

Data Warehouse Analisa Prestasi Akademik Siswa di SMP Roudlotul Jadid Lumajang

¹Yusi Dwi Dayati, ²Achmad Choiron, ³Slamet Kacung
Program Studi Teknik Informatika Universitas Dr. Soetomo Surabaya
yusidwidayati@yahoo.com, choironunitomo@gmail.com, kacungslamet@gmail.com

Abstract-The absence of a system that can provide fast analysis system is a hallmark of the assessment process is based on the transactional database, such as that which is now used in SMP Roudlotul Lumajang Jadid. This is due to the presence of the transactional data is not intended to be used in Online Analytical Processing (OLAP) that provides functions to analyze the data in order to improve the responsiveness and flexibility of the current system used in the school institution. Then the final project was creation of applications for analyzing OLAP junior high student academic achievement. The final task was done in order to obtain the system requirements, design, implement and test the performance of OLAP applications for analysis so as to provide benefits and correct the problem at school. Problem analysis is done by the system needs analysis, design and implementation, including design datawarehouse and OLAP design. In junior high student achievement analysis Roudlotul Jadid Lumajang can know that the level of mastery students has increased in the 2nd half, for the average value of the class and teacher attendance increase is most dominant in the 2nd half than 1st half.

Keyword: OLAP, Data Warehouse, Akademik, Data Mining

I. PENDAHULUAN

Perkembangan teknologi data warehouse telah mengalami kemajuan yang sangat pesat dalam berbagai bidang, salah satunya digunakan dalam bidang pendidikan. Salah satu tool untuk olap yang tersedia dapat digunakan untuk menganalisa data. Pembuatan data warehouse untuk menyatukan data yang beragam ke dalam sebuah tempat penyimpanan dimana pengguna dapat dengan mudah menjalankan query, menghasilkan laporan, dan melakukan analisis. Salah satu keuntungan yang diperoleh dari keberadaan data warehouse dapat meningkatkan efektifitas pembuatan keputusan.

Dalam penelitian ini, penulis menggunakan SMP Roudlotul Jadid Lumajang sebagai objek penelitian yang merupakan salah satu instansi pendidikan di kabupaten Lumajang. Instansi ini memiliki jumlah siswa 313 yang terdiri dari siswa kelas 1 sampai kelas 3 tahun ajaran 2011-2012.

Masalah yang sering terjadi adalah proses analisa prestasi akademik siswa yang masih dilakukan secara manual membuat para guru kerepotan untuk mengolah data rekap nilai semua siswa baik tingkat nilai ketuntasan siswa, nilai rata-rata kelas dan tingkat kehadiran guru. Apabila dibutuhkan menampilkan data salah satu siswa yang maka harus mencari dan membuka kembali data nilai siswa, hal ini sangat merepotkan tidak jarang guru melakukan kesalahan seperti kehilangan data siswa yang dibutuhkan.

Rumusan Masalah :

- 1) Belum adanya sistem yang mampu memberikan analisa prestasi akademik siswa SMP periode 2011-2012.

- 2) Manajemen sekolah belum bisa memberikan penilaian prestasi akademik siswa secara cepat dan akurat.

Tujuan:

- 1) Dapat menganalisa prestasi akademik siswa periode 2011-2012.
- 2) Mampu menganalisa faktor-faktor kenaikan dan penurunan nilai siswa.
- 3) Membuat sebuah sistem yang terintegrasi, sehingga bisa menampilkan data nilai siswa yang cepat dan akurat.

Metode

Metode penelitian yang penulis gunakan berdasarkan metode analisis dan desain sistem. Pada dasarnya metodologi terdiri dari analisa atas kebutuhan data, melakukan desain, mengimplementasi desain sistem, melakukan uji coba sistem dan pemeliharaan sistem.

- 1) Identifikasi Masalah
- 2) Analisa Sistem
- 3) Menentukan Kebutuhan Sistem
- 4) Perancangan sistem
- 5) Pengembangan dan dokumentasi sistem
- 6) Uji coba dan Maintenance Sistem
- 7) Penerapan dan Evaluasi Sistem

II. TINJAUAN PUSTAKA

1. Data Warehouse

Data warehouse merupakan database yang digunakan untuk pengambilan data dan menganalisa data. Jadi berbeda dengan database yang digunakan untuk keperluan pencatatan transaksi. Database ini berisi data sejarah yang sudah diringkas isinya. Data sejarah mengindikasikan bahwa data sudah di susun sebagai bagian dari kelompok yang lebih besar dibandingkan dengan data detail.

Data warehouse berfungsi untuk menyimpan data histories yang dibutuhkan untuk kepentingan analisis. Dalam pembuatan database ini perlu mempertimbangkan tentang bagaimana mengambil data yang cukup banyak dalam waktu sesingkat mungkin [1].

Dari uraian diatas bahwa pendekatan pembuatan data warehouse akan berbeda dengan pendekatan database transaksi. Bila dalam pembuatan database transaksi kita perlu menerapkan prinsip normalisasi, tetapi tidak dalam pembuatan data warehouse yang berbeda dengan database transaksi. Pada data warehouse dikenal dengan konsep staging area dimana Staging area merupakan tempat membersihkan (filter), mengubah, mengkombinasikan, dan menyediakan sumber data. Ada tiga fungsi penting yang ada didalam area ini:

a. Ekstaraksi Data

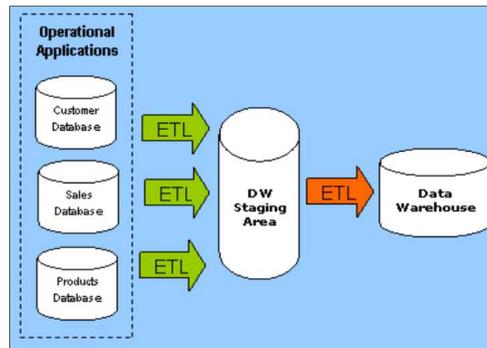
Sumber data mungkin berasal dari mesin yang berbeda-beda dengan format data yang berbeda pula atau mungkin akan digabungkan pula dengan data spreadsheet dan data per departemen. Ekstraksi data akan menjadi sangat kompleks oleh karena itu sumber data di ekstrak kedalam lingkungan fisik yang berbeda yang kemudian dimasukkan g data staging area. Setelah itu data dari staging area akan di ekstrak lagi kedalam data warehouse.

b. Transformasi Data

Transformasi data adalah membersihkan data yang di ekstrak dari masing-masing sumber. Maksud dari pembersihan ini adalah membenarkan kesalahan pengejaan kata, adanya perbedaan ukuran sumber data, atau untuk data yang hilang atau mengeliminasi data yang duplikat. Standarisasi tipe data dan ukuran field untuk setiap elemen data yang diambil dari banyak sumber adalah bagian yang paling besar dalam transformasi data.

c. Loading Data

Setelah menyelesaikan perancangan dan konstruksi data warehouse, dilakukan loading data kedalam data warehouse. Data di ekstrak dari data sumber, kemudian data tersebut di transform lalu di loading kedalam data staging area , begitu di staging area data akan dibersihkan, distandarkan kemudian di format ulang untuk siap di loadingkan kedalam data warehouse. Dengan munculnya data warehouse, konsep transformasi telah menyediakan tingkat tinggi kualitas dan keseragaman data.



Gambar 1: Staging Area

2. Karakteristik Data Warehouse

Sebenarnya karakteristik data warehouse sembilan karakteristik tetapi penulis akan menjelaskan empat dari sembilan karakteristik data warehouse. Empat sifat yang mencirikan data yang disimpan didalam data warehouse ini [2].

Empat karakteristik data warehouse adalah sebagai berikut;

1. Data berorientasi pada subyek yang khusus (Subyek Oriented)
Data yang dimasukkan dalam data warehouse adalah data yang benar-benar diperlukan dalam proses pengambilan dan menganalisis data dalam suatu subyek. Atau dengan kata lain semua data yang tidak relevan dengan tujuan pembuatan data warehouse walaupun ada dalam sumber harus dibuang sebelum masuk ke data warehouse.
2. Data konsisten dan terkonsolidasi (integrated)
Data yang berasal dari berbagai sumber sudah disusun dengan konversi bersama sehingga satu nama data warehouse memiliki arti dan format yang sama pada semua database sumber.
3. Time Variant
Data warehouse menyimpan sejarah (historical data). Melakukan perbandingan dengan kebutuhan sistem operasional yang hampir semuanya data mutakhir.
4. Non-Volatile
Sekali masuk kedalam data warehouse data-data tidak akan pernah di update atau dihapus.

3. OLAP (Online Analytical Processing)

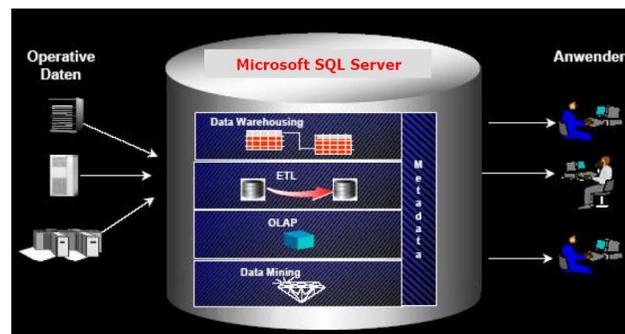
Mempunyai basis data relasional tidaklah berarti mendukung pembuatan keputusan, karena sebenarnya basis data relasional itu tidaklah ditunjukkan untuk menyediakan fungsi yang mampu melakukan sintesa data, analisis dan konsolidasi (kemudian secara umum dikenal sebagai analisis multidimensi atau OLAP). Basis data relasional memang cukup bagus untuk mengakses data yang besar melakukan ringkasannya seketika. Waktu respon yang lambat dan penggunaan sumber daya jaringan yang berlebihan adalah ciri aplikasi sistem pendukung keputusan yang dibangun diatas dasar teknologi basis data relasional [3].

Jika sebuah basis data relasional hanya mampu membaca 200 record dalam waktu satu detik, sebuah server OLAP dengan menggunakan aritmatika baris dan kolom mampu mengkonsolidasikan 20.000 sampai 30.000 sel dalam waktu satu detik. Kebutuhan fungsional sebuah sistem OLAP adalah akses dan kemampuan menghitung cepat, mampu melakukan analisis lintas dimensi, fleksibel dan mendukung multi-user.

4. Rancangan OLAP Security

Para pemakai sistem operasional tidak sama dengan pemakai data warehouse. Masalah utama adalah, bahwa relational model mendominasi dalam sistem operasional ketika sistem OLAP menggunakan nontraditional multidimensional model. Rencana akses kontrol tidak mudah dipetakan. Proteksi tidak digambarkan dalam kaitan dengan tabel, tetapi dimensi, alur hirarkis, granularas level. Sehingga dibutuhkan desain keamanan OLAP.

Telah dijelaskan bahwa perancangan akses OLAP harus dilakukan dengan teliti, ketika analisis ditolak atau hasil adalah salah. Apalagi kemampuan keamanan tool sangat proprietary dan sintak security tidak mungkin untuk didesain dan didokumentasi dari pembatasan akses seperti gambar dibawah ini [4].



Gambar 2: Olap Security

Dalam rangka mendekati topik dari sisi aplikasi, metodologi desain klasikal database (persyaratan analisis, konseptual, logis, dan phisik) harus diplikasikan pada kemanan OLAP. Harus diaplikasikan pada keamanan database reguler. Perbedaan yang penting, adalah konseptual multidimensional data model dan mekanisme keamanan OLAP yang jelas berbeda dengan kemampuan relational manajemen sistem database.

5. Metode Penyimpanan OLAP

Dalam penyimpanan aplikasi OLAP ada beberapa metode yang bisa dipilih yaitu MOLAP, ROLAP dan HOLAP. Penjelasan lebih lanjutnya lihat dibawah ini [5]:

a. MOLAP

Multidimensional online analytical processing (MOLAP) menyimpan data dan agregasi pada struktur data multidimensi. Struktur MOLAP ini tidak tersimpan pada datawarehouse tapi tersimpan pada OLAP server. Sehingga performa query yang dihasilkan olehnya sangat bagus.

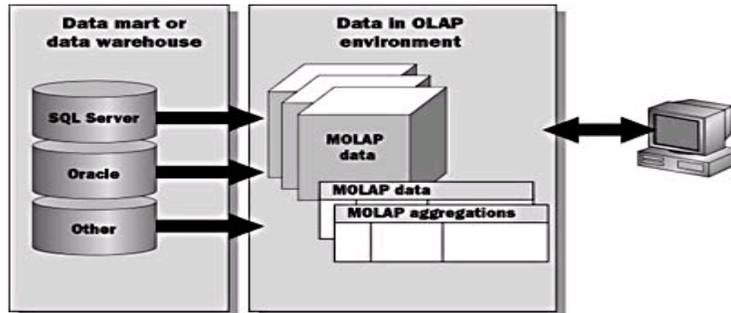
b. ROLAP

ROLAP (*Relational online analytical processing*) menggunakan tabel pada database relasional datawarehouse untuk menyimpan detail data dan agregasi kubus. Berbeda dengan MOLAP, ROLAP tidak menyimpan salinan database, ia mengakses langsung pada tabel fact ketika membutuhkan jawaban sebuah *query*. Sehingga *query* pada ROLAP mempunyai *response*

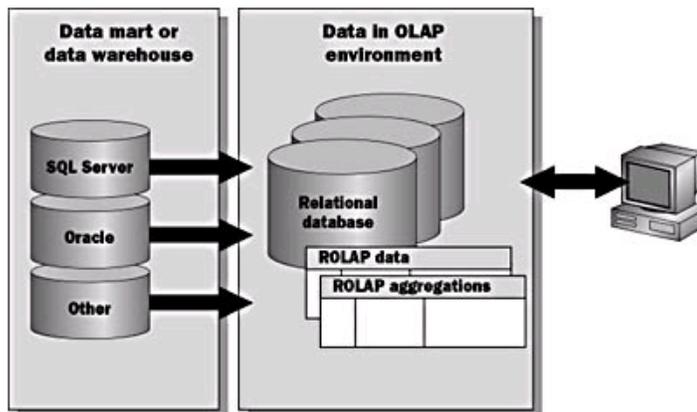
time yang lebih lambat dibandingkan ROLAP maupun HOLAP. Karakteristik model ini digunakan untuk menyimpan data yang besar dan jarang dilakukannya proses *query*.

c. HOLAP

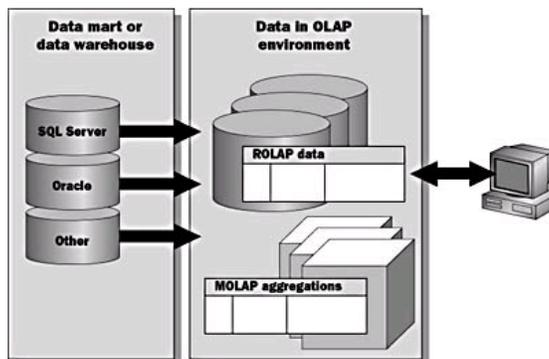
Gabungan model MOLAP dan ROLAP dapat kita peroleh dari HOLAP (*Hibrid online analytical processing*) Detil data tersimpan pada tabel relasional tapi agregasi data disimpan dalam format multidimensi. Misalkan proses *drill down* dilakukan pada sebuah tabel fakta, maka *retrive* data akan dilakukan dari tabel database relasional sehingga *query* tidak secepat MOLAP. Kubus HOLAP lebih kecil daripada kubus MOLAP tapi *response time query* masih lebih cepat jika dibandingkan dengan ROLAP. Model penyimpanan HOLAP ini biasanya sesuai untuk kubus yang membutuhkan performa *query* yang bagus dengan jumlah data yang besar.



Gambar 3: MOLAP



Gambar 4: ROLAP



Gambar 5: HOLAP

Deskripsi Sistem

Sistem analisa prestasi akademik siswa SMP ini menitik beratkan pada proses analisa yang cepat dalam menentukan tingkat ketuntasan siswa, rata-rata kelas dan tingkat kehadiran guru.

Database data warehouse ini dihasilkan dari database OLTP(database SMP Roudlotul Jadid Lumajang). Database operasional ini dilakukan proses *Extarcting- Transforming-Loading*. Setelah proses ETL berhasil data tersebut tersimpan dalam database data warehouse.

Setelah data warehouse terbentuk, maka dilanjutkan dengan pembuatan *On-line Analitical Prosesing* (OLAP) untuk melakukan analisa terhadap prestasi akademik siswa yang telah dicapai selama ini (sesuai dengan batas waktu yang ditentukan) pembuatan OLAP ini dilanjutkan dengan cara membuat data multidimensi (cube data) dan dilanjutkan dengan proses agregasi.

Proses analisa data yang dapat dilakukan sistem ini adalah analisa terhadap prestasi akademik siswa SMP Roudlotul Jadid Lumajang, yang meliputi indek prestasi siswa, tingkat kehadiran guru dan rata-rata kelas.

Kebutuhan Data

Sebelum melakukan pembuatan sistem terlebih dahulu dilakukan beberapa analisa terutama terhadap kebutuhan data yang nantinya akan dilakukan proses dan akhirnya menghasilkan suatu sistem. Data yang digunakan bisa dari sistem yang sudah ada di sekolah tersebut atau data mentah yang nantinya diproses oleh sistem, sistem menghasilkan informasi data, untuk dapat membuat sistem tersebut, data yang diperlukan adalah sebagai berikut:

2. Data Siswa

Data ini diperoleh dari database sekolah SMP Roudlotul Jadid Lumajang, data ini diperlukan untuk mengetahui prestasi siswa sehingga bisa digunakan untuk melakukan analisa indek prestasi siswa dan rata-rata nilai siswa dalam satu semester. Pengambilan data dimulai dari tahun 2011-2012 guna memenuhi syarat sebagai bahan analisa.

3. Data Guru

Data guru diperoleh dari pihak instansi sekolah dengan cara melihat daftar guru yang aktif mengajar di SMP Roudlotul Jadid Lumajang

4. Data Kelas

Data kelas diambil dari tiap-tiap kelas siswa yang ada di SMP

5. Data Pelajaran

Data Tahun ajaran langsung diambil dari data pelajaran disekolah.

6. Data Tahun Ajar

Data kelas diambil dari tiap-tiap tahun ajar siswa yang ada di SMP

7. Data Tuntas

Data tuntas diambil dari status ketuntasan siswa berdasarkan nilai dan kkm

Model Proses

1) Diagram Berjenjang

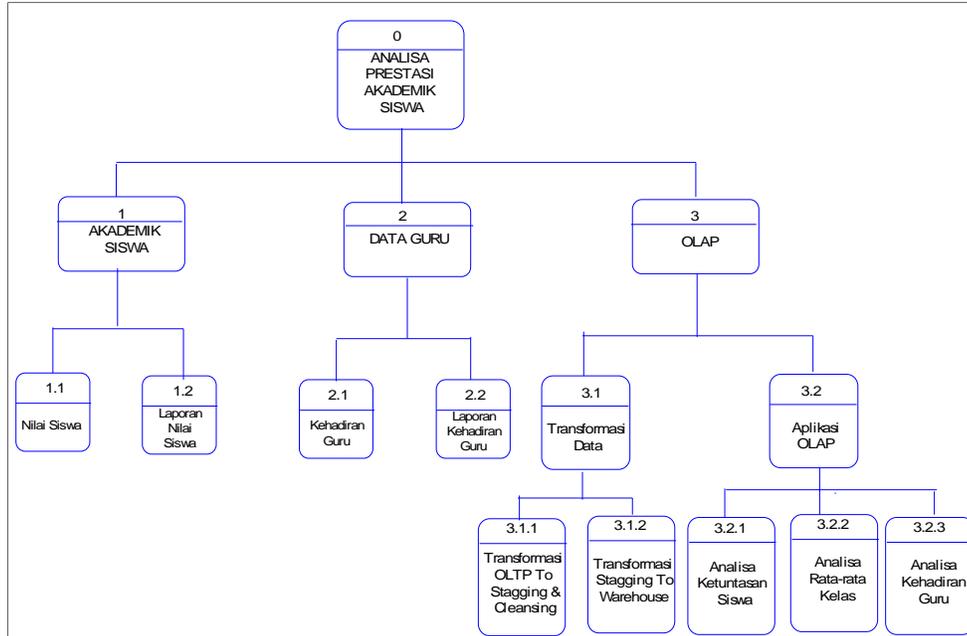
Merupakan suatu metode yang dapat digunakan untuk mendesain suatu sistem dengan cara menggambarkan tahapan-tahapannya dari setiap proses yang terjadi pada sistem mulai dari proses yang paling awal sampai pada akhir.

2) Data Flow Diagram

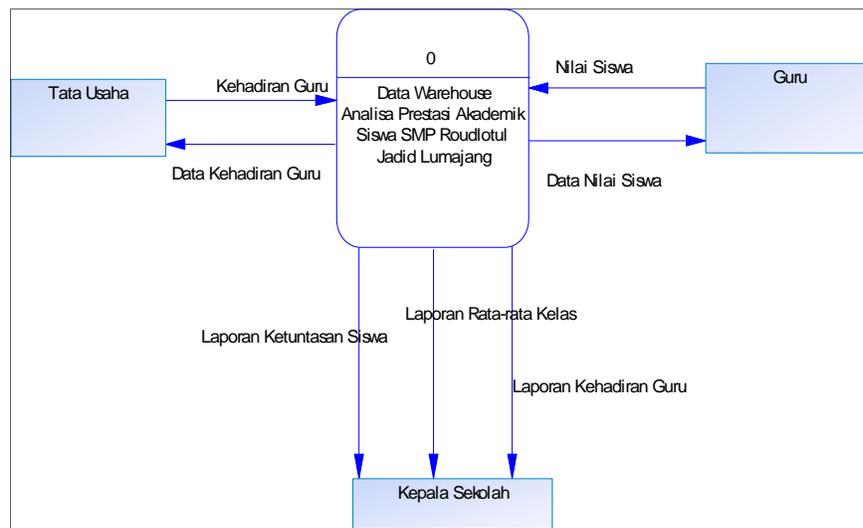
Data Flow Diagram (DFD) adalah Alat yang digunakan pada metodologi pengembangan sistem terstruktur (*Struktur Analys And Design*). Fungsi DFD untuk menggambarkan suatu sistem yang telah ada atau sistem yang baru akan dikembangkan secara logika tanpa mempertimbangkan lingkungan fisik dimana data tersebut mengalir dan data tersebut akan disimpan. Selain dapat menggambarkan arus data dalam sistem dengan terstruktur, DFD juga merupakan dokumentasi dari sistem yang baik.

Model Data

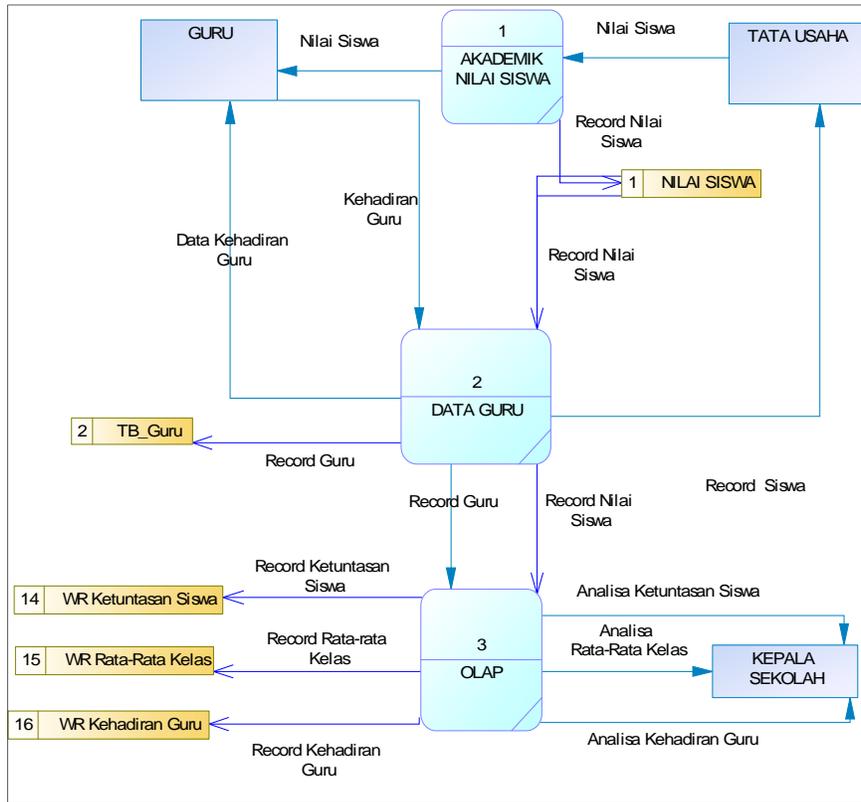
Untuk desain basis data pertama dibuat gambaran real time dari sistem analisa prestasi akademik siswa SMP Roudlotul Jadid Lumajang. Gambaran real time tersebut akan tergambarkan dalam bentuk *Conceptual Data Model* (CDM) sedangkan tool yang digunakan adalah power designer 15, selanjutnya akan tergambarkan dalam bentuk komputerisasi yang lebih dikenal *Physical Data Model* (PDM).



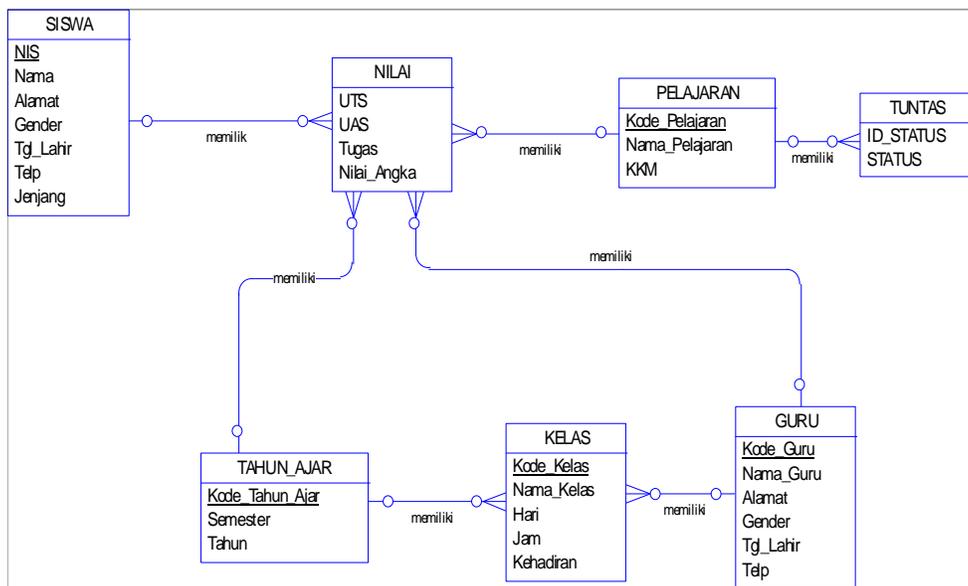
Gambar 6: Diagram Berjenjang



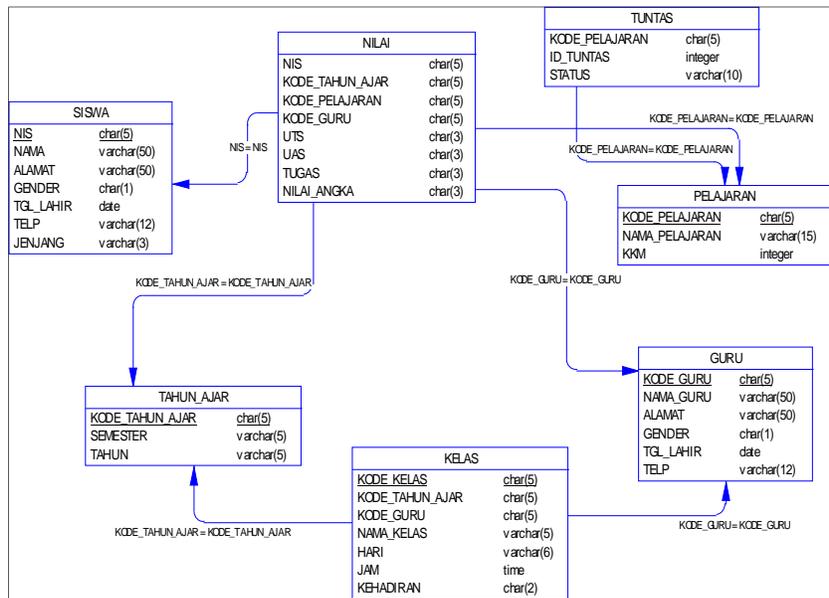
Gambar 7: Context Diagram



Gambar 8: DFD Level 1



Gambar 9: CDM



Gambar 10: PDM

Software dan Hardware

1) Pengembangan Sistem

Berikut ini adalah Software dan Hardware yang digunakan pada saat melakukan pengembangan sistem :

A. Software

- Sistem Operasi Windows Xp Profesional park 3
- Microsoft SQL Server 2005
- Microsoft Visual Studio 2005

B. Hardware

- CPU: Intel Pentium Dual Core Processor T4300
- RAM 1 Gb

2) Membangun Star Schema

Query transformasi staging ke data warehouse

1. Dimensi Siswa

```
select distinct
NIS, nama FROM [Staging].[dbo].['AKADEMIK SMP$']
```

2. Dim Guru

```
select distinct
Kode_Guru, Nama_Guru from [Staging].[dbo].['AKADEMIK SMP$']
```

3. Dim kelas

```
select distinct
kode_kelas, nama_kelas FROM [Staging].[dbo].['AKADEMIK SMP$']
```

4. Dim pelajaran

```
select distinct
kode_pelajaran, nama_matapelajaran FROM [Staging].[dbo].['AKADEMIK SMP$']
```

5. Dim Tahun Ajar

```
select distinct
KODE_TAHUN_AJAR ,semester FROM [Staging].[dbo].['AKADEMIK SMP$']
```

6. Dim Tiuntas

```
Select distinct
Id_Status, status FROM [Staging].[dbo].['AKADEMIK SMP$']
```

7. Fakta Ketuntasan Siswa

Select distinct
Nis,Kode_pelajaran,Kode_kelas,kode_tahun_ajar,nilai,kkm,id_status FROM
[Staging].[dbo].['AKADEMIK SMP\$'] Fakta Rata-rata kelas

8. Fakta Rata-rata kelas

Select distinct
NIS,Kode_guru,kode_pelajaran,Kode_kelas,Kode_tahun_ajar,nilai FROM
[Staging].[dbo].['AKADEMIK SMP\$']

9. Fakta Kehadiran Guru

Select distinct
Kode_tahun_ajar,kode_pelajaran,kode_guru,kode_kelas,jumlah_kehadiran FROM
[Staging].[dbo].['AKADEMIK SMP\$']

Hasil Pengujian

1) Data Olap Ketuntasan Siswa

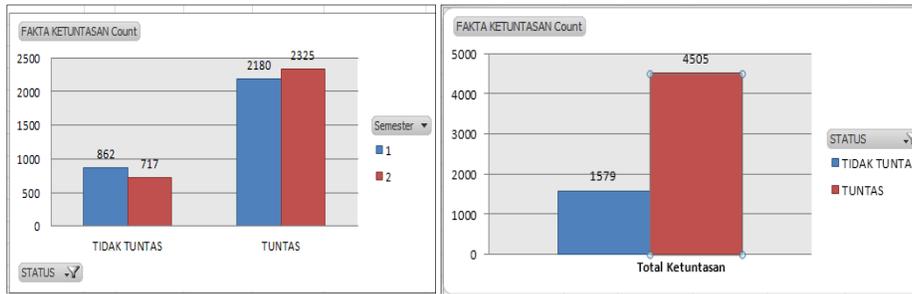
Uji kasus melakukan analisa ketuntasan siswa berdasarkan data dari SMP Roudlotul jadid Lumajang .

Nama Mata Pelajaran	Nama Kelas	STATUS	Semester		Grand Total
			1	2	
Agama	7A	TIDAK TUNTAS	4	10	14
		TUNTAS	33	27	60
		Total	37	37	74
	7B	TIDAK TUNTAS	7	12	19
		TUNTAS	32	27	59
		Total	39	39	78
	8A	TIDAK TUNTAS	21	14	35
		TUNTAS	23	30	53
		Total	44	44	88
	8B	TIDAK TUNTAS	14	13	27
		TUNTAS	30	31	61
		Total	44	44	88
9A	TIDAK TUNTAS	37	37	74	
	TUNTAS	39	39	78	
	Total	36	36	72	
9D	TIDAK TUNTAS	37	37	74	
	TUNTAS	37	37	74	
	Total	313	313	626	
Bahasa Indonesia	7A	TIDAK TUNTAS	37	37	74
	7B	TIDAK TUNTAS	39	39	78
	8A	TIDAK TUNTAS	44	44	88
	8B	TIDAK TUNTAS	44	44	88
	9A	TIDAK TUNTAS	37	37	74
	9B	TIDAK TUNTAS	39	39	78
Bahasa Inggris	7A	TIDAK TUNTAS	37	37	74
	7B	TIDAK TUNTAS	39	39	78
	8A	TIDAK TUNTAS	44	44	88
	8B	TIDAK TUNTAS	44	44	88
	9A	TIDAK TUNTAS	37	37	74
	9B	TIDAK TUNTAS	39	39	78

Gambar 11: Data Olap Ketuntasan Siswa

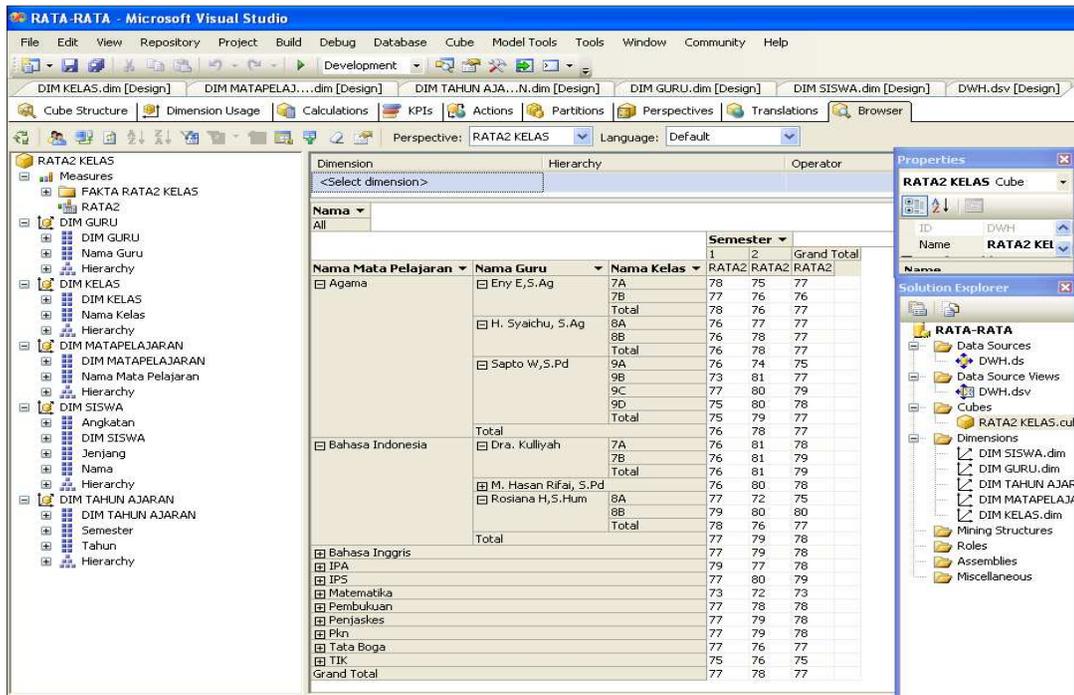
2) Grafik Analisa Ketuntasan Siswa

Gambar 11 menunjukkan nilai perbandingan 3 mata pelajaran yaitu agama, bahasa indonesia dan matematika. Data menunjukkan bahwa untuk pelajaran agama dan bahasa indonesia ketuntasan siswa mengalami kenaikan dari semester 1 ke semester 2 sedangkan matematika mengalami penurunan semester 1 sebesar 232 menjadi 226 pada semester 2.



Gambar 12: Ketuntasan Siswa

3) Data Analisa rata-rata kelas

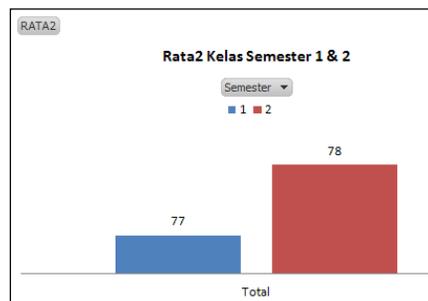


Gambar 13: Analisa Rata-Rata Kelas

Dengan menempatkan nama siswa *didrop* filter, nama mata pelajaran, nama guru dan nama kelas di drop row sedangkan untuk mengetahui perbandingan antara semester 1 dan 2 maka dimensi semester di taruh di drop column.

4) Grafik Analisa rata-rata kelas

Menunjukkan bahwa nilai rata-rata kelas siswa tertinggi terdapat pada semester 2 sebesar 78 dibandingkan dengan semester 1 sebesar 77.



Gambar 14: Grafik rata-rata kelas

DAFTAR PUSTAKA

- [1] Sing, Harry. *Interaktive Datawarehousing*. Prentice-Hall.Inc. 1999
- [2] Inmon,W.H. *Building the Data Warehouse*. Jonh Wiley. 1992
- [3] Forsman, Sarah. *OLAP Council White Paper*. Olap Council. 1997
- [4] Supawi. *Data Warehouse dan Keamanan OLAP*. Informasi. Bandung. 2004
- [5] Nurjanah. *Analisa Penjualan Furniture*. Informatika. Surabaya. 2007
- [6] Kenneth E. Kendal, Julie E. Kendal. *System Analysis and Design. 6 Edition*. Pearson Education International
- [7] Turley, Paul. *SQL Server Reporting Service*. Wiley Publishing.Inc. 2004
- [8] Taufik. *Implementasi Data Warehouse*. Informatika. Surabaya. 2009