BAB 1
PENDAHULUAN

1.1 Latar Belakang

Dalam menghadapi era digital yang semakin menuntut efisiensi dan skalabilitas, perusahaan-perusahaan besar seperti Smartfren membutuhkan sistem pengelolaan data yang mampu mendukung pengambilan keputusan secara cepat, akurat, dan berbasis *real-time* [1]. Sebelumnya, Smartfren menggunakan sistem basis data Greenplum. Greenplum secara tradisional dikenal sebagai sistem data warehouse OLAP yang memiliki keterbatasan dalam menangani beban kerja OLTP secara *real-time*, sehingga diperlukan pendekatan yang mampu menggabungkan kedua jenis beban kerja tersebut secara efisien [2], namun sistem tersebut memiliki keterbatasan dalam hal performa, skalabilitas, serta integrasi dengan layanan *cloud* modern.

Big Data dan *Cloud Computing* merupakan dua teknologi utama yang kini banyak dimanfaatkan di berbagai sektor, karena keduanya memungkinkan penyimpanan, pemrosesan, dan analisis data dalam jumlah besar secara efisien dan hemat biaya[3]. Migrasi dari data warehouse lama seperti Teradata ke *platform cloud-native* seperti Google BigQuery merupakan langkah transformatif dalam memodernisasi manajemen dan analitik data[4]. Sebagai bagian dari strategi modernisasi infrastruktur data, Smartfren memutuskan untuk melakukan migrasi basis data dari Greenplum ke Google BigQuery. Proyek migrasi ini dilaksanakan dengan dukungan dari PT Devoteam, yang bertindak sebagai mitra resmi dan penyedia layanan Google Cloud. Migrasi ini tidak hanya bertujuan untuk meningkatkan kinerja pemrosesan data, tetapi juga memungkinkan Smartfren memanfaatkan berbagai fitur canggih dalam ekosistem Google Cloud, seperti analisis data berbasis *AI*, pemodelan *machine learning*, visualisasi dengan Data Studio, serta pengelolaan beban kerja secara dinamis dan efisien melalui BigQuery [5].

Google BigQuery terus mengalami evolusi untuk memenuhi kebutuhan pelanggan di berbagai spektrum beban kerja analitik dan *AI/ML*, berkat arsitektur *cloud-native* yang *disaggregated*. *Machine Learning* dapat menjadi kunci untuk membuka nilai dari data perusahaan dan pelanggan, karena memungkinkan bisnis memanfaatkan data tersebut untuk membuat prediksi dan keputusan yang lebih

akurat[6]. Salah satu kebutuhan utama pelanggan terhadap BigQuery adalah kemampuannya dalam menyatukan beban kerja *data lake* dan *enterprise data warehouse* dalam satu sistem terpadu. Pendekatan ini menggabungkan kebutuhan akan fitur inti pengelolaan data seperti keamanan, tata kelola, metadata *runtime* bersama, akselerasi performa, dan transaksi *ACID*, dengan fleksibilitas ekosistem *open-source* serta dukungan terhadap beban kerja modern seperti *AI/ML* yang dijalankan di atas data tidak terstruktur pada penyimpanan objek (*object storage*) [7]. Selain itu, BigQuery juga mendukung implementasi *multi-cloud*, sejalan dengan tren perusahaan modern yang mengadopsi pendekatan *multi-cloud footprint* secara default guna menghindari ketergantungan pada satu penyedia layanan saja.

Google BigQuery sendiri merupakan solusi *data warehouse* berbasis *cloud* yang dirancang untuk menganalisis jumlah data besar secara cepat, dengan kemampuan eksekusi kueri dalam hitungan detik serta skalabilitas otomatis sesuai beban kerja [8]. *Platform* ini menghilangkan kebutuhan pengelolaan infrastruktur fisik, sehingga memungkinkan organisasi seperti Smartfren untuk fokus pada pemanfaatan data, bukan pada pengelolaan server dan sumber daya teknis lainnya.

Dalam implementasinya, proses migrasi basis data skala besar seperti ini bukanlah hal yang sederhana. Sebagaimana dijelaskan oleh Kansara [9], migrasi dari sistem *on-premise* ke *cloud* menghadirkan tantangan kompleks, termasuk perbedaan struktur data, keterbatasan dokumentasi sistem lama, serta risiko performa dan keamanan yang tinggi. Untuk mengatasi hal tersebut, diperlukan pendekatan sistematis yang mencakup tahapan seperti asesmen, perencanaan, *proof of concept*, validasi, *cutover*, dan optimalisasi pasca-migrasi. Dengan mengikuti pendekatan ini, Smartfren mampu meminimalkan *downtime*, memastikan integritas data, dan menjaga kelangsungan operasional selama transisi berlangsung.

Selain itu, penting untuk memahami dasar-dasar dari sistem basis data itu sendiri. Sejak tahun 1960-an, sistem manajemen basis data telah berkembang dari *file processing system* (FPS) menjadi pendekatan basis data terintegrasi yang lebih efisien dan terstruktur. Basis data modern merupakan kumpulan data yang disimpan dalam format tertentu, serta memanfaatkan model seperti *Entity-Relationship* (ER) untuk menggambarkan skema konseptual [10]. Pemahaman mengenai arsitektur dan perbedaan antar *platform*, seperti Oracle dan SQL Server, turut berperan penting dalam proses perencanaan dan kompatibilitas migrasi. Hal ini menegaskan pentingnya fondasi teori basis data dalam menyusun strategi migrasi yang efektif dan efisien.

Di tengah percepatan digitalisasi, jumlah data yang dihasilkan dari berbagai

sumber baik daring maupun luring meningkat secara eksponensial setiap detiknya. Data berskala besar atau *Big Data* ini memiliki karakteristik *volume*, *velocity*, dan *variety* yang menantang untuk ditangani dengan pendekatan tradisional [1]. *Cloud computing* muncul sebagai solusi ideal karena menyediakan sumber daya komputasi, penyimpanan, dan layanan yang fleksibel tanpa bergantung pada perangkat lokal pengguna. *Platform* seperti Google Cloud, Amazon Web Services (AWS), IBM Cloud, dan Microsoft Azure telah menyediakan infrastruktur yang memungkinkan analisis data besar secara *real-time*. Salah satu layanan unggulan dalam kategori ini adalah Google BigQuery, yang dirancang untuk menganalisis *dataset* dalam jumlah besar dan memberikan hasil cepat guna mendukung pengambilan keputusan secara efisien [4].

Melalui proses migrasi ini, Smartfren dapat mengakses data dalam jumlah besar secara lebih cepat, aman, dan terintegrasi, sekaligus membuka peluang untuk mengembangkan layanan-layanan baru berbasis data. Implementasi ini menunjukkan kolaborasi strategis antara penyedia layanan *cloud* dan klien korporat dalam mengadopsi solusi teknologi berbasis *cloud* guna mendukung transformasi digital yang berkelanjutan.

1.2 Maksud dan Tujuan Magang

Adapun maksud dan tujuan dari pelaksanaan magang di PT. Devoteam adalah sebagai berikut:

Maksud dari pelaksanaan kegiatan magang ini adalah untuk menambah pengalaman serta memberikan kontribusi nyata dalam proyek migrasi data yang ditangani oleh Devoteam sebagai penyedia layanan *cloud* bagi perusahaan Smartfren. Melalui keterlibatan ini, diharapkan peserta magang dapat memahami proses kerja profesional serta memperoleh wawasan mendalam terkait penerapan teknologi *cloud computing* dalam skala industri.

Tujuan dari kegiatan magang ini adalah untuk mengimplementasikan proses migrasi basis data dari sistem on-premise Greenplum ke Google BigQuery. Proses migrasi ini bertujuan untuk meningkatkan performa, skalabilitas, dan efisiensi dalam pengelolaan data perusahaan, serta mendukung strategi modernisasi infrastruktur data yang tengah dijalankan oleh Smartfren dengan dukungan Devoteam.

1.3 Waktu dan Prosedur Pelaksanaan Kerja Magang

Kegiatan magang di PT. Devoteam berlangsung mulai dari tanggal 3 Februari hingga 2 Agustus dan dilaksanakan dengan sistem hybrid. Selama pelaksanaan magang, peserta akan dibimbing langsung oleh kepala tim produk dengan pembimbing lapangan yaitu Bapak Heryudi Ganesha. Berikut adalah rincian prosedur pelaksanaan kerja magang:

1. Hari kerja: Magang dilakukan dari Senin hingga Jumat, dimulai pukul 08:00 hingga 18:00 WIB.
2. Durasi magang: Magang berlangsung mulai tanggal 3 Februari hingga 2 Agustus.
3. Total jam kerja mingguan: Setiap minggu peserta magang menjalani aktivitas magang dengan total 45 jam kerja.
4. Sistem pelaksanaan: Magang dilaksanakan secara *hybrid*, yaitu kombinasi antara kerja daring *(online)* pada hari Selasa dan Rabu dan kerja langsung di kantor *(onsite)* pada hari Senin, Kamis dan Jumat.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA