

## DAFTAR ISI

HALAMAN PENGESAHAN .....	ii
PERNYATAAN BEBAS PLAGIASI .....	iii
HALAMAN PERSEMBAHAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI .....	vi
DAFTAR TABEL .....	x
DAFTAR GAMBAR .....	xi
DAFTAR SINGKATAN .....	xiv
INTISARI .....	xv
ABSTRACT .....	xvi
BAB I Pendahuluan .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Penelitian .....	4
1.5 Manfaat Penelitian .....	5
1.6 Sistematika Penulisan .....	5
BAB II Tinjauan Pustaka dan Dasar Teori .....	7
2.1 Tinjauan Pustaka .....	7
2.1.1 Teknologi Pengolahan <i>Big Data</i> .....	7
2.1.2 Pengolahan dan Analisis Geospasiotemporal pada Data Mobilitas Manusia .....	18
2.2 Dasar Teori .....	23
2.2.1 <i>Big Data</i> .....	23
2.2.2 <i>Passive Mobile Positioning Data</i> .....	24
2.2.3 <i>Haversine Formula</i> .....	25
2.2.4 Python .....	26
2.2.5 Komputasi Paralel Terdistribusi .....	26
2.2.6 Kubernetes .....	27
2.2.6.1 Peran Kubernetes dalam Komputasi Terdistribusi .....	27
2.2.6.2 Arsitektur Kubernetes .....	27
2.2.7 Apache Hadoop .....	29
2.2.7.1 <i>Hadoop Distributed File System</i> (HDFS) .....	29
2.2.8 Apache Spark .....	30
2.2.8.1 Spark pada Kluster Kubernetes .....	30
2.2.9 ILUM .....	31

2.3	Prometheus .....	32
2.4	Node Exporter.....	33
2.5	cAdvisor .....	33
2.6	Grafana .....	34
2.7	MinIO .....	35
2.8	Analisis Perbandingan Metode .....	35
BAB III Metode Penelitian.....		37
3.1	Alat dan Bahan Tugas Akhir .....	37
3.1.1	Alat Tugas akhir.....	37
3.1.2	Bahan Tugas akhir .....	39
3.2	Metode yang Digunakan.....	40
3.3	Alur Tugas Akhir .....	40
3.3.1	Pengumpulan Data .....	42
3.3.2	Pengembangan Infrastruktur.....	42
3.3.3	Pengembangan Kode Pemrosesan Data .....	45
3.3.4	Pengujian Performa Mesin Komputasi Pengolah Data .....	47
3.3.5	Pengumpulan Data Hasil Pengujian.....	50
3.4	Keterbatasan Penelitian .....	50
BAB IV Hasil dan Pembahasan.....		52
4.1	Implementasi Teknologi Pemrosesan MPD Pasif .....	52
4.1.1	<i>Control-plane</i> .....	54
4.1.1.1	ambari .....	54
4.1.2	<i>Worker Nodes</i> .....	55
4.1.2.1	dn001.hpc.ugm.....	55
4.1.2.2	dn002 .....	56
4.1.2.3	dn003 .....	57
4.1.2.4	nn001 .....	58
4.1.2.5	nn002 .....	58
4.2	Perbandingan Lingkungan Komputasi Lokal Python, Lokal Spark, dan Kluster Spark dalam Pemrosesan serta Analisis MPD Pasif.....	60
4.2.1	Lingkungan Komputasi Lokal Python .....	60
4.2.2	Lingkungan Komputasi Lokal Spark.....	60
4.2.3	Lingkungan Komputasi Kluster Spark .....	60
4.3	Perbandingan Profil Kinerja Lingkungan Python dan Spark dalam Pemrosesan serta Analisis MPD Pasif .....	62
4.3.1	Perbandingan Profil Kinerja Lingkungan Komputasi untuk Data 10.000 Rows .....	63
4.3.1.1	Performa Berdasarkan Waktu Eksekusi .....	63
4.3.1.2	Performa Berdasarkan Waktu CPU .....	64

4.3.1.3	Performa Berdasarkan Penggunaan Memori .....	65
4.3.2	Perbandingan Profil Kinerja Lingkungan Komputasi untuk Data 100.000 Rows .....	66
4.3.2.1	Performa Berdasarkan Waktu Eksekusi .....	67
4.3.2.2	Performa Berdasarkan Waktu CPU .....	68
4.3.2.3	Performa Berdasarkan Penggunaan Memori .....	69
4.3.3	Perbandingan Profil Kinerja Lingkungan Komputasi untuk Data 1.000.000 Rows .....	70
4.3.3.1	Performa Berdasarkan Waktu Eksekusi .....	71
4.3.3.2	Performa Berdasarkan Waktu CPU .....	72
4.3.3.3	Performa Berdasarkan Penggunaan Memori .....	73
4.3.4	Perbandingan Profil Kinerja Lingkungan Komputasi untuk Data 10.000.000 Rows .....	74
4.3.4.1	Performa Berdasarkan Waktu Eksekusi .....	75
4.3.4.2	Performa Berdasarkan Waktu CPU .....	76
4.3.4.3	Performa Berdasarkan Penggunaan Memori .....	77
4.3.5	Perbandingan Profil Kinerja Lingkungan Komputasi untuk Data 1 GB .....	78
4.3.5.1	Performa Berdasarkan Waktu Eksekusi .....	79
4.3.5.2	Performa Berdasarkan Waktu CPU .....	80
4.3.5.3	Performa Berdasarkan Penggunaan Memori .....	81
4.3.6	Perbandingan Profil Kinerja lingkungan komputasi untuk Data 5 GB .....	82
4.3.6.1	Performa Berdasarkan Waktu Eksekusi .....	83
4.3.6.2	Performa Berdasarkan Waktu CPU .....	84
4.3.6.3	Performa Berdasarkan Penggunaan Memori .....	85
4.3.7	Perbandingan Profil Kinerja lingkungan komputasi untuk Data 15 GB .....	86
4.3.7.1	Performa Berdasarkan Waktu Eksekusi .....	87
4.3.7.2	Performa Berdasarkan Waktu CPU .....	88
4.3.7.3	Performa Berdasarkan Penggunaan Memori .....	89
4.3.8	Perbandingan Profil Kinerja Lingkungan Komputasi Berdasarkan Waktu Eksekusi.....	91
4.3.9	Perbandingan Profil Kinerja Lingkungan Komputasi Berdasarkan Waktu CPU .....	92
4.3.10	Perbandingan Profil Kinerja Lingkungan Komputasi Berdasarkan Penggunaan Memori .....	93
BAB V	Kesimpulan dan Saran.....	96
5.1	Kesimpulan.....	96

3.2	Saran .....	96
	DAFTAR PUSTAKA .....	98
	LAMPIRAN .....	L-1
L.1	Data Hasil Pencatatan Performa Mesin Menggunakan Pustaka 'psutils' ...	L-1
L.1.1	Data Performa Kinerja Mesin Lingkungan Komputasi Berdasarkan Waktu Eksekusi .....	L-1
L.1.2	Data Performa Kinerja Mesin Lingkungan Komputasi Berdasarkan Waktu CPU .....	L-2
L.2	Data Hasil Pencatatan Performa Mesin Menggunakan Grafana .....	L-3
L.2.1	Data Performa Kinerja Mesin Lingkungan Komputasi Berdasarkan Total Alokasi Memori pada Lingkungan Lokal .....	L-3
L.2.2	Data Performa Kinerja Mesin Lingkungan Komputasi Berdasarkan Total Alokasi Memori pada Lingkungan Kluster Terdistribusi .....	L-3
L.3	Pemantauan Kluster Kubernetes .....	L-4
L.4	Kode Pemindahan Data Sampel .....	L-5
L.5	Tangkapan Layar Pemrosesan Data pada Jupyter Notebook .....	L-6
L.6	Sampel Hasil Pemrosesan MPD Pasif .....	L-7
L.6.1	Data Pengguna yang Terdeteksi Berhenti di Suatu Lokasi untuk Waktu Tertentu .....	L-7
L.6.2	Data Pengguna yang Melakukan Mobilisasi .....	L-7