

Makul

Manajemen dan Infrastruktur Data

No.

Date Selasa, 8 Oktober

Lathif Ramadhan\_5231811022\_Saif Data 23 2024

Soal 1. Konsep Dasar Arsitektur Data.

Jelaskan apa yang dimaksud dengan arsitektur data dan sebutkan 3 komponen utama yang membentuk arsitektur data dalam sebuah organisasi!

Jawab:

Arsitektur data adalah kerangka kerja yang menentukan bagaimana data di kumpulkan, disimpan, diproses, dan diakses dalam sebuah organisasi. Arsitektur ini menggabungkan teknologi, proses, dan standar untuk memastikan data dapat dimanfaatkan secara efisien dan aman.

Tiga komponen utama arsitektur data:

1) Data Pipeline

→ Mekanisme tempat data mengalir dari satu titik ke titik lain. Alur data menangani pengumpulan, transformasi, dan penyimpanan data, yang pada akhirnya membentangi cacah untuk dianalisis.

2) Model Kecerdasan Buatan (AI) dan Pembelajaran Mesin (ML)

→ Mengidentifikasi pola dan wawasan dalam kumpulan data untuk hasil yang lebih cepat dan akurat tanpa campur tangan manusia.

3) Real-time Analysis

→ Proses menganalisis dan memperolah wawasan dari data segera setelah data tersebut masuk ke dalam basis data.



Soal 2: Infrastruktur Teknologi untuk Manajemen Data.  
(Apa saja infrastruktur teknologi yang diperlukan untuk mendukung arsitektur data di dalam suatu organisasi?)  
Sebutkan dan jelaskan 4 jenis infrastruktur teknologi yang penting dalam manajemen data!

Jawab:

1) Data base

→ Sistem untuk menyimpan dan mengelola data.

Database relasional seperti MySQL atau NoSQL seperti MongoDB sering digunakan.

2) Data Warehouse

→ Penyimpanan khusus yang digunakan untuk analisis data skala besar. Contohnya adalah Amazon Redshift atau Google BigQuery.

3) ETL Tools

→ Alat untuk mengekstrak data dari sumber, ~~dan~~ mengubahnya menjadi format yang standar, dan memuatnya ke dalam penyimpanan. Alat seperti Apache Nifi atau Talend sering digunakan.

4) Cloud Computing

→ Infrastruktur komputasi berbasis cloud. Seperti AWS, Azure, atau Google Cloud, memungkinkan pengolahan dan penyimpanan data yang skalabel dan fleksibel.



### Soal 3: Manajemen Data Terdistribusi

Sebuah perusahaan besar memiliki data yang tersebar di beberapa lokasi geografis, dengan kantor pusat dan beberapa cabang di seluruh dunia. Jelaskan dan Pendekatan untuk mengelola data yang terdistribusi dan Sebutkan tantangan yang mungkin dihadapi dalam masing-masing Pendekatan!

Jawab:

Untuk mengelola data yang tersebar di beberapa lokasi geografis, ada dua Pendekatan umum:

#### 1). Data Centralization (Sentralisasi Data)

→ Semua data dikumpulkan dan disimpan di satu lokasi pusat. Ini memudahkan manajemen dan analisis data, tapi membutuhkan bandwidth yang tinggi dan bisa menyebabkan latency (Penundaan).

Tantangan: Masalah utama adalah kecepatan akses, terutama jika kantor cabang jauh dari pusat data.

#### 2). Data Replication (Replikasi Data)

→ Data disimpan di beberapa lokasi berbeda, dengan Salinan (replica) yang sama di setiap lokasi. Ini mempercepat akses data, tetapi memerlukan Sinkronisasi yang konstan untuk menjaga konsistensi.

Tantangan: Sinkronisasi antar lokasi bisa kompleks dan meningkatkan risiko terjadinya konflik data (Data inconsistency).



Soal 4: Peran Data Governance dalam Arsitektur Data Apa yang dimaksud dengan data governance, dan mengapa penting dalam arsitektur data Sebuah organisasi? Jelaskan 3 elemen utama yang menjadi fokus dalam Penerapan data Governance!

Jawab:

Data Governance adalah praktik untuk mengelola kualitas, keamanan, dan akses data dalam sebuah organisasi. Ini memastikan bahwa data dikelola dengan benar dan sesuai regulasi.

Tiga elemen utama data governance:

1) Data Quality

Memastikan data yang disimpan akurat, konsisten, dan dapat diandalkan

2) Data Security

Melindungi data dari akses yang tidak sah atau kebocoran, misalnya dengan enkripsi atau kontrol akses.

3) Compliance

Memastikan bahwa data dikelola sesuai dengan regulasi dan hukum yang berlaku, seperti GDPR atau HIPAA.



## Soal 5: Infrastruktur Cloud dalam Arsitektur Data Modern.

Sebuah Organisasi Sedang Mempertimbangkan Untuk Beralih dari infrastruktur lokal (on-premise) ke cloud dalam Manajemen datanya. Jelaskan 3 Keuntungan Utama menggunakan cloud infrastructure untuk arsitektur data organisasi, dan 2 tantangan yang mungkin dihadapi dalam transisi ke cloud.

Jawab: Menggunakan cloud untuk arsitektur data menawarkan banyak keuntungan. Seperti:

- 1). Scalability → Kapasitas penyimpanan dan pemrosesan dapat ditingkatkan atau diturunkan dengan cepat sesuai kebutuhan.
- 2). Cost Efficiency → Hanya membayar sesuai dengan penggunaan, sehingga lebih hemat biaya dibandingkan dengan infrastruktur On-Premise.
- 3). Accessibility → Data dapat diakses dari mana saja, selama ada koneksi internet.

Namun, ada juga tantangan dalam transisi ke cloud:

- 1) Data Security → Keamanan data menjadi perhatian utama, karena data berada di luar infrastruktur internal perusahaan.
- 2) Migration Complexity → Proses migrasi data dari sistem on-premise ke cloud bisa rumit dan membutuhkan waktu serta sumber daya yang cukup besar.