

## DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN.....	iv
KATA PENGANTAR.....	v
DAFTAR ISI .....	vi
DAFTAR TABEL .....	viii
DAFTAR GAMBAR.....	ix
INTISARI .....	xi
ABSTRACT .....	xii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penelitian .....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 Konsep <i>Big Data</i> .....	11
2.2 <i>Bus Rapid Transit</i> (BRT) .....	12
2.3 Apache Kafka .....	12
2.4 Apache Spark .....	16
2.5 Apache Cassandra .....	21
2.6 Matriks Asal-Tujuan (AT).....	23
2.7 Jenis <i>Query</i> .....	25
2.8 <i>Java Regular Expressions</i> .....	26
2.9 Hipotesis .....	27
BAB III BAHAN DAN METODE PENELITIAN .....	28
3.1 Bahan.....	28
3.2. Peralatan .....	29
3.3 Prosedur Penelitian.....	29
3.4 Pengujian Hipotesis Penelitian.....	41
BAB IV HASIL PENELITIAN DAN PEMBAHASAN .....	57
4.1 Hasil Pengujian Pengiriman <i>Dataset</i> menggunakan Media Kabel LAN .....	58
4.2 Hasil Pengiriman <i>Dataset</i> menggunakan Media <i>Wireless</i> .....	65
4.3 Hasil Pengujian Operasi <i>Join</i> Data Ekstraksi dengan Tabel Refrensi.....	72

4.4	Hasil Matriks AT .....	80
BAB V PENUTUP .....		86
5.1	Kesimpulan.....	86
5.2	Saran.....	87
DAFTAR PUSTAKA.....		88

## DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian.....	9
Tabel 3.1 Spesifikasi PC – <i>Host</i> .....	29
Tabel 3.2 Spesifikasi PC – <i>Master node</i> .....	29
Tabel 3.3 Struktur data <i>dataset</i> transaksi BRT .....	39
Tabel 3.4 Tabel Pengalamatan IP <i>Host</i> dan <i>Master node</i> .....	44
Tabel 3.5 Tabel Pengalamatan IP <i>Host</i> dan <i>Master node</i> .....	48
Tabel 3.6 Tabel <i>Script producer</i> untuk mengirim <i>dataset</i> (LAN).....	49
Tabel 3.7 Tabel Pengalamatan IP <i>Host</i> dan <i>Master node</i> pada skenario kedua.....	50
Tabel 3.8 Tabel konfigurasi <i>broker list</i> Kafka .....	51
Tabel 3.9 Tabel <i>Script producer</i> untuk mengirim <i>dataset</i> ( <i>wireless</i> ).....	52
Tabel 3.10 Sampel tabel refrensi gate-halte .....	53
Tabel 4.1 Tabel pengujian penelitian .....	57
Tabel 4.2 Sampel catatan waktu pengiriman 500000 transaksi 1 detik per <i>batch</i> (LAN)...	58
Tabel 4.3 Catatan waktu Pengiriman 500000 transaksi 5 detik per <i>batch</i> (LAN).....	59
Tabel 4.4 Catatan waktu Pengiriman 500000 transaksi 10 detik per <i>batch</i> (LAN).....	60
Tabel 4.5 Sampel catatan waktu Pengiriman 3500000 transaksi 1 detik per <i>batch</i> (LAN)	61
Tabel 4.6 Sampel catatan waktu Pengiriman 3500000 transaksi 5 detik per <i>batch</i> (LAN)	62
Tabel 4.7 Sampel catatan waktu pengiriman 3500000 transaksi 10 detik per <i>batch</i> (LAN)	63
Tabel 4.8 Sampel catatan waktu pengiriman 500000 transaksi 1 detik per <i>batch</i> ( <i>wireless</i> )	65
Tabel 4.9 Sampel catatan waktu pengiriman 500000 transaksi 5 detik per <i>batch</i> ( <i>wireless</i> )	66
Tabel 4.10 Catatan waktu pengiriman 500000 transaksi 10 detik per <i>batch</i> ( <i>wireless</i> ).....	67
Tabel 4.11 Sampel catatan waktu pengiriman 3500000 transaksi 1 detik per <i>batch</i> ( <i>wireless</i> ) .....	68
Tabel 4.12 Sampel catatan waktu pengiriman 3500000 transaksi 5 detik per <i>batch</i> ( <i>wireless</i> ) .....	69
Tabel 4.13 Sampel catatan waktu pengiriman 3500000 transaksi 10 detik per <i>batch</i> ( <i>wireless</i> ) .....	70

## DAFTAR GAMBAR

Gambar 2.1 Diagram cara kerja Kafka .....	15
Gambar 2.2 Arsitektur Spark .....	17
Gambar 2.3 Fitur RDD .....	18
Gambar 2.4 Fitur Spark <i>Streaming</i> .....	20
Gambar 2.5 Cara kerja DStream .....	20
Gambar 2.6 <i>Batch interval</i> pada DStream .....	21
Gambar 2.7 Arsitektur Cassandra .....	23
Gambar 2.8 Representasi matriks AT .....	24
Gambar 3.1 Bagan Alir Metode Penelitian .....	30
Gambar 3.2 <i>Package</i> Apache Kafka .....	32
Gambar 3.3 Direktori Apache Kafka .....	32
Gambar 3.4 <i>Package</i> Apache Spark .....	32
Gambar 3.5 Direktori Apache Spark .....	33
Gambar 3.6 Folder winutils .....	33
Gambar 3.7 <i>Environment Variables</i> Windows .....	34
Gambar 3.8 Konfigurasi SPARK_HOME .....	34
Gambar 3.9 Konfigurasi JAVA_HOME .....	35
Gambar 3.10 Konfigurasi HADOOP_HOME .....	35
Gambar 3.11 Konfigurasi <i>Path</i> .....	35
Gambar 3.12 Tampilan Spark-shell .....	36
Gambar 3.13 Instalasi Datastax DevCenter .....	37
Gambar 3.14 <i>Package</i> Apache Cassandra .....	37
Gambar 3.15 File cassandra.bat .....	38
Gambar 3.16 Tampilan <i>Shell Server</i> Cassandra .....	38
Gambar 3.17 File cqlsh.bat .....	38
Gambar 3.18 Tampilan <i>command prompt</i> CQL .....	39
Gambar 3.19 Represensati dataset transaksi BRT .....	41
Gambar 3.20 Alur simulasi pengiriman <i>dataset</i> dari halte ke <i>datacenter</i> hingga terbentuk matriks AT .....	43
Gambar 3.21 Topologi pengiriman <i>dataset</i> dengan media LAN .....	44
Gambar 3.22 Folder Konfigurasi <i>Broker</i> Pertama Kafka .....	45
Gambar 3.23 Folder Konfigurasi <i>Broker</i> Kedua Kafka .....	45
Gambar 3.24 Folder bin Kafka .....	46
Gambar 3.25 Script untuk menjalankan Zookeeper .....	46
Gambar 3.26 Jendela Zookeeper <i>Server</i> .....	46
Gambar 3.27 <i>Script</i> untuk menjalankan <i>broker</i> pertama .....	46
Gambar 3.28 Jendela <i>broker</i> pertama .....	47
Gambar 3.29 <i>Script</i> untuk menjalankan <i>broker</i> kedua .....	47
Gambar 3.30 <i>Script</i> untuk membuat topik .....	47
Gambar 3.31 <i>Package dependencies</i> Spark .....	48
Gambar 3.32 Konfigurasi Kafka parameter, DirectStream dan topik pada Spark .....	49
Gambar 3.33 Konfigurasi penyimpanan data ke Cassandra .....	49
Gambar 3.34 Perintah <i>script</i> pada salah satu <i>producer</i> .....	50
Gambar 3.35 Topologi pengiriman <i>dataset</i> dengan media <i>wireless</i> .....	50
Gambar 3.36 Konfigurasi Kafka parameter pada skenario kedua .....	52

Gambar 3.37 Data hasil ekstraksi di Cassandra .....	53
Gambar 3.38 <i>Query</i> untuk membuat tabel <i>gatehalte_id</i> .....	54
Gambar 3.39 <i>Query</i> untuk <i>copy</i> tabel dari csv ke cassandra .....	54
Gambar 3.40 Kode program untuk <i>join</i> data ekstraksi dengan tabel <i>gate-halte</i> .....	54
Gambar 3.41 Sampel hasil <i>join</i> data ekstraksi dengan tabel <i>gate-halte</i> .....	55
Gambar 3.42 Spark-submit untuk menjalankan algoritma <i>trip chaining</i> .....	55
Gambar 3.43 Sampel hasil matriks AT dalam format csv .....	56
Gambar 4.1 Rerata streaming data (LAN).....	64
Gambar 4.2 Rerata streming data (wireless) .....	71
Gambar 4.3 Catatan waktu operasi <i>join</i> 500000 transaksi .....	72
Gambar 4.4 Catatan waktu operasi <i>join</i> 1000000 transaksi .....	73
Gambar 4.5 Catatan waktu operasi <i>join</i> 1500000 transaksi .....	74
Gambar 4.6 Catatan waktu operasi <i>join</i> 2000000 transaksi .....	75
Gambar 4.7 Catatan waktu operasi <i>join</i> 2500000 transaksi .....	76
Gambar 4.8 Catatan waktu operasi <i>join</i> 2500000 transaksi .....	77
Gambar 4.9 Catatan waktu operasi <i>join</i> 3500000 transaksi .....	78
Gambar 4.10 Rerata performa Spark dalam operasi <i>Join</i> .....	79
Gambar 4.11 Catatan waktu Spark menjalankan algoritma pada data 500000 .....	80
Gambar 4.12 Catatan waktu Spark menjalankan algoritma pada data 1000000 .....	80
Gambar 4.13 Catatan waktu Spark menjalankan algoritma pada data 1500000 .....	81
Gambar 4.14 Catatan waktu Spark menjalankan algoritma pada data 2000000 .....	81
Gambar 4.15 Catatan waktu Spark menjalankan algoritma pada data 2500000 .....	82
Gambar 4.16 Catatan waktu Spark menjalankan algoritma pada data 3000000 .....	82
Gambar 4.17 Catatan waktu Spark menjalankan algoritma pada data 3500000 .....	82
Gambar 4.18 Rerata performa Spark mengeksekusi algoritma <i>trip chaining</i> .....	83
Gambar 4.19 Visualisasi Matriks AT .....	85