

# SKALABILITAS DAN KINERJA DALAM BIG DATA (Lanjutan)

Mata Kuliah Infrastruktur dan Teknologi Big Data

Oleh: Galih Hermawan, S.Kom., M.T.  
Program Studi Teknik Informatika. FTIK.  
Universitas Komputer Indonesia

# Faktor-Faktor yang Memengaruhi Kinerja

- ❖ Volume: Jumlah data yang besar dapat memengaruhi kinerja sistem Big Data. Semakin besar volume data, semakin sulit untuk memproses dan menganalisis data.
- ❖ Velocity: Kecepatan data yang masuk ke dalam sistem dapat memengaruhi kinerja sistem Big Data. Semakin cepat data masuk ke dalam sistem, semakin cepat pula data harus diproses.
- ❖ Value: Nilai data yang dihasilkan dapat memengaruhi kinerja sistem Big Data. Data yang bernilai dapat membantu meningkatkan kinerja sistem Big Data.
- ❖ Veracity: Keakuratan data dapat memengaruhi kinerja sistem Big Data. Data yang tidak akurat dapat mempengaruhi hasil analisis data dan kinerja sistem secara keseluruhan.
- ❖ Variation: Keragaman data dapat memengaruhi kinerja sistem Big Data. Data yang beragam dapat mempersulit proses analisis data.

# Contoh Kasus

- ❖ Sebuah sistem Big Data yang digunakan untuk memproses data transaksi penjualan ritel akan memiliki kinerja yang lebih rendah jika volume data transaksi penjualan ritel meningkat secara signifikan.
- ❖ Sebuah sistem Big Data yang digunakan untuk memproses data *streaming* video akan memiliki kinerja yang lebih rendah jika kecepatan *streaming* video meningkat secara signifikan.
- ❖ Sebuah sistem Big Data yang digunakan untuk memproses data sensor industri akan memiliki kinerja yang lebih rendah jika keanekaragaman data sensor industri meningkat secara signifikan.
- ❖ Sebuah sistem Big Data yang digunakan untuk memproses data medis akan memiliki kinerja yang lebih rendah jika kesalahan data medis meningkat secara signifikan.

# Analisis dan Evaluasi

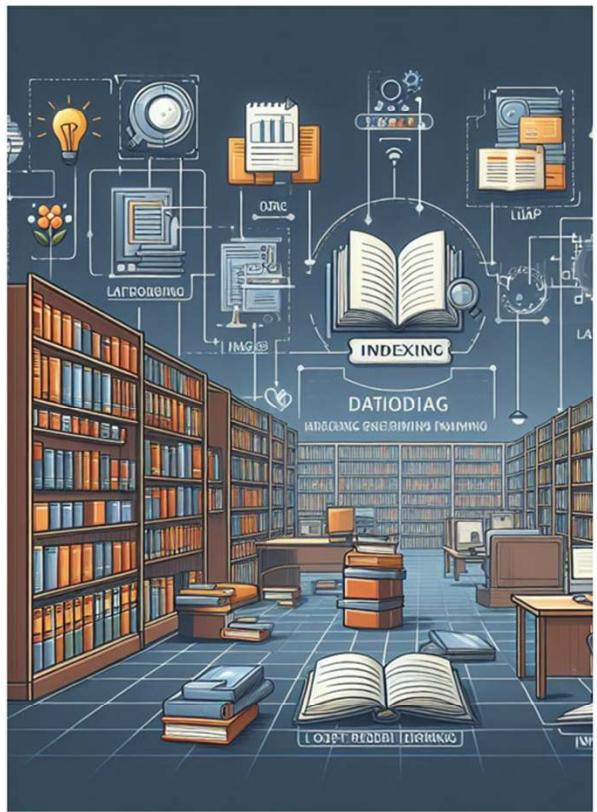
- ❖ Analisis dan evaluasi dilakukan untuk mengetahui sejauh mana faktor-faktor tersebut memengaruhi kinerja sistem Big Data.
- ❖ Analisis dan evaluasi dapat dilakukan dengan menggunakan teknik-teknik seperti analisis regresi, analisis korelasi, dan analisis faktor.
- ❖ Hasil analisis dan evaluasi dapat digunakan untuk meningkatkan kinerja sistem Big Data.

# Contoh Pengukuran Kinerja

- ❖ Pengukuran Volume Data: grafik pertumbuhan data dari waktu ke waktu.
- ❖ Pengukuran Kecepatan Data: diagram alur data dan waktu respons sistem.
- ❖ Penanganan Keanekaragaman Data: efektivitas algoritma dan metode pemrosesan yang digunakan.
- ❖ Manajemen Kesalahan Data: tingkat kesalahan dan dampaknya terhadap hasil analisis.

# Teknik Optimasi untuk Kinerja Sistem Big Data

1. Pengindeksan Data
2. Pengoptimalan Query
3. Penggunaan Caching
4. Monitoring dan Diagnosa
5. Contoh Implementasi di Lingkungan Produksi
6. Studi Kasus: Perbaikan Kinerja Sistem Big Data



# Pengindeksan Data

- ❖ Pengindeksan data adalah proses pembuatan indeks untuk data.
- ❖ Indeks adalah struktur data yang menyimpan informasi tentang data untuk mempercepat pencarian data.
- ❖ Manfaat pengindeksan data:
  - Meningkatkan kecepatan pencarian data
  - Mengurangi jumlah I/O yang diperlukan untuk mencari data
  - Mengurangi penggunaan memori

# Pengindeksan Data (lanjutan)

## ❖ Jenis-jenis indeks:

- Indeks *hash*: Indeks *hash* mengasosiasikan setiap nilai kunci dengan lokasi data yang sesuai. Indeks *hash* sangat cepat untuk pencarian data, tetapi tidak dapat digunakan untuk mencari data yang berada di antara dua nilai kunci tertentu.
- Indeks *tree*: Indeks *tree* mengurutkan nilai kunci dalam struktur pohon. Indeks *tree* dapat digunakan untuk mencari data yang berada di antara dua nilai kunci tertentu, tetapi lebih lambat daripada indeks *hash*.
- Indeks *spatial*: Indeks *spatial* mengurutkan nilai kunci dalam struktur spasial. Indeks *spatial* digunakan untuk mencari data yang berada dalam suatu wilayah geografis atau spasial.

## ❖ Contoh:

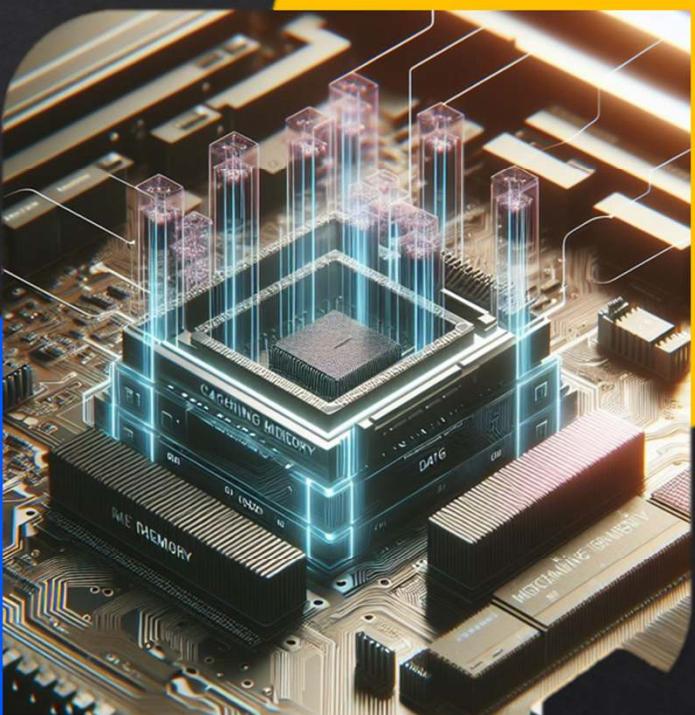
- Sistem *e-commerce* dapat menggunakan indeks untuk mempercepat pencarian produk berdasarkan nama, kategori, atau harga.
- Sistem media sosial dapat menggunakan indeks untuk mempercepat pencarian *posting* (tulisan/pesan) berdasarkan tanggal, penulis, atau tagar.
- Sistem pemetaan dapat menggunakan indeks untuk mempercepat pencarian lokasi berdasarkan alamat atau koordinat.

# Pengoptimalan Query

- ❖ Pengoptimalan *query* adalah proses meningkatkan efisiensi *query*.
- ❖ Pengoptimalan *query* dapat dilakukan dengan berbagai cara, seperti:
  - Memilih algoritma *query* yang tepat
  - Mengubah struktur data
  - Menggunakan teknik-teknik lain, seperti *caching*, *indexing*, dan *load balancing*
- ❖ Manfaat pengoptimalan *query*:
  - Meningkatkan kecepatan pemrosesan *query*
  - Mengurangi penggunaan sumber daya
  - Meningkatkan kepuasan pengguna
- ❖ Contoh:
  - Sebuah sistem Big Data yang digunakan untuk memproses data transaksi penjualan ritel dapat menggunakan pengoptimalan *query* untuk meningkatkan kecepatan pemrosesan *query* yang melibatkan banyak data.



# Penggunaan Caching



- ❖ *Caching* adalah proses menyimpan data yang sering digunakan dalam memori *cache*.
- ❖ Memori *cache* adalah memori yang lebih cepat daripada memori utama.
- ❖ Manfaat penggunaan *caching*:
  - Meningkatkan kecepatan akses data
  - Mengurangi jumlah I/O yang diperlukan untuk mengakses data
  - Mengurangi penggunaan memori utama
- ❖ Contoh:
  - Sebuah sistem Big Data yang digunakan untuk memproses data transaksi penjualan ritel dapat menggunakan *caching* untuk menyimpan data transaksi yang sering digunakan.

# Monitoring dan Diagnosa

- ❖ *Monitoring* adalah proses mengumpulkan data tentang kinerja sistem.
- ❖ Diagnosa adalah proses mengidentifikasi penyebab masalah kinerja sistem.
- ❖ Manfaat *monitoring* dan diagnosa:
  - Membantu mengidentifikasi masalah kinerja sistem
  - Membantu memecahkan masalah kinerja sistem
  - Meningkatkan kinerja sistem secara keseluruhan
- ❖ Contoh:
  - Sebuah sistem Big Data dapat menggunakan *monitoring* dan diagnosa untuk mengidentifikasi masalah kinerja yang disebabkan oleh kesalahan data.

# Contoh Implementasi di Lingkungan Produksi

- ❖ Google menggunakan pengindeksan data untuk mempercepat pencarian data di mesin pencarinya.
- ❖ Netflix menggunakan pengoptimalan *query* untuk meningkatkan kecepatan rekomendasi film dan acara TV.
- ❖ Amazon menggunakan *caching* untuk meningkatkan kecepatan pemrosesan transaksi belanja *online*.

# Studi Kasus: Perbaikan Kinerja Sistem Big Data

- ❖ Sebuah perusahaan ritel memiliki sistem Big Data yang digunakan untuk memproses data transaksi penjualan.
- ❖ Sistem tersebut memiliki kinerja yang buruk, sehingga menyebabkan keterlambatan dalam pengiriman laporan penjualan.
- ❖ Perusahaan tersebut melakukan analisis dan evaluasi terhadap faktor-faktor yang mempengaruhi kinerja sistem.
- ❖ Hasil analisis menunjukkan bahwa volume data transaksi penjualan meningkat secara signifikan.
- ❖ Perusahaan tersebut menerapkan teknik pengindeksan data untuk meningkatkan kinerja sistem.
- ❖ Hasilnya, kinerja sistem meningkat secara signifikan dan keterlambatan dalam pengiriman laporan penjualan dapat dihilangkan.

# Ringkasan

Studi Kasus - Implementasi Skalabilitas dan Optimasi

- Pengalaman Praktis dalam Mengimplementasikan Skalabilitas
- Hasil Implementasi Teknik Optimasi
- Analisis Dampak terhadap Kinerja Sistem



# Pengalaman Praktis dalam Mengimplementasikan Skalabilitas

- ❖ Skala horizontal:
  - Menambah node ke cluster
  - Mendistribusikan data dan aplikasi ke beberapa node
- ❖ Skala vertikal:
  - Menambah kapasitas sumber daya ke node yang ada
- ❖ Tantangan dalam mengimplementasikan skalabilitas:
  - Memilih arsitektur skalabilitas yang tepat
  - Mengelola infrastruktur yang kompleks
  - Menjaga ketersediaan dan ketahanan sistem

# Hasil Implementasi Teknik Optimasi

- ❖ Pengindeksan data:

- Meningkatkan kecepatan pencarian data
  - Mengurangi jumlah I/O yang diperlukan untuk mencari data

- ❖ Pengoptimalan *query*:

- Meningkatkan kecepatan pemrosesan *query*
  - Mengurangi penggunaan sumber daya

- ❖ Penggunaan *caching*:

- Meningkatkan kecepatan akses data
  - Mengurangi jumlah I/O yang diperlukan untuk mengakses data

# Analisis Dampak terhadap Kinerja Sistem

- ❖ Peningkatan kinerja:

- Mengurangi waktu pemrosesan
- Meningkatkan kapasitas sistem
- Meningkatkan kepuasan pengguna

- ❖ Efisiensi biaya:

- Mengurangi kebutuhan sumber daya
- Mengurangi biaya operasional

# Kesimpulan

- ❖ Implementasi skalabilitas dan optimasi dapat meningkatkan kinerja sistem Big Data secara signifikan.
- ❖ Pemilihan teknik skalabilitas dan optimasi yang tepat harus disesuaikan dengan kebutuhan dan anggaran perusahaan.