

DATA WAREHOUSE PENJUALAN PT XYZ

Minda Mora Purba

karo_727@yahoo.com

Abstract

Data warehouse is a decision support system (decision support system) which use and manufacture independent in one part of certain business environment at company, for example sales department. Problems faced by this sales is the difficulty in data collection products and customers, and the acquisition of sales reports are fast and accurate. With these problems the company needs a data warehouse that can be a trusted data source for the activities of this company.

The process of developing the data warehouse includes the analysis phase, the design stage and implementation.

Keywords: warehouse, data, DSS, ETL

PENDAHULUAN

Berkembangnya teknologi dan informasi saat ini telah melahirkan “gunungan” data di bidang ilmu pengetahuan, bisnis dan pemerintah. Kemampuan teknologi informasi untuk mengumpulkan dan menyimpan berbagai tipe data jauh meninggalkan kemampuan untuk menganalisis, meringkas dan mengekstraksi “pengetahuan” dari data. Metodologi tradisional untuk menganalisis data yang ada, tidak dapat menangani data dalam jumlah besar. Para peneliti melihat peluang untuk melahirkan sebuah teknologi baru yang menjawab kebutuhan ini, yaitu *data warehouse*. Teknologi ini sekarang sudah ada dan diaplikasikan oleh perusahaan-perusahaan untuk memecahkan berbagai permasalahan bisnis.

Data warehouse merupakan metode dalam perancangan database yang menunjang DSS (Decision Support System) dan EIS (Executive Information System). Secara fisik *data warehouse* adalah database, tapi perancangan data warehouse dan database sangat berbeda.

Dalam perancangan database tradisional menggunakan normalisasi, sedangkan pada *data warehouse* normalisasi bukanlah cara yang terbaik.

LANDASAN TEORI

Pengertian Data Warehouse

Data warehouse adalah suatu konsep dan kombinasi teknologi yang memfasilitasi organisasi untuk mengelola dan memelihara data historis yang diperoleh dari sistem atau aplikasi operasional. Pemakaian teknologi data warehouse hampir dibutuhkan oleh semua organisasi, tidak terkecuali perpustakaan. Data warehouse memungkinkan integrasi berbagai macam jenis data dari berbagai macam aplikasi atau sistem. Hal ini menjamin mekanisme akses “satu pintu bagi manajemen untuk memperoleh informasi dan menganalisisnya untuk pengambilan keputusan”.

Beberapa konsep dasar tentang data warehouse :

- a. Data warehouse adalah data-data yang berorientasi subjek, terintegrasi, memiliki dimensi waktu,

- serta merupakan koleksi tetap (non-volatile), yang digunakan dalam mendukung proses pengambilan keputusan oleh para manajer di setiap jenjang (namun terutama pada jenjang manajerial yang memiliki peringkat tinggi).
- b. Data warehouse adalah suatu paradigma baru dilingkungan pengambilan keputusan strategik. Data warehouse bukan suatu produk tetapi suatu lingkungan dimana user dapat menemukan informasi strategik. Data warehouse adalah kumpulan data-data logik yang terpisah dengan database operasional dan merupakan suatu ringkasan.
 - c. Data warehouse adalah data yang diperoleh dari proses dimana organisasi mengekstraksi makna dari aset informasi yang mereka miliki. Data warehouse adalah inovasi baru dalam hal teknologi informasi. Sejak dimulai sekitar 15 tahun lalu, konsep data warehouse ini berkembang secara cepat sehingga saat ini konsep data warehouse ini adalah konsep yang paling banyak dibicarakan oleh para ahli di bidang teknologi informasi.
 - d. Data Warehouse adalah pusat repositori informasi yang mampu memberikan database berorientasi subyek untuk informasi yang bersifat historis yang mendukung DSS (Decision Support System) dan EIS (Executive Information System).

- e. Salinan dari transaksi data yang terstruktur secara spesifik pada query dan analisa.
- f. Salinan dari transaksi data yang terstruktur spesifik untuk query dan laporan.

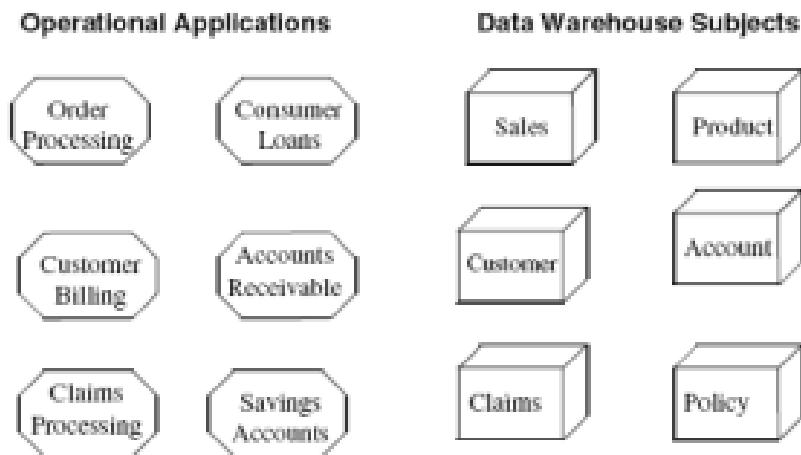
Karakteristik Data Warehouse

1. Berorientasi Subjek.

Data warehouse terorganisasi di seputar subjek kunci (atau entitas-entitas peringkat tinggi) dalam perusahaan. Data warehouse adalah tempat penyimpanan berdasarkan subyek bukan berdasarkan aplikasi. Subyek merupakan bagian dari suatu perusahaan. Contoh subyek pada perusahaan manufaktur adalah penjualan, konsumen, inventori, dan lain sebagainya.

Contoh lain misalnya di bank, aplikasi kredit mengotomasi fungsi-fungsi: verifikasi lamaran dan credit checking, pemeriksaan kolateral, approval, pendanaan, tagihan, dan seterusnya. Didalam data warehouse data-data yang dihasilkan dari proses kredit ini, diatur kembali (dikelompokkan) dan diintegrasikan (digabung) dengan data-data dari fungsi-fungsi lain, agar berorientasi pada misalnya nasabah dan produk.

Gambar dibawah ini merupakan perbedaan mengenai data warehouse dan database operasional.



Gambar 1

2. Terintegrasi

Data yang tersimpan dalam data warehouse didefinisikan menggunakan konversi penamaan yang konsisten, format-format, struktur terkodekan, serta karakteristik-karakteristik yang berhubungan. Sumber data yang ada dalam data warehouse tidak hanya berasal dari database operasional (internal source) tetapi juga berasal dari data diluar sistem (external source). Data pada sumber berbeda dapat di-encode dengan cara yang berbeda. **Sebagai contoh, data jenis kelamin dapat di-enkode sebagai 0 dan 1 di satu tempat dan "m" dan "f" di tempat lain.**

Contoh lain misalnya : Data dari macam-macam aplikasi transaksi (untuk bank misalnya: tabungan, kredit, rekening koran) semua mengandung data nasabah, ada yang sama ada yang spesifik (yang sama misalnya: nama dan alamat, yang spesifik misalnya: untuk kredit ada kolateral, untuk rekening koran ada overdraft) didalam data warehouse data-data yang sama harus diintegrasikan disatu data-

base, termasuk misalnya diseragamkan formatnya (sederhana tetapi paling sering terjadi – aplikasi-aplikasi sering dibeli vendor berbeda, dibuat dengan/ dijalankan di teknologi berbeda-beda)

3. Memiliki dimensi waktu (Time variant)

Data yang tersimpan dalam data warehouse mengandung dimensi waktu yang mungkin digunakan sebagai rekaman bisnis untuk tiap waktu tertentu. Data warehouse menyimpan sejarah (historical data). Bandingkan dengan kebutuhan sistem operasional yang hampir semuanya adalah data mutakhir. Waktu merupakan tipe atau bagian data yang sangat penting didalam data warehouse. Di dalam data warehouse sering disimpan macam-macam waktu, seperti waktu suatu transaksi terjadi/dirubah/dibatalkan, kapan efektifnya, kapan masuk ke komputer, kapan masuk ke data warehouse; juga hampir selalu disimpan versinya, **misalnya terjadi perubahan definisi kode pos, maka yang lama dan yang baru**

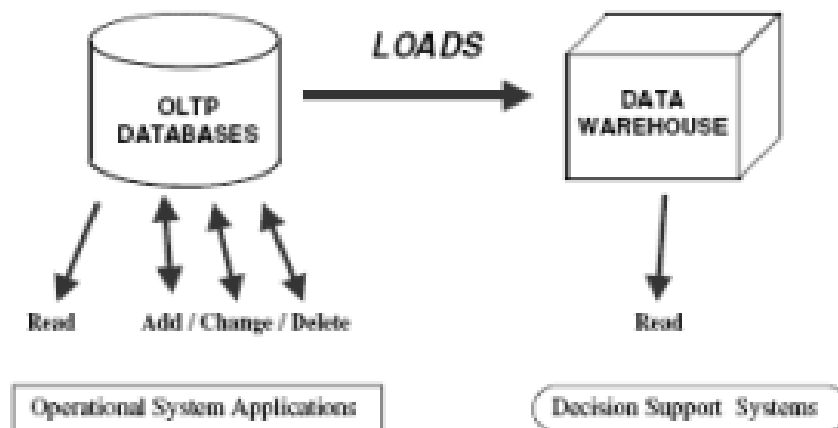
ada semua didalam data warehouse kita. Sekali lagi, data warehouse yang bagus adalah yang menyimpan sejarah.

4. Non-volatile

Data yang tersimpan dalam data warehouse diambil dari sistem operasional yang sedang berjalan, tetapi tidak dapat diperbaharui (di-update) oleh pengguna (bersifat 'hanya-baca'). Sekali masuk ke dalam data warehouse, data-data, terutama data tipe transaksi, tidak akan pernah di update atau dihapus (delete).

Terlihat, bahwa keempat karakteristik ini saling terkait kesemuanya harus diimplementasikan agar

suatu data warehouse bisa efektif memiliki data untuk mendukung pengambilan-keputusan. Dan, implementasi keempat karakteristik ini membutuhkan struktur data dari data warehouse yang berbeda dengan database sistem operasional. Data dalam database operasional akan secara berkala atau 48bisa dipindahkan kedalam data warehouse sesuai dengan jadwal yang sudah ditentukan. Misal perhari, perminggu, perbulan, dan lain sebagainya. Sekali masuk ke dalam data warehouse, data adalah read-only . Pada gambar 2 dibawah ini bisa dilihat bahwa database OLTP bisa dibaca, diupdate dan dihapus. Tetapi pada database data warehouse hanya bisa dibaca.



Gambar 2

5. Ringkas

Jika diperlukan, data operasional dikumpulkan ke dalam ringkasan-ringkasan.

6. Granularity

Pada sistem operasional data dibuat secara real-time sehingga untuk mendapatkan informasi langsung dilakukan proses query. Pada data

warehouse untuk menganalisis harus memperhatikan level-of-detail misalkan perhari, ringkasan perbulan, ringkasan per-tiga-bulan.

7. Tidak ternormalisasi

Data di dalam sebuah data warehouse biasanya tidak ternormalisasi dan sangat redundan. Dasar dari suatu data warehouse adalah suatu

data yang besar yang mengandung informasi bisnis. Data-data yang ada di dalam data warehouse bisa berasal dari banyak sumber, misalkan dari database operasional atau transaksional dan sumber dari luar misalkan dari web, penyedia jasa informasi, dari perusahaan lain dan lain sebagainya.

Keuntungan Data Warehouse

1. Potential high returns on investment

Suatu organisasi harus menjalankan sejumlah besar sumberdaya untuk memastikan kesuksesan implementasi dari data warehouse dengan biaya yang sangat variatif. Nilai investasi yang ditanamkan lebih cepat kembali dengan menggunakan data warehouse.

2. Competitive advantage

Keuntungan kompetitif dapat dicapai dengan memungkinkannya akses pengambil keputusan ke data yang sebelumnya sangat tidak mungkin.

3. Increased productivity of corporate decision-makers

Data warehousing dapat meningkatkan produktivitas pengambil keputusan dengan membangun database terintegrasi yang konsisten, berorientasi subjek, dan data history. Dengan memindahkan data kedalam bentuk yang lebih berarti, data warehouse memungkinkan manajer bisnis untuk menampilkan analisa yang lebih substantif, akurat dan konsisten.

Sejarah Data Warehouse dan perkembangannya

Sejak awal 1990-an, data gudang yang berada di garis depan aplikasi teknologi informasi sebagai cara bagi organisasi untuk secara efektif menggunakan informasi digital untuk perencanaan bisnis dan pengambilan keputusan. Oleh karena itu, pemahaman tentang sistem arsitektur data warehouse akan menjadi penting dalam peran dan tanggung jawab dalam pengelolaan informasi.

- **1960** – General

Mills dan Dartmouth College, dalam proyek penelitian bersama, mengembangkan *dimensi* syarat dan *fakta*. [2]

- **1970** – ACNielsen dan IRI menyediakan data mart dimensi untuk penjualan eceran.

- **1983** – Teradata memperkenalkan sistem manajemen database yang khusus dirancang untuk mendukung keputusan.

- **1988** – Barry Devlin dan Paul Murphy mempublikasikan artikel *arsitektur An untuk dan sistem informasi bisnis* di *IBM Systems Journal* mana mereka memperkenalkan istilah “data bisnis” gudang.

- **1990** – memperkenalkan Sistem Bata Merah Red Brick Warehouse, sebuah sistem manajemen database khusus untuk data warehouse.

- **1991** – memperkenalkan Prism Prism Solusi Gudang Manager, perangkat lunak untuk mengembangkan gudang data.

- **1991** – Bill Inmon menerbitkan buku *Membangun Data Warehouse*.

- **1995** – Data Warehousing Institute, sebuah organisasi nirlaba yang

mempromosikan data warehouse, didirikan.

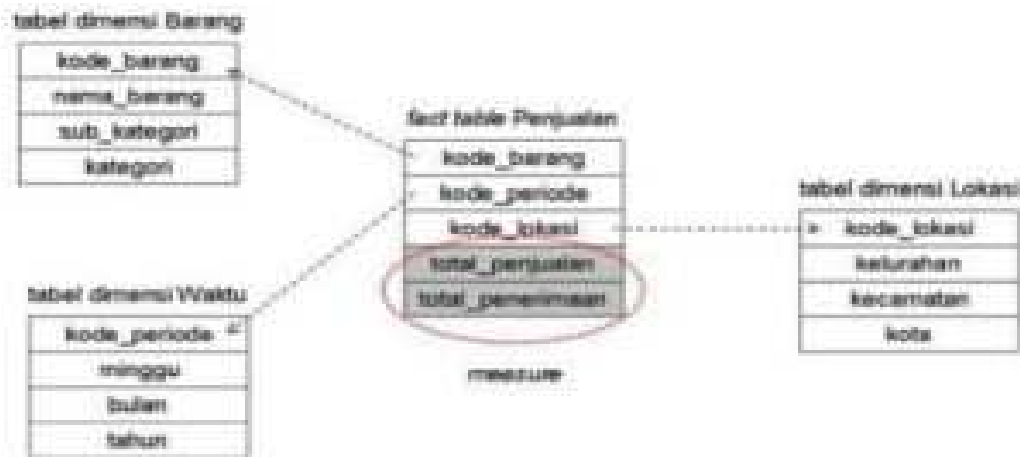
- **1996** – Ralph Kimball menerbitkan buku *The Data Warehouse Toolkit*.
- **2000** – Daniel Linstedt melepaskan *Vault Data*, memungkinkan real time diaudit Data Warehouse.

Skema Data Warehouse

Dalam menggambarkan relasi database pada data warehouse dapat digunakan 2

pendekatan model skema yaitu star schema atau snowflake schema. Disebut starschema karena Entity Relationship Diagram atau ERD-nya yang menyerupai bintang, tabel fakta berada di tengah dengan dikelilingi tabel dimensi di sampingnya.

Berikut adalah contoh gambar star schema:

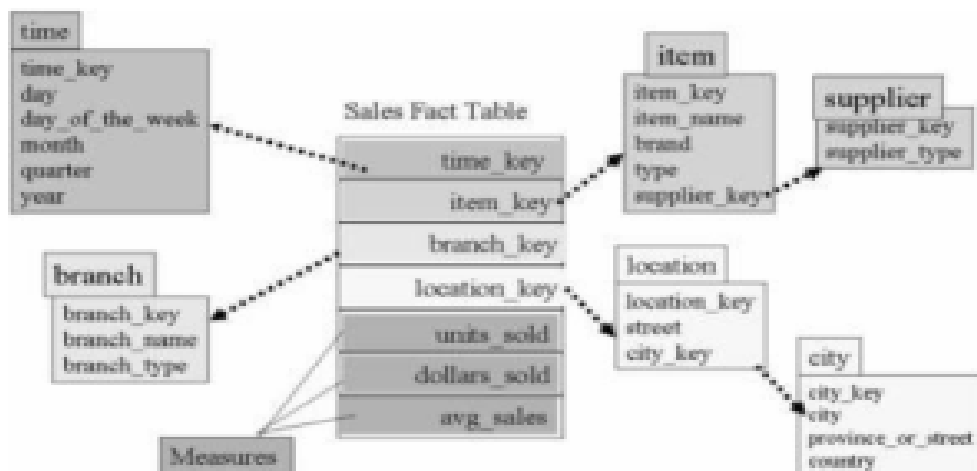


Gambar 3

Sedangkan snowflake schema lebih kompleks dibandingkan dengan star schema karena merupakan pengembangan dari star schema, karena tabel-tabel dimensinya merupakan hasil

normalisasi dari beberapa tabel yang berhubungan.

Berikut adalah contoh gambar snowflake schema:



Gambar 4

Berdasarkan hasil penelitian yang dilakukan oleh Paiman dkk. (2010) yang berjudul “Analisis Perbandingan Antara Model Dimensi *Star Schema* Dan *Snowflake*”, didapat kesimpulan mengenai kelebihan dan kekurangan baik dari *star schema* maupun *snowflake schema*. Berikut adalah kelebihan dan kekurangan model dimensi *star schema*.

Kelebihan model dimensi *star schema* :

1. cenderung mudah dipahami karena modelnya yang lebih sederhana,
2. memudahkan mencari isi karena kesederhanaannya dengan cara melihat *step by step* dari masing-masing dimensinya,
3. proses *query* lebih cepat pada saat proses OLAP.

Kekurangan model dimensi *star schema* :

1. ukuran data lebih besar karena ada data yang disimpan ulang,
2. *maintenance* dan *update* lebih sulit.

Berikut adalah kelebihan dan kekurangan model dimensi *snowflake schema*. Kelebihan model dimensi *snowflake schema*:

1. ukuran data lebih kecil di dalam tempat penyimpanan,
2. lebih mudah dilakukan *maintenance* dan *update*,
3. proses *query* lebih cepat pada saat proses ETL.

Kekurangan model dimensi *snowflake schema*:

1. cenderung lebih sulit dipahami karena kompleksitasnya,
2. sulit mencari isi karena melihat strukturnya yang kompleks dan bercabang-cabang.

Istilah-istilah yang berhubungan dengan data warehouse

1. Data Mart

Adalah suatu bagian pada data warehouse yang mendukung pembuatan laporan dan analisa data pada suatu unit, bagian atau operasi pada suatu perusahaan.

2. On-Line Analytical Processing (OLAP)

Merupakan suatu pemrosesan database yang menggunakan tabel fakta dan dimensi untuk dapat menampilkan berbagai macam bentuk laporan, analisis, query dari data yang berukuran besar.

3. On-Line Transaction Processing (OLTP)

Merupakan suatu pemrosesan yang menyimpan data mengenai kegiatan operasional transaksi sehari-hari.

4. Dimension Table

Tabel yang berisikan kategori dengan ringkasan data detail yang dapat dilaporkan. Seperti laporan laba pada tabel fakta dapat dilaporkan sebagai dimensi waktu (yang berupa perbulan, perkwartal dan pertahun).

5. Fact Table

Merupakan tabel yang umumnya mengandung angka dan data history dimana *key* (kunci) yang dihasilkan sangat unik, karena *key* tersebut terdiri dari *foreign key* (kunci asing) yang merupakan *primary key* (kunci utama) dari beberapa dimension table yang berhubungan.

6. DSS

Merupakan sistem yang menyediakan informasi kepada pengguna yang menjelaskan bagaimana sistem ini dapat menganalisa situasi dan mendukung suatu keputusan yang baik.

PERANCANGAN DATA WAREHOUSE

Pengembangan data warehouse yang ditujukan untuk menganalisis data dilakukan dengan mengikuti beberapa fase desain yaitu spesifikasi kebutuhan, desain konseptual, desain logis, dan desain fisik.

Analisa Kebutuhan

Langkah-langkah analisa data hingga diperoleh logikal *data warehouse* adalah sebagai berikut :

1. Menterjemahkan kebutuhan dari proses bisnis ke dalam model dimensional.
2. Membuat suatu logikal *data warehouse* yang disebut sebagai *star-schema* ataupun *snowflake*.
3. Melakukan test terhadap model dimensional yang sudah dibuat, dan melakukan analisis hingga pemetaan antara target data dengan sumber data.

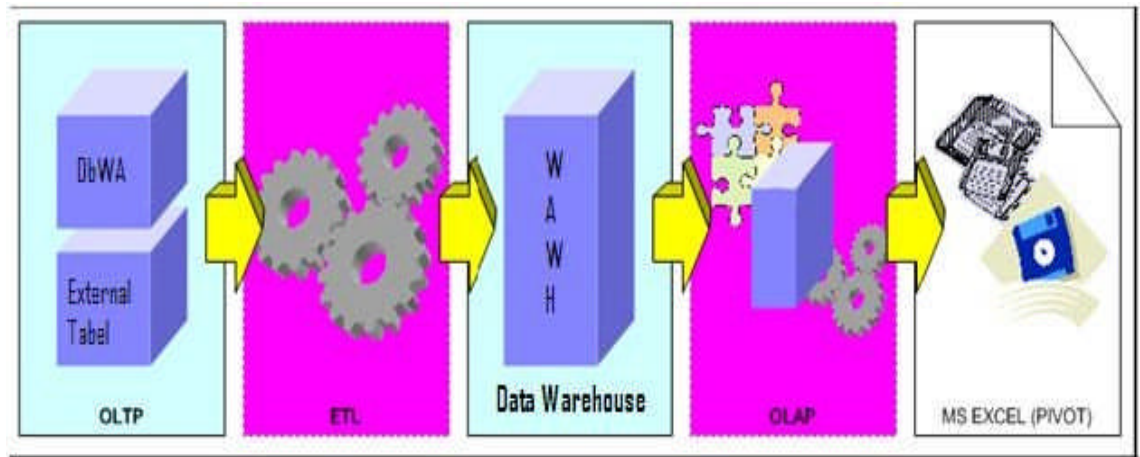
Arsitektur data warehousenya dimulai dengan mengumpulkan data yang berkaitan dengan penjualan yang bersumber dari berbagai *database* seperti *database* penjualan, promosi, *customer*, dan produksi. Pada proses kedua data – data dari data source tersebut di gabung menjadi satu dan di Extraction dan hasil dari data itulah yang digunakan lalu data tersebut di simpan ke dalam format .xls (excelsheet). Proses selanjutnya adalah Transformation yaitu mengambil data –

data yang sudah di Extraction itu tadi lalu di perbaiki, maksud di perbaiki adalah merapikan data – data mentah dari hasil Extraction tersebut lalu di rapikan agar bisa digunakan untuk proses selanjutnya. Proses keempat adalah Loading yaitu memasukkan data – data yang sudah benar dan rapi yang berasal dari hasil proses Transformation itu tadi kedalam data warehousenya. Data warehouse itu sendiri yang berisi data – data yang berasal dari data source yang sudah melalui proses – proses di atas sehingga menjadi kumpulan data yang bisa diolah menjadi informasi. Terakhir adalah proses output yaitu keluaran dari hasil olahan data yang diambil dari data warehouse yang selanjutnya akan di gunakan untuk kebutuhan perusahaan.

Perancangan dan Identifikasi Kelayakan Sistem

Pada perancangan ini dibutuhkan data yang sesuai dan menunjang hasil yang diinginkan. Untuk melakukan identifikasi kelayakan sistem pertama kali dilakukan dengan percobaan terhadap desain proses ETL (Ekstraksi, Transformasi, dan *Load* data), yang dilanjutkan dengan melakukan pencetakan laporan yang diinginkan.

Adapun elemen kompetensi kriteria unjuk kerja dalam membangun desain data warehouse yang digunakan sebagai konsep dalam pembangunan data warehouse adalah sebagai berikut:



Gambar 5. Alur Pembuatan Data Warehouse

Perancangan Database untuk *Data Warehouse*

Ada sembilan tahap dalam perancangan database untuk *data warehouse*, yaitu :

- Langkah 1 : Pemilihan proses
- Langkah 2 : Pemilihan sumber
- Langkah 3 : Mengidentifikasi dimensi
- Langkah 4 : Pemilihan fakta
- Langkah 5 : Menyimpan pre-kalkulasi di tabel fakta
- Langkah 6 : Melengkapi tabel dimensi
- Langkah 7 : Pemilihan durasi database
- Langkah 8 : Menelusuri perubahan dimensi yang perlahan
- Langkah 9 : Menentukan prioritas dan mode *query*

Implementasi

Setelah dilakukan perancangan maka untuk dapat digunakan *user* dilakukan implementasi. Implementasi beracuan pada perancangan yang telah disusun dan diidentifikasi kelayakannya.

Uji coba dan Evaluasi Kelayakan Antar Muka

Uji coba dan evaluasi antarmuka berguna untuk pengguna bahwa *data warehouse* dan antarmuka yang dibuat telah sesuai dengan yang diinginkan.

Pada evaluasi ini akan terlihat kemampuan *data warehouse* dan antarmukanya dalam menyajikan informasi yang cepat dan akurat, sesuai dengan perancangan yang telah dibuat.

PENUTUP

Data warehouse merupakan suatu cara/metode dari suatu database yang berorientasi kepada subjek, non-volatile, time-variance dan terintegrasi yang digunakan untuk mempermudah para pengambil keputusan dalam memecahkan masalah.

Data warehouse yang dirancang adalah hasil dari extract, transfer dan load dari sistem transaksional yang sudah berjalan saat ini, dimana data sumber akan direstrukturisasi terlebih dahulu. Dengan mempertimbangkan bentuk normalisasi data sumber yang ada, *data warehouse* ini mampu melakukan pengkoleksian semua datanya. Dengan adanya konversi struktur data dan konversi type data, maka data-data sumber dapat terlebih dahulu diseragamkan baik struktur datanya maupun type-type datanya, sehingga proses ETL menjadi lebih mudah.

Dengan menggunakan pivot table dari Ms. Excel, pembuatan laporan-laporan dengan bentuk dan format yang diinginkan oleh user menjadi lebih mudah. Dengan adanya data warehouse ini, query tidak dilakukan berulang-ulang dan prosedur pembuatan laporan juga menjadi lebih sederhana dan mudah sehingga dalam pembuatan laporannya menjadi lebih cepat

Keberadaan data warehouse sangat penting sebagai tools dari DSS, karena data warehouse memang digunakan untuk itu. Dengan adanya data warehouse, diharapkan suatu perusahaan dapat lebih unggul dari kompetitornya dan lebih jeli lagi dalam melihat peluang pasar.

DAFTAR PUSTAKA

<https://opistation.wordpress.com/2013/10/15/karakteristik-data-warehouse/>

<https://levinbosz.wordpress.com/2014/01/02/star-schema-dan-snowflake-schema-baru/#more-245>

<http://dundungismyblog.blogspot.co.id/2014/09/analisis-kebutuhan-data-warehouse.html>

Conolly,Thomas dan Begg Carolyn(2002). *Database systems – A Practical Approach to Design, Implementation and Management*, edisi-3. Addison Wesley Longman.Inc., USA

Fatansyah(2002). *Buku Teks Ilmu Komputer – Basis Data*, cetakan-4. Informatika

Inmon, W.H.(2002). *Building the Data Warehouse*,edisi-3. Wiley Computer Publishing.

Kimball,R.,Merz, R (1998). *The Data Warehouse Lifecycle Toolkit. Expert Methods for Designing, Developing and Deploying Data Warehouses*. Wiley Computer Publishing, Canada.

McLeod,Raymond(1996). *Sistem Informasi Manajemen*, Jilid-1. Terjemahan Teguh,H. PT. Prenhallindo, Jakarta.

Poe, Vidette(1998). *Building Data Warehouse for Decision Support*, edisi-2. Prentice Hall.