

# Tugas Big Data Architecture and Infrastructure

## Pertemuan 6

### *Hadoop Platform Management*

---

<b>Nama</b>	<b>: Lukas Febrian Laufra</b>
<b>Kelas/Nim</b>	<b>: TI22J/20220040076</b>
<b>Dosen Pengajar</b>	<b>: Pak Ir. Somantri, S.T, M.Kom</b>
<b>Waktu Pengerjaan</b>	<b>: 03.52/22 November 2024</b>

---

#### Soal

##### **Kasus: Analisis Data Besar di Perusahaan Retail**

1. Perusahaan retail **MegaMart** ingin meningkatkan pemahaman mereka tentang perilaku pelanggan dan pola pembelian produk. Saat ini, perusahaan memiliki banyak data dari berbagai sumber, seperti transaksi penjualan, data pelanggan, dan interaksi digital melalui aplikasi seluler. Mereka ingin memanfaatkan **Hadoop** untuk menyimpan, memproses, dan menganalisis data dalam skala besar ini.
2. MegaMart ingin menganalisis data transaksi mereka untuk memahami pola pembelian produk. MapReduce adalah fitur Hadoop yang memungkinkan pemrosesan data besar secara paralel. Jelaskan bagaimana mekanisme MapReduce bekerja, dan bagaimana MapReduce dapat mempermudah MegaMart dalam melakukan analisis pola pembelian pada skala besar.

#### Jawaban

##### **1. Mekanisme MapReduce**

MapReduce adalah model pemrograman yang memungkinkan pemrosesan data besar secara paralel. Prosesnya terdiri dari dua fase utama:

###### **a. Fase Map:**

- i. Data dibaca dari HDFS dan dibagi menjadi beberapa bagian.
- ii. Setiap bagian diproses di node yang berbeda untuk menghasilkan pasangan kunci-nilai. Contohnya, untuk analisis transaksi, setiap transaksi bisa diubah menjadi pasangan kunci-nilai di mana kunci adalah ID produk dan nilai adalah jumlah yang dibeli.

###### **b. Fase Reduce:**

- i. Pasangan kunci-nilai dari fase Map yang memiliki kunci sama dikelompokkan.
- ii. Data ini kemudian diproses untuk menghasilkan output akhir, seperti total penjualan per produk atau pola pembelian berdasarkan waktu.

##### **2. Manfaat untuk MegaMart**

Dengan menggunakan MapReduce, MegaMart dapat:

- a. Menganalisis data dalam volume besar: Proses paralel akan mengurangi waktu yang diperlukan untuk menganalisis data transaksi.
- b. Menemukan pola pembelian: Hasil dari proses Reduce akan memberikan wawasan tentang produk yang sering dibeli bersama, tren musiman, dan perilaku pelanggan.
- c. Meningkatkan pengambilan keputusan: Analisis ini membantu MegaMart mengoptimalkan inventaris, penawaran produk, dan strategi pemasaran berdasarkan pola pembelian yang teridentifikasi.

Dengan demikian, MapReduce memungkinkan MegaMart untuk melakukan analisis data besar secara efisien dan efektif.