

BAB I PENDAHULUAN

1.1 Latar Belakang Masalah

Lembaga Layanan Pendidikan Tinggi Wilayah XIV Papua sebelumnya hanya menangani perguruan tinggi yang ada di dua propinsi yaitu, propinsi papua dan papua barat. Namun dalam upaya usaha percepatan pembangunan di propinsi papua dan papua barat melalui Kementerian Dalam Negeri di mekarkan empat wilayah provinsi baru yaitu, papua barat daya, papua tengah, papua pegunungan dan papua selatan. Hal ini tentunya akan menjadi tantangan dan tanggung jawab yang besar bagi Lembaga Layanan Pendidikan Tinggi Wilayah XIV dalam mengakomodir dan mengelolah seluruh perguruan tinggi yang tersebar di enam provinsi yang ada.

Agar dapat tercipta data yang akurat, tidak tumpang tindih, dan konsisten, diperlukan sistem yang terintegrasi. Perencanaan pengembangan sebuah sistem informasi manajemen sangat diperlukan agar sistem yang dikembangkan dapat berjalan sesuai dengan kebutuhan. Sistem informasi manajemen merupakan kumpulan interaksi antar sistem informasi yang berfungsi mengolah data untuk menyediakan informasi yang akan digunakan oleh semua tingkat manajemen. Sistem informasi manajemen memungkinkan sebuah organisasi untuk lebih mengontrol seluruh aktivitas menjadi lebih terorganisir sehingga memberikan hasil yang maksimal demi keberlangsungan organisasi. Penerapan sistem informasi manajemen sudah menjadi hal biasa karena merupakan jaminan kesuksesan perusahaan. Sistem informasi manajemen tidak hanya mencakup aktivitas-aktivitas yang menggunakan komputer tetapi seluruh aktivitas baik berbasis komputer maupun tidak, karena sistem informasi harus mengikat seluruh elemen dalam sebuah perusahaan demi kelancaran proses bisnis (Nawawi & Rubedo, 2021). Saat ini rekayasa data menjadi semakin penting dalam beberapa tahun terakhir karena ledakan data yang dihasilkan oleh bisnis, pemerintah, individu maupun Pendidikan. Dalam membangun sebuah sistem informasi yang terintegrasi, perencanaan arsitektur data *pipelines* adalah langkah awal yang harus

dilakukan. Arsitektur data *pipelines* merupakan suatu rangkaian set elemen pemrosesan data yang dihubungkan secara berseri sehingga hasil keluaran satu elemen menjadi masukan bagi elemen berikutnya (Science & Science, 2023). Secara konseptual, data *pipelines* adalah jalur dari sumber ke tujuan melalui beberapa transformasi untuk berbagai aplikasi analitik. Langkah pertama adalah menghubungkan ke sumber data mentah dan mengambil beban data menggunakan layanan REST API. Sistem memproses beban, untuk menangani masalah integritas data, seperti data redundan, data yang dilewati, data yang diperbarui dan dihapus, dan perubahan tipe data untuk bidang tertentu berdasarkan skema yang tersedia di sumbernya. Langkah terakhir adalah memuat data bersih ke database berbasis cloud (Cottur & Gadad, 2020). Tiga proses integrasi data yang saling bergantung disebut "Extract-Transform-Load (ETL) digunakan untuk mengambil data dari satu database dan melakukan proses tertentu untuk mendapatkan hasil yang ingin dicapai.

Dengan adanya pemekaran empat provinsi baru peran teknologi sangat di butuhkan dalam membangun suatu infrastruktur data antara kampus Akademi Teknik Biak dengan Lembaga Layanan Pendidikan Tinggi Wilayah XIV. Dari permasalahan yang ada penulis mempunyai ide untuk membangun suatu sistem infrastruktur data terpadu antara Lembaga Layanan Pendidikan Tinggi Wilayah XIV dengan Kampus Akademi Teknik Biak.

2.1 Rumusan Masalah

Dalam penelitian ini terdapat beberapa rumusan masalah diantaranya adalah sebagai berikut:

- a. Bagaimana merancang arsitektur data *pipelines* sistem terpadu pada Kampus Akademi Teknik Biak dengan Lembaga Layanan Pendidikan Tinggi Wilayah XIV.
- b. Bagaimana mendapatkan data dalam jumlah yang besar secara real-time.

3.1 Batasan Masalah

Dengan rumusan masalah tersebut akan dilakukan pembatasan penelitian sebagai berikut:

- a) Menggunakan data internal kampus.
- b) Menggunakan data eksternal Lembaga Layanan Pendidikan Tinggi Wilayah XIV
- c) Menggunakan teknologi *Apache Spark* sebagai piranti pengembangan infrastruktur data pipelines Akademi Teknik Biak.

4.1 Tujuan

Tujuan yang ingin dicapai pada penelitian ini:

- a) Merancang arsitektur data *pipelines* sistem terpadu pada Kampus Akademi Teknik Biak menggunakan teknologi *Apache spark* sebagai *software* piranti pengembangan untuk rekayasa data.
- b) Mendapatkan data dalam jumlah yang besar secara real-time dengan menggunakan proses Extract Transform Load (ETL) dengan memanfaatkan teknologi apache spark streaming.

5.1 Manfaat

Penelitian ini diharapkan dapat memberi manfaat berupa:

1. Penulis mendapatkan pemahaman dan wawasan terkait pengembangan teknologi *data engineering*.
2. Kampus Akademi Teknik Biak memiliki sebuah arsitektur sistem data terpadu.
3. Membantu Lembaga Layanan Pendidikan Tinggi Wilayah XIV dalam mengelola dan mengakomodir semua perguruan tinggi swasta maupun negeri di provinsi papua.