya teknologi yang ada, maka ukuran data yang diolah juga akan semakin besar. Hadoop adalah salah satu kerangka kerja perangkat lunak untuk server yang menangani data besar. Hadoop itu sendiri adalah sebuah proyek open source berkembang dengan Apache, yang telah berlangsung selama kurang lebih sembilan tahun, di mana dua tahun setelahnya, Hadoop digunakan oleh Google untuk pengolahan data. Metode dalam penelitian ini akan menggunakan lima server dengan Hadoop terinstal, satu *master* dan empat server *slave*. Dalam penelitian ini akan merumuskan beberapa masalah yang akan diamati. Sehingga akan dilakukan beberapa percobaan untuk melihat karakteristik Hadoop. Penelitian ini akan dilakukan tiga percobaan, mencabut *master client* dan melihat keadaan *client* server, melakukan pengukuran penerimaan data pada *client* server dan penambahan *client* server. Lalu dari percobaan yang dilakukan, data akan dikumpulkan dan dibuat grafik. Sehingga dapat ditarik kesimpulan dari percobaan tersebut.

Kata kunci: Hadoop, Apache, Google, *Master Server*, *Client Server*, log.

ABSTRACT

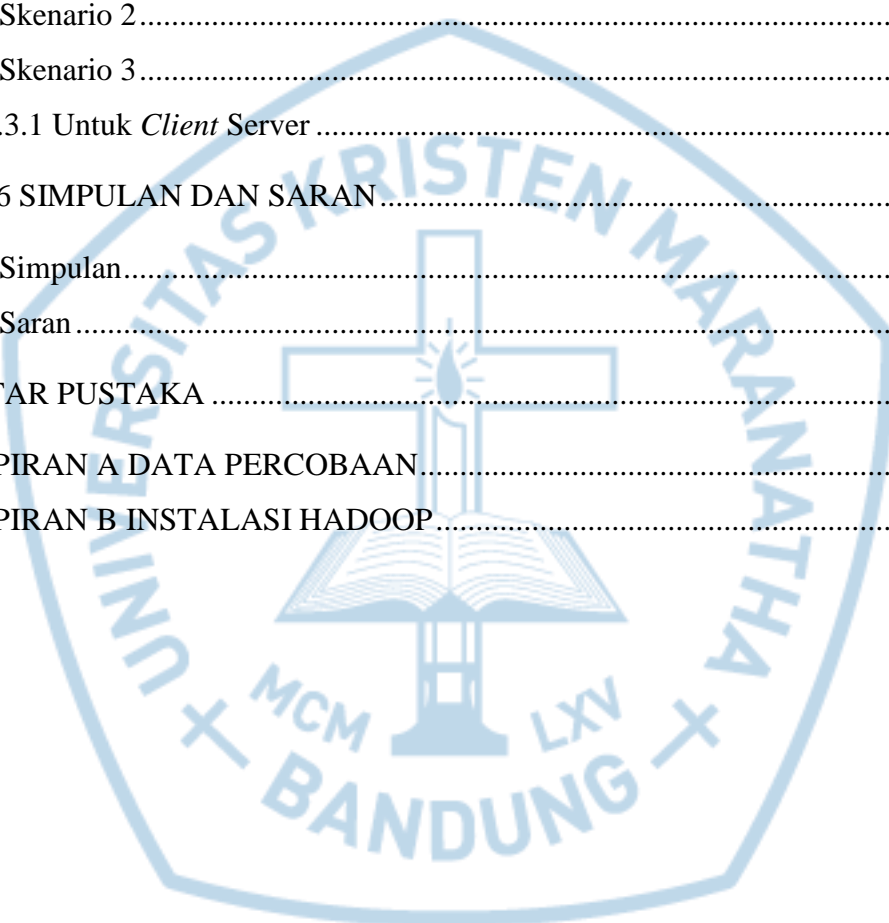
This point is the era where technology highly evolved quickly and rapidly. Along with the development of anything, so the size of the data process will also bigger. Hadoop is the one of the framework for server software that handles data large. Hadoop itself is a project open source developing by Apache. Who has run for nine years, where two years later, Hadoop used by Google to data processing. Methods in this research will use of five server with Hadoop installed, one a master and four server slave. In this research will formulate a few problems to be observed. So that will be carried out several experiments to see the characteristics of Hadoop. This research will be conducted three experiments. Deprive master client and assess the situation client server, take measurements of the reception client data on servers and the addition of client server. Then from experiments conducted, data will be collected and made a chart. So that conclusions may be drawn from these experiments.

Keywords: Hadoop, Apache, Google, Master Server, Client Server, log.

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 Metodologi Penelitian	3
1.6 Sistematika Laporan	3
BAB 2 KAJIAN TEORI	5
2.1 Hadoop	5
2.2 Hadoop Single-Node	8
2.3 Hadoop Multi-Node	9
2.4 Penelitian Terkait	10
BAB 3 ANALISIS DAN RANCANGAN SISTEM	11
3.1 Pemodelan	11
3.2 Skenario	11
3.2.1 Skenario 1	12
3.2.2 Skenario 2	13
3.2.3 Skenario 3	14
BAB 4 IMPLEMENTASI	Error! Bookmark not defined.

4.1 Konfigurasi <i>Client</i> Server Hadoop	16
4.1.1 Host dan hostname	16
4.2 Hadoop	19
4.3 SSH(<i>Secure Shell</i>)	21
4.4 Hue	21
BAB 5 PENGUJIAN	22
5.1 Skenario 1	22
5.2 Skenario 2	23
5.3 Skenario 3	40
5.3.1 Untuk <i>Client</i> Server	40
BAB 6 SIMPULAN DAN SARAN	43
6.1 Simpulan	43
6.2 Saran	44
DAFTAR PUSTAKA	45
LAMPIRAN A DATA PERCOBAAN	A-1
LAMPIRAN B INSTALASI HADOOP	B-1



DAFTAR GAMBAR

GAMBAR 2.1 ECOSYSTEM YANG TERDAPAT PADA HADOOP [4].....	5
GAMBAR 2.2 CARA KERJA HDFS	6
GAMBAR 2.3 SECURE SHELL (SSH) [6].....	7
GAMBAR 2.4 ARSITEKTUR HDFS SINGLE-NODE	8
GAMBAR 2.5 ARSITEKTUR HDFS MULTI-NODE	9
GAMBAR 2.6 DATA PENELITIAN TERKAIT.....	10
GAMBAR 3.1 ANALYSIS MODELING EXAMPLE.....	11
GAMBAR 3.2 SKENARIO 1	12
GAMBAR 3.3 SKENARIO 2	13
GAMBAR 3.4 SKENARIO 3	15
GAMBAR 4.1 HOSTNAME	16
GAMBAR 4.2 HOSTS	17
GAMBAR 4.3 HADOOP CONFIG FOLDER	17
GAMBAR 4.4 CORE-SITE.XML	18
GAMBAR 4.5 YARN-SITE.XML.....	19
GAMBAR 4.6 HDFS-SITE.XML	19
GAMBAR 4.7 MASUK KE BASHRC	20
GAMBAR 4.8 FILE .BASHRC	20
GAMBAR 4.9 FILE HADOOP-ENV.SH.....	20
GAMBAR 4.10 FILE HUE.INI.....	21
GAMBAR 5.1 JPS DARI MASTER SERVER.....	22
GAMBAR 5.2 MASTER SERVER TELAH DI REBOOT.....	22
GAMBAR 5.3 SERVER HADOOP ONLINE KEMBALI.....	23
GAMBAR 5.4 FILE INFORMATION DARI PORT 50070	24
GAMBAR 5.5 DATA NODE LOGS.....	24
GAMBAR 5.6 FILE INFORMATION DAN FOLDER TEMPAT FILE DISIMPAN	25
GAMBAR 5.7 GRAFIK TOTAL DURASI PADA SLAVE 1 DENGAN MENGGUNAKAN REGRESI LINEAR	26
GAMBAR 5.8 HASIL TRAFFIC CACTI PERCOBAAN 1 PADA SLAVE 1 [127 MB].....	26
GAMBAR 5.9 GRAFIK TOTAL DURASI PADA SLAVE 2 DENGAN MENGGUNAKAN REGRESI LINEAR	27
GAMBAR 5.10 HASIL TRAFFIC CACTI PERCOBAAN 1 PADA SLAVE 2 [127 MB].....	27
GAMBAR 5.11 GRAFIK TOTAL DURASI PADA SLAVE 3 DENGAN MENGGUNAKAN REGRESI LINEAR	28
GAMBAR 5.12 HASIL TRAFFIC CACTI PERCOBAAN 1 PADA SLAVE 3 [127 MB].....	28
GAMBAR 5.13 GRAFIK TOTAL DURASI PADA SLAVE 4 DENGAN MENGGUNAKAN REGRESI LINEAR	29
GAMBAR 5.14 HASIL TRAFFIC CACTI PERCOBAAN 1 PADA SLAVE 4 [127 MB].....	29
GAMBAR 5.15 RATA-RATA PER SLAVE PERCOBAAN 1 [127 MB] DENGAN MENGGUNAKAN REGRESI	

LINEAR	30
GAMBAR 5.16 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 1 DENGAN MENGGUNAKAN REGRESI LINEAR	31
GAMBAR 5.17 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 2 PADA <i>SLAVE</i> 1 [235 MB].....	31
GAMBAR 5.18 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 2 DENGAN MENGGUNAKAN REGRESI LINEAR	32
GAMBAR 5.19 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 2 PADA <i>SLAVE</i> 2 [235 MB].....	32
GAMBAR 5.20 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 3 DENGAN MENGGUNAKAN REGRESI LINEAR	33
GAMBAR 5.21 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 2 PADA <i>SLAVE</i> 3 [235 MB].....	33
GAMBAR 5.22 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 4 DENGAN MENGGUNAKAN REGRESI LINEAR	34
GAMBAR 5.23 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 2 PADA <i>SLAVE</i> 4 [235 MB].....	34
GAMBAR 5.24 RATA-RATA PER <i>SLAVE</i> PERCOBAAN 2 [235 MB] DENGAN MENGGUNAKAN REGRESI LINEAR	35
GAMBAR 5.25 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 1 DENGAN MENGGUNAKAN REGRESI LINEAR	35
GAMBAR 5.26 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 3 PADA <i>SLAVE</i> 1 [314 MB].....	36
GAMBAR 5.27 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 2 DENGAN MENGGUNAKAN REGRESI LINEAR	36
GAMBAR 5.28 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 3 PADA <i>SLAVE</i> 2 [314 MB].....	37
GAMBAR 5.29 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 3 DENGAN MENGGUNAKAN REGRESI LINEAR	37
GAMBAR 5.30 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 3 PADA <i>SLAVE</i> 3 [314 MB].....	38
GAMBAR 5.31 GRAFIK TOTAL DURASI PADA <i>SLAVE</i> 4 DENGAN MENGGUNAKAN REGRESI LINEAR	38
GAMBAR 5.32 HASIL <i>TRAFFIC</i> CACTI PERCOBAAN 3 PADA <i>SLAVE</i> 4 [314 MB].....	39
GAMBAR 5.33 RATA-RATA PER <i>SLAVE</i> PERCOBAAN 3 [314 MB] DENGAN MENGGUNAKAN REGRESI LINEAR	39
GAMBAR 5.34 DATA DAN HASIL RATA-RATA PEMROSESAN DATA DENGAN <i>BLOCK SIZE</i> 128 MB. 40	
GAMBAR 5.35 CLONE MESIN	41
GAMBAR 5.36 TYPE CLONE	41
GAMBAR 5.37 EDIT <i>FILE</i> HOST.....	41
GAMBAR 5.38 EDIT HOSTNAME.....	42
GAMBAR 5.39 EDIT <i>FILE</i> INTERFACES.....	42
GAMBAR 5.40 HAPUS DAN BUAT KEMBALI FOLDER <i>DATANODE</i>	42
GAMBAR 5.41 <i>FORMAT</i> <i>DATANODE</i>	42