

HALAMAN PENGESAHAN	ii
PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR SINGKATAN.....	xiii
INTISARI.....	xiv
ABSTRACT	xv
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan.....	4
BAB II Tinjauan Pustaka dan Dasar Teori	6
2.1 Tinjauan Pustaka	6
2.1.1 Teknologi untuk Analisis Mahadata dan Solusi Sistem Terdistribusi	6
2.1.2 Prapemrosesan Mahadata Geo-Spasiotemporal pada <i>Mobile Po-</i> <i>sitioning Data</i> untuk Analisis Mobilitas Manusia	12
2.2 Dasar Teori	18
2.2.1 Big Data	19
2.2.2 Mobile Positioning Data	20
2.2.3 <i>Trajectory Filtering</i> : Penyaringan Data Berdasarkan Kecepatan ...	21
2.2.4 <i>Reverse Geocoding</i>	22
2.2.5 The Great Circle Distance	23
2.2.6 Python	24
2.2.6.1 Python Global Interpreter Lock	24
2.2.7 Komputasi Paralel.....	25
2.2.8 Komputasi Terdistribusi	27
2.2.9 Docker	27
2.2.9.1 Docker Swarm	29
2.2.10 Portainer	30
2.2.11 Apache Hadoop	32

2.2.11.1	HDFS	33
2.2.12	Apache Spark	33
2.2.12.1	Pyspark	37
2.2.13	Apache Sedona	37
2.2.14	Prometheus	38
2.2.15	Grafana	40
2.2.16	NodeExporter	40
2.2.17	cAdvisor	41
2.2.18	Pandas	41
2.2.19	Geopandas	42
2.3	Analisis Perbandingan Metode	43
BAB III Metode Penelitian		44
3.1	Alat dan Bahan Tugas akhir	44
3.1.1	Alat Tugas akhir	44
3.1.2	Bahan Tugas akhir	46
3.2	Metode yang Digunakan	46
3.3	Alur Tugas Akhir	47
3.3.1	Koleksi Data	48
3.3.2	Implementasi dan Pengembangan Infrastruktur	49
3.3.3	Pembangunan Kode Pemrosesan Mahadata <i>Spatiotemporal</i> dari dataset MPD aktif GPS	52
3.3.4	Pengembangan Kode Pengambilan Data Performa	58
3.3.5	Skema Pengujian	59
3.4	Limitasi Penelitian	59
BAB IV Hasil dan Pembahasan		61
4.1	Pengembangan Solusi Teknologi Analisis Mahadata untuk Pengolahan Mahadata MPD	61
4.1.1	Arsitektur Kluster dari Perspektif Perangkat Keras dan Jaringan	62
4.1.1.1	Arsitektur Kluster pada HPC-06	62
4.1.1.2	Arsitektur Kluster pada HPC-07	63
4.1.1.3	Arsitektur Kluster pada HPC-08	64
4.1.1.4	Arsitektur Jaringan pada Kluster DockerSwarm	64
4.1.2	Arsitektur Komputer dari Perspektif Perangkat Lunak	66
4.1.2.1	<i>Stack Big Data Cluster</i>	66
4.1.2.2	<i>Stack Server Monitoring</i>	66
4.2	Analisis Perbandingan Antara Lingkungan Python dan Spark dalam Mengeksekusi Fungsi Analisis Mahadata MPD	69
4.2.1	Analisis Perbandingan Kinerja Pemrosesan untuk Data Berukuran 1GB	69

4.2.1.1	Performa Waktu Eksekusi	69
4.2.1.2	Kinerja Sumber Daya	70
4.2.2	Analisis Perbandingan Kinerja Pemrosesan untuk Data Berukuran 5GB	75
4.2.2.1	Performa Waktu Eksekusi	75
4.2.2.2	Kinerja Sumber Daya	75
4.2.3	Analisis Perbandingan Kinerja Pemrosesan untuk Data Berukuran 10GB	81
4.2.3.1	Performa Waktu Eksekusi	81
4.2.3.2	Kinerja Sumber Daya	81
4.2.4	Analisis Perbandingan Kinerja Pemrosesan untuk Data Berukuran 26GB	87
4.2.4.1	Performa Waktu Eksekusi	87
4.2.4.2	Kinerja Sumber Daya	87
4.2.5	Analisis Perbandingan Waktu Eksekusi Secara Keseluruhan	93
4.2.6	Analisis Perbandingan Penggunaan Sumber Daya Secara Keseluruhan	95
4.2.6.1	Penggunaan CPU	95
4.2.6.2	Penggunaan Memori	95
4.2.6.3	Data Terkirim dan Diterima	96
BAB V	Kesimpulan dan Saran	98
5.1	Kesimpulan	98
5.2	Saran	99
DAFTAR PUSTAKA	100
LAMPIRAN	L-1
L.1	Data Total Waktu Eksekusi pada Setiap Tahapan	L-1
L.2	Data Rerata dan Standar Deviasi dari Penggunaan Sumber Daya Pada setiap Tahapan.	L-2
L.3	Tangkapan Layar WebUI dari Teknologi Mahadata yang Digunakan	L-6