



UNIVERSITAS INDONESIA

**EKSPERIMENT *END-TO-END* UNTUK KAFKA RESOURCE
MANAGEMENT PADA PENELITIAN “EFFICIENT TOPIC
PARTITIONING OF APACHE KAFKA FOR HIGH-RELIABILITY
REAL-TIME DATA STREAMING APPLICATIONS” DI LAB CSL
FAKULTAS ILMU KOMPUTER UI**

LAPORAN KERJA PRAKTIK

**BRYAN RAIHAN 'ILMAN
2106704351**

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
DEPOK
AGUSTUS 2024**



UNIVERSITAS INDONESIA

**EKSPERIMENT END-TO-END UNTUK KAFKA RESOURCE
MANAGEMENT PADA PENELITIAN “EFFICIENT TOPIC
PARTITIONING OF APACHE KAFKA FOR HIGH-RELIABILITY
REAL-TIME DATA STREAMING APPLICATIONS” DI LAB CSL
FAKULTAS ILMU KOMPUTER UI**

LAPORAN KERJA PRAKTIK

Diajukan sebagai salah satu syarat kelulusan mata kuliah
Kerja Praktik

**BRYAN RAIHAN 'ILMAN
2106704351**

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
DEPOK
AGUSTUS 2024**

HALAMAN PERSETUJUAN DOSEN KERJA PRAKTIK

Laporan Kerja Praktik ini diajukan oleh:

Nama : Bryan Raihan 'Ilman

NPM : 2106704351

Program Studi : Ilmu Komputer

Judul Kerja Praktik : Eksperimen *End-To-End* untuk *Kafka Resource Management* pada penelitian “Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications” di lab CSL Fakultas Ilmu Komputer UI

Telah berhasil diselesaikan laporan kerja praktik untuk Fakultas Ilmu Komputer dan dipresentasikan hasil kerja praktiknya sebagai persyaratan yang harus dipenuhi dalam mata kuliah Kerja Praktik.

DOSEN MATA KULIAH KERJA PRAKTIK

Ichlasul Affan, M.Kom.

Ditetapkan di : Depok

Tanggal : 25 Agustus 2024

ABSTRAK

Nama : Bryan Raihan 'Ilman
Program Studi : Ilmu Komputer
Judul : Eksperimen *End-To-End* untuk *Kafka Resource Management* pada penelitian “Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications” di lab CSL Fakultas Ilmu Komputer UI
Pembimbing : Ichlasul Affan, M.Kom.

Streaming data secara *real-time* adalah pengiriman data kontinu dari sumber ke penerima, sering digunakan dalam aplikasi yang memerlukan respons cepat dan aliran data besar. Apache Kafka mengelola proses ini dengan mempartisi data menjadi topik untuk pemrosesan paralel. Kerja praktik ini berfokus pada penelitian optimisasi partisi topik Apache Kafka, sebagaimana dijelaskan dalam paper berjudul “Efficient Topic Partitioning of Apache Kafka for High-Reliability Real-Time Data Streaming Applications”. Pelaksana kerja praktik berperan sebagai *Research Assistant* di Computer Systems Laboratory (CSL) Fakultas Ilmu Komputer Universitas Indonesia, sebuah laboratorium yang fokus pada penelitian komputasi kinerja tinggi dan manajemen *big data*. Tugas utamanya adalah menjalankan eksperimen *end-to-end*, yang dalam konteks ini berarti mengamati aliran data dari produser hingga konsumen untuk mengukur latensi dari di seluruh proses. Metodologi yang digunakan dalam kerja praktik ini mencakup beberapa langkah penting. Di awal, pelaksana menyiapkan VM di Google Cloud Platform dan menginstal semua dependensi yang dibutuhkan untuk eksperimen Kafka. Setelah itu, eksperimen dijalankan dengan satu VM yang bertindak sebagai klien Kafka dan satu lagi sebagai klaster Kafka. Eksperimen difokuskan pada pengujian algoritma BroMax dan BroMin untuk memantau perilaku partisi topik. Algoritma tersebut digunakan untuk secara matematis menentukan jumlah partisi yang optimal demi mencapai *throughput* maksimal. Selain itu, *resource monitoring* dilakukan secara rutin untuk melacak penggunaan CPU, RAM, dan *disk space* pada VM. Eksperimen lalu diulang pada infrastruktur UI. Metodologi eksperimental kuantitatif diterapkan dalam proyek ini, menggunakan berbagai teknologi termasuk Apache Kafka untuk streaming data, Docker dan Linux untuk containerization, serta Google Cloud Platform untuk infrastruktur. Selain itu, Python dan Bash digunakan dalam automasi proses eksperimen. Hasil penelitian ini mencakup pengembangan *framework* berbasis *containerization* yang mengotomatisasi proses *experiment setup* sehingga mempermudah pengujian tanpa perlu konfigurasi manual, dan pemahaman mendalam tentang perilaku *latency* dan *throughput* di bawah berbagai *application constraints*.

Kata kunci:

Fog Computing, Data Streaming, Topic Partitioning, Distributed Systems

ABSTRACT

Name : Bryan Raihan 'Ilman
Study Program : Computer Science
Title : End-To-End Experiments for Kafka Resource Management in the research "Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications" di lab CSL Fakultas Ilmu Komputer UI
Counselor : Ichlasul Affan, M.Kom.

Real-time data streaming is the continuous transmission of data from the source to the receiver, often used in applications requiring fast responses and large data flows. Apache Kafka manages this process by partitioning data into topics for parallel processing. This internship focused on researching the optimization of Apache Kafka topic partitioning, as outlined in the paper titled "Efficient Topic Partitioning of Apache Kafka for High-Reliability Real-Time Data Streaming Applications." The intern served as a Research Assistant at the Computer Systems Laboratory (CSL) of the Faculty of Computer Science, Universitas Indonesia, a laboratory focused on high-performance computing and big data management research. The primary task was to conduct end-to-end experiments, which in this context means observing the flow of data from producers to consumers to measure latency throughout the process. The methodology used in this internship included several key steps. Initially, the intern set up VMs on Google Cloud Platform and installed all necessary dependencies for Kafka experiments. After that, experiments were run with one VM acting as a Kafka client and another as the Kafka cluster. The experiments focused on testing the BroMax and BroMin algorithms to monitor the behavior of topic partitioning. These algorithms were used to mathematically determine the optimal number of partitions to achieve maximum throughput. Additionally, resource monitoring was regularly conducted to track CPU, RAM, and disk space usage on the VMs. The experiments were then repeated on UI infrastructure. A quantitative experimental methodology was applied in this project, using various technologies including Apache Kafka for data streaming, Docker and Linux for containerization, and Google Cloud Platform for infrastructure. Additionally, Python and Bash were used to automate the experimental processes. The research outcomes include the development of a containerization-based framework that automates the experiment setup, simplifying testing without manual configuration, and a deep understanding of latency and throughput behavior under different application constraints.

Keywords:

Fog Computing, Data Streaming, Topic Partitioning, Distributed Systems

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN DOSEN KERJA PRAKTIK	ii
ABSTRAK	iii
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	viii
DAFTAR LAMPIRAN	viii
BAB 1 PENDAHULUAN	1
1.1 Proses Pencarian Kerja Praktik	1
1.2 Tempat Kerja Praktik	1
1.2.1 Profil Tempat Kerja Praktik	2
1.2.2 Posisi Penempatan Pelaksana Kerja Praktik dalam Struktur Organisasi	2
BAB 2 ISI	3
2.1 Latar Belakang Pekerjaan	3
2.2 Deskripsi Pekerjaan	3
2.3 Tinjauan Pustaka	4
2.3.1 Apache Kafka	4
2.3.2 Docker Containerization	4
2.3.3 Apache Zookeeper	4
2.3.4 Distributed System	5
2.4 Metodologi Pekerjaan	5
2.5 Teknologi	7
2.6 Hasil Pekerjaan	8
2.6.1 Framework Berbasis Containerization dengan Pendekatan Docker-in-Docker (DinD)	8
2.6.2 Analisis Eksperimen End-To-End	9
2.7 Aspek Non Teknis	10
2.7.1 Komunikasi	11
2.7.2 Kerja Sama Tim	12
2.8 Analisis Pelaksanaan Kerja Praktik	12
2.8.1 Ulasan kesesuaian dan perbedaan dengan KAKP	13
2.8.2 Ulasan tentang kendala dan cara menanganinya	16
2.8.2.1 Tidak Memiliki Akses Root pada Infrastruktur UI	16
2.8.2.2 Dokumentasi dari Peneliti Asli Kurang Lengkap atau Usang	17
2.8.2.3 Penggunaan Resource yang Besar	17
2.8.3 Penilaian Individu Terhadap Tempat Kerja Praktik	18
2.8.4 Relevansi dengan Perkuliahan	18
2.8.4.1 Sistem Operasi	18
2.8.4.2 Jaringan Komputer	18
2.8.4.3 Pemrograman Lanjut	18

BAB 3 PENUTUP	20
3.1 Kesimpulan	20
3.2 Saran	20
3.2.1 Pelaksana Kerja Praktik Berikutnya	20
3.2.2 Fakultas Ilmu Komputer UI	21
DAFTAR REFERENSI	22

DAFTAR GAMBAR

Gambar 1.1.	Struktur Organisasi Pusat Penelitian Fakultas Ilmu Komputer UI	2
Gambar 2.1.	Eksekusi Eksperimen <i>End-To-End</i>	6
Gambar 2.2.	Analisis Eksperimen <i>End-To-End</i>	9
Gambar 2.3.	Email Pertama dari Profesor Claudio	11

DAFTAR TABEL

Tabel 2.1. Perbandingan Rencana dan Realisasi Kerja Praktik 13

DAFTAR LAMPIRAN

Lampiran 1. KERANGKA ACUAN KERJA PRAKTIK	24
Lampiran 2. LOG KERJA PRAKTIK	30
Lampiran 3. DOKUMENTASI REPRODUKSI EKSPERIMEN	49

BAB 1

PENDAHULUAN

Pada bab ini dijelaskan mengenai tempat pelaksanaan kerja praktik. Pertama dijelaskan mengenai proses pencarian kerja praktik. Kemudian, dijelaskan mengenai profil tempat pelaksanaan kerja praktik. Selain itu, dijelaskan juga mengenai posisi penempatan di tempat kerja praktik.

1.1 Proses Pencarian Kerja Praktik

Proses pencarian kerja praktik dimulai dengan pelaksana kerja praktik meminta izin kepada dosen pembimbing akademik untuk mengisi liburan musim panas dengan menjadi *Research Assistant* di salah satu lab di Fakultas Ilmu Komputer Universitas Indonesia. Dosen pembimbing akademik mendukung penuh keputusan ini dan menekankan bahwa ini adalah cara yang produktif untuk memanfaatkan liburan dan penting bagi pengembangan keahlian teknis serta kemampuan analisis. Dengan dorongan tersebut, pelaksana kerja praktik mulai mencari lab yang sedang membuka lowongan.

Pelaksana kerja praktik mengajukan lamaran untuk menjadi *Research Assistant* di salah satu lab di Fakultas Ilmu Komputer UI. Proses seleksi dimulai dengan wawancara *online* yang dilakukan bersama dosen peneliti dari lab yang dituju. Pada saat wawancara, dosen peniliti menjelaskan beberapa topik penelitian yang sedang berlangsung di lab tersebut, termasuk topik-topik yang berfokus pada *computer networks protocols*, *high-performance computing*, dan *real-time distributed systems*. Pelaksana kerja praktik memutuskan untuk terlibat dalam penelitian terkait *real-time distributed systems*, sebuah topik yang sangat relevan dengan pengelolaan data streaming dalam skala besar.

Setelah memilih topik ini, pelaksana kerja praktik mengikuti proses administrasi lebih lanjut untuk secara resmi diterima sebagai *Research Assistant* di lab tersebut.

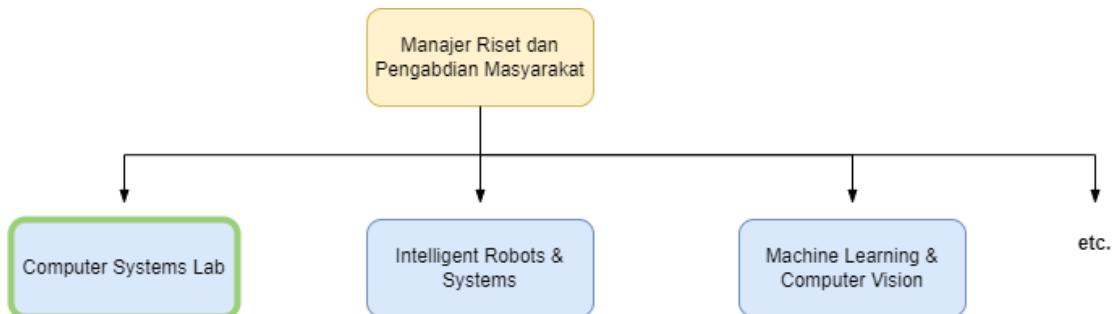
1.2 Tempat Kerja Praktik

Bagian ini menjelaskan tempat pelaksanaan kerja praktik secara rinci, meliputi profil lab, struktur organisasi, dan penempatan pelaksana kerja praktik selama periode kerja praktik.

1.2.1 Profil Tempat Kerja Praktik

Kerja praktik dilaksanakan di Computer Systems Laboratory (CSL) yang berlokasi di Fakultas Ilmu Komputer, Universitas Indonesia. CSL merupakan pusat penelitian yang fokus pada pengembangan teknologi komputasi kinerja tinggi, manajemen *big data*, dan aplikasi berbasis textitcloud. Lab ini mendukung berbagai penelitian yang sedang dijalankan mahasiswa baik untuk mendapatkan gelar Sarjana ataupun Magister. CSL tidak hanya menjadi tempat bagi penelitian akademik, tetapi juga berperan dalam kolaborasi lintas disiplin untuk menghasilkan inovasi di dunia teknologi.

1.2.2 Posisi Penempatan Pelaksana Kerja Praktik dalam Struktur Organisasi



Gambar 1.1: Struktur Organisasi Pusat Penelitian Fakultas Ilmu Komputer UI

Dijelaskan pada Gambar 1.1 bahwa Computer Systems Laboratory (CSL) berada di bawah supervisi Manajer Riset dan Pengabdian Masyarakat, bersama dengan beberapa pusat penelitian lainnya, seperti Intelligent Robots & Systems (IRoS) dan Machine Learning & Computer Vision (MLCV). Struktur organisasi di CSL sendiri bersifat flat, di mana setiap penelitian dilakukan secara independen oleh masing-masing tim atau individu. Namun ada dosen peneliti yang bertanggung jawab memantau seluruh proyek untuk memastikan semua penelitian berjalan sesuai standar akademik.

Pelaksana kerja praktik ditempatkan dalam tim yang mengerjakan penelitian manajemen cluster Kafka, yang berfokus pada optimasi performa sistem distribusi data *real-time*. Sebagai *Research Assistant*, pelaksana terlibat langsung dalam berbagai tahap eksperimen, mulai dari penyiapan infrastruktur hingga analisis hasil.

BAB 2

ISI

Pada bab ini dijelaskan secara rinci mengenai detil pekerjaan yang dilakukan selama program kerja praktik berlangsung. Latar belakang pekerjaan hingga hasil pekerjaan dijelaskan pada bab ini.

2.1 Latar Belakang Pekerjaan

Pelaksana kerja praktik terlibat dalam penelitian terkait manajemen klaster Kafka di Computer Systems Laboratory (CSL) Fakultas Ilmu Komputer Universitas Indonesia. Penelitian ini berfokus pada optimalisasi partisi topik untuk meningkatkan performa *real-time distributed systems* berbasis Kafka. Proyek ini merupakan bagian dari penelitian bertajuk *Efficient Topic Partitioning of Apache Kafka for High-Reliability Real-Time Data Streaming Applications*. Tugas pelaksana kerja praktik adalah menjalankan eksperimen yang bertujuan untuk menguji algoritma yang digunakan untuk menentukan jumlah partisi optimal, berupa BroMax dan BroMin (Raptis et al., 2024).

2.2 Deskripsi Pekerjaan

Pelaksana kerja praktik bertanggung jawab atas konfigurasi dan pengelolaan eksperimen *end-to-end*. Proses dimulai dengan penyiapan VM di Google Cloud Platform (GCP) dan infrastruktur Universitas Indonesia (DGX dan BMKGCS), yang digunakan untuk menjalankan simulasi klaster Kafka. Tugas utama meliputi instalasi dependencies, pengaturan Kafka broker, serta analisis hasil eksperimen untuk mengukur performa sistem, khususnya *throughput* dan latensi, saat algoritma BroMax dan BroMin diterapkan.

Pekerjaan dilakukan secara *hybrid* dari pukul 08.00 a.m. WIB hingga pukul 12.00 p.m. WIB, dengan kehadiran di lab hanya pada hari Jumat, sedangkan hari-hari lainnya bekerja secara *remote*. Hasil eksperimen dianalisis dan dilaporkan kepada dosen peneliti untuk evaluasi lebih lanjut.

2.3 Tinjauan Pustaka

Penelitian yang dilakukan selama kerja praktik ini melibatkan sejumlah konsep penting dalam pengelolaan *real-time distributed systems* dengan menggunakan Apache Kafka dan teknologi pendukung lainnya. Konsep-konsep ini menjadi fondasi utama dalam penelitian dan pengembangan yang dilakukan.

2.3.1 Apache Kafka

Apache Kafka adalah platform *open-source* yang dirancang untuk menangani *data streaming* secara *real-time* dengan *throughput* tinggi dan latensi rendah. Data dibagi menjadi topik-topik yang dipartisi, sehingga memungkinkan pemrosesan paralel yang lebih efisien di seluruh klaster (Raptis et al., 2024). Partisi pada topik Kafka sangat penting karena memungkinkan distribusi beban kerja yang merata antar *broker*, sehingga meningkatkan kemampuan sistem untuk menangani data dalam jumlah besar tanpa menimbulkan *bottleneck* (Raptis et al., 2024). Selain itu, dengan optimalisasi jumlah partisi yang tepat, *throughput* sistem dapat dimaksimalkan, sehingga menjaga kinerja tinggi meskipun volume data yang diproses sangat besar (Raptis et al., 2024).

2.3.2 Docker Containerization

Docker Containerization adalah teknologi yang memungkinkan aplikasi dijalankan dalam *container* yang terisolasi, sehingga dapat memisahkan lingkungan eksekusi aplikasi dari sistem *host*. Docker memungkinkan konsistensi lingkungan, meminimalkan konflik konfigurasi, serta meningkatkan portabilitas sistem (Kul et al., 2024). Dengan Docker, seluruh proses *deployment* Kafka menjadi lebih efisien karena setiap *node* dalam klaster dapat dijalankan dengan konfigurasi yang seragam, sehingga mengurangi kesalahan operasional (Kul et al., 2024).

2.3.3 Apache Zookeeper

Apache Zookeeper adalah layanan koordinasi terpusat yang digunakan untuk mengelola *metadata* dan menjaga konsistensi di dalam klaster Kafka. Zookeeper bertugas mengatur eleksi pemimpin partisi, menyinkronkan data antar broker, serta memastikan bahwa setiap perubahan dalam klaster Kafka ditangani dengan cepat dan efisien (Pham et al., 2014). Dalam konteks Kafka, Zookeeper memainkan peran penting dalam memastikan

bahwa sistem tetap *fault-tolerant*. Jika terjadi kegagalan pada salah satu broker, Zookeeper secara otomatis memilih broker lain untuk menjadi pemimpin, sehingga menjaga kestabilan sistem (Pham et al., 2014). Zookeeper juga menyimpan informasi konfigurasi mengenai topik dan partisi Kafka, dan memastikan bahwa produser dan konsumen dapat berinteraksi dengan sistem secara konsisten (Pham et al., 2014).

2.3.4 Distributed System

Distributed System adalah sistem yang terdiri dari beberapa *node* yang terhubung dan bekerja bersama untuk mencapai tujuan tertentu. Dalam konteks Kafka, sistem terdistribusi ini memastikan bahwa data dapat diolah secara paralel dan skalabel di berbagai broker yang tersebar dalam klaster. Keunggulan dari sistem terdistribusi seperti Kafka adalah kemampuannya untuk tetap berfungsi meskipun terjadi kegagalan pada satu atau beberapa *node* dalam klaster (Ledmi et al., 2018). Dalam sistem terdistribusi Kafka, partisi topik didistribusikan ke berbagai broker, sehingga meningkatkan ketersediaan data dan mempercepat proses pengiriman data dari produser ke konsumen (Ledmi et al., 2018). Sistem ini juga memungkinkan replikasi data secara otomatis, di mana salinan data disimpan di beberapa broker untuk meningkatkan keandalan dan redundansi sistem (Ledmi et al., 2018).

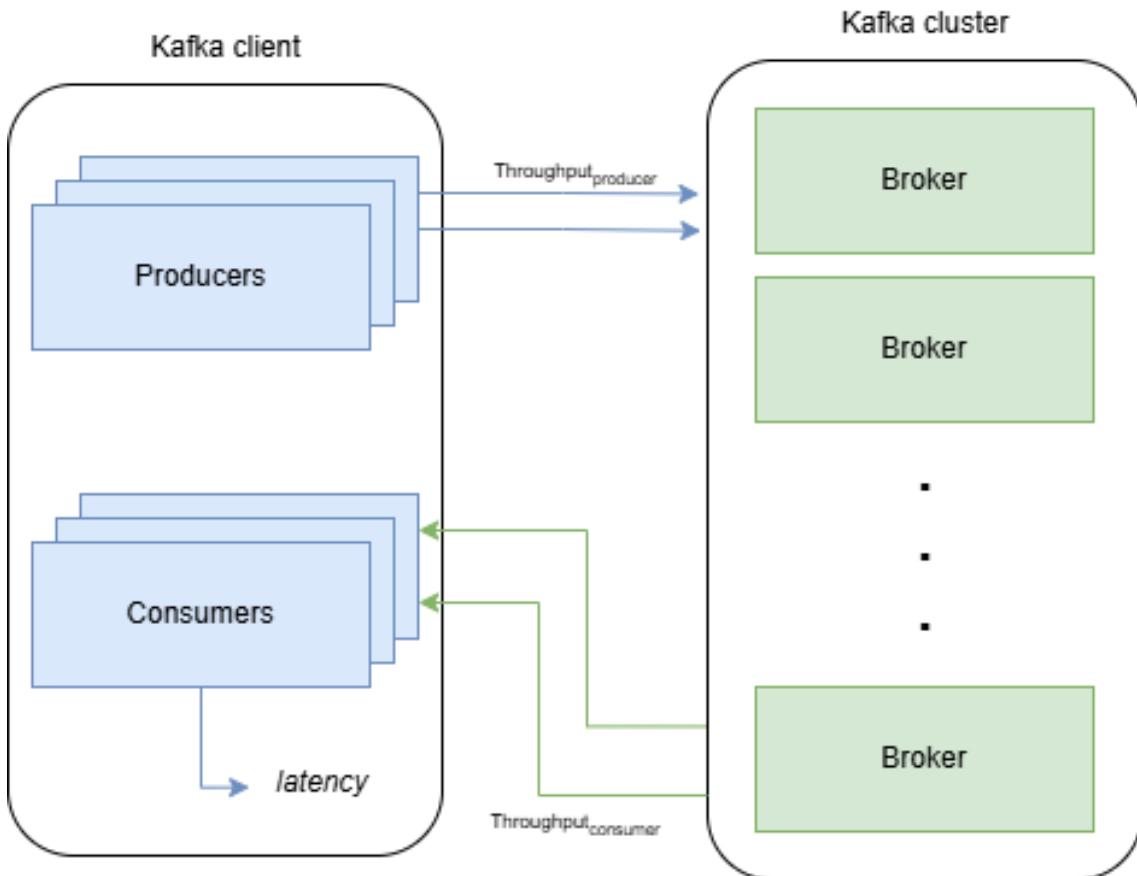
2.4 Metodologi Pekerjaan

Metodologi eksperimental kuantitatif dalam proyek ini melibatkan pengumpulan dan analisis data numerik secara end-to-end untuk mengukur performa Apache Kafka pada berbagai konfigurasi sistem. Tahapan pekerjaan yang dilaksanakan secara bertahap sesuai dengan proses yang dirancang, mulai dari pengaturan infrastruktur hingga analisis hasil eksperimen. Berikut adalah tahapan metodologi pekerjaan yang dilakukan selama kerja praktik:

1. **Pengaturan VM.** Langkah awal dalam kerja praktik ini adalah menyiapkan VM (Virtual Machine) di Google Cloud Platform (GCP). VM tersebut dikonfigurasi menggunakan SSH dengan akses *root*. Setelah VM disiapkan, seluruh dependensi yang diperlukan untuk menjalankan eksperimen Kafka diinstal. Konfigurasi ini memastikan bahwa VM siap untuk menjalankan berbagai eksperimen Kafka, di mana satu VM bertindak sebagai Kafka *client*, dan VM lainnya berperan sebagai Kafka

cluster. Seluruh pengaturan ini bertujuan untuk mendukung kelancaran eksperimen yang dilakukan.

2. Eksekusi Eksperimen End-To-End.



Gambar 2.1: Eksekusi Eksperimen *End-To-End*

Setelah pengaturan infrastruktur selesai, eksperimen dilakukan dengan fokus pada simulasi *end-to-end* yang bertujuan untuk memantau algoritma BroMax dan BroMin dalam menentukan jumlah partisi dan broker pada sistem Apache Kafka. Terlihat pada Gambar 2.1 bahwa eksperimen ini melibatkan pengaturan dua VM dengan fungsi yang berbeda: satu VM bertindak sebagai Kafka client yang mengirimkan dan menerima data, sedangkan VM lainnya bertindak sebagai Kafka cluster berisi broker yang mengelola partisi. Eksperimen ini penting untuk memahami perilaku *throughput* dan latensi sistem secara keseluruhan dalam kondisi yang mendekati *real-time streaming*. Hasil dari eksperimen ini memberikan wawasan tentang cara kerja algoritma BroMax dan BroMin dalam optimasi partisi.

3. Pemantauan Sumber Daya.

Selama eksperimen berlangsung, pemantauan sumber

daya dilakukan secara berkala untuk memantau penggunaan CPU, RAM, dan *disk space* pada VM. Hal ini penting untuk memastikan bahwa eksperimen berjalan dengan efisien tanpa ada *bottleneck* dalam pemanfaatan sumber daya. Data yang dihasilkan dari pemantauan sumber daya ini juga menjadi acuan dalam mengevaluasi performa algoritma BroMax dan BroMin.

4. **Eksperimen Infrastruktur UI.** Setelah eksperimen di GCP berhasil dilaksanakan, eksperimen serupa diulang pada infrastruktur UI menggunakan mesin DGX dan BMKGCS tanpa akses *root*. Eksperimen ini dilakukan untuk memastikan bahwa hasil eksperimen dapat direplikasi di lingkungan yang berbeda dengan kapasitas sumber daya yang lebih besar.
5. **Implementasi Containerization dengan Docker.** Dalam eksperimen ini, Docker digunakan untuk membangun dan mengelola *container* yang berisi aplikasi Apache Kafka. Pelaksana kerja praktik mempelajari cara kerja Docker, mulai dari membangun *image* hingga pengaturan komunikasi antar *container*. Dengan *containerization*, eksperimen dapat dijalankan pada *host* tanpa akses *root* sekalipun. Eksperimen ini juga mengeksplorasi penggunaan teknologi Docker-in-Docker (DinD) ketika membuat Kafka *cluster container* yang di dalamnya berisi beberapa *sbroker containers*.
6. **Analisis Data dan Pelaporan.** Setelah seluruh eksperimen selesai dijalankan, data hasil eksperimen dianalisis untuk membuat plot yang menggambarkan performa *throughput* dan latensi sistem. Plot ini kemudian digunakan untuk mendeskripsikan perilaku algoritma BroMax dan BroMin dalam kondisi eksperimen yang telah dilakukan. Hasil analisis ini kemudian disajikan dalam laporan kepada dosen peneliti, serta didiskusikan dengan para peneliti lainnya untuk mendapatkan umpan balik dan evaluasi lebih lanjut mengenai hasil eksperimen.

2.5 Teknologi

Selama pengerjaan proyek pada kerja praktik, beberapa teknologi utama digunakan untuk mendukung penelitian. Berikut adalah teknologi-teknologi tersebut:

- **Apache Kafka.** Sebuah platform yang digunakan untuk memproses data secara paralel dan *real-time* melalui topik yang dipartisi. Kafka memungkinkan penanganan volume data yang sangat besar dengan *throughput* tinggi dan latensi rendah, serta mendukung replikasi data untuk menjaga keandalan sistem.

- **Docker**, Merupakan platform *containerization* yang memungkinkan pengembang menjalankan aplikasi dalam lingkungan terisolasi yang disebut container. Docker digunakan untuk menyederhanakan *deployment* sistem Apache Kafka dan komponen lainnya selama eksperimen. Docker memastikan lingkungan yang seragam dan memudahkan migrasi antar infrastruktur.
- **Linux**, Sistem operasi *open-source* yang digunakan dalam menjalankan berbagai komponen pada infrastruktur server. Linux menawarkan stabilitas dan kontrol penuh terhadap manajemen sumber daya, yang sangat penting untuk menjalankan aplikasi terdistribusi seperti Kafka dan Docker.
- **Bash**, Bash adalah bahasa skrip yang digunakan untuk mengotomatisasi berbagai tugas di lingkungan Linux. Dalam proyek ini, Bash digunakan untuk menjalankan skrip yang mengelola proses *deployment*, *system monitoring*, dan pemeliharaan sumber daya.
- **Python**, Sebuah bahasa pemrograman yang digunakan untuk menulis skrip otomasi dan pengolahan data. Python sangat fleksibel dan mendukung banyak *library* untuk analisis data, pemantauan sistem, dan integrasi dengan platform lain seperti Kafka.
- **Google Cloud Platform (GCP)**, GCP adalah platform komputasi awan yang menyediakan infrastruktur untuk menjalankan VM yang mendukung eksperimen Kafka. GCP digunakan untuk skala dan fleksibilitas, memungkinkan pengembang menguji aplikasi di lingkungan cloud yang efisien.
- **WhatsApp**, Sebagai aplikasi perpesanan, WhatsApp digunakan dalam konteks komunikasi informal antara tim dan dosen peneliti selama proyek.

2.6 Hasil Pekerjaan

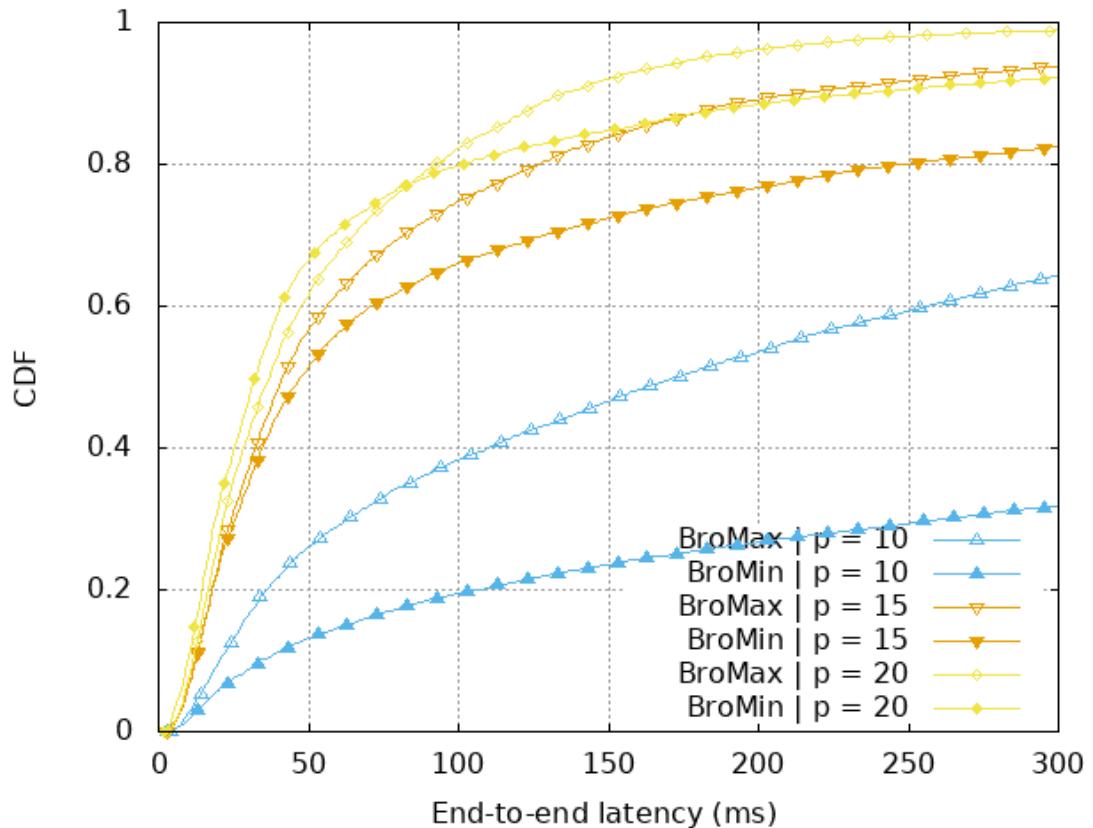
Selama periode kerja praktik, pelaksana terlibat dalam tugas-tugas yang menuntut analisis mendalam terkait *real-time data streaming* menggunakan Apache Kafka. Pada bagian ini dijelaskan tugas apa saja yang sudah diselesaikan pada periode kerja praktik berlangsung.

2.6.1 Framework Berbasis Containerization dengan Pendekatan Docker-in-Docker (DinD)

Pelaksana kerja praktik mengusulkan ide kepada dosen peneliti untuk mengembangkan *framework* berbasis *containerization* dengan pendekatan Docker-in-Docker (DinD). *Framework* ini dirancang agar dapat mengotomatisasi seluruh proses *setup*, termasuk pembuatan Kafka *cluster container* yang di dalamnya berisi beberapa *broker containers*,

tanpa perlu akses *root* pada *host*, sehingga pengguna hanya perlu menjalankan sedikit skrip tanpa perlu mengulang pengaturan secara manual. Dengan adanya *framework* ini, pelaksana kerja praktik dapat langsung fokus pada analisis eksperimen tanpa harus khawatir tentang masalah teknis terkait lingkungan yang berulang kali harus disesuaikan. *Framework* ini dapat diakses oleh siapa saja melalui <https://github.com/bryan-ilman-on-github/kafka-part-exp>.

2.6.2 Analisis Eksperimen End-To-End



Gambar 2.2: Analisis Eksperimen End-To-End

Salah satu temuan utama dari eksperimen ini adalah bahwa metode penentuan jumlah partisi sangat memengaruhi *throughput* dan latensi. Gambar 2.2 menampilkan CDF, yaitu garis kurva yang menunjukkan persentase kumulatif latency pada *millisecond* tertentu. Garis yang lebih tinggi di awal menunjukkan lebih banyak latency rendah, yang berarti performa lebih baik. Terlihat pada Gambar 2.2 bahwa algoritma BroMax, yang aktif memaksimalkan penggunaan semua sumber daya yang tersedia untuk mendistribusikan beban kerja, berhasil mempertahankan latensi yang rendah sehingga menghasilkan

throughput yang relatif lebih tinggi. Sebaliknya, algoritma BroMin, yang lebih konservatif dalam penggunaan sumber daya, cenderung memiliki latensi yang sedikit lebih tinggi, sehingga *throughput*-nya relatif lebih rendah.

Selain itu, terlihat pula pada Gambar 2.2 bahwa jumlah produsen yang lebih banyak justru mengurangi latensi secara keseluruhan, berlawanan dengan intuisi awal yang menganggap bahwa lebih banyak produsen akan meningkatkan beban klaster dan menyebabkan latensi lebih tinggi. Setelah diskusi dengan tim, didapat penjelasan untuk temuan ini bahwa tingginya *overhead* pada klaster dengan sedikit produsen tidak ideal untuk beban kerja yang sebenarnya. Sebaliknya, pada klaster dengan lebih banyak produsen, *overhead* menjadi tidak signifikan dibandingkan dengan beban kerja yang ada, sehingga performanya lebih stabil. Hasil eksperimen ini memberikan wawasan yang lebih mendalam tentang pengoptimalan konfigurasi klaster Kafka.

2.7 Aspek Non Teknis

Dalam pelaksanaan kerja praktik, terdapat berbagai pembelajaran non-teknis yang turut berperan penting dalam mendukung keberhasilan tugas. Aspek-aspek seperti komunikasi dan kerja sama tim menjadi bagian integral dari keseluruhan pengalaman kerja praktik.

2.7.1 Komunikasi



Gambar 2.3: Email Pertama dari Profesor Claudio

Selama kerja praktik, komunikasi memainkan peranan yang sangat penting dalam memastikan kelancaran proyek dan kesesuaian hasil dengan tujuan yang ditetapkan. Salah satu contoh penting dalam hal ini adalah komunikasi yang dilakukan melalui email dengan Profesor Claudio Cicconetti dari University of Pisa, penulis penelitian yang eksperimennya sedang pelaksana kerja praktik coba untuk replikasi. Terlihat pada Gambar 2.3 bahwa pelaksana kerja praktik pertama kali menginformasikan kepada Profesor Cicconetti bahwa eksperimennya sedang direproduksi. Setelah itu, email-email selanjutnya berfokus pada klarifikasi mengenai beberapa aspek teknis dari penelitiannya,

dan beliau dengan ramah memberikan tanggapan serta penjelasan yang sangat membantu. Hal ini menekankan pentingnya komunikasi yang efektif dalam lingkungan akademik dan profesional, terutama saat bekerja dengan pihak eksternal yang memiliki keahlian khusus dalam bidang tertentu.

2.7.2 Kerja Sama Tim

Kerja sama tim merupakan aspek lain yang sangat penting dalam pelaksanaan proyek selama kerja praktik. Dalam mengerjakan proyek, tidak ada tugas yang dilakukan sepenuhnya secara mandiri. Setiap anggota tim memiliki peran spesifik yang saling berhubungan, dan keberhasilan satu bagian proyek sering kali tergantung pada penyelesaian bagian lain yang dikerjakan oleh anggota tim lain.

2.8 Analisis Pelaksanaan Kerja Praktik

Pada bagian ini dijelaskan analisis menyeluruh mengenai pelaksanaan kerja praktik. Penjelasan dimulai dengan mengkaji kesesuaian pelaksanaan kerja praktik terhadap kerangka acuan kerja praktik yang telah disusun. Selain itu, kendala-kendala yang muncul selama kerja praktik berlangsung juga akan diuraikan, beserta solusi yang diterapkan untuk mengatasi hambatan tersebut.

2.8.1 Ulasan kesesuaian dan perbedaan dengan KAKP

Tabel 2.1: Perbandingan Rencana dan Realisasi Kerja Praktik

Waktu	Rencana Kerja	Realisasi Kerja
Minggu 1 (3 Juni 2024 – 7 Juni 2024)	<p>Meninjau jurnal dari Elsevier dan mencari penelitian terbaru mengenai Apache Kafka yang dapat diterapkan kembali di CSL.</p> <ul style="list-style-type: none"> – Menganalisis metodologi dan hasil dari paper berjudul "Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications". <p>Bergabung dengan Komunitas FogVerse.</p> <ul style="list-style-type: none"> – Bergabung dengan komunitas FogVerse di WhatsApp dan Discord. 	Sesuai dengan KAKP
Minggu 2 (10 Juni 2024 – 14 Juni 2024)	<p>Merumuskan langkah-langkah untuk menerapkan kembali isi paper.</p> <ul style="list-style-type: none"> – Membaca dan memahami paper yang akan diterapkan kembali. – Merancang langkah-langkah secara detail. <p>Mengikuti perkembangan proyek flagship CSL, yaitu FogVerse.</p> <ul style="list-style-type: none"> – Mempelajari kode FogVerse dari GitHub. – Mengunduh kode FogVerse. 	Sesuai dengan KAKP
Minggu 3 (17 Juni 2024 – 21 Juni 2024)	<p>Mengkaji kode Kafka Topic Partitioning yang ada di paper.</p> <ul style="list-style-type: none"> – Memeriksa kode untuk eksperimen, memastikan semua baris kode telah terkumpul. – Merumuskan seluruh dependensi yang diperlukan. 	Sesuai dengan KAKP

Minggu 4 (24 Juni 2024 – 28 Juni 2024)	Mengatur pertemuan dan menyiapkan presentasi awal. <ul style="list-style-type: none"> – Mengatur pertemuan dengan Brandon dan Ikram, yang merupakan penerus penelitian flagship tentang Kafka di CSL. – Menyiapkan file serta pertanyaan untuk presentasi awal kepada Bapak Hilman tentang progres. 	Sesuai dengan KAKP
Minggu 5 (1 Juli 2024 – 5 Juli 2024)	Merencanakan dan mengatur spesifikasi VM di GCP. <ul style="list-style-type: none"> – Mengatur dan menghubungkan dua VM tempat eksperimen di GCP. kedua VM akan digunakan untuk Kafka Cluster dan Kafka Clients. – Melakukan eksperimen kalibrasi (bukan eksperimen utama). 	Sesuai dengan KAKP
Minggu 6 (8 Juli 2024 – 12 Juli 2024)	Membagi tugas eksperimen. <ul style="list-style-type: none"> – Membagi tugas eksperimen dengan Rifqi dan Anin, yang merupakan rekan dalam menerapkan ulang paper. – Memodifikasi spesifikasi VM untuk mengatasi masalah selama eksperimen. 	Sesuai dengan KAKP
Minggu 7 (15 Juli 2024 – 19 Juli 2024)	Menjalankan eksperimen. <ul style="list-style-type: none"> – Menjalankan eksperimen utama yang merupakan inti dari penerapan ulang paper. – Memeriksa performa VM. – Menyampaikan kebutuhan sumber daya kepada Bapak Hilman jika diperlukan. 	Sesuai dengan KAKP

Minggu 8 (22 Juli 2024 – 26 Juli 2024)	Membuat plot dan menganalisis hasil eksperimen. – Membuat plot dari hasil eksperimen. – Menganalisis hasil eksperimen. – Berdiskusi dengan tim tentang kemajuan eksperimen dan hasil yang diperoleh.	Sesuai dengan KAKP
Minggu 9 (29 Juli 2024 – 2 Agustus 2024)	Melakukan ulang eksperimen pada infrastruktur UI. – Mengakses mesin DGX dan mesin BMKGCS milik UI, yang memiliki sumber daya jauh lebih besar dibandingkan VM murah di GCP, untuk mendekati kondisi penelitian asli. – Melakukan eksperimen hingga jalan.	Sesuai dengan KAKP
Minggu 10 (5 Agustus 2024 – 9 Agustus 2024)	Melakukan ulang eksperimen dengan Setup Docker-in-Docker pada infrastruktur UI. – Melakukan ulang eksperimen pada infrastruktur UI dengan batasan VM seperti yang dijelaskan dalam paper, agar lebih mirip dengan penelitian asli. Untuk ini, diperlukan pengaturan Docker-in-Docker agar container yang berisi Kafka cluster dan Kafka clients dapat dikonfigurasi tanpa perlu root access pada hostnya. Membuat plot dan menganalisis hasil eksperimen. – Membuat plot dari hasil eksperimen. – Menganalisis hasil eksperimen. – Berdiskusi dengan tim tentang kemajuan eksperimen dan hasil yang diperoleh.	Sesuai dengan KAKP

Minggu 11 (12 Agustus 2024 – 16 Agustus 2024)	Mengatur pertemuan dengan peneliti asli. – Menghubungi Profesor Claudio Cicconetti melalui email untuk merencanakan pertemuan online. – Mengadakan pertemuan dengan Claudio Cicconetti untuk klarifikasi hal-hal yang masih diragukan dan meminta evaluasinya.	Sesuai dengan KAKP
Minggu 12 (19 Agustus 2024 – 23 Agustus 2024)	Melakukan presentasi akhir. – Menyampaikan laporan hasil kegiatan kepada Bapak Hilman. Menyusun laporan akhir kerja praktik untuk keperluan fakultas.	Sesuai dengan KAKP

Seperti yang dapat dilihat pada Tabel 2.1, secara keseluruhan, pelaksanaan kerja praktik telah berjalan sesuai dengan yang tercantum dalam Kerangka Acuan Kerja Praktik (KAKP). Semua target yang telah ditetapkan terlaksana dengan baik, dan tidak ada permintaan yang melebihi apa yang telah diuraikan dalam KAKP.

2.8.2 Ulasan tentang kendala dan cara menanganinya

Selama kerja praktik, beberapa kendala teknis muncul, berkaitan langsung dengan pekerjaan yang dilakukan. Setiap kendala yang ada diatasi dengan solusi yang sesuai untuk memastikan eksperimen tetap berjalan lancar.

2.8.2.1 Tidak Memiliki Akses Root pada Infrastruktur UI

Salah satu kendala utama yang muncul selama kerja praktik adalah keterbatasan akses *root* pada infrastruktur Universitas Indonesia (UI). Akses *root* sangat penting dalam eksperimen, terutama untuk mengunduh *library* atau melakukan konfigurasi yang lebih mendalam pada sistem. Tanpa akses ini, beberapa langkah teknis yang memerlukan izin tingkat sistem tidak dapat dilakukan secara langsung, sehingga menghambat kelancaran eksperimen. Kondisi ini menuntut solusi yang inovatif untuk mengatasi batasan tersebut.

Untuk mengatasi kendala ini, diterapkan solusi *containerization* dengan dibuatnya Kafka *client container* dan Kafka *cluster container*. Lebih spesifiknya, pendekatan yang diambil adalah Docker-in-Docker (DinD) karena di dalam Kafka *cluster container*, terdapat banyak *broker containers* yang berjalan secara paralel. Pendekatan ini memungkinkan pengunduhan dan instalasi *library* secara independen di dalam lingkungan yang terisolasi tanpa memerlukan akses *root* langsung pada mesin *host*.

2.8.2.2 Dokumentasi dari Peneliti Asli Kurang Lengkap atau Usang

Kendala lain yang dihadapi adalah dokumentasi penelitian asli yang tidak lengkap atau sudah usang. Dokumentasi dari penelitian yang direplikasi sering kali tidak menyertakan penjelasan yang detail mengenai setiap langkah yang dilakukan, atau sudah tidak relevan dengan perkembangan teknologi terbaru. Hal ini menimbulkan tantangan dalam mengikuti alur eksperimen dan mencapai hasil yang sama seperti yang dicapai oleh peneliti sebelumnya.

Untuk mengatasi masalah ini, dilakukan improvisasi terhadap beberapa langkah eksperimen. Pelaksana kerja praktik memanfaatkan referensi tambahan dan hasil riset terbaru untuk menyesuaikan langkah-langkah yang tidak dijelaskan dalam dokumentasi. Selain itu, pelaksana kerja praktik juga mengandalkan hasil komunikasi dengan Profesor Claudio Cicconetti untuk memperoleh klarifikasi lebih lanjut mengenai metode yang diterapkan.

2.8.2.3 Penggunaan Resource yang Besar

Salah satu tantangan terbesar yang pelaksana kerja praktik hadapi selama kerja praktik adalah kebutuhan sumber daya komputasi yang sangat tinggi, terutama saat menjalankan eksperimen di Google Cloud Platform (GCP). Awalnya, konfigurasi awal VM yang dibuat tidak memiliki cukup kapasitas untuk menangani eksperimen berat yang melibatkan proses *real-time data streaming* menggunakan Apache Kafka. Keterbatasan ini menyebabkan VM sering kali kehabisan memori atau mengalami penurunan performa, yang berdampak pada tidak optimalnya hasil eksperimen.

Untuk mengatasi masalah ini, dilakukan beberapa iterasi dalam penyesuaian alokasi sumber daya pada VM di GCP. Setelah beberapa percobaan dan pemantauan terhadap penggunaan CPU dan RAM, ditemukan bahwa eksperimen ini memerlukan alokasi

sebesar 128 GB RAM dan 64 core CPU agar dapat berjalan dengan lancar tanpa hambatan. Dengan peningkatan sumber daya ini, eksperimen berhasil berjalan dengan performa optimal.

2.8.3 Penilaian Individu Terhadap Tempat Kerja Praktik

CSL merupakan lingkungan kerja yang dinamis dan menyenangkan. Banyak inovasi baru yang dihasilkan oleh mahasiswa yang sedang mengerjakan tesis atau skripsi, sehingga menciptakan atmosfer yang kreatif dan kolaboratif. Selain itu, CSL memberikan kesempatan untuk menerapkan teori-teori yang didapat dari perkuliahan ke dalam praktik nyata, sehingga menjadi tempat yang ideal untuk mengembangkan keterampilan teknis.

2.8.4 Relevansi dengan Perkuliahan

Selama kerja praktik, beberapa tugas yang diberikan memiliki relevansi langsung dengan pembelajaran yang didapatkan dari beberapa mata kuliah di Fakultas Ilmu Komputer UI. Berikut adalah mata kuliah yang relevan:

2.8.4.1 Sistem Operasi

Mata kuliah Sistem Operasi memberikan pemahaman tentang penggunaan GNU/Linux, *bash scripting*, serta manajemen proses dan memori. Pengetahuan ini sangat berguna selama kerja praktik, di mana *bash scripting* digunakan untuk mengotomatisasi tugas-tugas tertentu dan mengkonfigurasi virtual machine.

2.8.4.2 Jaringan Komputer

Mata kuliah ini memberikan dasar yang kuat tentang prinsip-prinsip jaringan, mulai dari lapisan aplikasi hingga lapisan fisik, serta aspek keamanan jaringan. Selama kerja praktik, pemahaman ini diterapkan untuk memastikan aliran data berjalan lancar, mengatur koneksi SSH, dan mengelola setup VM di platform seperti GCP dan infrastruktur UI, sehingga komunikasi jaringan dan pengalokasian sumber daya berjalan dengan optimal.

2.8.4.3 Pemrograman Lanjut

Pada mata kuliah Pemrograman Lanjut, dipelajari konsep manajemen pesan dengan RabbitMQ. Di tempat kerja praktik, konsep ini diterapkan menggunakan Apache Kafka,

sebuah platform yang lebih sesuai untuk kebutuhan *real-time data streaming* yang kompleks.

BAB 3

PENUTUP

Pada bab ini dijelaskan kesimpulan dari kerja praktik yang telah dilaksanakan di Computer Systems Laboratory (CSL). Penjelasan dimulai dengan kesimpulan secara menyeluruh tentang pengalaman yang diperoleh selama kerja praktik, diikuti dengan saran untuk beberapa pihak.

3.1 Kesimpulan

Pelaksanaan kerja praktik di Computer Systems Laboratory (CSL) Fakultas Ilmu Komputer UI memberikan banyak pengalaman berharga, baik dari sisi teknis maupun non-teknis. Pelaksana kerja praktik mendapatkan wawasan mendalam tentang pengelolaan data *real-time*, pengoptimalan sumber daya, serta penerapan *containerization* menggunakan Docker.

Dari sisi non-teknis, pelaksana kerja praktik juga mempelajari pentingnya komunikasi efektif, baik dengan dosen peneliti maupun tim di CSL, serta dengan pihak eksternal yang terkait dengan penelitian. Salah satu pengalaman penting adalah pertukaran email dengan Profesor Claudio Cicconetti dari Universit di Pisa, penulis dari penelitian yang eksperimennya sedang direplikasi.

3.2 Saran

Berdasarkan pelaksanaan kerja praktik secara keseluruhan, terdapat beberapa saran yang ingin pelaksana sampaikan kepada pelaksana kerja praktik berikutnya dan Fakultas Ilmu Komputer UI. Adapun mengenai tempat kerja praktik, pelaksana menilai bahwa pelaksanaannya sudah berjalan dengan baik dan efektif, sehingga tidak ada saran yang perlu disampaikan terkait hal tersebut.

3.2.1 Pelaksana Kerja Praktik Berikutnya

Bagi mahasiswa yang akan melaksanakan kerja praktik di CSL, disarankan untuk mempersiapkan manajemen waktu yang baik, terutama jika kerja praktik dilakukan bersamaan dengan perkuliahan aktif. Kerja praktik di CSL menawarkan fleksibilitas

dalam skema *hybrid*, namun tetap membutuhkan fokus yang tinggi dalam menjalankan eksperimen dan memenuhi tanggung jawab akademik secara bersamaan. Penting untuk selalu berkomunikasi dengan dosen pembimbing atau anggota tim jika mengalami kesulitan teknis agar kendala dapat segera diatasi dan waktu tidak terbuang percuma. Selain itu, mahasiswa harus siap dengan tantangan teknis terkait infrastruktur dan manajemen sumber daya, serta mampu beradaptasi dengan solusi-solusi teknologi terbaru yang mungkin berbeda dari yang dipelajari di perkuliahan.

3.2.2 Fakultas Ilmu Komputer UI

Secara umum, pelaksanaan kerja praktik di Fakultas Ilmu Komputer UI sudah berjalan dengan baik, terutama dalam hal waktu dan fleksibilitas bagi pelaksana. Namun, disarankan agar Fakultas Ilmu Komputer UI memperluas kerja sama dengan berbagai institusi dan perusahaan untuk membuka lebih banyak kesempatan kerja praktik bagi mahasiswa. Hal ini penting terutama dalam menghadapi tantangan *tech winter*, di mana peluang kerja praktik di bidang teknologi lebih terbatas. Menjalin kemitraan yang lebih kuat dengan berbagai perusahaan akan memberikan mahasiswa peluang lebih besar untuk mengaplikasikan pengetahuan mereka dalam lingkungan kerja nyata.

DAFTAR REFERENSI

- Kul, S., Kumcu, S., and Sayar, A. (2024). Docker container-based framework of apache kafka node ecosystem: Vehicle tracking system by license plate recognition on surveillance camera feeds. *International Journal of Intelligent Transportation Systems Research*.
- Ledmi, A., Bendjenna, H., and Hemam, S. M. (2018). Fault tolerance in distributed systems: A survey. *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)*.
- Pham, C. M., Dogaru, V., Wagle, R., Venkatramani, C., Kalbarczyk, Z., and Iyer, R. (2014). An evaluation of zookeeper for high availability in system s. *ICPE '14: Proceedings of the 5th ACM/SPEC international conference on Performance engineering*.
- Raptis, T. P., Cicconetti, C., and Passarella, A. (2024). Efficient topic partitioning of apache kafka for high-reliability real-time data streaming applications. *Future Generation Computer Systems*.

LAMPIRAN

LAMPIRAN 1: KERANGKA ACUAN KERJA PRAKTIK



Kerangka Acuan Kerja Praktik Semester Ganjil 2024/2025

FAKULTAS
ILMU
KOMPUTER

Bryan Raihan 'Ilman
2106704351
Ilmu Komputer
bryanilman20@gmail.com
081311235125

- Judul Kerja Praktik : Eksperimen End-To-End untuk Kafka Resource Management pada penelitian “Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications” di lab CSL Fakultas Ilmu Komputer UI.**

Tempat Kerja Praktik	Nama Perusahaan : Fakultas Ilmu Komputer UI Alamat : Ruangan A520 lantai 5, Gedung Baru Fasilkom UI, Pondok Cina, Kecamatan Beji, Kota Depok, Jawa Barat 16424 Telepon/Fax : +62 21 786 3419 Website : https://cs.ui.ac.id/laboratorium-pusat-penelitian/
Waktu Pelaksanaan KP	Periode : 3 Juni 2024 – 25 Agustus 2024 Hari dan Jam Kerja : Senin - Jumat, pukul 08.00 – 12.00 WIB Dengan ini saya menyatakan bahwa Kerja Praktik tidak akan mengganggu perkuliahan yang saya ikuti. Fakultas tidak akan memberikan izin khusus pada perkuliahan lainnya yang diakibatkan oleh kegiatan Kerja Praktik.
Penyelia Kerja Praktik	Nama : Muhammad Hafizhuddin Hilman Jabatan : Director of Computer Systems Laboratory (CSL) Telepon : +62 812-9067-3770 e-mail : muhammad.hilman@ui.ac.id
Deskripsi Tempat KP	<i>Computer System Lab</i> (CSL) di Fakultas Ilmu Komputer Universitas Indonesia (Fasilkom UI) adalah laboratorium dan pusat penelitian yang fokus pada pengembangan protokol, algoritma, dan framework untuk manajemen <i>big data</i> , jaringan komputer, dan komputasi kinerja tinggi. CSL berkolaborasi dengan berbagai laboratorium riset dan pusat kajian di fakultas untuk menghasilkan produk dan kajian terbaru dalam bidang sistem komputer. Divisi yang menjadi

	tempat kerja praktik saya di CSL terlibat dalam penelitian <i>flagship</i> Smart Fog Computing Platform.
Ruang Lingkup Kerja Praktik	<p>Nama peran : Research Assistant</p> <p>Dekripsi peran:</p> <ul style="list-style-type: none"> • Memahami <i>framework</i> Smart Fog Computing yang dikembangkan oleh Muhammad Ariq Basyar, bernama FogVerse. • Membaca dan memahami paper "Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications" dan memahami teknik <i>efficient topic partitioning</i> menggunakan BroMax dan BroMin yang diklaim lebih baik daripada industry best practice. • Melakukan set up VM di Google Cloud Platform (GCP) dan menghubungkan VM dengan SSH serta akses root. • Menyiapkan VM dengan menginstal dependencies yang diperlukan. • Menjalankan eksperimen yang ada di repositori GitHub terkait Kafka, dengan satu VM bertindak sebagai Kafka clients (producer dan consumer) dan satu VM sebagai Kafka cluster (broker dan zookeeper). • Memfokuskan eksperimen pada simulasi 'end-to-end experiment' untuk memantau perilaku algoritma BroMax dan BroMin dalam menentukan jumlah partisi dan broker serta efeknya terhadap <i>throughput</i>. • Mempelajari cara kerja Docker, termasuk build image dan setup container. • Menangani troubleshooting komunikasi antar container jika ada masalah. • Memahami penggunaan resource di setiap VM, termasuk CPU utilization, RAM usage, dan hard disk space required. • Mengkonfigurasi VM hingga eksperimen berjalan dengan baik. • Mengulangi eksperimen di infrastruktur UI dengan VM bernama 'DGX' dan 'BMKGCS'. • Mencari cara menjalankan eksperimen tanpa root access menggunakan pendekatan Docker in Docker (DinD). • Membuat framework untuk menjalankan eksperimen di infrastruktur UI. • Menganalisis dan membuat plot hasil eksperimen 'end-to-end experiment' dan mendeskripsikan perilaku BroMax dan BroMin. • Menghubungi peneliti melalui email untuk bertanya dan mendapatkan evaluasi. • Menyampaikan temuan kepada Bapak Muhammad Hafizhuddin Hilman.

2. Rencana Jadwal Kerja Praktik

Tabel 1 Rencana Jadwal Kerja Praktik Pelaksana

Waktu	Rencana Kerja
Minggu 1 3 Juni 2024 – 7 Juni 2024	<p>Meninjau jurnal dari Elsevier dan mencari penelitian terbaru mengenai Apache Kafka yang dapat diterapkan kembali di CSL.</p> <ul style="list-style-type: none"> - Menganalisis metodologi dan hasil dari paper berjudul Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications. <p>Bergabung dengan Komunitas FogVerse</p> <ul style="list-style-type: none"> - Bergabung dengan komunitas FogVerse di WhatsApp dan Discord
Minggu 2 10 Juni 2024 – 14 Juni 2024	<p>Merumuskan langkah-langkah untuk menerapkan kembali isi paper.</p> <ul style="list-style-type: none"> - Membaca dan memahami paper yang akan diterapkan kembali. - Merancang langkah-langkah secara detail. <p>Mengikuti perkembangan proyek flagship CSL, yaitu FogVerse.</p> <ul style="list-style-type: none"> - Mempelajari kode FogVerse dari GitHub. - Mengunduh kode FogVerse.
Minggu 3 17 Juni 2024 – 21 Juni 2024	<p>Mengkaji kode Kafka Topic Partitioning yang ada di paper.</p> <ul style="list-style-type: none"> - Memeriksa kode untuk eksperimen, memastikan semua baris kode telah terkumpul. - Merumuskan seluruh dependensi yang diperlukan.
Minggu 4 24 Juni 2024 – 28 Juni 2024	<p>Mengatur pertemuan dan menyiapkan presentasi awal.</p> <ul style="list-style-type: none"> - Mengatur pertemuan dengan Brandon dan Ikram,

	<p>yang merupakan penerus penelitian <i>flagship</i> tentang Kafka di CSL.</p> <ul style="list-style-type: none"> - Menyiapkan file serta pertanyaan untuk presentasi awal kepada Bapak Hilman tentang progres.
Minggu 5 1 Juli 2024 – 5 Juli 2024	<p>Merencanakan dan mengatur spesifikasi VM di GCP.</p> <ul style="list-style-type: none"> - Mengatur dan menghubungkan dua VM tempat eksperimen di GCP. kedua VM akan digunakan untuk Kafka Cluster dan Kafka Clients. - Melakukan eksperimen kalibrasi (bukan eksperimen utama).
Minggu 6 8 Juli 2024 – 12 Juli 2024	<p>Membagi tugas eksperimen.</p> <ul style="list-style-type: none"> - Membagi tugas eksperimen dengan Rifqi dan Anin, yang merupakan rekan dalam menerapkan ulang paper. - Memodifikasi spesifikasi VM untuk mengatasi masalah selama eksperimen.
Minggu 7 15 Juli 2024 – 19 Juli 2024	<p>Menjalankan eksperimen.</p> <ul style="list-style-type: none"> - Menjalankan eksperimen utama yang merupakan inti dari penerapan ulang paper. - Memeriksa performa VM. - Menyampaikan kebutuhan sumber daya kepada Bapak Hilman jika diperlukan.
Minggu 8 22 Juli 2024 – 26 Juli 2024	<p>Membuat plot dan menganalisis hasil eksperimen.</p> <ul style="list-style-type: none"> - Membuat plot dari hasil eksperimen. - Menganalisis hasil eksperimen. - Berdiskusi dengan tim tentang kemajuan eksperimen dan hasil yang diperoleh.
Minggu 9 29 Juli 2024 – 2 Agustus 2024	<p>Melakukan ulang eksperimen pada infrastruktur UI.</p> <ul style="list-style-type: none"> - Mengakses mesin DGX dan mesin BMKGCS milik UI, yang memiliki sumber daya jauh lebih besar dibandingkan VM murah di GCP, untuk mendekati

	<p>kondisi penelitian asli.</p> <ul style="list-style-type: none"> - Melakukan eksperimen hingga jalan.
Minggu 10 5 Agustus 2024 – 9 Agustus 2024	<p>Melakukan ulang eksperimen dengan Setup Docker-in-Docker pada infrastruktur UI.</p> <ul style="list-style-type: none"> - Melakukan ulang eksperimen pada infrastruktur UI dengan batasan VM seperti yang dijelaskan dalam paper, agar lebih mirip dengan penelitian asli. Untuk ini, diperlukan pengaturan Docker-in-Docker agar container yang berisi Kafka cluster dan Kafka clients dapat dikonfigurasi tanpa perlu <i>root access</i> pada <i>hostnya</i>. <p>Membuat plot dan menganalisis hasil eksperimen.</p> <ul style="list-style-type: none"> - Membuat plot dari hasil eksperimen. - Menganalisis hasil eksperimen. - Berdiskusi dengan tim tentang kemajuan eksperimen dan hasil yang diperoleh.
Minggu 11 12 Agustus 2024 – 16 Agustus 2024	<p>Mengatur pertemuan dengan peneliti asli.</p> <ul style="list-style-type: none"> - Menghubungi Profesor Claudio Cicconetti melalui email untuk merencanakan pertemuan online. - Mengadakan pertemuan dengan Claudio Cicconetti untuk klarifikasi hal-hal yang masih diragukan dan meminta evaluasinya.
Minggu 12 19 Agustus 2024 – 23 Agustus 2024	<p>Melakukan presentasi akhir.</p> <ul style="list-style-type: none"> - Menyampaikan laporan hasil kegiatan kepada Bapak Hilman. <p>Menyusun laporan akhir kerja praktik untuk keperluan fakultas.</p>

3. Daftar Pustaka

- [1] Basyar, M. A. (2022). FogVerse: Sistem cerdas untuk smart-CCTV berbasis arsitektur fog computing [Skripsi, UI - Skripsi Membership]. Fakultas Ilmu Komputer Universitas Indonesia.
<https://lib.ui.ac.id/detail.jsp?id=9999920519875#digital#digital>

[2] Raptis, T. P., Cicconetti, C., & Passarella, A. (2024). Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications. *Future Generation Computer Systems*, vol. 154, 173-188. <https://doi.org/10.1016/j.future.2023.12.028>.

[3] Cicconetti, C. (2024). kafka-hdd [GitHub repository]. Diamambil dari <https://github.com/ccicconetti/kafka-hdd/>.

Depok, 1 Juni 2024

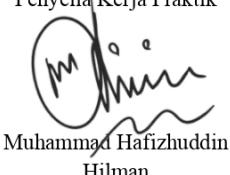
Menyetujui,

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Penyelia Kerja Praktik



Muhammad Hafizhuddin
Hilman

Dosen Mata Kuliah Kerja
Praktik

Ichlasul Affan

LAMPIRAN 2: LOG KERJA PRAKTIK



Log Kerja Praktik Semester Gasal 2021/2022

3 Juni 2024 – 7 Juni 2024

NPM / Nama : 2106704351 / Bryan Raihan ‘Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 3 Juni 2024	10:00 – 15:00	<p>Researching and studying scientific papers and journals related to Apache Kafka</p> <ul style="list-style-type: none"> • “Kafka-based Architecture in Building Data Lakes for Real-time Data Streams” • “Real-Time Network Monitoring: A Big Data Approach” • “Towards flexible data stream collaboration: Federated Learning in Kafka-ML”
Selasa, 4 Juni 2024	10:00 – 15:00	<p>Getting to know FogVerse and related terms</p> <ul style="list-style-type: none"> • Watching Muhammad Ariq’s Thesis Proposal on Youtube and frequently pausing to look up technical terms on Google.
Rabu, 5 Juni 2024	10:00 – 15:00	<p>Gaining a deeper and more formal understanding of the FogVerse</p> <ul style="list-style-type: none"> • Reading FogVerse original Thesis by Muhammad Ariq Basyar
Kamis, 6 Juni 2024	10:00 – 15:00	<p>Reviewing understanding of the FogVerse</p> <ul style="list-style-type: none"> • Writing notes and going over the original thesis again
Jum’at, 7 Juni 2024	10:00 – 15:00	<p>Staying connected to the FogVerse community</p> <ul style="list-style-type: none"> • Making sure I have joined FogVerse Whatsapp and Discord groups in CSL • Making sure I have access to the Github repository of FogVerse

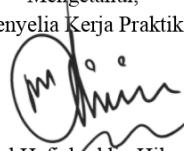
Depok, 7 Juni 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizhuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

10 Juni 2024 – 14 Juni 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 10 Juni 2024	09.00 -12.00; 13.00 - 15.00	Deciding which scientific paper or journal to replicate <ul style="list-style-type: none"> After reviewing Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications, I intend to pursue reproducing its steps.
Selasa, 11 Juni 2024	10:00 – 15.00	Studying the FogVerse code <ul style="list-style-type: none"> I learned how FogVerse technically works by reviewing Github repository of FogVerse.
Rabu, 12 Juni 2024	10:00 – 15.00	Revisiting the FogVerse code <ul style="list-style-type: none"> I deepened my understanding about FogVerse code by reading it again.
Kamis, 13 Juni 2024	10:00 – 15.00	Cloning the FogVerse code <ul style="list-style-type: none"> I downloaded the FogVerse code to my computer to play around with some of its functionalities.
Jum'at, 14 Juni 2024	10:00 – 15.00	Studying the paper I intend to reproduce <ul style="list-style-type: none"> I started reading Efficient topic partitioning of Apache Kafka for high-reliability real-time data streaming applications.

Depok, 14 Juni 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizbiddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

17 Juni 2024 – 21 Juni 2024

NPM / Nama : 2106704351 / Bryan Raihan ‘Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 17 Juni 2024		Libur Idul Adha
Selasa, 18 Juni 2024		Libur Cuti Bersama
Rabu, 19 Juni 2024	10:00 – 19.00	<p>Reviewing the paper I intend to reproduce</p> <ul style="list-style-type: none"> • I refreshed my understanding about the paper by reading it again.
Kamis, 20 Juni 2024	10:00 – 18.00	<p>Studying the Kafka topic partitioning code</p> <ul style="list-style-type: none"> • After understanding the paper, I began reading the code, especially the part about experiments.

Jum'at, 21 Juni 2024	10:00 – 18.00	Revisiting the Kafka topic partitioning code
		<ul style="list-style-type: none">I made sure I understood all parts of the code.

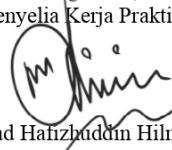
Depok, 21 Juni 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

24 Juni 2024 – 28 Juni 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 24 Juni 2024	13:30 – 18:30	<p>Setting up meeting with Brandon Ivander and Ikram</p> <ul style="list-style-type: none"> chat via Discord <div style="background-color: black; color: white; padding: 10px;"> <p>24 June 2024</p> <p>Bryan 'Ilman 24/06/2024 13:37 Selamat siang Kak Brandon dan Ikram. Anin, Rifqi, dan aku ingin belajar tentang cara kerja FogVerse dan konsep2 pentingnya. Kami ingin melihat demonstrasi FogVerse agar lebih memahami cara kerjanya. Kapankah kemungkinan kita bisa meet?</p> <p> Ikram 24/06/2024 17:18 Okee malam sabtu gmn Jam" 9</p> </div>
Selasa, 25 Juni 2024	10:00 – 15.00	<p>Reviewing Linux commands</p> <ul style="list-style-type: none"> I read some popular articles from Hostingter and DigitalOcean to refresh my memories about Linux commands.
Rabu, 26 Juni 2024	10:00 – 15.00	<p>Interacting with the research code</p> <ul style="list-style-type: none"> I began cloning the github repo and followed the steps in the README on GitHub and ensured all the requirements were met before running the experiment.
Kamis, 27 Juni 2024	10:00 – 15.00	<p>Preparing for tomorrow's meetups</p> <ul style="list-style-type: none"> I sorted my files and photos to present to Mr. Hilman I began creating plenty of questions to ask Brandon and Ikram to widen my understanding about Kafka and FogVerse. I also collected questions from Anin and Rifqi

Jum'at, 28 Juni 2024	11:00 – 14.00 21:00 – 23.00	Attending meetups <ul style="list-style-type: none"> • I met Mr. Hilman at CSL, together with Anin and Rifqi. • I talked via Discord with Brandon and Ikram, together with Anin and Rifqi.
----------------------------	--	--

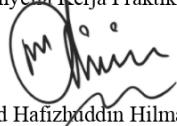
Depok, 28 Juni 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizhuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

1 Juli 2024 – 5 Juli 2024

NPM / Nama : 2106704351 / Bryan Raihan ‘Ilman

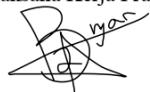
Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 1 Juli 2024	10.00 -12.00	Planning VM specs
	14.00 - 17.00	<ul style="list-style-type: none"> • team meeting • I calculated how much space, how many CPU cores and RAM gigabytes ‘kafka-cluster’ VM and ‘kafka-clients’ VM are going to need to complete the experiments.
Selasa, 2 Juli 2024	10.00 -12.00	Setting up and connecting VMs
	14.00 - 17.00	<ul style="list-style-type: none"> • team meeting • I set up two VMs on GCP and figured out how to gain root access via SSH from one VM to another.
Rabu, 3 Juli 2024	10.00 -12.00	Preparing VMs
	14.00 - 17.00	<ul style="list-style-type: none"> • team meeting • Installing all dependencies required for the experiments using ‘apt’.
Kamis, 4 Juli 2024	10.00 -12.00	Experimenting with VMs
	14.00 - 17.00	<ul style="list-style-type: none"> • team meeting • I ran calibration experiments and learned how to generate plots in PNG format using ‘gnuplot’.
Jum’at, 5 Juli 2024	10.00 -12.00	Assignment of Tasks
	14.00 - 17.00	<ul style="list-style-type: none"> • Our team discussed who should do what tasks. • I reviewed the paper to refresh my understanding.

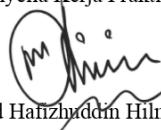
Depok, 5 Juli 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizhuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

8 Juli 2024 – 12 Juli 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 8 Juli 2024	10.00 -12.00 14.00 - 17.00	Assignment of Tasks <ul style="list-style-type: none"> We as a team agreed that Rifqi will do experiment number one, Anin will do number two, and I will do number three and four.
Selasa, 9 Juli 2024	10.00 -12.00 14.00 - 17.00	Running experiment number one <ul style="list-style-type: none"> Initial Attempt: <ul style="list-style-type: none"> Duration: Ran for 3 hours. Outcome: Ended in failure due to being stuck in a while-loop. Subsequent Attempts: <ul style="list-style-type: none"> Result: Experiment failed again after multiple tries. Issues Encountered: Consistently faced the same problem with the while-loop, causing the experiment to get stuck and ultimately fail.
Rabu, 10 Juli 2024	10.00 -12.00 14.00 - 17.00	Modifying VM specs <ul style="list-style-type: none"> In order to resolve the error, I tried many combination of specs on the two VMs
Kamis, 11 Juli 2024	10.00 -12.00 14.00 - 17.00	Running experiment number three <ul style="list-style-type: none"> The experiment had made no progress because the cause of the failure remained unidentified.
Jum'at, 12 Juli 2024	10.00 -12.00 14.00 - 17.00	Identifying the error <ul style="list-style-type: none"> Finally, I discovered that the error occurred because the Zookeeper was not running when the brokers attempted to connect. Therefore, I added a delay between setting up Zookeeper and starting the brokers.

Depok, 12 Juli 2024

Pelaksana Kerja Praktik

Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik

Muhammad Hafizhuddin Hilman, Ph.D



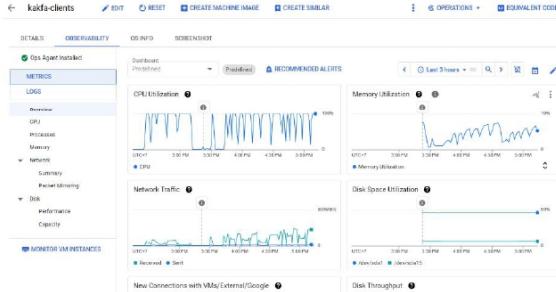
Log Kerja Praktik Semester Gasal 2021/2022

15 Juli 2024 – 19 Juli 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

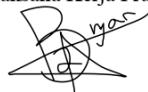
Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 15 Juli 2024	10.00 -12.00 14.00 - 17.00	<p>Running experiment number four</p> <ul style="list-style-type: none"> After two hours, the experiment ended in a failure with message 'KafkaTimedoutError'. Experiment failed again after many tries.
Selasa, 16 Juli 2024	10.00 -12.00 14.00 - 17.00	<p>Trying to solve the error</p> <ul style="list-style-type: none"> I checked if any of the broker or the zookeeper itself was down but nothing seems wrong
Rabu, 17 Juli 2024	10.00 -12.00 14.00 - 17.00	<p>Hypothesizing a solution</p> <ul style="list-style-type: none"> I checked the health and performance of my VMs in GCP and found that 'kafka-clients' maxed out CPU utilization so probably I need to add more cores to handle more workloads. 
Kamis, 18 Juli 2024	10.00 -12.00 14.00 - 17.00	<p>Implementing the hypothesis</p> <ul style="list-style-type: none"> I increased the CPU cores of 'kafka-clients' to 16 and ran the experiment again. It completed successfully
Jum'at, 19 Juli 2024	10.00 -12.00 14.00 - 17.00	<p>Redoing previous experiments</p> <ul style="list-style-type: none"> I met Mr. Hilman at CSL and explained to him how the experiments required enormous resources. I ran both experiments number three and number four one more time.

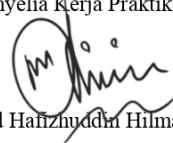
Depok, 19 Juli 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik



Muhammad Hafizuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

22 Juli 2024 – 26 Juli 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 22 Juli 2024	10.00 -12.00 14.00 - 17.00	Creating plots from experiment results <ul style="list-style-type: none"> I obtained 'dat' files from running the experiments and converted them to PNG using 'gnuplot'
Selasa, 23 Juli 2024	10.00 -12.00 14.00 - 17.00	Analyzing the results of the experiment <ul style="list-style-type: none"> I tried to understand the results of the experiments that I ran and how it relates with the paper.
Rabu, 24 Juli 2024	10.00 -12.00 14.00 - 17.00	Analyzing the results of the experiment <ul style="list-style-type: none"> I tried to understand the results of the experiments that I ran and how it relates with the paper.
Kamis, 25 Juli 2024	10.00 -12.00 14.00 - 17.00	Team meeting <ul style="list-style-type: none"> I consulted with my team regarding their progress with their respective experiments and discussed how my results correlated with theirs.
Jum'at, 26 Juli 2024	10.00 -12.00 14.00 - 17.00	Team meeting <ul style="list-style-type: none"> I consulted with my team regarding their progress with their respective experiments and discussed how my results correlated with theirs.

Depok, 26 Juli 2024

Pelaksana Kerja Praktik

Bryan Raihan 'Ilman'

Mengetahui,
Penyelia Kerja Praktik

Muhammad Hafizhuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

29 Juli 2024 – 2 Agustus 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 29 Juli 2024	10.00 -12.00	Accessing UI's DGX and BMKGCS machines
	14.00 - 17.00	<ul style="list-style-type: none"> • I accessed DGX and BMKGCS machines owned by UI. • I verified that the machines meet the required specifications and resources.
Selasa, 30 Juli 2024	10.00 -12.00	Configuring UI's machines for experiments
	14.00 - 17.00	<ul style="list-style-type: none"> • I set up the necessary software and dependencies on DGX and BMKGCS machines.
Rabu, 31 Juli 2024	10.00 -12.00	Running experiments on UI's infrastructure
	14.00 - 17.00	<ul style="list-style-type: none"> • I executed experiment number three on the UI infrastructure. • I've found that BMKGCS performs better than DGX, so I'll be sticking with BMKGCS and discontinuing the use of DGX from now on.
Kamis, 1 Agustus 2024	10.00 -12.00	Troubleshooting experiment results
	14.00 - 17.00	<ul style="list-style-type: none"> • I analyzed experiment number three results and identified any problems. • I made necessary adjustments to configurations or experiment parameters based on findings.
Jum'at, 2 Agustus 2024	10.00 -12.00	Finalizing experiments on UI's Infrastructure
	14.00 - 17.00	<ul style="list-style-type: none"> • I ran both experiments number three and four on the UI infrastructure. • I confirmed that experiments were completed successfully and analyzed the results.

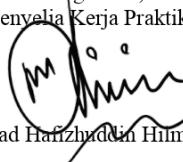
Depok, 2 Agustus 2024

Pelaksana Kerja Praktik



Bryan Raihan 'Ilman

Mengetahui,
Penelia Kerja Praktik



Muhammad Hafizhuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

5 Agustus 2024 – 9 Agustus 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman
 Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)
 Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 5 Agustus 2024	10.00 -12.00 14.00 - 17.00	Setting up Docker-in-Docker on UI's infrastructure <ul style="list-style-type: none"> I configured Docker-in-Docker setup on the UI's machine. I ensured proper containerization of Kafka cluster and Kafka clients without requiring root access on the host.
Selasa, 6 Agustus 2024	10.00 -12.00 14.00 - 17.00	Running experiments with Docker-in-Docker setup <ul style="list-style-type: none"> I re-ran experiments on UI's infrastructure using the Docker-in-Docker setup. I ensured VM limitations align with the constraints described in the paper for accurate results.
Rabu, 7 Agustus 2024	10.00 -12.00 14.00 - 17.00	Creating plots from experiment results <ul style="list-style-type: none"> I generated plots from the latest experiment data. I utilized tools like gnuplot or similar for visual representation.
Kamis, 8 Agustus 2024	10.00 -12.00 14.00 - 17.00	Analyzing experiment results <ul style="list-style-type: none"> I analyzed the plots and data to understand the outcomes of the experiments. I compared results with expected outcomes and previous experiments.
Jum'at, 9 Agustus 2024	10.00 -12.00 14.00 - 17.00	Discussing results with the team <ul style="list-style-type: none"> I held a discussion with the team regarding the progress and results of the experiments. I reviewed findings and correlated with team members' experiments for comprehensive insights.

Depok, 9 Agustus 2024

Pelaksana Kerja Praktik

Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik

Muhammad Hafizbuddin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

12 Agustus 2024 – 16 Agustus 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 12 Agustus 2024	10.00 -12.00 14.00 - 17.00	Attempting to schedule a meeting with the researcher <ul style="list-style-type: none"> I sent multiple emails to Professor Claudio Cicconetti to arrange an online meeting. I attempted to reach him via WhatsApp, Line, and Skype with no response.
Selasa, 13 Agustus 2024	10.00 -12.00 14.00 - 17.00	Reviewing findings and refreshing understanding <ul style="list-style-type: none"> I utilized time to review the findings from recent experiments. I revisited relevant literature to refresh understanding of key concepts.
Rabu, 14 Agustus 2024	10.00 -12.00 14.00 - 17.00	Documenting experiment results and insights <ul style="list-style-type: none"> I compiled notes and insights from recent experiments.
Kamis, 15 Agustus 2024	10.00 -12.00 14.00 - 17.00	Preparing questions and discussion points <ul style="list-style-type: none"> I created a list of questions and topics for discussion with Professor Cicconetti. I focused on clarifying doubts and seeking evaluation of experimental results.
Jum'at, 16 Agustus 2024	10.00 -12.00 14.00 - 17.00	Reevaluating experiment approach <ul style="list-style-type: none"> I analyzed the current experiment approach and considered any necessary adjustments.

Depok, 16 Agustus 2024

Pelaksana Kerja Praktik

Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik

Muhammad Hafizul Nidin Hilman, Ph.D



Log Kerja Praktik Semester Gasal 2021/2022

19 Agustus 2024 – 23 Agustus 2024

NPM / Nama : 2106704351 / Bryan Raihan 'Ilman

Tempat KP : Fakultas Ilmu Komputer Universitas Indonesia / Computer System Lab (CSL)

Judul KP : FogVerse: Fog Computing and Resource Management Platform based on Kafka

Hari, Tanggal	Waktu	Deskripsi Pekerjaan
Senin, 19 Agustus 2024	10.00 -12.00 14.00 - 17.00	Creating presentation slides <ul style="list-style-type: none"> I started designing and structuring the presentation slides. I focused on key points and data from the experiments for inclusion.
Selasa, 20 Agustus 2024	10.00 -12.00 14.00 - 17.00	Developing presentation content <ul style="list-style-type: none"> I continued adding detailed content to the slides. I ensured clarity and coherence in the presentation of findings. I spoke with my team about assigning different parts of the presentation to each member.
Rabu, 21 Agustus 2024	10.00 -12.00 14.00 - 17.00	Finalizing presentation slides <ul style="list-style-type: none"> I refined and polished slides, including visual aids and data representations. I conducted a review to ensure the presentation is complete and effective.
Kamis, 22 Agustus 2024	10.00 -12.00 14.00 - 17.00	Writing the final report <ul style="list-style-type: none"> I compiled detailed results and observations into the final report. I ensured the report meets faculty requirements and accurately reflects the work done.
Jum'at, 23 Agustus 2024	10.00 -12.00 14.00 - 17.00	Delivering the final presentation <ul style="list-style-type: none"> My team presented the report and findings to Bapak Hilman. I addressed any questions or feedback from the presentation.

Depok, 23 Agustus 2024

Pelaksana Kerja Praktik

Bryan Raihan 'Ilman

Mengetahui,
Penyelia Kerja Praktik

Muhammad Hafizhuddin Hilman, Ph.D

LAMPIRAN 3: DOKUMENTASI REPRODUKSI EKSPERIMEN

Dokumentasi Reproduksi Eksperimen

EFFICIENT TOPIC PARTITIONING OF APACHE KAFKA FOR HIGH-RELIABILITY REAL-TIME DATA STREAMING APPLICATIONS

Catatan: Panduan ini mengasumsikan Anda menggunakan sistem Ubuntu atau distribusi Linux lain yang menggunakan `apt` package manager.

1. Lakukan clone terhadap Kafka Partitioning Experiment Github Repository

Langkah pertama adalah mengkloning GitHub repository yang berisi kode untuk menjalankan eksperimen Kafka Topic Partitioning. Pastikan bahwa pada sistem Anda telah terpasang `git`. Jika belum, Anda dapat menginstalnya dengan perintah berikut:

```
sudo apt update
sudo apt install -y git
```

Kemudian, kloning repository dengan perintah berikut:

```
git clone https://github.com/bryan-ilman-on-github/kafka-part-exp.git
```

2. Navigasi ke Repo Eksperimen

Selanjutnya, kita navigasi ke repo github yang baru saja didownload dan pastikan bahwa script yang dibutuhkan mempunyai permission yang tepat untuk dijalankan. Jalankan perintah berikut:

```
cd kafka-part-exp
chmod +x run-experiment.sh
chmod +x collect-experiment.sh
```

3. Persiapan Docker pada Host

Pastikan pada host yang akan digunakan sudah terpasang **Docker** dan **Docker Compose**. Jika belum terpasang, ikuti langkah berikut untuk menginstalnya:

```
sudo apt update
sudo apt install -y docker.io docker-compose
```

Setelah Docker diinstal, aktifkan dan jalankan Docker daemon dengan perintah berikut:

```
sudo systemctl enable docker
sudo systemctl start docker
```

4. Build Docker Images

Jalankan perintah berikut untuk membangun beberapa Docker *images* yang diperlukan:

```
docker build -t kafka-cluster-image -f kafka-cluster-dockerfile .
docker build -t kafka-clients-image -f kafka-clients-dockerfile .
```

5. Jalankan Eksperimen

Sebelum menjalankan eksperimen, pastikan sistem Anda memiliki minimal **128 GB RAM** dan **180 GB disk space** yang tersedia. Kebutuhan ini penting untuk memastikan eksperimen berjalan lancar tanpa hambatan. Jalankan eksperimen dengan perintah berikut:

```
./run-experiment.sh -n <number>
```

atau

```
./run-experiment.sh --number <number>
```

Isi `<number>` dengan nomor eksperimen yang ingin dijalankan (1-4). Opsi `-n` atau `--number` bersifat wajib dan digunakan untuk menentukan nomor eksperimen.

6. Ambil Hasil Eksperimen

Setelah eksperimen selesai, ambil hasilnya dengan perintah berikut:

```
./collect-results.sh -n <number>
```

atau

```
./collect-results.sh --number <number>
```

Isi `<number>` dengan nomor eksperimen yang hasilnya ingin diambil (1-4). Opsi `-n` atau `--number` bersifat opsional. Jika tidak diisi, maka hasil dari semua nomor eksperimen akan diambil. Folder bernama `experiment-<number>` akan dibuat untuk menyimpan hasil eksperimen.