

# *DATA WAREHOUSE*

# Outline

- Konsep dan Arsitektur Data Warehouse
- Alur Data Warehouse
- Teknologi dan Peralatan Data Warehouse
- Perancangan Data Warehouse
- Penggunaan Oracle pada Data Warehouse

# Konsep dan Arsitektur Data Warehouse

- Data warehouse : kumpulan data yang berorientasi subjek, terintegrasi, *time-variant*, dan *non volatile* untuk mendukung proses pengambilan keputusan.

*Subject oriented* :

Data warehouse mengorganisasi-kan subjek utama perusahaan (pe-  
langgan, produk, dan penjualan),  
bukan area aplikasi utama (faktur  
pelanggan, pengawasan stock, dan  
penjualan produk). Hal ini meng-  
gambarkan kebutuhan untuk me-  
nyimpan data pendukung keputus-  
ab daripada aplikasi yang berorien-  
tasi data.

- *Integrated*

pengambilan secara bersamaan sumber data yang berasal dari sistem aplikasi berbagai perusahaan besar yang berbeda. Sumber data sering tidak konsisten, misal berbeda format. Sumber data yang terintegrasi harus dapat dibuat konsisten untuk menggambarkan view gabungan data ke pemakai.

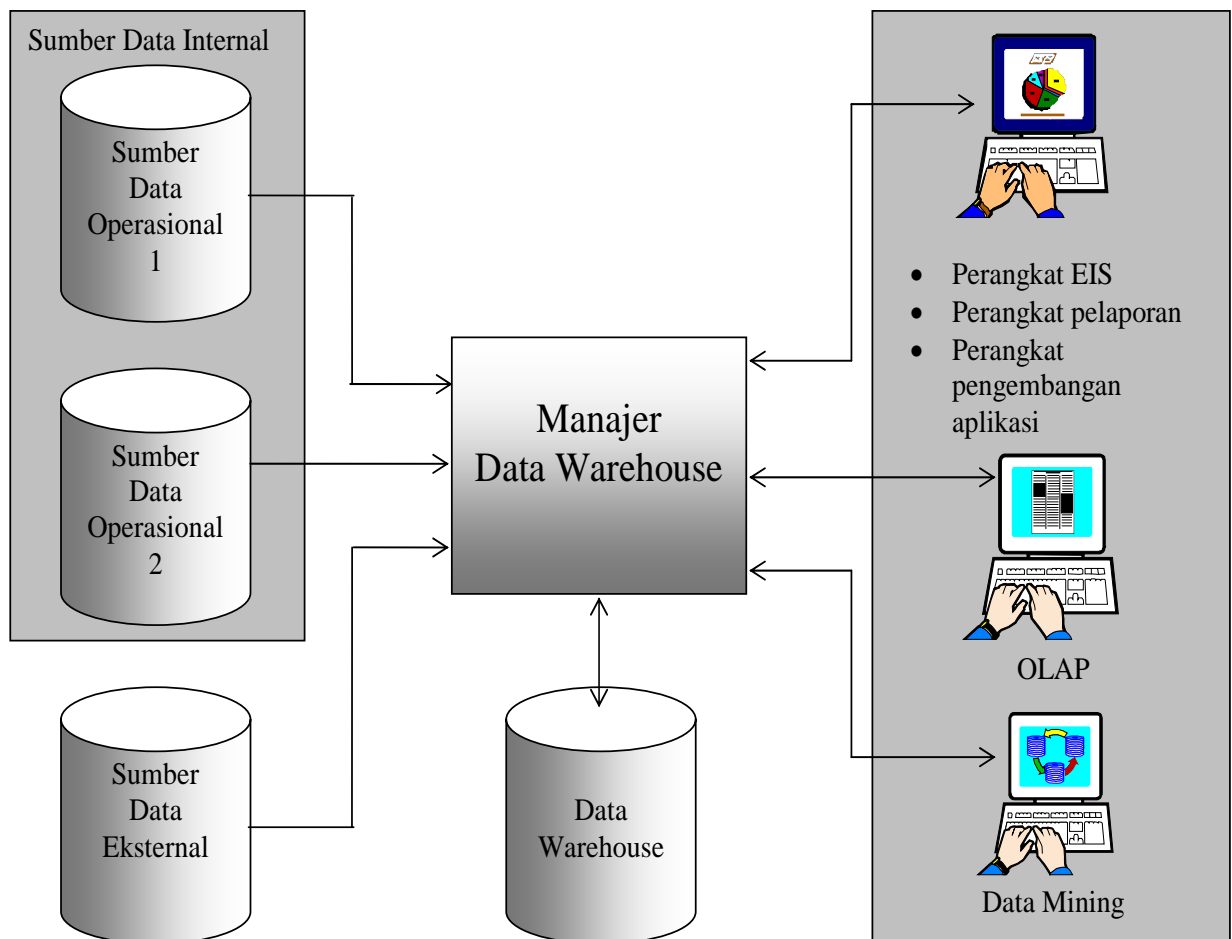
- Time variant

data warehouse hanya akurat dan valid pada saat tertentu atau beberapa interval waktu tertentu.

- *Non volatile*  
data tidak diperbaharui secara *real time* tetapi diperbaharui dari sistem operasional secara regular. Data baru selalu ditambahkan sebagai lampiran pada basis data, bukan menggantikan data lama.
- Keuntungan data warehouse
  - Pengembalian investasi yang tinggi
  - Keuntungan yang kompetitif
  - Meningkatkan produktifitas para pembuat keputusan korporasi

- *Data warehouse* digunakan untuk mendukung pengambilan keputusan, bukan untuk melaksanakan pemrosesan transaksi
- *Data warehouse* hanya berisi informasi-informasi yang relevan bagi kebutuhan pemakai yang dipakai untuk pengambilan keputusan.

# Arsitektur Data Warehouse



Sumber data untuk data warehouse :

1. Data operasional dalam organisasi, misalnya basis data pelanggan dan produk.
2. Sumber eksternal yang diperoleh misalnya melalui Internet, basis data komersial, basis data pemasok atau pelanggan.



- Berbagai data yang berasal dari sumber digabungkan dan diproses lebih lanjut oleh manajer *data warehouse* dan disimpan dalam basis data tersendiri.
- Selanjutnya, perangkat lunak seperti OLAP dan *data mining* dapat digunakan oleh pemakai untuk mengakses *data warehouse*

# Alur Data Warehouse

- Inflow

pengekstrasian, pembersihan, dan pemuatan sumber data.

- Upflow

Menambahkan nilai data ke dalam warehouse melalui ringkasan, pemaketan, dan pendistribusian data.

- Downflow

pengarsipan dan *back up* data ke dalam warehouse.

- **Outflow**  
membuat data agar tersedia untuk pemakai akhir.
- **Meta-flow**  
pengaturan meta-data

## Teknologi dan Peralatan Data Warehouse

Peralatan untuk pengekstrasian,  
pembersihan, dan transformasi :

- Code generator  
membuat program transforma-  
si 3GL/4GL yang dapat disesu-  
aikan berdasarkan sumber dan  
target pendefinisian data.

- Database data replication tools menggunakan database trigger dan recovery log untuk merekam perubahan terhadap sumber data tunggal pada satu sistem dan menggunakan perubahan tsb ke suatu salinan sumber data yang dialokasikan pada sebuah sistem yang berbeda.
- Dynamic transformation engines merekam data dari sumber sistem pada interval yang telah definisikan pemakai, mentransformasikan data kemudian mengirim dan memuat hasilnya kedalam target lingkungan

# Perancangan Data Warehouse

Petunjuk membangun data warehouse :

- Menentukan misi dan sasaran bisnis bagi pembentukan *data warehouse*.
- Mengidentifikasi data dari basis data operasional dan sumber lain yang diperlukan bagi *data warehouse*.
- Menentukan item-item data dalam perusahaan dengan melakukan standarisasi penamaan data dan artinya

- Merancang basis data untuk *data warehouse*
- Membangun kebijakan dalam mengarsipkan data lama sehingga ruang penyimpanan tak menjadi terlalu besar dan agar pengambilan keputusan tidak menjadi terlalu lamban.
- Menarik data produksi (operasional) dan meletakkan ke basis data milik *data warehouse*

# Penggunaan Oracle untuk Data Warehouse

Fasilitas yang disediakan Oracle untuk data warehouse :

## 1. Summary management

mekanisme untuk menyimpan data multi dimensi dan ringkasan perhitungan pada sebuah tabel.

## 2. Analytical functions

penggunaan fungsi analitikal SQL (ranking, moving aggregate, dll) untuk keperluan *business intelligence* dan aplikasi data warehouse.



### 3. Bitmapped indexes

meningkatkan kinerja yang sangat baik terhadap aplikasi data warehouse.

### 4. Advanced join methods

dengan *partition wise-join*, secara dramatis meningkatkan kinerja join yang melibatkan tabel yang telah dipartisi pada *join key*.

5. Sophisticated SQL optimizer optimizer secara dinamik menentukan *path* dan *join* yang paling efisien untuk setiap query
6. Resource management menyediakan pengawasan sumber-sumber sistem yang telah ditentukan pemakai.

# Referensi

- Connolly, Thomas; Begg, Carolyn; Strachan, Anne;  
**Database Systems : A Practical Approach to Design, Implementation and Management**, 3rd edition, Addison Wesley, 2003.
- <http://mti.ugm.ac.id/~adji/courses/resources/lectures/InformSystem/AKA/si-6-2.ppt>

# Kesimpulan

- Data warehouse merupakan kumpulan data yang berorientasi subjek, terintegrasi, *time-variant*, dan *non volatile* untuk mendukung proses pengambilan keputusan.
- Sumber data untuk data warehouse : data operasional dalam organisasi, dan sumber eksternal
- Alur data warehouse : inflow, upflow, downflow, outflow, meta flow.
- Peralatan data warehouse : code generator, database data replication tools, dynamic transformation engines