

Implementasi Data Mart Kepegawaian Menggunakan Tiga Domain

(Studi Kasus di Dinas Pertanian Kabupaten Bandung)

Rahmadi Wijaya, Bambang Pudjoatmodjo
Prodi D3 Teknik Informatika, Fakultas Ilmu Terapan
Universitas Telkom
rakit2272@yahoo.com, b.pudjoatmodjo@gmail.com

Abstrak— Dinas Pertanian Kabupaten Bandung mempunyai banyak data yang harus diolah, seperti data keuangan, data kepegawaian, data proyek, data manajemen perlengkapan, dan data pertanian serta data ekspor-impor. Fokus penelitian dibuat yaitu pembangunan data mart untuk database kepegawaian. Sampai saat ini database kepegawaian pada Dinas Pertanian Kabupaten Bandung harus mengekstraksi data dari tiap domain yang ada secara manual untuk kemudian digunakan sebagai informasi dalam mengambil keputusan. Data mart merupakan bagian dari data warehouse yang membagi area tunggal dari data dan diatur untuk melakukan analisis dengan cepat. Data mart biasanya lebih kecil dan lebih fokus ke bagian-bagian departemen dibandingkan data warehouse. Data yang disimpan di Data Mart berasal dari sumber data operasional database pegawai, database fingerprint dan database sms gateway di Dinas Pertanian Kabupaten Bandung. Prototipe Data Mart dapat digunakan sebagai acuan dalam mengimplementasi rancangan Data Mart pada kondisi sebenarnya.

Pembangunan data mart untuk subjek database kepegawaian dilakukan dengan menggunakan SQL Server 2000. Dengan adanya data mart beserta report-nya diharapkan dapat membantu kinerja kepegawaian dalam pembuatan maupun penyediaan informasi-informasi dengan cepat dan akurat.

Keyword: Data Mart, Database Kepegawaian, Ekstraksi Data.

I. PENDAHULUAN

Kunci sukses bagi sebuah perusahaan untuk bertahan pada masa sekarang adalah kemampuan untuk menganalisa, merencanakan dan bereaksi terhadap perubahan lingkungan bisnis secara cepat dan akurat. Kemampuan ini hanya dapat dipenuhi dengan tersedianya informasi yang memadai bagi para manajer, direktur dan para pengambilan keputusan lainnya. Informasi yang dibutuhkan tersebut seringkali adalah berupa data operasional dan sulit untuk didapatkan. Walaupun tersedia akses ke data tersebut, seringkali format dan struktur data yang ada tidak seperti yang diinginkan. Adanya Sistem Integrasi database dapat dijadikan sebuah platform teknologi yang memungkinkan organisasi mengintegrasikan dan mengkoordinasikan proses bisnis yang

mereka miliki serta dapat dilakukan analisa atas data tersebut, dan mengeluarkan report-report yang dibutuhkan bagi para analis dan pengambilan keputusan. [1]

Sebagai salah satu Instansi Pemerintahan, Dinas Pertanian Kabupaten Bandung mempunyai banyak data yang harus diolah, seperti data keuangan, data kepegawaian, data proyek, data manajemen perlengkapan, dan data pertanian serta data ekspor-impor. Fokus penelitian dibuat yaitu pembangunan data mart untuk database kepegawaian. Sampai saat ini database kepegawaian pada Dinas Pertanian Kabupaten Bandung harus mengekstraksi data dari tiap domain yang ada secara manual untuk kemudian digunakan sebagai informasi dalam mengambil keputusan.

Pembangunan data mart data kepegawaian berfokus revisi pada pendukung keputusan. Fokus pada pengambilan keputusan yang dilakukan berupa proses keputusan, dirancang dengan mudah, sederhana, dapat diterapkan dengan cepat dan mudah diubah, dirancang dan dioperasikan oleh manajer, memberikan informasi yang berguna bagi analisis kegiatan manajerial, berkaitan dengan hanya bagian kecil dari masalah besar, memiliki logika yang serupa dengan cara manajer menganalisis situasi yang sama, memiliki basis data berisi informasi yang disarikan dari file dan informasi lain organisasi yang berasal dari lingkungan eksternal, memungkinkan manajer untuk menguji hasil yang mungkin dari serangkaian alternatif. [2]

Kondisi Departemen Pertanian pada saat ini dalam melakukan analisis dan pengolahan data sebagai berikut yaitu setiap sistem informasi memiliki basis data masing-masing yang berbeda-beda bahkan masih ada yang dalam bentuk flat file (.xls); data tersebar di setiap sistem informasi dan ada data yang mengalami redundansi; dan dalam analisis dan pengambilan keputusan, pihak eksekutif Departemen Pertanian harus melakukan rekap ulang dari data setiap sistem informasi.

A. Rumusan Masalah

Memperhatikan latar belakang yang telah dipaparkan sebelumnya, terdapat beberapa permasalahan yang muncul antara lain :

Kondisi Dinas Pertanian Kabupaten Bandung pada saat ini dalam melakukan analisis dan pengolahan data sebagai berikut :

- Setiap sistem informasi memiliki basis data masing-masing yang berbeda-beda bahkan masih ada yang dalam bentuk flat file (.xls).
- Data tersebar di setiap sistem informasi dan ada data yang mengalami redundansi.
- Dalam analisis dan pengambilan keputusan, pihak eksekutif Dinas Pertanian Kabupaten Bandung harus melakukan rekap ulang dari data setiap sistem informasi.

II. TINJAUAN PUSTAKA

A. Data Mart

Data mart adalah suatu bagian pada data warehouse yang mendukung pembuatan laporan dan analisa data pada suatu unit, bagian atau operasi pada suatu perusahaan. Dalam beberapa implementasi data warehouse, data mart adalah miniature data warehouse. Data mart sering digunakan untuk memberikan informasi kepada segmen fungsional organisasi.

Data Mart adalah fasilitas penyimpanan data yang berorientasi pada Subject tertentu atau berorientasi pada Departemen tertentu dari suatu organisasi, fokus pada kebutuhan Departemen tertentu seperti Sales, Marketing, Operation atau Collection. Sehingga suatu Organisasi bisa mempunyai lebih dari satu Data Mart. Data Mart pada umumnya di organisasikan sebagai suatu Dimensional Model, seperti Star-Schema (OLAP Cube) yang tersusun dari sebuah tabel Fact dan beberapa tabel Dimension.

Contoh umum data mart adalah untuk departemen penjualan, departemen persediaan dan pengiriman, departemen keuangan, manajemen tingkat atas, dan seterusnya. Data mart juga dapat digunakan untuk gudang data segmen data untuk mencerminkan bisnis secara geografis terletak di mana masing-masing daerah relatif otonom. Sebagai contoh, sebuah organisasi layanan yang besar mungkin memperlakukan pusat operasi regional sebagai unit usaha perorangan, masing-masing dengan data mart sendiri yang memberikan kontribusi untuk gudang data master.

B. Perbedaan Data Mart dan Data Warehouse

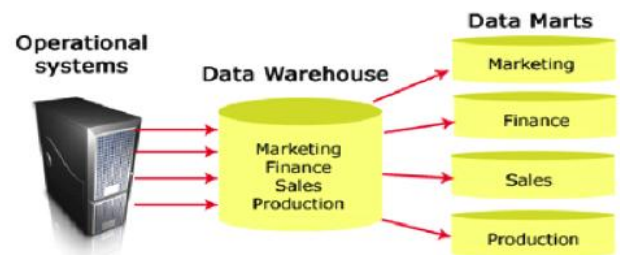
Kadang kala kita sulit untuk membedakan antara data warehouse dan data mart karena keduanya hampir sama. Namun, jika dikaji lebih jauh ada beberapa perbedaan yang dimiliki keduanya.

Sebenarnya Data Mart memang tidak sama dengan Data Warehouse ada banyak perbedaannya, seperti ditunjukkan pada tabel dibawah ini:

Category	Data Mart	Data Warehouse
Scope	Departemental	Corporate
Granularity	High level	Lowest Level
Subject	Single	Multiple
Data Sources	Sedikit	Banyak
Index	Highly Indexed	Lightly Indexed
Size (typical)	Ratusan GB	Beberapa TB
Age	Beberapa Hari atau Bulan	Beberapa Tahun

Gambar 1. DM vs DW

Arsitektur Data Mart bisa dibedakan menjadi dua, yaitu : Dependent Data Mart dan Independent Data Mart. Perbedaan dari kedua arsitektur tersebut hanya terletak pada ketergantungan sumber datanya terhadap data warehouse.



Gambar 2. Dependent Data Mart

Dependent Data Mart (Inmon advocated) berlaku sebagai komponen atau suatu bagian dari enterprise Data Warehouse, Data Mart dibangun dengan cara extract data dari Data Warehouse.

Di lain pihak pada Independent Data Mart (Kimball advocated) dibangun dengan cara extract langsung data dari berbagai Source System. Independent Data Mart tidak tergantung pada pusat penyimpanan data seperti Data Warehouse arsitektur ini biasa juga disebut sebagai "Data Warehouse Bus structure".



Gambar 3. Independent Data Mart

Kedua arsitektur diatas menentukan bagaimana Data Mart dibangun, karena itu bisa dibedakan menjadi dua pendekatan, yaitu :

a. Top-Down approach

Awalnya dibangun Enterprise Data Warehouse lebih dahulu, belakangan baru diturunkan per LOB atau departemen untuk menjadi Data Mart.

b. Bottom-Up approach

Awalnya dibangun beberapa Data Mart, belakangan beberapa Data Mart yang mempunyai Conform Dimension bisa dirangkai menggunakan jalur bersama yang disebut Arsitektur Data Warehouse BUS (Ralph Kimball).

Beberapa keuntungan dalam membangun Data Mart lebih dulu dibanding langsung membangun Data Warehouse :

- Waktu yang diperlukan untuk membangun Data Mart adalah lebih sedikit.
- Volume Data pada Data Mart lebih sedikit
- Waktu Query lebih cepat
- Biaya membangun Data Mart lebih murah.

III. PEMBAHASAN

A. Alur Proses Data Mart

Alur proses perancangan Data Mart di atas adalah sebagai berikut :

- Mengidentifikasi tabel-tabel pada data sumber yang memuat data-data yang dibutuhkan untuk metrik dan dimensi yang telah ditetapkan sesuai kebutuhan user (user requirements).
- Data-data tersebut kemudian dimasukkan ke dalam staging area.

Hal ini dimaksudkan agar proses Extract, Transform and Load (ETL) yang akan dilakukan terhadap data-data tersebut tidak mengganggu proses transaksi Online Transaction Processing (OLTP) yang sedang berlangsung.

Staging area yang digunakan pada kasus ini adalah SQL Server, dan databasenya adalah pertanian. Dalam database pegawai tersebut dilakukan proses ETL.

Pertama, data data dalam tabel diekstraksi, yaitu dipilih field-field yang dibutuhkan yang sesuai dengan metrik dan dimensi pada user requirements, sedangkan field-field yang tidak dibutuhkan diabaikan atau dibuang.

Kedua, dilakukan proses transformasi terhadap data-data terpilih itu. Data-data tersebut diproses menjadi bentuk yang diinginkan untuk kemudian dimasukkan ke data mart.

Adapun proses transformasi yang dilakukan yaitu :

- Memformat data, yaitu menentukan tipe data pada setiap field.

- Proses joining antara tabel Duk dan tabel Dmp agar dapat diketahui hasil pangan mana yang sudah masa panen dan yang masih dibiakkan di hutan.
- Membuat tabel-tabel fakta dan tabel-tabel dimensi sesuai dengan user requirements yang telah ditentukan.
- Kemudian pada tabel-tabel fakta dan dimensi tersebut ditambahkan beberapa field turunan, yaitu field-field yang diturunkan (dihasilkan) dari perhitungan field-field lain, misalnya kota_lahir. Selain itu juga ditambahkan surrogate key (key pengganti).

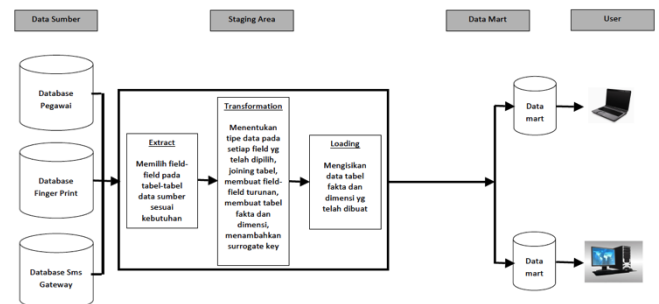
Ketiga, setelah ditransform, data-data tersebut lalu diisikan ke dalam tabel-tabel fakta dan dimensi yang telah dibuat.

- Data-data yang telah tersimpan dalam tabel fakta dan dimensi itu kemudian diubah menjadi bentuk data yang multidimensi yang dapat dilihat dari segala sudut sehingga dapat diakses oleh client untuk keperluan analisis.

Proses ini menggunakan query-query SQL dan divisualisasikan menjadi aplikasi sehingga memudahkan user dalam melihat data-data di data mart

B. Analysis User Requirements

Pada perancangan Data Mart ini, di Dinas Pertanian dibagi menjadi 3 domain pada sistem informasi kepegawaian yaitu database pegawai, database finger print dan database sms gateway. Adapun proses ETL yang dilakukan seperti Gambar 1.



Gambar 1. Proses ETL

1) Extraction

Ekstraksi dilakukan dari sumber data yang digunakan melalui proses pemilihan data yang kemudian disimpan pada database Data Warehouse Temporary (DWTemp). Penempatan DWTemp diletakkan pada penyimpanan database, mesin dan platform yang sama dengan data warehouse. Adapun kumpulan tabel dapat dilihat di lampiran. Untuk lebih jelasnya mengenai perbedaan karakteristik sumber data dan tujuan dapat dilihat pada Tabel 1.

TABEL I. HASIL EKSTRAKSI

	Source	Destination
Engine	Pegawai Engine	Microsoft SQL Server 2005
Platform	HP UX	Microsoft Winfows Server 2003
Sumber Data	database pegawai	DWTemp
Jumlah Total table	6	1
Jumlah Total Field	52	40

Pada Tabel 2 berikut ini dapat dilihat juga contoh dari sejumlah field yang digunakan pada proses ekstraksi dari sumber data yaitu source I (database pegawai) dan source II (excel) yang kemudian diekstrak ke tempat tujuan yaitu DWTemp.

TABEL II. CONTOH FIELD PADA PROSES EKSTRAKSI

Tabel	Field	Digunakan	
		Ya	Tidak
Duk	Nip	✓	
	Nama	✓	
	Kota_lahir		x
	Tgl_lahir	✓	
	Jenis_kel	✓	
	Agama	✓	
	Stat_nikah	✓	
	Alamat	✓	
	RT		x
	RW		x
	Kel		x
	Kec		x
	Kota		x
	Prov		x
	Telepon	✓	

Proses berikutnya adalah dari data tersebut dilakukan penyeragaman data dari segi tipe data yang akan digunakan. Contoh dari ketidakkonsistenan data dapat dilihat pada Tabel 3 di bawah ini :

TABEL III. PERBEDAAN FORMAT DATA

Ketidakkonsistenan Format Data		
Tabel	Contoh	Format
Duk	27-12-2014	dd-mm-yyyy
Dmp	12-22-2014	mm-dd-yyyy
Dgt	08-31-2014	mm-dd-yyyy
Dph	08-31-2014	mm-dd-yyyy
Dpt	08-31-2014	mm-dd-yyyy
Dkp	08-31-2014	mm-dd-yyyy

Untuk mengatasi permasalahan pada tabel 3 di atas maka dilakukan penyeragaman format data waktu mengikuti format yang banyak digunakan yaitu mm-dd-yyyy, hal ini dilakukan untuk menghindari data yang tidak konsisten yang dapat menimbulkan data yang bias.

Proses Cleansing

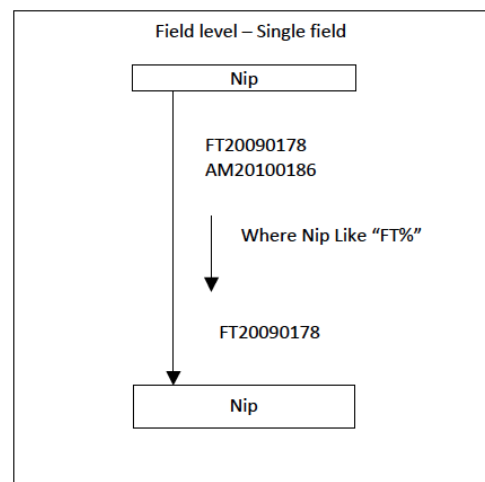
Proses yang akan dilakukan selanjutnya adalah melakukan pembersihan data (cleansing) dari hasil yang diperoleh melalui proses ekstraksi. Proses yang dilakukan pada proses cleansing ini adalah proses untuk membersihkan data yang redundant dan data yang tidak konsisten satu sama lain. Contoh dari ketidakkonsistenan data adalah seperti data yang rusak (corrupt) sehingga isi darinya tidak benar, data kosong, data dengan spasi yang berlebihan dan contoh seperti ini terjadi pada data di field telepon. Isi dari telepon yang ditampilkan mengalami perubahan sesuai perubahan yang terjadi pada konfigurasi yang ada sehingga ketika data baru yang telah berubah ditambahkan ke database maka akan terjadi perbedaan pada field telepon tersebut karena itu data telepon tersebut harus diubah agar menjadi konsisten.

TABEL IV. KETIDAKKONSISTENAN DATA

Telepon
(022) 5943564
(022)5946712

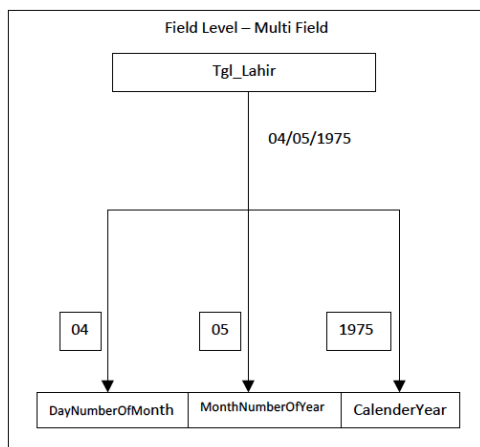
2) Transformation

Proses ini dilakukan setelah data yang ada sudah melewati proses ekstraksi dan pembersihan. Proses transformasi yang dilakukan dibagi berdasarkan level record-level dan field-level, pada proses ini dilakukan proses pemilihan, penggabungan dan agregasi untuk mendapatkan data ringkasan sesuai dengan dimensi yang akan dibuat. Proses transformasi ke field yang baru dapat dilakukan dengan menggunakan suatu fungsi tertentu untuk melakukannya. Contoh dari proses transformasi dengan record-level dan field-level dapat dilihat pada Gambar 2.



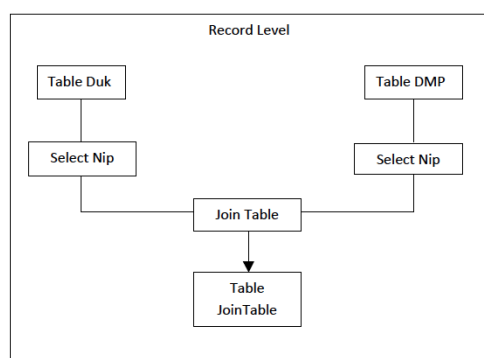
Gambar 2. Field Level – Single Field

Pada gambar di atas menunjukkan proses transformasi single-field, yang terjadi pada proses ini adalah data diambil dari field sumber (Nip) ke field tujuan (Nip) dengan kriteria bahwa dua huruf pertama dari setiap record harus diawali dengan karakter “FT”, sehingga Nip FT20090178 akan digunakan sedangkan AM20100186 tidak akan digunakan dalam proses tersebut.



Gambar 3. Field Level – Multiple-Field

Pada gambar di atas menunjukkan proses transformasi multiple-field, proses yang dilakukan adalah mengambil data dari satu field sumber (Tgl_Lahir) kemudian ditransformasikan ke dalam tiga field tujuan (DayNumberOfMonth, MonthNumberOfYear, MonthNumberOfYear).



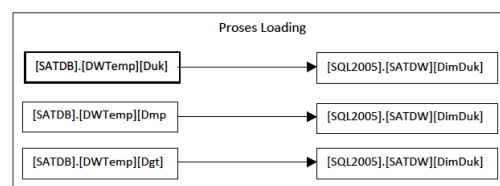
Gambar 4. Record Field

Sedangkan pada Gambar 3 ditampilkan proses transformasi record field, pada proses diambil beberapa field dari tabel sumber (tabel Duk dan tabel Dmp) lalu digabungkan ke dalam tabel tujuan (tabel JointTable).

3) Loading

Proses yang dilakukan pada tahap akhir adalah proses pemuatan data (loading). Data yang digunakan pada tahap ini merupakan data dari proses-proses yang dilakukan sebelumnya yaitu ekstraksi, pembersihan (cleansing) dan transformasi ke dalam data warehouse. Cara pemuatan data

ke dalam data warehouse adalah menggunakan script yang dijalankan secara periodik. Alur proses loading yang dilakukan dapat dilihat pada Gambar 5.



Gambar 5. Proses Loading

Keterangan :

[1].[2].[3]

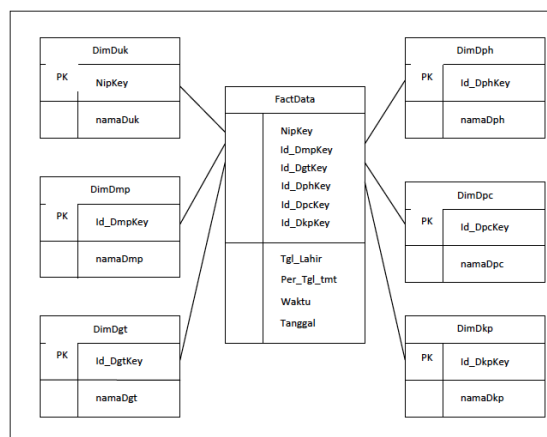
Kolom 1 = Nama Server

Kolom 2 = Nama Database

Kolom 3 = Nama Table

C. Model Skema Data

Model skema yang akan digunakan pada perancangan data warehouse adalah skema bintang. Dengan perancangan ini, skema bintang yang digunakan terdiri satu tabel fakta dan enam tabel dimensi. Model dari skema bintang ini dapat dilihat pada Gambar 6.



Gambar 6. Star Schema FactData

IV. KESIMPULAN

Berdasarkan hasil data yang diperoleh maka penulis dapat menarik kesimpulan sebagai berikut :

1. Di Dinas Pertanian memiliki data histori kepegawaian yang belum dimanfaatkan secara maksimal.
2. Data kepegawaian yang bersifat historis dapat disimpan ke dalam data mart pegawai, sehingga dapat menghasilkan informasi yang berguna bagi pihak top management yang mendukung proses pengambilan keputusan.

3. User dapat melihat data dengan pemahaman yang baik karena kemampuan data mart yang mampu mengerjakan peringkasan (filtering) dan perbandingan data dengan baik.
4. Dalam implementasi data mart, kebutuhan informasi yang diperlukan dan siapa yang menjadi pengguna akhir sangat mempengaruhi pembangunannya
5. Desain star schema mendukung proses rancang bangun data mart kepegawaian. Desain star schema juga memudahkan proses pengambilan keputusan di setiap domain.

DAFTAR PUSTAKA

- [1] Ibrahim, Nugroho Setyabudhi, Takariyana Heni A., 2004, Perancangan Data Warehouse Pada Pusat Data dan Informasi
- [2] Inmon, W.H., 2002, Building The Sistem integrasi database, Third Edition, John Wiley and Sons, Inc., New York
- [3] O'Neil, Bonnie, Michael Schrader, John Dakin, Kieron Hardy, Matthew Townsend, Michael Whitmer, 1997, Oracle® Data Warehousing, Norhern Lights Software, Ltd., SAMS Publishing
- [4] Pressman, Roger S., Ph.D., 2002, Rekayasa Perangkat Lunak Pendekatan Praktisi (Buku II), Andi, Yogyakarta
- [5] Turban, Efram, Jay E. Aronson, and Ting Peng Liang, 2005, Decision Support Systems and Intelligent Systems (Sistem Pendukung Keputusan dan Sistem Cerdas), Edisi 7 Jilid 1, Andi, Yogyakarta
- [6] Rainardi V 2010 Building a Data warehouse With Examples in SQL Server New York

Scan Bardcode dibawah untuk mengecek keabsahan dokumen anda

