

# LECTURE NOTES

## 7023T Advanced Database System

### Enrichment Activity

### Session 03

### Technical Architecture

## LEARNING OUTCOMES

- Peserta diharapkan mampu memahami model arsitektur sistem DW/BI.
- Peserta diharapkan dapat menjelaskan komponen *back room*, *presentation server*, dan *front room*.
- Peserta diharapkan mampu memahami infrastruktur teknis yang dapat digunakan untuk membangun sebuah sistem DW/BI.

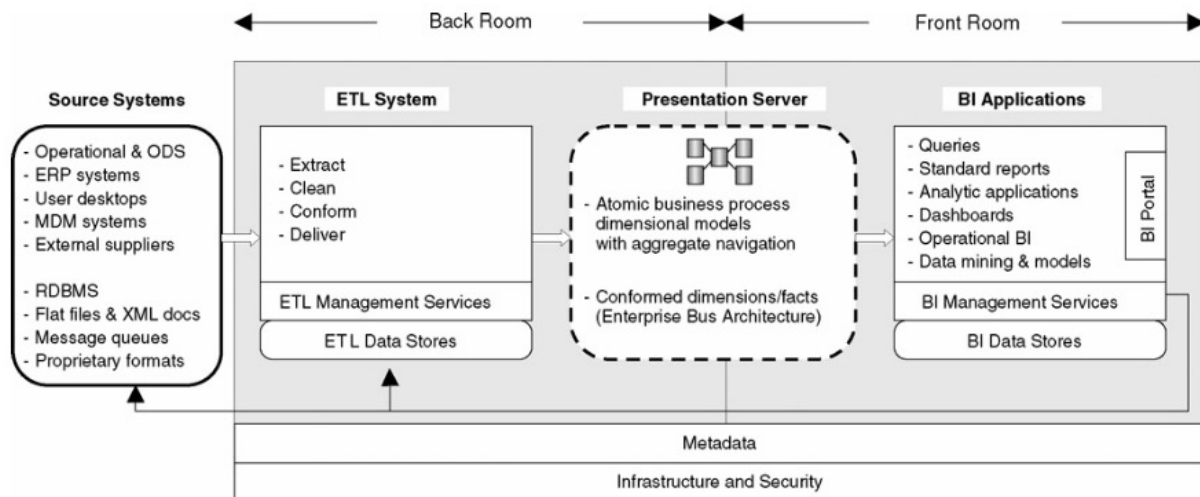
### OUTLINE MATERI (Sub-Topic):

1. *Technical Architecture*
2. *Back Room Architecture*
3. *Presentation Server Architecture*
4. *Front Room Architecture*
5. *Technical Infrastructure*

## Technical Architecture

*Technical architecture* atau arsitektur teknis adalah kerangka kerja yang menjelaskan aturan-aturan dan struktur dari desain keseluruhan dari sebuah sistem. Contoh *technical architecture* seperti *data architecture*, *client-server architecture*, *networking architecture*, atau *networking architecture*. *Data architecture* mendefinisikan bagaimana data mengalir dalam sistem. *Technical infrastructure* adalah cara untuk mengimplementasikan *technical architecture*, yang terdiri dari berbagai komponen seperti teknologi, platform, database, dan komponen lainnya.

Arsitektur teknis dari sebuah data warehouse membantu *project manager* dalam mendesain kekeluruhan proyek serta seluruh anggota tim untuk memahami gambaran besar dari proyek, juga tugas tanggung-jawabnya masing-masing. *Top-level view* dari sebuah arsitektur DW/BI dapat digambarkan seperti diagram pada gambar 1. Data mengalir dari *Source Systems* menuju ke sistem ETL (disebut juga sebagai *Back Room* dari data warehouse), kemudian berlanjut ke *Presentation Server* yang memuat dimensional model, terakhir ke aplikasi BI (disebut juga sebagai *Front Room*).



**Gambar 1.** Model arsitektur *high-level* DW/BI

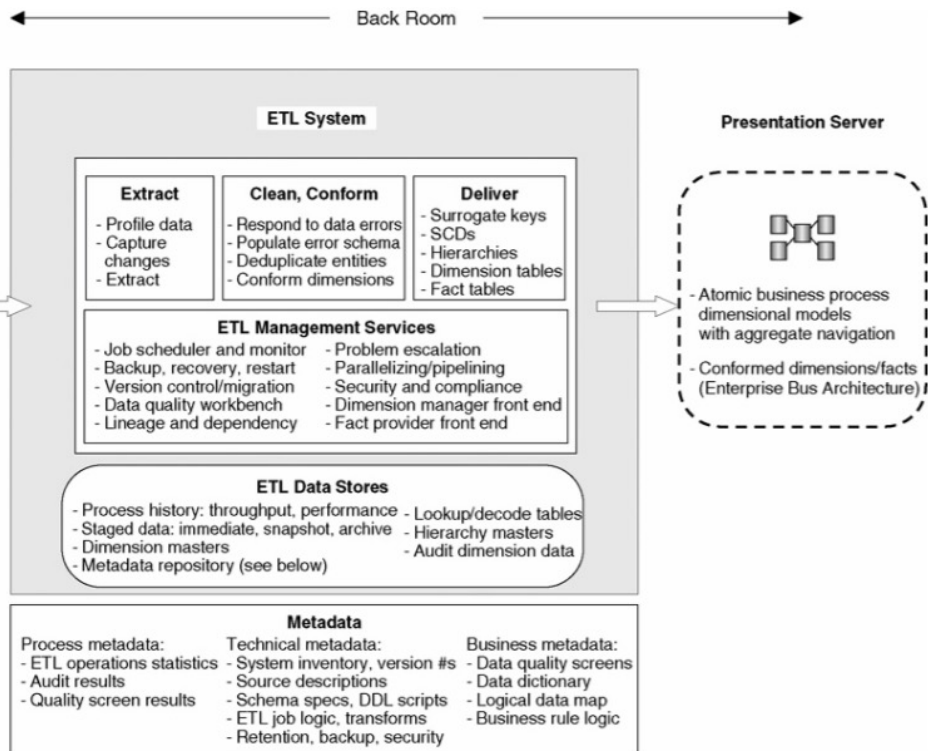
Pada gambar 1 dapat dilihat bahwa metadata melingkupi keseluruhan arsitektur. Metadata atau data mengenai data, adalah informasi yang menjelaskan struktur, operasi-operasi, dan isi dari suatu sistem DW/BI. Metadata memegang peranan penting dalam arsitektur DW/BI. Semua metadata perlu disimpan pada satu lokasi agar semua subsistem DW/BI dapat berbagi pakai metadata tersebut. Dengan menggunakan strategi ini maka

apabila terjadi perubahan pada metadata, semua perubahan tersebut dapat langsung diadopsi oleh semua subsistem DW/BI. Strategi ini juga sangat membantu dalam proses audit serta menjamin kualitas dan integritas data. Namun pada prakteknya, sangat sedikit perangkat lunak yang mengizinkan berbagi pakai metadata dengan perangkat lunak yang berasal dari vendor lain. Salah satu pendekatan untuk menghindari masalah tersebut adalah dengan hanya menggunakan produk yang berasal dari satu vendor. Beberapa vendor juga mendukung standar *Common Warehouse Metamodel* (CWM) bersama *Meta Object Facility* (MOF) yang memungkinkan penggunaan metadata pada sistem yang terdiri dari multi vendor.

## ***Back Room Architecture***

Komponen utama dari *back room architecture* seperti digambarkan pada gambar 2 adalah sistem ETL. Sistem ETL terdiri dari tiga operasi, yakni (E) mengekstrak data dari sumber data, (T) melakukan transformasi seperti *cleansing*, dan (L) yang akan melakukan loading data hasil transformasi ke *presentation server*. Sistem ETL pada umumnya merupakan elemen standar dari sebuah sistem perangkat lunak DW/BI. Fase ETL patut mendapat perhatian khusus dalam proyek pengembangan sistem DW/BI karena fase ini dapat menghabiskan sekitar 70% waktu pengembangan. Perlu diyakinkan bahwa sistem ETL yang digunakan dapat mengekstrak data yang akan digunakan pada sistem DW/BI, jika tidak maka proses ETL untuk data yang tidak didukung perlu dilakukan secara manual dengan mengembangkan kode sendiri.

Sumber data yang umum dari sistem DW/BI seperti sistem *Enterprise Resource Planning* (ERP), *Operational Data Store* (ODS), sistem *Master Data Management* (MDM), file *spreadsheet*, maupun XML. Sistem ERP tersusun atas beberapa modul mencakup fungsi-fungsi utama dari bisnis seperti *human resources*, *sales*, dan *manufacturing*. Proses integrasi data yang berasal dari sistem ERP tidak mudah dilakukan karena jumlah tabel pada sumber data bisa sekitar ribuan. Mencari tabel-tabel yang relevan diantara ribuan tabel tersebut tentu bukanlah hal yang mudah untuk dilakukan. Selain itu, sistem ERP tidak dirancang untuk *query* untuk kebutuhan analitis. Beberapa vendor besar ERP menyediakan solusi untuk mengatasi permasalahan ini. Beberapa vendor dari sistem ETL juga menyediakan solusi untuk mengekstrak data dari sistem ERP standar.



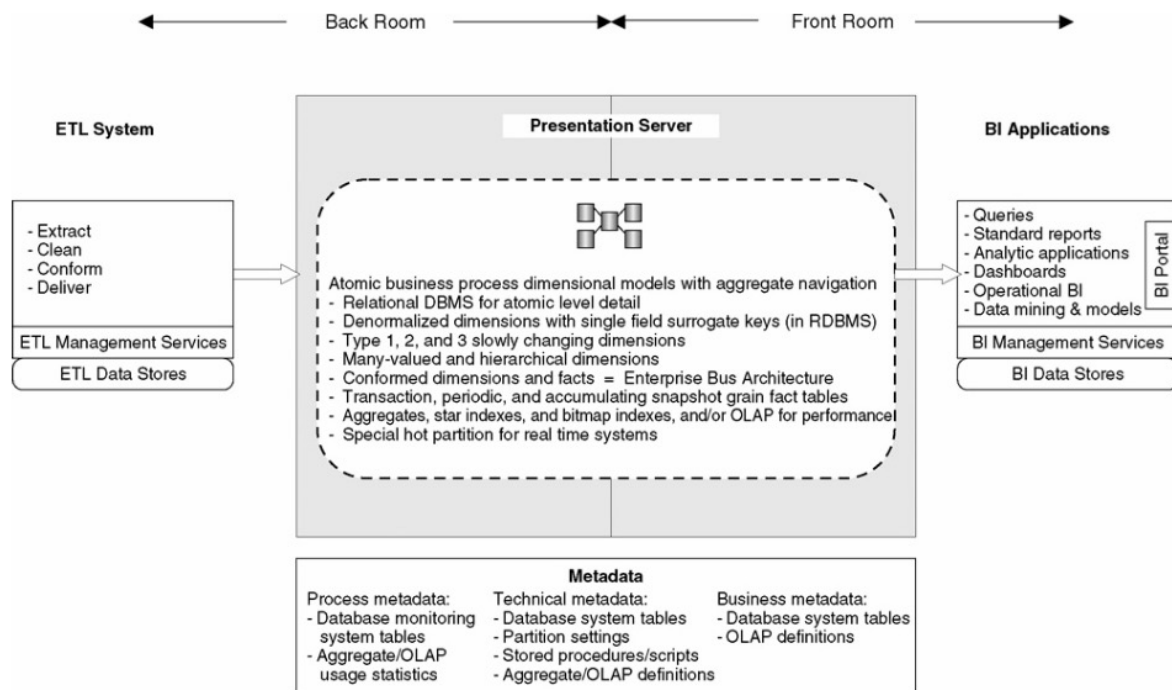
**Gambar 2.** Model arsitektur *back room*

Kimball mendefinisikan sistem ODS dengan cara yang berbeda dengan *Corporate Information Factory* dari Inmon. Menurut Kimball, ODS adalah salinan dari database operasional untuk mendukung kebutuhan data yang berumur kurang dari 24 jam. ODS perlu diintegrasikan ke dalam sistem DW/BI. MDM merupakan salinan utama dari entitas kunci seperti *customer* atau *product*. MDM berperan dalam integrasi data yang berasal dari beberapa sumber berbeda. Oleh karena itu, MDM juga merupakan komponen yang bermanfaat dalam DW/BI, khususnya jika diimplementasikan dengan konsep SOA (*software as service*).

## ***Presentation Server Architecture***

*Presentation server* seperti diilustrasikan pada gambar 3 menyimpan data yang akan diakses oleh pengguna dari kalangan bisnis melalui *query*. *Query* untuk kebutuhan analitik dapat dikatakan bersifat tidak dapat diprediksi dalam banyak aspek, oleh karena itu *presentation server* harus bersifat fleksibel agar dapat mengikuti perubahan yang terjadi. Struktur database pada *presentation server* yang berupa *star scheme* dirancang sesuai dengan

*enterprise bus matrix*. Untuk dataset berukuran besar, data atomik dihasilkan dari operasi agregat. Agregat dibangun dengan basis harian, dan nilainya selalu berubah sesuai dengan kebutuhan penggunaanya. Agregat yang lama tidak digunakan akan dihapus dan digantikan dengan agregat yang baru. Teknik ini disebut *usage-based optimization* yang sangat bermanfaat untuk dataset berukuran besar.



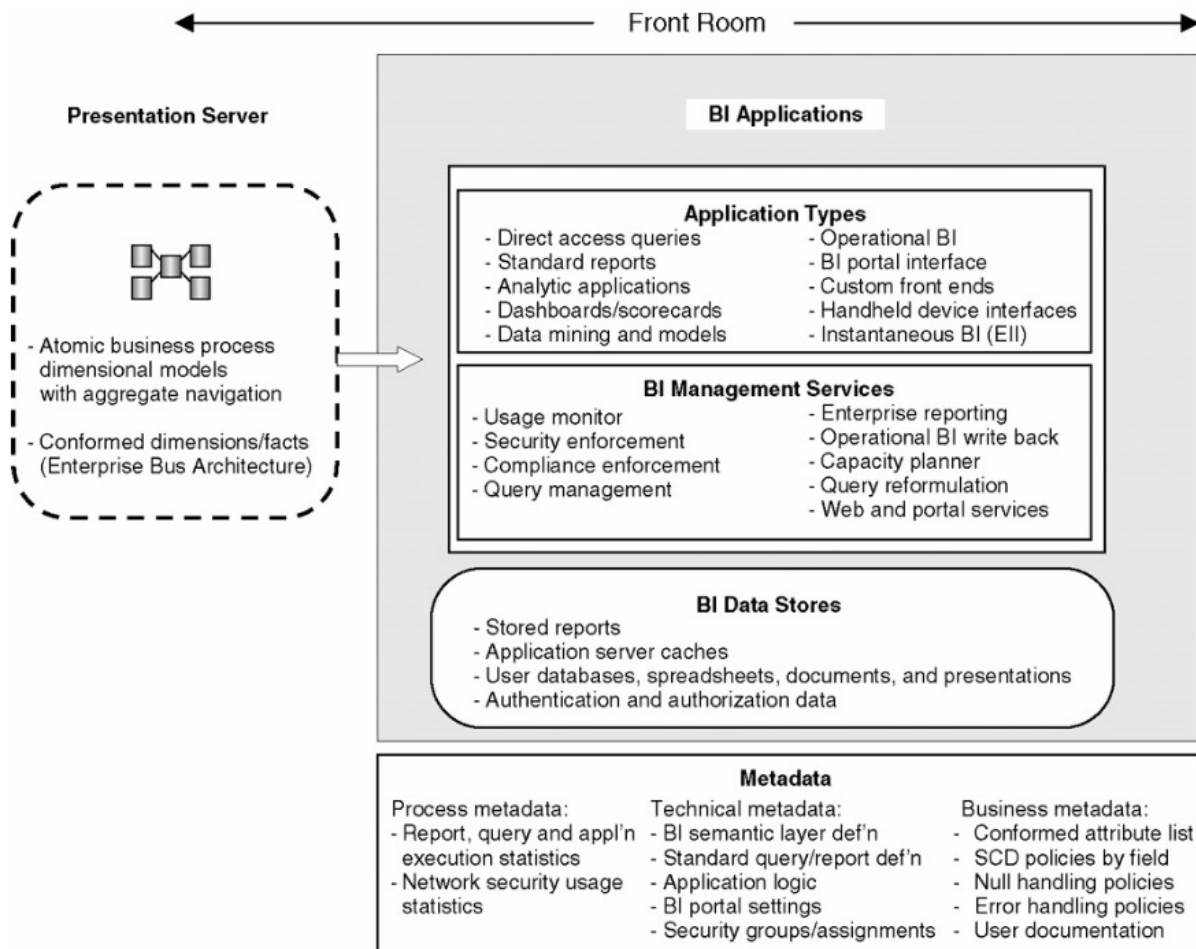
**Gambar 3.** Model arsitektur *presentation server*

## ***Front Room Architecture***

Mayoritas pengguna bisnis memanfaatkan data warehouse pada bagian *front room*. *Front room* terdiri dari beberapa jenis aplikasi BI yang dapat diakses oleh pengguna bisnis melalui berbagai jenis *application interface* dan *BI management services*, yang mendefinisikan pengguna desktop dengan *shared services* seperti diilustrasikan pada gambar 4. *Shared services* dapat berupa *metadata services*, *security services*, *usage monitoring*, *query management*, *enterprise reporting services*, *web access*, atau jenis servis lainnya. Saat ini terdapat beberapa jenis aplikasi BI seperti *direct access queries*, *standard reports*, *analytic applications*, *dashboard* dan *scorecards*, *data mining* dan *model*, *operational BI*. Aplikasi-



aplikasi BI tersebut dapat diakses melalui berbagai macam antarmuka seperti BI portal, *mobile devices*, maupun *instantaneous BI*.

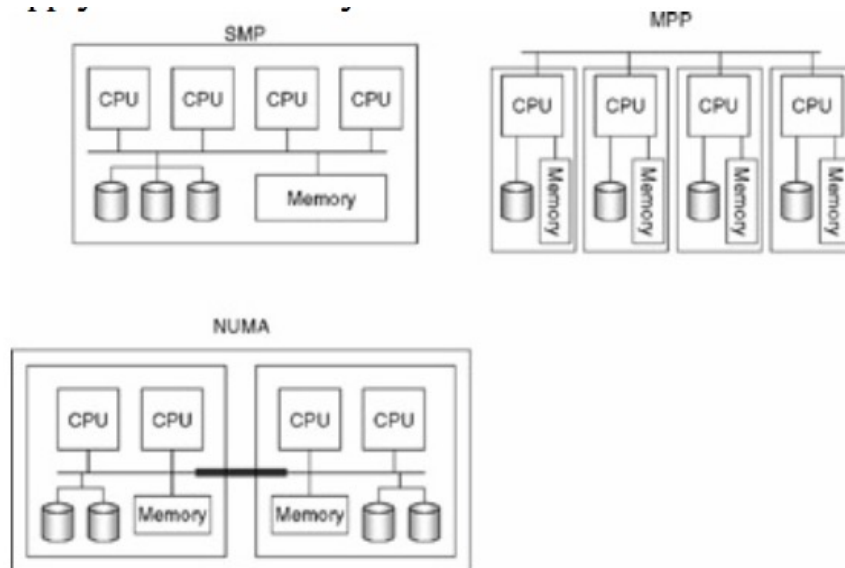


**Gambar 4.** Model arsitektur *front room*.

## ***Technical Infrastructure***

Infrastruktur DW/BI terdiri dari perangkat keras, jaringan, dan fungsi tingkat rendah. Arsitektur perangkat keras untuk pemrosesan paralel mencakup *Symmetric Multiprocessing* (SMP), *Massive Parallel Processing* (MPP), and *Non-Uniform Memory Architecture* (NUMA) seperti diilustrasikan pada gambar 5. SMP adalah sebuah komputer dengan multi prosesor yang semuanya dikelola oleh satu sistem operasi dan berbagi *disk* dan *memory*. Arsitektur SMP sangat cocok untuk *query ad-hoc*. MPP adalah sekumpulan komputer independen dengan masing-masing *disk* dan *memory*, dan sistem operasi. Komputer-komputer tersebut bekerja sama dalam menjalankan suatu operasi dengan cara saling berkomunikasi satu sama

lainnya. Arsitektur ini baik untuk menyimpan data atomik dan menjalankan pendekatan *brute-force* dalam mengeksekusi *query*. NUMA merupakan cluster yang dibangun dari sekumpulan komputer SMP dengan *bandwidth* besar dengan kemampuan koordinasi yang lebih baik dibanding arsitektur MPP. Cluster adalah sekumpulan komputer yang saling terkoneksi dan dikonfigurasi sedemikian sehingga tampak sebagai sebuah server (bukan multi komputer) oleh *client*. Cluster komputer bersifat *scale out* atau diekspansi dengan cara menambahkan beberapa komputer baru, maupun *scale up* atau diupgrade menggunakan mesin yang lebih tinggi spesifikasinya.



**Gambar 5.** Arsitektur pemrosesan paralel



## SIMPULAN

- Model arsitektur DW/BI dapat dibagi menjadi tiga komponen utama, yaitu *back room*, *presentation server*, dan *front room*.
- Komponen utama dari arsitektur *back room* adalah sistem ETL
- *Presentation server* menyediakan data dalam bentuk yang optimal untuk operasi *query* analitik.
- *Back room* menyediakan aplikasi BI bagi pengguna bisnis yang akan memanfaatkan sistem DW/BI untuk mendukung proses pengambilan keputusan.

## DAFTAR PUSTAKA

1. Kimball, R. (2008). *The Data Warehouse Lifecycle Toolkit*. John Wiley & Sons.
2. Kimball, R., & Ross, M. (2011). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons.
3. Inmon, W. H. (2005). *Building the Data Warehouse*. John Wiley & Sons.