

SKRIPSI

**ANALISIS PROFIL KINERJA DAN BATAS STABILITAS ARSITEKTUR
KLUSTER CEPH MELALUI PENDEKATAN BENCHMARKING DI
LABORATORIUM INFORMATIKA UNIVERSITAS ISLAM NEGERI
SUNAN KALIJAGA**



DIAJUKAN OLEH:

RAVI AKVARIAN KUSUMA PUTRA

NIM. 220106050030

PROGRAM STUDI INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

**UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2026

LEMBAR PENGESAHAN TUGAS AKHIR



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1171/Un.02/DST/PP.00.9/06/2026

Tugas Akhir dengan judul : ANALISIS PROFIL KINERJA DAN BATAS STABILITAS ARSITEKTUR
KLUSTER CEPH MELALUI PENDEKATAN BENCHMARKING DI
LABORATORIUM INFORMATIKA UNIVERSITAS ISLAM NEGERI SUNAN
KALIJAGA

yang dipersiapkan dan disusun oleh:

Nama : RAVI AKVARIAN KUSUMA PUTRA
Nomor Induk Mahasiswa : 22106050030
Telah diujikan pada : Selasa, 26 Mei 2026
Nilai ujian Tugas Akhir : A/B

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang
Prof. Dr. Ir. Shofwatul Uyun, S.T., M.Kom., IPM., ASEAN Eng.
SIGNED

Valid ID: 6a20c5951f23



Penguji I
Dr. Ir. Bambang Sugiantoro, M.T., IPU.,
ASEAN Eng.
SIGNED

Valid ID: 6a1fca872f0a



Penguji II
Eko Hadi Gunawan, M.Eng.
SIGNED

Valid ID: 6a20bc14cbdb



Yogyakarta, 26 Mei 2026
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 6a2130c02b1ef

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

LEMBAR KEASLIAN TUGAS AKHIR

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Ravi Akvarian Kusuma Putra
NIM : 22106050030
Program Studi : Informatika
Fakultas : Sains Dan Teknologi

Menyatakan dengan sesungguhnya, bahwa skripsi saya yang berjudul Analisis Profil Kinerja Dan Batas Stabilitas Arsitektur Kluster Ceph Melalui Pendekatan Benchmarking Di Laboratorium Informatika Universitas Islam Negeri Sunan Kalijaga merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 18.1.5/2026



Ravi Akvarian Kusuma Putra
22106050030

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

SURAT PERSETUJUAN TUGAS AKHIR

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp : -

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Ravi Akvarian Kusuma Putra
NIM : 22106050030
Judul Skripsi : Analisis Profil Kinerja Dan Batas Stabilitas Arsitektur
Kluster Ceph Melalui Pendekatan Benchmarking Di
Laboratorium Informatika Universitas Islam Negeri
Sunan Kalijaga

sudah dapat diajukan kembali kepada Program Studi Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Informatika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 10 Mei 2026

Pembimbing



Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom.,

IPM., ASEAN Eng

NIP 19820511 200604 2 002

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

LEMBAR PERSETUJUAN SKRIPSI/TUGAS AKHIR

Tugas akhir ini tidak dipublikasikan secara umum, namun dapat diakses di perpustakaan dalam lingkungan UIN Sunan Kalijaga Yogyakarta. Dokumen ini diperbolehkan untuk dijadikan referensi kepustakaan, dengan ketentuan bahwa setiap pengutipan harus mendapatkan izin dari penulis serta mencantumkan sumber sesuai kaidah penulisan ilmiah. Hak kepemilikan dokumen tugas akhir ini berada pada UIN Sunan Kalijaga Yogyakarta.



INTISARI

Penelitian ini mengevaluasi stabilitas dan performa sistem penyimpanan terdistribusi Ceph pada infrastruktur komoditas heterogen di bawah spesifikasi minimum (*sub-minimum*) guna menjawab kendala anggaran pengadaan perangkat keras di Laboratorium Informatika UIN Sunan Kalijaga. Pengujian performa mentah kluster yang terdiri dari campuran node prosesor Intel Core i7 Generasi 2 dan 4, RAM 2–8 GB, serta HDD mekanis pada jaringan 100 Mbps dilakukan langsung pada lapisan objek menggunakan utilitas RADOS Bench dengan eskalasi konkurensi hingga 128 *thread*. Hasil pengukuran mengungkap bahwa *bottleneck* absolut sistem murni dikendalikan oleh limitasi mekanis HDD yang memicu kondisi *starved network* (utilisasi jaringan efektif hanya 2,9%). Titik jenuh operasional (*knee point*) teridentifikasi secara presisi pada 16 *thread* ($t=16$), di mana penambahan beban kerja setelah titik ini hanya menghasilkan pembengkakan latensi ekstrem hingga 59,52 detik tanpa eskalasi *throughput* material. Analisis fenomena *straggler* mengungkap anomali performa: OSD dengan prosesor lawas dan disk yang mengalami *bad sector* masif justru mencatatkan *commit latency* tercepat (62 ms berbanding 107 ms) akibat agresivitas algoritma revisi *firmware* CC76 yang mengkompensasi keausan mekanis, sehingga menghasilkan *Heterogeneity Penalty Factor* sebesar 0,58. Kesimpulannya, kluster Ceph terbukti fungsional beroperasi pada *hardware* bekas dengan batas aman $t=16$ dan mengonfirmasi bahwa homogenitas *firmware controller* disk jauh lebih determinan terhadap stabilitas latensi dibandingkan keseragaman generasi komputasi.

Kata Kunci: Ceph, sistem penyimpanan terdistribusi, *benchmarking*, infrastruktur heterogen, *knee point*, *straggler*, perangkat keras komoditas.

ABSTRACT

This research evaluates the stability and performance of the Ceph distributed storage system on a heterogeneous, sub-minimum commodity infrastructure to address hardware procurement budget constraints at the Informatics Laboratory of UIN Sunan Kalijaga. Raw performance testing of the cluster comprising a mix of Intel Core i7 2nd and 4th Generation processor nodes, 2–8 GB of RAM, and mechanical HDDs over a 100 Mbps network was conducted directly at the object layer using the RADOS Bench utility with concurrency scaling up to 128 threads. Empirical results reveal that the absolute system bottleneck is dictated entirely by the mechanical limitations of the HDDs, triggering a starved network condition with an effective network utilization of only 2.9%. The operational saturation point (knee point) was precisely identified at 16 threads ($t=16$); increasing the workload beyond this threshold only resulted in severe latency inflation (peaking at 59.52 seconds) without any material throughput gains. Furthermore, an analysis of the straggler effect uncovered a critical performance anomaly: the Object Storage Daemon (OSD) running on a legacy processor with a heavily degraded disk (massive bad sectors) unexpectedly recorded the fastest commit latency (62 ms versus 107 ms). This paradox was driven by the aggressive caching algorithms of the CC76 firmware revision, which masked the mechanical wear and ultimately contributed to a Heterogeneity Penalty Factor of 0.58. In conclusion, this study demonstrates that a Ceph cluster can remain functionally viable on repurposed hardware up to a safe operational limit of $t=16$, and confirms that disk controller firmware homogeneity is a significantly more decisive factor for latency stability than the uniformity of CPU generations.

Keywords: *Ceph, distributed storage system, benchmarking, heterogeneous infrastructure, knee point, straggler, commodity hardware.*

HALAMAN MOTTO

كُلُّ نَفْسٍ ذَائِقَةُ الْمَوْتِ ۗ ثُمَّ إِلَيْنَا تُرْجَعُونَ

"Every soul will taste death. Then to Us will you be returned."

~Al-Ankabut, Ayat 57

"Memento Mori sed memento vivere."

Remember to die, but remember to live.

~Persona 3

"And then one day you find ten years have got behind you.

No one told you when to run, you missed the starting gun."

~Time, Pink Floyd

"Nothing in life is promised except death. If you have the opportunity to play this game of life, you need to appreciate every moment. A lot of people don't appreciate the moment until it's passed"

~Kanye West

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Puji syukur ke hadirat Allah SWT, Sang Pemilik segala ilmu, yang telah memberikan kekuatan dan ketabahan hingga langkah terakhir dalam penyusunan skripsi ini dapat terselesaikan. Karya ini, yang dibangun di atas dedikasi setiap fase eksperimen dan ketelitian dalam sepuluh kali replikasi pengujian, saya persembahkan dengan penuh cinta untuk:

1. Kedua Orang Tua Tercinta, yang doa-doanya mengalir lebih deras daripada throughput jutaan requests. Terima kasih atas dukungan moral dan materi yang menjadi bahan bakar utama dalam setiap fase ramp-up kehidupan saya.
2. Dosen Pembimbing, yang telah memberikan arahan berharga serta dukungan penuh dalam menjaga kelancaran proses akademik penelitian ini hingga tahap penyelesaian.
3. Diri Saya Sendiri, Ravi Akvarian Kusuma Putra. Terimakasih karena telah berusaha menyelesaikan sesuai dengan apa yang saya mulai

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Segala puji bagi Allah SWT, Tuhan Semesta Alam, yang telah melimpahkan rahmat, hidayah dan kekuatan-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Performa dan Stabilitas Sistem Penyimpanan Terdistribusi Berbasis Ceph: Studi Kasus Laboratorium Informatika UIN Sunan Kalijaga". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta.

Penulis menyadari bahwa proses penyelesaian karya ini bukanlah perjalanan yang singkat. Dibutuhkan ketekunan dalam setiap tahap eksperimen, terutama dalam konfigurasi *cluster* dan pengujian pada infrastruktur heterogen, untuk memastikan setiap data yang diambil memiliki landasan yang kuat. Oleh karena itu, dengan kerendahan hati, penulis menyampaikan ucapan terima kasih yang tulus kepada:

1. Prof. Noorhaidi Hasan S.Ag., M.A., M.Phil., Ph.D. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Bapak Dr. Muhammad Mustakim, S.T. M.T. selaku Ketua Program Studi Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Dr. Agus Mulyanto, S.Si., M.Kom., ASEAN Eng. selaku dosen penasihat akademik yang telah memberikan arahan dan bimbingan selama masa perkuliahan penulis.
5. Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom., IPM., ASEAN Eng., selaku dosen pembimbing yang telah memberikan bimbingan, masukan yang berharga, serta dukungan penuh dengan penuh kesabaran dalam setiap tahapan penyusunan Tugas Akhir ini.
6. Seluruh dosen dan karyawan Program Studi Informatika UIN Sunan Kalijaga Yogyakarta yang telah memberikan ilmu, bantuan, serta pelayanan selama penulis menempuh pendidikan.

7. Pencapaian ini saya dedikasikan sepenuhnya kepada kedua orang tua saya yang telah memberikan segalanya demi masa depan penulis. Kepada almarhum Ayah, meskipun raga tak lagi bersisi, namun semangat yang pernah Ayah tanamkan tetap menjadi kompas yang memandu saya untuk terus mencoba bertahan di tengah badai tantangan akademis. Kepada Ibu, pendukung nomor satu yang tak pernah lelah menunjukkan arti kesabaran yang sesungguhnya. Terima kasih karena selalu percaya pada kemampuan saya dan terus membimbing saya melalui nasihat-nasihat bijak agar saya tidak hanya menjadi seorang sarjana, tetapi juga manusia yang lebih baik. Tanpa doa, dukungan dan pengorbanan kalian yang mengalir tanpa henti, perjalanan ini tidak akan pernah mencapai muaranya.
8. Sembah sujud dan rasa rindu yang mendalam saya persembahkan kepada almarhumah kakak saya, Lindiana Puspitasari. Beliau adalah sumber inspirasi abadi yang senantiasa menitipkan pesan melalui kenangan untuk tidak pernah berhenti memperjuangkan apa yang layak diperjuangkan. Semoga karya ini menjadi salah satu bentuk penghormatan atas semangat yang pernah ia berikan.
9. Rasa terima kasih yang tulus juga saya sampaikan kepada kakak-kakak tersayang: Febrian Adiputra, Amelia Septyaningtyas, Nugraha Sofyan Aji, Kumalasari Lubis, Bobby Darmawan dan Eka Kurnia Putri. Terimakasih atas dukungan moral, doa dan kasih sayang yang menjadi pilar kekuatan saya dalam menyelesaikan perjalanan panjang ini.
10. Terima kasih kepada keponakan-keponakan saya: Lantunan Madiha Nadian, Dea Samawapina dan Abian Janu Kautsarrazky, Bian. Kehadiran dan keceriaan kalian adalah penawar lelah yang memberikan warna tersendiri di sela-sela pengerjaan tugas akhir ini.
11. Apresiasi setinggi-tingginya saya tujukan kepada rekan-rekan *party* gim saya, Gabut Bang, serta sahabat-sahabat karib yang senantiasa hadir menemani saya mencoba melampaui batas kemampuan diri. Terima kasih atas kesabaran kalian yang tidak pernah lelah memberikan dukungan moral, bantuan teknis yang sangat berarti, hingga limpahan semangat di saat-saat

tersulit yang hampir mematahkan tekad. Keberadaan kalian di sisi saya adalah bukti nyata bahwa perjalanan panjang dan berliku ini tidak pernah saya lalui sendirian

12. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan, dukungan dan doa dalam penyelesaian Tugas Akhir ini.
13. Secara khusus, saya ingin berterima kasih kepada Kanye West. Melalui album *Graduation*, ia memberikan narasi musikal tentang kelulusan yang mendorong saya untuk segera menyelesaikan proses *benchmarking* dan merayakan pencapaian ini dengan penuh keyakinan.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Akhir kata, penulis berharap karya ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya di bidang sistem penyimpanan terdistribusi (*Distributed Storage System*) dan optimasi infrastruktur IT.

Yogyakarta, 12 Mei 2026

Penyusun,

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Ravi Akvarian Kusuma Putra

NIM. 22106050030

DAFTAR ISI

SKRIPSI.....	i
LEMBAR PENGESAHAN TUGAS AKHIR	ii
LEMBAR KEASLIAN TUGAS AKHIR	iii
SURAT PERSETUJUAN TUGAS AKHIR	iv
INTISARI.....	vi
<i>ABSTRACT</i>	vii
HALAMAN MOTTO	viii
HALAMAN PERSEMBAHAN	ix
KATA PENGANTAR	x
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR	xvii
BAB I	18
PENDAHULUAN	18
1.1 Latar Belakang	18
1.2 Rumusan Masalah	20
1.3 Batasan Masalah.....	20
1.4 Tujuan Penelitian.....	21
1.5 Manfaat Penelitian.....	21
BAB II.....	23
TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	23
2.1 Tinjauan Pustaka	23
2.2 Landasan Teori	28
2.2.1 Tinjauan Umum Sistem Terdistribusi	28
2.2.2 Sistem Penyimpanan Terdistribusi (<i>Distributed Storage System</i>)	31
2.2.3 Arsitektur Ceph dan Komparasi Implementasi pada Sistem Terdistribusi	33
2.2.4 Pengaruh Algoritma CRUSH Pada Arsitektur.....	35
2.2.5 Benchmarking dan Stress Testing pada Infrastruktur Penyimpanan	36
2.2.6 Model Skalabilitas dan Batas Performa Sistem	38

2.2.6	Karakteristik Perangkat Keras dan Fenomena Straggler pada Kluster Heterogen.....	40
BAB III	42
METODE PENELITIAN	42
3.1	Alat dan Bahan Penelitian	42
3.2	Alur Metodologi Penelitian	44
3.2.1.	Persiapan Lingkungan.....	45
3.2.2.	Desain Skenario Uji	47
3.2.3.	Proses Implementasi dan Pengumpulan Data	49
3.2.4.	Analisis Data	50
BAB IV	52
HASIL PENELITIAN DAN PEMBAHASAN	52
4.1	Alur Metodologi Penelitian	52
4.1.1	Persiapan Lingkungan.....	52
4.1.2	Desain Skenario Uji	53
4.1.3	Proses Implementasi Dan Pengumpulan Data	54
4.1.3.1	Hasil Pemetaan Performa Mentah (Titik Nol)	54
4.1.3.2	Analisis Throughput Sekuensial (Blok 4MB).....	57
4.1.3.2.1	Data Hasil Pengujian Multithread	58
4.1.3.2.2	Analisis Skalabilitas dan Stagnasi Throughput	58
4.1.3.3	Analisis IOPS Acak (Blok 4KB).....	59
4.1.3.3.1	Justifikasi Blok 4KB dan Karakteristik Beban Kerja Acak .	60
4.1.3.3.2	Hasil Pengukuran IOPS Acak Multithread.....	60
4.1.3.3.3	Asimetri Skalabilitas: <i>Write IOPS</i> Naik, <i>Read IOPS</i> Turun	63
4.1.3.3.4	Fenomena Cache Exhaustion pada Jalur Baca	64
4.1.3.3.5	Perbandingan Perilaku IOPS terhadap Throughput Sekuensial	65
4.1.4	Analisis Data	67
4.1.4.1	Analisis Komparatif dan Identifikasi Bottleneck Fisik.....	67
4.1.4.2	Identifikasi Knee Point dan Stabilitas	71
4.1.4.2.1	Kerangka Analisis Knee Point dan Marginal Gain	71
4.1.4.2.2	Knee Point pada Workload Sekuensial 4MB	72
4.1.4.2.3	Knee Point pada Workload Acak 4KB.....	73

4.1.4.2.4	Peta Operasional: Sintesis Knee Point Lintas Workload	74
4.1.4.2.5	Analisis Stabilitas Pasca Knee Point	75
4.1.4.3	Fenomena <i>Straggler</i> pada Infrastruktur Heterogen.....	77
4.1.4.3.1	Definisi dan Mekanisme Straggler Effect dalam Sistem Replikasi	77
4.1.4.3.2	Pemetaan Node Fisik ke OSD dan Profil Hardware	78
4.1.4.3.3	Analisis Commit Latency per OSD pada Kondisi Baseline .	80
4.1.4.3.4	Evolusi Straggler Effect di Bawah Beban Multithread	81
4.1.4.3.5	Kuantifikasi Penalti Heterogenitas dan Implikasi Arsitektural.....	83
4.1.4.4	Identifikasi <i>Bottleneck</i> Fisik.....	84
4.1.4.4.1	Hierarki Bottleneck Sistem.....	84
4.1.4.4.2	Batas Operasional Aman (Safe Operating Threshold)	86
4.1.4.4.3	Implikasi Heterogenitas Perangkat Keras	87
4.1.4.4.4	Rekomendasi Teknis	89
A.	Penetapan Batas Konkurensi Operasional pada $t=16$	89
B.	Homogenitas Revisi Firmware antar-OSD.....	89
C.	Pemantauan Aktif Metrik Straggler dan Fenomena Burst-Stall.....	90
4.2	Pembahasan	91
BAB V	95
KESIMPULAN	95
5.1	Kesimpulan.....	95
5.2	Saran	96
5.2.1	Saran untuk Penelitian Selanjutnya	96
5.2.2	Saran untuk Praktisi	96
DAFTAR PUSTAKA	98
LAMPIRAN	I

DAFTAR TABEL

Tabel 2. 1 Gap Penelitian Terdahulu.....	25
Tabel 4. 1 Hasil Pengujian Baseline Performa Mentah (Replikasi 3x, Menggunakan 1 Thread).....	55
Tabel 4. 2 Hasil Pengujian Sequential Write 4MB	58
Tabel 4. 3 Hasil Pengujian Sequential Read 4MB.....	58
Tabel 4. 4 Hasil Pengujian Random Write 4KB (Lengkap)	61
Tabel 4. 5 Hasil Pengujian Random Read 4KB	61
Tabel 4. 6-Total Objek Dibaca Dan Estimasi Working Set.....	64
Tabel 4. 7 Perbandingan Karakteristik Skalabilitas: Sekuensial vs Acak.....	66
Tabel 4. 8 Straggler Node	67
Tabel 4. 9 Analisis Degradasi Performa dan Batas Teoritis	69
Tabel 4. 10 Marginal Gain Analysis: Sequential Write 4MB.....	72
Tabel 4. 11 Marginal Gain Analysis: Sequential Read 4MB.....	72
Tabel 4. 12 Marginal Gain Analysis: Random Write 4KB.....	73
Tabel 4. 13 Marginal Gain Analysis: Random Read 4KB.....	73
Tabel 4. 14 Ringkasan Knee Point dan Rekomendasi Operasional.....	74
Tabel 4. 15 Metrik Stabilitas Sequential Write 4MB.....	76
Tabel 4. 16 Spesifikasi Hardware Pada Osd	78
Tabel 4. 17 Commit Latency.....	80
Tabel 4. 18 Evolusi Commit latency per OSD.....	81
Tabel 4. 19 Dekomposisi Penalti Performa Infrastruktur	83

DAFTAR GAMBAR

Gambar 2. 1 Skenario Distributed Computing System.....	29
Gambar 2. 2 Skenario Distributed Storage System.....	30
Gambar 3. 1 Skenario Kluster Ceph.....	43
Gambar 3. 2 Alur Metodologi Penelitian.....	45
Gambar 4. 1 Perbandingan Random Read dan Write IOPS (Blok 4KB).....	62
Gambar 4. 2 Gambar Rata-rata Latensi Random Read dan Write (Blok 4KB).....	62
Gambar 4. 3 Perbandingan Bandwidth Random Read dan Write (Blok 4KB).....	63
Gambar 4. 4 Heatmap Commit Latency per OSD(ms).....	66

BAB I

PENDAHULUAN

1.1 Latar Belakang

Peningkatan kebutuhan pengelolaan data di lingkungan institusi pendidikan [1] saat ini tidak lagi terbatas pada penyimpanan dokumen berbasis teks, tetapi juga mencakup pengelolaan dataset penelitian berukuran besar, penyediaan image sistem operasi untuk kebutuhan virtualisasi laboratorium, serta repositori proyek mahasiswa yang memuat konten multimedia dan kode sumber dalam jumlah signifikan. Di Laboratorium Informatika UIN Sunan Kalijaga, kebutuhan tersebut menjadi semakin krusial seiring meningkatnya aktivitas akademik yang menuntut ketersediaan data secara cepat dan berkelanjutan, seperti penerapan sistem pencadangan otomatis untuk tugas akhir maupun penyediaan media penyimpanan jaringan (network storage) yang mampu melayani banyak komputer laboratorium secara bersamaan. Akan tetapi, upaya modernisasi infrastruktur kerap terkendala oleh keterbatasan anggaran, sementara di sisi lain masih terdapat berbagai perangkat keras dengan spesifikasi beragam yang belum dimanfaatkan secara optimal dan belum terintegrasi dalam satu sistem terpadu. Oleh karena itu, pemanfaatan kembali perangkat-perangkat tersebut sebagai server data terdistribusi menjadi solusi strategis yang tidak hanya mampu menekan biaya pengadaan, tetapi juga meningkatkan efisiensi pemanfaatan sumber daya melalui transformasi perangkat keras komoditas menjadi infrastruktur penyimpanan yang andal dan berkelanjutan.

Peralihan menuju sistem penyimpanan terdistribusi dengan memanfaatkan perangkat lunak seperti Ceph pada perangkat keras bekas menghadirkan tantangan teknis yang cukup kompleks [2],[3], khususnya berkaitan dengan keterbatasan kapasitas memori (RAM) dan kemampuan pemrosesan CPU. Secara konseptual, platform tersebut pada umumnya dioptimalkan untuk dijalankan pada perangkat keras kelas enterprise yang memiliki sumber daya komputasi besar. Setiap Object Storage Device (OSD), misalnya, membutuhkan alokasi memori yang signifikan untuk mengelola struktur pemetaan data, sementara mekanisme algoritmik CRUSH

memerlukan kestabilan siklus CPU agar proses penentuan lokasi penyimpanan data dapat berlangsung secara cepat dan efisien. Pemanfaatan prosesor generasi lama, seperti lini awal Intel Core i7, memiliki keterbatasan pada rendahnya jumlah instruksi per siklus (*Instructions Per Cycle*) serta performa pengendali memori yang tertinggal dibandingkan standar perangkat keras modern. Keterbatasan tersebut berpotensi memicu kondisi *CPU throttling*, yaitu penurunan frekuensi kerja akibat peningkatan suhu ketika sistem menghadapi beban tinggi. Dampaknya, fenomena *straggler* dapat terjadi, yakni ketika satu node dengan performa menurun menjadi hambatan bagi kinerja keseluruhan jaringan penyimpanan. Oleh sebab itu, diperlukan proses validasi kelayakan operasional melalui pengukuran metrik latensi dan throughput guna memastikan sistem tidak hanya mampu beroperasi secara fungsional, tetapi juga tetap stabil, sehat dan andal saat menghadapi tekanan beban kerja yang berat.

Studi terdahulu telah banyak mengevaluasi performa penyimpanan terdistribusi, seperti analisis efisiensi biaya oleh Parashar [4], optimasi kluster heterogen oleh Lim dan Park[5], serta algoritma *read/write* oleh Waleed Reeda [6]. Namun, terdapat celah penelitian yang signifikan di mana mayoritas literatur tersebut fokus pada perangkat keras standar industri atau variasi spesifikasi yang masih dalam ambang batas operasional yang direkomendasikan oleh produsen. Penelitian ini membedakan diri dengan mengambil sudut pandang *Extreme Resource Constraint* (keterbatasan sumber daya ekstrem). Fokus utama riset ini adalah membedah secara empiris bagaimana sistem bereaksi ketika dijalankan pada perangkat keras yang secara spesifik berada di bawah standar spesifikasi minimum yang direkomendasikan oleh pengembang (*sub-minimum requirements*). Hal ini penting untuk menghilangkan subjektivitas dalam menilai kelayakan sistem pada infrastruktur "PC bekas".

Berangkat dari celah riset tersebut[7], penelitian ini difokuskan pada korelasi antara identifikasi *Knee Point* sebagai titik jenuh di mana kecepatan transfer data terhenti dan jeda waktu (*latency*) meningkat tajam dengan profil stabilitas sistem secara keseluruhan. Melalui metodologi *benchmarking* yang presisi menggunakan RADOS Benchmark, batasan fisik infrastruktur di

Laboratorium Informatika UIN Sunan Kalijaga dapat dipetakan guna menentukan ambang batas aman operasional. Langkah ini krusial untuk meminimalkan potensi kegagalan sistem saat menghadapi beban kerja riil di masa depan. Hasil akhir dari riset ini diproyeksikan menjadi dasar penyusunan pedoman teknis atau *Standard Operating Procedure (SOP)*[8] dalam mendayagunakan perangkat keras komoditas menjadi sistem penyimpanan terdistribusi yang mumpuni. Pada akhirnya, kontribusi riset ini terletak pada penguatan keandalan infrastruktur digital kampus, memastikan pengelolaan data akademik tetap stabil, serta mencegah terjadinya penurunan performa yang tidak terantisipasi.

1.2 Rumusan Masalah

1. Bagaimana profil performa mentah kluster Ceph pada infrastruktur perangkat keras heterogen yang berada di bawah spesifikasi minimum?
2. Pada tingkat konkurensi berapa kluster Ceph mencapai titik jenuh (knee point) dan bagaimana karakteristik degradasi stabilitasnya setelah titik tersebut?
3. Faktor fisik perangkat keras apa yang secara dominan menentukan batas stabilitas kluster Ceph pada infrastruktur heterogen?

1.3 Batasan Masalah

1. Penelitian ini hanya berfokus pada analisa arsitektur penyimpanan kluster terhadap performa dan stabilitas sistem Ceph di lingkungan laboratorium Informatika Universitas Islam Sunan Kalijaga dan tidak mencakup implementasi Ceph pada skala industri atau cloud.
2. Penelitian ini dilakukan dalam kerangka bantuan teknis untuk implementasi yang sedang dilakukan oleh rekan peneliti yang berfokus pada optimasi performa Ceph di laboratorium akademik, sehingga hanya mencakup konfigurasi kluster skala kecil dengan sumber daya terbatas.
3. Analisis performa dan stabilitas pada metrik yang dapat diukur secara langsung melalui pengujian benchmarking pada lapisan RADOS, mencakup

throughput, IOPS, commit latency per-OSD dan variabilitas bandwidth, tanpa membahas pengaruh faktor eksternal seperti load balancing dinamis atau optimasi algoritma distribusi data.

4. Konfigurasi kluster yang diuji terbatas pada setup node dasar Ceph (termasuk OSD, monitor dan MDS) dengan topologi standar, tanpa memperhitungkan optimasi berbasis perangkat keras khusus atau kustomisasi sistem secara mendalam.

1.4 Tujuan Penelitian

1. Memetakan profil performa mentah kluster Ceph pada infrastruktur heterogen sub-minimum ditinjau dari metrik *throughput* dan IOPS.
2. Mengidentifikasi titik jenuh operasional dan menganalisis karakteristik degradasi stabilitas sistem pasca-saturasi.
3. Mengidentifikasi dan mengkuantifikasi faktor fisik perangkat keras yang menentukan batas stabilitas kluster pada lingkungan heterogen.

1.5 Manfaat Penelitian

Secara praktis, penelitian ini memberikan kontribusi yang signifikan bagi pengelolaan sistem penyimpanan terdistribusi di lingkungan laboratorium komputasi, antara lain:

1. Validasi Performa pada Infrastruktur Heterogen: Membantu pengelola laboratorium memetakan dan memahami pengaruh variabilitas perangkat keras terhadap capaian *throughput*, stabilitas latensi dan batas aman operasional kluster Ceph.
2. Optimalisasi Lingkungan *Resource-Constrained*: Memberikan rekomendasi operasional dan ambang batas aman guna memaksimalkan utilitas perangkat keras yang ada (termasuk PC lawas) tanpa mengorbankan performa sistem secara keseluruhan.
3. Panduan Mitigasi Saturasi Sistem: Menyediakan rujukan praktis untuk mempertahankan stabilitas kluster dan reliabilitas data ketika

infrastruktur menghadapi lonjakan beban kerja (*workload spikes*) hingga mencapai titik jenuh.

4. Menjadi acuan empiris bagi praktisi DevOps dan administrator infrastruktur dalam memahami batas performa dan faktor-faktor fisik yang menentukan stabilitas infrastruktur penyimpanan terdistribusi berbasis perangkat keras komoditas.



BAB V

KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil pengujian benchmarking yang telah dilaksanakan terhadap kluster Ceph pada infrastruktur Laboratorium Informatika UIN Sunan Kalijaga, dapat dirumuskan kesimpulan sebagai berikut:

1. Profil performa mentah kluster pada kondisi *single-thread* menunjukkan throughput sekuensial tulis sebesar 3,39 MB/s dan baca sebesar 9,685 MB/s nilai yang secara empiris membuktikan bahwa batas absolut sistem ditentukan sepenuhnya oleh kecepatan mekanis HDD, bukan kapasitas jaringan. Utilisasi jaringan efektif yang hanya mencapai 2,9% dari kapasitas teoritis mengkonfirmasi kondisi *starved network* yang disebabkan oleh bottleneck lapisan penyimpanan, diperparah oleh mekanisme replikasi sinkronus Ceph yang mengikat seluruh operasi tulis pada *commit latency* node paling lambat.
2. Kluster secara konsisten mencapai titik jenuh (*knee point*) pada $t=16$ thread untuk seluruh beban kerja sekuensial. Setelah titik ini, sistem menunjukkan tiga karakteristik degradasi: fenomena *burst-stall* pada throughput, degradasi IOPS baca sebesar 41,3% dari puncak hingga $t=128$ dan peningkatan *tail latency* yang tidak proporsional seluruhnya konsisten dengan prediksi model USL dan Hukum Little bahwa overhead koordinasi telah mendominasi kapasitas layanan.
3. Tiga faktor fisik secara hierarkis menentukan batas stabilitas kluster: keterbatasan throughput mekanis HDD sebagai bottleneck primer, overhead koordinasi replikasi $\times 3$ sebagai bottleneck sekunder dan heterogenitas revisi firmware antar-OSD sebagai bottleneck tersier. Temuan orisinal penelitian ini adalah bahwa heterogenitas firmware terbukti lebih determinan dibandingkan heterogenitas generasi CPU dibuktikan oleh anomali OSD 2 yang mencatatkan *commit latency* terendah (62ms) meskipun berusia paling tua dan memiliki 8.448 *reallocated sectors* menghasilkan *Heterogeneity*

Penalty Factor sebesar 0,58 yang mengindikasikan sistem hanya beroperasi pada 58% potensi optimalnya.

Secara keseluruhan, penelitian ini membuktikan bahwa kluster Ceph dapat beroperasi secara fungsional pada infrastruktur sub-minimum dengan batas operasional aman pada $t=16$ thread dan bahwa homogenitas revisi firmware antar-OSD merupakan faktor optimasi paling efektif yang dapat diterapkan tanpa biaya perangkat keras tambahan.

5.2 Saran

5.2.1 Saran untuk Penelitian Selanjutnya

Penelitian ini membuka beberapa arah pengembangan yang dapat ditindaklanjuti oleh peneliti selanjutnya. Pertama, pengujian serupa dapat dilakukan menggunakan media penyimpanan NVMe SSD sebagai pengganti HDD mekanis pada konfigurasi infrastruktur yang identik, guna mengisolasi dan mengkuantifikasi secara empiris seberapa besar kontribusi bottleneck storage terhadap batas performa yang ditemukan dalam penelitian ini. Kedua, penelitian lanjutan dapat mengeksplorasi konfigurasi *erasure coding* sebagai alternatif replikasi faktor $\times 3$ mengingat *erasure coding* berpotensi menghasilkan karakteristik bottleneck yang berbeda pada infrastruktur heterogen sub-minimum. Ketiga, validasi temuan pada skala kluster yang lebih besar (minimum 8 node OSD) diperlukan untuk menguji apakah pola *straggler migration* dan *Heterogeneity Penalty Factor* yang ditemukan bersifat konsisten atau berubah seiring penambahan node khususnya pada lingkungan dengan tingkat heterogenitas perangkat keras yang lebih tinggi. Keempat, penggunaan *tool* benchmarking pada lapisan VFS seperti *fio* dan *dd* dapat memberikan perspektif komparatif terhadap hasil RADOS Bench yang beroperasi pada lapisan RADOS memungkinkan isolasi overhead yang disumbangkan oleh lapisan abstraksi Ceph terhadap performa aktual disk.

5.2.2 Saran untuk Praktisi

Bagi administrator laboratorium dan praktisi infrastruktur yang merencanakan implementasi sistem penyimpanan terdistribusi berbasis perangkat keras komoditas, penelitian ini merekomendasikan tiga hal utama. Pertama, lakukan verifikasi homogenitas revisi firmware seluruh disk OSD sebelum kluster

dioperasikan perbedaan revisi firmware pada model disk yang sama terbukti menghasilkan disparitas *commit latency* yang signifikan dan tidak dapat dieliminasi melalui penambahan node. Kedua, tetapkan batas konkurensi operasional pada $t=16$ thread untuk infrastruktur berbasis HDD pengoperasian di atas ambang ini tidak menghasilkan peningkatan throughput yang berarti namun meningkatkan risiko degradasi latensi secara signifikan. Ketiga, implementasikan pemantauan metrik *ceph osd perf* secara berkala untuk mendeteksi dini potensi *straggler* sebelum berdampak pada ketersediaan layanan disparitas *commit latency* antar-OSD di atas 50% merupakan indikator awal yang perlu diinvestigasi. Apabila peningkatan kapasitas diperlukan di masa mendatang, prioritaskan penggantian HDD ke SSD sebagai langkah pertama sebelum mempertimbangkan upgrade CPU atau RAM, mengingat bottleneck primer sistem bersifat *storage-bound*.



DAFTAR PUSTAKA

- [1] A. Santos, J. Martins, P. Pestana, R. Gonçalves, and H. São, “Factors affecting cloud computing adoption in the education context - Systematic Literature Review,” 2024, doi: 10.1109/ACCESS.2024.3400862.
- [2] B. Turkkan, E. Rodrigues, T. Kosar, A. Charapko, A. Ailijiang, and M. Demirbas, “How to Evaluate Distributed Coordination Systems?—A Survey and Analysis,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 37, no. 1, pp. 198–212, 2026, doi: 10.1109/TPDS.2025.3631614.
- [3] C. Yang, “Theoretical Analysis of Distributed Systems and Their Scalability,” *Adv. Comput. Signals Syst.*, vol. 9, no. 1, pp. 21–24, 2025, doi: 10.23977/acss.2025.090104.
- [4] P. P. -, “Comparative Analysis of Distributed Storage Systems: Architectural Design, Performance, and Cost Trade-offs in Modern Cloud Environments,” *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.*, vol. 13, no. 3, pp. 1–11, 2025, doi: 10.37082/ijirmps.v13.i3.232624.
- [5] S. Lim and D. Park, “Improving Hadoop MapReduce performance on heterogeneous single board computer clusters,” *Futur. Gener. Comput. Syst.*, vol. 160, pp. 752–766, Nov. 2024, doi: 10.1016/J.FUTURE.2024.06.025.
- [6] W. Reda, *Accelerating Distributed Storage in Heterogeneous Settings*. 2022.
- [7] B. Schwartz, “Universal Scalability Law,” pp. 1–52, 2015.
- [8] A. J. Peters and D. C. van der Ster, “Evaluating CephFS Performance vs. Cost on High-Density Commodity Disk Servers,” *Comput. Softw. Big Sci.*, vol. 5, no. 1, pp. 3–12, 2021, doi: 10.1007/s41781-021-00071-1.
- [9] S. B. Dada, E. G. and Joseph and Department, “A COMPARATIVE STUDY OF THE ARCHITECTURES AND APPLICATIONS OF SCALABLE HIGH-PERFORMANCE DISTRIBUTED FILE SYSTEMS,” vol. 3, no. 32, pp. 1–44, 2018.
- [10] J. Y. Lee, M. H. Kim, S. A. R. Shah, S. U. Ahn, H. Yoon, and S. Y. Noh, “Performance evaluations of distributed file systems for scientific big data

- in fuse environment,” *Electron.*, vol. 10, no. 12, pp. 1–16, 2021, doi: 10.3390/electronics10121471.
- [11] M. Vaidya and S. Deshpande, “Study and Experimental Setup of Testing of Performance Parameters on Various Distributed File Systems,” *Int. J. Comput. Appl.*, vol. 159, no. 9, pp. 23–27, 2017, doi: 10.5120/ijca2017913077.
- [12] M. S. Aslanpour, A. N. Toosi, M. A. Cheema, M. B. Chhetri, and M. A. Salehi, “Load balancing for heterogeneous serverless edge computing: A performance-driven and empirical approach,” *Futur. Gener. Comput. Syst.*, vol. 154, pp. 266–280, May 2024, doi: 10.1016/j.future.2024.01.020.
- [13] Z. Li and Y. Wang, “An adaptive read/write optimized algorithm for Ceph heterogeneous systems via performance prediction and multi-attribute decision making,” *Cluster Comput.*, vol. 26, no. 2, pp. 1125–1146, 2023, doi: 10.1007/s10586-022-03764-3.
- [14] L. Carracciuolo, D. Bottalico, D. Michelino, G. Sabella, and B. Spisso, “Benchmarking a High Performance Computing Heterogeneous Cluster,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13827 LNCS, pp. 101–114, 2023, doi: 10.1007/978-3-031-30445-3_9.
- [15] Z. Wang, E. Adamiak, and A. Aiken, “A Model for Query Execution Over Heterogeneous Instances,” *Cidr*, 2024.
- [16] J. Monniot, F. Tessier, M. Robert, and G. Antoniu, “Supporting dynamic allocation of heterogeneous storage resources on HPC systems,” *Concurr. Comput. Pract. Exp.*, vol. 35, no. 28, pp. 1–16, 2023, doi: 10.1002/cpe.7890.
- [17] J. Marquez and O. H. Mondragon, “Modelling the Impact of Cloud Storage Heterogeneity on HPC Application Performance,” *Computation*, vol. 12, no. 7, 2024, doi: 10.3390/computation12070150.
- [18] S. Chum, H. Park, and J. Choi, “Supporting sla via adaptive mapping and heterogeneous storage devices in ceph,” *Electron.*, vol. 10, no. 7, pp. 1–14, 2021, doi: 10.3390/electronics10070847.

- [19] M. Oh, “Design of Efficient and Fast Distributed Storage System for Low Latency Storage Devices,” 2020.
- [20] A. S. Tanenbaum, *Distributed Systems: Principles and Paradigms*. 2006. [Online]. Available: http://www.amazon.de/Distributed-Systems-Principles-Paradigms-Tanenbaum/dp/B006V36H4S/ref=sr_1_5?ie=UTF8&qid=1352386184&sr=8-5
- [21] S. A. Weil, S. A. Brandt, E. L. Miller, D. D. E. Long, and C. Maltzahn, “CEPH: A scalable, high-performance distributed file system,” *OSDI 2006 - 7th USENIX Symp. Oper. Syst. Des. Implement.*, pp. 307–320, 2006.
- [22] D. Gudu, M. Hardt, and A. Streit, “Evaluating the performance and scalability of the Ceph distributed storage system,” *Proc. - 2014 IEEE Int. Conf. Big Data, IEEE Big Data 2014*, pp. 177–182, Jan. 2015, doi: 10.1109/BigData.2014.7004229.
- [23] T. N. Mouratidis, “Optimizing the recovery of data consistency gossip algorithms on distributed object-store systems (CEPH) Βελτιστοποίηση αλγορίθμου φλυαρίας για συνεπή δεδομένα σε καταναμημένα συστήματα αποθήκευσης αντικειμένων (CEPH),” no. December, 2020.
- [24] G. Donvito, G. Marzulli, and D. Diacono, “Testing of several distributed file-systems (HDFS, Ceph and GlusterFS) for supporting the HEP experiments analysis,” *J. Phys. Conf. Ser.*, vol. 513, no. TRACK 4, 2014, doi: 10.1088/1742-6596/513/4/042014.
- [25] C. Proceedings, V. A. City, and G. M. Amdahl, “Validity of the Single Processor Approach to,” vol. 30, pp. 19–20, 2007.
- [26] J. D. C. Little and S. C. Graves, “Chapter 5 Little ’ s Law,” pp. 81–100, 2008, doi: 10.1007/978-0-387.
- [27] J. Blanchette, M. Summerfield, M. Gorman, A. J. Massa, and N. Mcfarlane, *Understanding the Linux* ®.
- [28] D. A. Patterson, G. Gibson, R. H. Katz, and E. Hall, “A Case for Redundant Arrays of Inexpensive Disks (RAID),” 1988.
- [29] M. Zaharia, A. Konwinski, A. D. Joseph, R. Katz, and I. Stoica,

- “Improving MapReduce Performance in Heterogeneous Environments,” pp. 29–42.
- [30] H. Botal and D. J. Dewitt, “A methodology for database system performance evaluation,” *Proc. ACM SIGMOD Int. Conf. Manag. Data*, pp. 176–185, 1984, doi: 10.1145/602259.602283.
- [31] S. Verlag, C. Wohlin, M. Höst, and K. Henningsson, “Engineering ”, In *Lecture Notes in Computer Science : Empirical Methods and Studies in Software Engineering : Experiences from ESERNET* , edited by A . I . Wang and R . Empirical Research Methods in Software Engineering,” 2003.
- [32] M. Simulation, W. C. Publishing, and J. Wiley, *Preface Acknowledgments PART I — AN OVERVIEW OF PERFORMANCE EVALUATION CHAPTER 1 — INTRODUCTION CHAPTER 2 — COMMON MISTAKES AND HOW TO AVOID CHAPTER 3 — SELECTION OF TECHNIQUES AND METRICS FURTHER READING FOR PART I PART II — MEASUREMENT TECHNIQUES AND TOOLS CHAPTER 4 — TYPES OF WORKLOADS CHAPTER 6 — WORKLOAD CHARACTERIZATION.*
- [33] J. Dean and S. Ghemawat, “MapReduce : Simplified Data Processing on Large Clusters,” pp. 1–13.