BAB IV
PERANCANGAN INTEGRASI DATA

4.1 *Data Track*

Data track merupakan pendekatan dalam merancang basis data yang bertujuan untuk menggabungkan data dari berberapa sumber data masing-masing Divisi di Telkom. Perancangan basis data dapat dimulai dengan membuat suatu model berdasarkan informasi yang digunakan oleh perusahaan tersebut berupa *Conceptual Design,* lalu mendesain model dimensi atau *dimensional modeling* yang dikatagorikan kedalam *Logical Design*, dan selanjutnya model dimensi dikonversi menjadi *Physical design.*

4.1.1 Perancangan *Data Warehouse*

Perancangan *data warehouse* merupakan kelanjutan dari proses perencanaan pada dan analisis kebutuhan pada tahap sebelumnya, dengan menitikberatkan pada *Conceptual Design*. Agar proses pembuatan *data warehouse* menjadi sistematis tahapan ini menggunakan standar *Nine Step Methodology.*

1) Memilih Proses (*Choosing the process*)

Pada tahap pemilihan proses ini, didasarkan pada proses bisnis dan fungsi yang terlibat pada unit *War Room* Telkom, antara lain dengan melakukan identifikasi dari beberapa proses bisnis dan kebutuhan yang sesuai dengan peta peran dalam organisasi *War Room* Telkom tersebut. Hasil identifikasi proses yang akan digunakan dalam perancangan sistem integrasi data dapat dideskripsikan pada table 4.1.

Tabel 4.1 Pemilihan Proses

Bisnis Proses	Deskripsi	Fungsi Yang Terlibat
Akuisisi Pelanggan	Merupakan proses bisnis yang berhubungan dengan pasang baru pelanggan dan pencabutan untuk berhenti berlangganan	Unit Marketing, Operasional di Wilayah dan Divisi
Billing	Merupakan proses bisnis yang berhubungan dengan pendapatan dari tagihan biaya jasa pelayanan dan proses yang berhubungan dengan belum adanya realisasi pembayaran tagihan.	Unit Finance, Operasional di Wilayah dan Divisi
Potensi Infrastruktur	Merupakan proses bisnis yang berhubungan dengan penyediaan data infrastruktur telekomunikasi yang siap jual	Unit Network Management System Divisi
Analisis Eksternal	Merupakan proses bisnis yang berhubungan dengan analisa data terhadap data potensi kompetitor dan data gangguan masal	Unit Consumer, Agen Data Kompetitor dan Gamas

2) Memilih *Grain* (*Choosing the grain*)

Setelah proses bisnis ditentukan, selanjutnya adalah menentukan *grain* dari proses tersebut yang nantinya akan dijadikan table fakta. Berikut merupakan data dari calon tabel fakta yang dapat dianalisis (*grain)*, yang dideskripsikan pada tabel 4.2.

Tabel 4.2 Pemilihan Sumber

Grain	Deskripsi	Proses Bisnis Terkait
Transaksi Lis Pelanggan	Jumlah pelanggan yang melakukan pasang baru	Pasang Baru Pelanggan
Transaksi Pendapatan	Nilai tagihan yang dapat diterima menjadi pendapatan diluar tunggakan	*Billing System*
Transaksi Potensi	Jumlah potensi infrastruktur yang dapat dipasarkan	Potensi Infrastruktur
Transaksi eksternal	Jumlah data dari kompetitor dan data Gamas	Analisis Data Eksternal

3) Identifikasi dan Penyesuaian Dimensi (*Identifying & Conforming the Dimensions*)

Tahap berikutnya adalah mengidentifikasi dimensi yang terkait dengan grain tersebut. Tabel dimensi yang digunakan dapat diuraikan pada table 4.3.

Tabel 4.3 Identifikasi dan Penyesuaian Dimensi

Dimensi	Deskripsi	Grain
Pelanggan	Data pelanggan yang melakukan transaksi	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Produk	Jenis produk layanan yang dipasarkan oleh Telkom	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Lokasi	Lokasi wilayah pelayanan tempat terjadinya transaksi	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
LineCat	Katagori pelanggan yang melakukan transaksi	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Waktu	Waktu transaksi dilaksanakan dengan hirarkhi tahun, bulan	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Harga	Data harga abonemen, feature produk yang dipasarkan	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Kompetisi	Data eksternal kompetitor untuk mengetahui status market share	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
Gamas	Data gangguan yang digunakan sebagai early warning	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal
NMS	Data infrastruktur yang digunakan sebagai data potensi	Transaksi LIS Pelanggan, Pendapatan, Potensi dan Eksternal

4) Memilih Fakta (*Choosing the facts*)

Setelah menentukan dimensi terkait, selanjutnya adalah memastikan table fakta yang diperoleh dalam proses *grain*.

Tabel 4.4 Pemilihan Fakta

Fakta	Deskripsi	Isi
fakta_lis_pelanggan	Fakta LIS (*Line In Service*) Pelanggan meliputi jumlah pelanggan yang merupakan hasil perhitungan jumlah pelanggan pasang baru	jumlah_pelanggan dim_pelanggan_sk_pelanggan, dim_produk_sk_produk, dim_lokasi_sk_lokasi, dim_linecat_sk_linecat, dim_waktu_sk_waktu
fakta_pendapatan	Fakta Pendapatan meliputi jumlah Pendapatan yang merupakan hasil perhitungan Nilai penjumlahan pendapatan	jumlah_pendapatan dim_pelanggan_sk_pelanggan, dim_produk_sk_produk, dim_lokasi_sk_lokasi, dim_linecat_sk_linecat, dim_waktu_sk_waktu dim_harga_sk_harga
fakta_potensi	Fakta Potensi infrastruktur meliputi status yang siap dipasarkan	jumlah_potensi dim_pelanggan_sk_pelanggan, dim_produk_sk_produk, dim_lokasi_sk_lokasi, dim_waktu_sk_waktu dim_nms_sk_nms
fakta_eksternal	Fakta data eksternal Kompetitor dan Gamas	jumlah_lis_kompetitor jumlah_gamas dim_pelanggan_sk_pelanggan, dim_produk_sk_produk, dim_lokasi_sk_lokasi, dim_waktu_sk_waktu dim_kompetisi_sk_komp dim_gamas_sk_gamas

5) Menyimpan Prekalkulasi Tabel Fakta (*Storing Pre-calculations in Fact Table*)

a. Agregasi pada Fakta LIS Pelanggan

Adalah jumlah pelanggan baru yang melakukan transaksi pemasangan produk telekomunikasi pada masing-masing wilayah dan divisi. Jumlah pelanggan aktif adalah dikurangi jumlah pelanggan yang mengakhiri masa berlangganan atas permintaan sendiri atau karena tunggakan terhadap tagihan pada masing-masing wilayah dan divisi.

b. Agregasi pada Fakta Pendapatan

Adalah jumlah pendapatan yang merupakan nilai pendapatan dari tagihan pemakaian dan pendapatan lainnya dari biaya pasang baru, mutasi, balik nama seluruh pelanggan sesuai jenis produknya pada masing-masing wilayah dan divisi. Jumlah pendapatan tersebut dikurangi jumlah tunggakan merupakan nilai pendapatan yang belum dapat diterima sebagai pendapatan.

c. Agregasi pada Fakta Potensi

Adalah data potensi infrastruktur yang memanfaatkan data dari sistem NMS (*Network Management System*) yang menunjukkan data siap jual dan terisi.

d. Agregasi pada Fakta Eksternal

Adalah data eksternal yang merupakan data dari *market share* terhadap kompetitor dan data Gamas sebagai *early warning*.

6) Melengkapi Tabel Dimensi (*Rounding out the dimension tables*)

Tahapan ini menentukan kelengkapan tabel-tabel dimensi yang akan digunakan untuk membuat skema *database*. Kelengkapan tabel dimensi berisi data antara lain nama *field*, tipe data dan panjang data. Sedangkan tabel-tabel dimensi yang akan dibuat kelengkapan data nantinya dapat dilihat pada tabel 4.5.

Tabel 4.5 Tabel Dimensi

Nama Dimensi	Deskripsi	Isi
Pelanggan (dim_pelanggan)	Berisi detail data pelanggan aktif	sk_pelanggan, regno, notelp, nama, alamat, kota, paket
Produk (dim_produk)	Berisi macam produk Telkom yang dipasarkan	sk_produk,cprod, nama_produk
Lokasi (dim_lokasi)	Berisi detail kota proses pemasaran	sk_lokasi,kd_lokasi, nama_lokasi
Harga (dim_harga)	Berisi detail harga untuk masing-masing produk.	sk_harga, kd_harga, nama_harga, harga
Line Category (dim_linecat)	Berisi detail kelas katagori pelanggan.	sk_linecat, kd_linecat,nama_ linecat
Waktu (dim_waktu)	Berisi detail waktu	sk_waktu, tahun, bulan, tanggal
NMS (dim_nms)	berisi data potensi infrastruktur telekomunikasi.	sk_nms, kd_nms, nama_infra, dapot_infra
Kompetisi (dim_kompetisi)	Berisi data potensi kompetitor	sk_komp, kd_komp, nama_komp, lis, produk, sales, churn
Ganguan Massal (dim_gamas)	Berisi data gangguan massal.	sk_gamas, kd_gamas, nama_gamas, lok_gamas

7) Memilih durasi dari database (*Choose the durations of the database*)

Pemilihan durasi data histori dapat dilakukan sesuai dengan kebutuhan informasi pihak eksekutif. Pada umumnya semakin banyak data yang disimpan dalam *data warehouse,* maka semakain lengkap pula informasi yang bisa dihasilkan. Walaupun sebenarnya semakin lama data yang disimpan juga menimbulkan beberapa permasalahan, antara lain permasalahan menginterpretasikan file lama dan versi dimensinya adalah versi lama bukan versi terbaru dan akhirnya hanya menjadi data sampah yang tidak dapat dimanfaatkan. Sehingga *database* yang akan dibuat perlu ditentukan durasi penyimpanannya yang paling ideal, hal ini bertujuan untuk menjaga tingkat akurasi data. PT.Telkom memberikan rekomendasi untuk penyimpanan data sampai dengan 5 (lima) tahun.

8) Melacak perubahan dimensi secara perlahan (*Determine the need to track slowly changing dimensions*)

Untuk menentukan perubahan yang terjadi pada dimensi agar data tetap konsisten, maka menurut Kimball ada 3 (tiga) tipe SCD, yaitu SCD tipe 1 apabila ada data yang berubah, maka data sebelumnya juga berubah. SCD tipe 2 apabila ada data yang berubah, maka akan memunculkan *record* baru. SCD tipe 3 apabila ada data yang berubah, maka akan menciptakan atribut alternatif sehingga nilai dapat diakses pada *record* yang sama. PT Telkom untuk menjaga keakuratan data menggunakan type SCD tipe 2, yaitu apabila ada atribut dimensi yang berubah, maka akan menyebabkan pembentukan record baru.

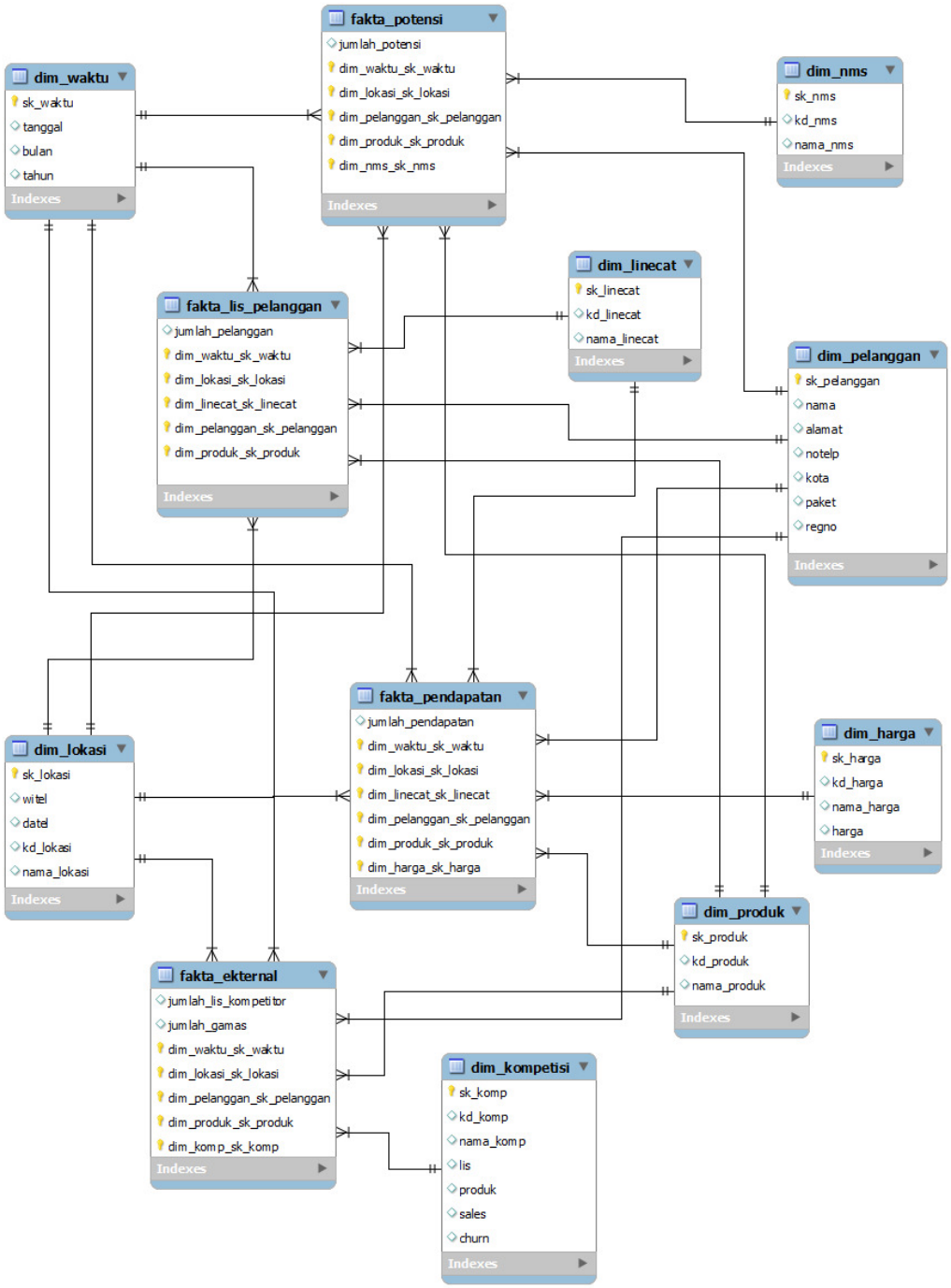
9) Memilih prioritas query dan mode query (*Decide the query priorities and the query modes*)

Dalam tahap ini dibahas mengenai proses *Extract Transform* dan *Load* (ETL) dari sumber data internal berupa tabel dari system *legacy* atau dari sumber data eksternal berupa data kompetitor dan data gangguan menjadi tabel dimensi dan tabel fakta yang akan digunakan.

4.1.2 *Dimensional Modelling*

Dimensional Modelling merupakan suatu model informasi yang digunakan untuk *logical design,* merupakan pembuatan skema *local logical model* sampai dengan *global logical design*. Skema yang dipilih adalah skema bintang (*star schema*), skema ini dipilih karena merupakan skema yang paling mudah dipahami dan digunakan, bentuknya yang tidak terlalu rumit, memudahkan dalam hal *query* dan cukup untuk memenuhi kebutuhan skema perancangan.

Pada gambar 4.1 adalah model dimensional *star schema* secara lengkap untuk memodelkan integrasi data yang dirancang pada *data warehouse*, yang terdiri dari 4 (empat) tabel fakta dan 9 (Sembilan) tabel dimensi.



Gambar 4.1 *Star Schema* Arsitektur Data

4.1.3 *Physical Design*

Pada tahap ini spesifikasi logikal diubah ke dalam detail teknologi, *physical design* yang akan dilakukan analisis adalah *metadata* untuk masing-masing tabel dimensi dan tabel fakta pada model logikal yang telah dibuat, dan juga melakukan analisis perancangan kapasitas data penyimpanan *hard disk* yang harus disediakan.

4.1.3.1 *Metadata*

Metadata adalah data tentang data atau data yang menjelaskan tentang data, metadata pada umumnya berisi informasi mengenai karakteristik data dan memegang peranan penting di dalam mekanisme pencarian maupun pertukaran data. Metadata yang ditentukan meliputi beberapa informasi berupa nama tabel beserta detail informasinya yaitu field, tipe, ukuran, sumber data dan penentuan *surrogate key*.

Penggunaan *surrogate key* dalam *data warehouse* fungsinya hampir sama dengan *primary key*, namun perbedannya pada *data warehouse* menggunaan type SCD 2 (*slowly changing dimension type two*) yaitu teknik perubahan data yang dilakukan dengan menambah record baru.

a. Metadata Tabel Dimensi Pelanggan :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_pelanggan
Deskripsi Tabel : Tabel dim_pelanggan berisi data pelanggan yang diambil dari data sumber transaksi pasang baru DTT dan DTB

Tabel 4.6 Metadata Tabel Dimensi Pelanggan

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_pelanggan	Int	4	Surrogate Key	-	-	-	-	Create
regno	varchar	15		sub_id	varchar	15	tm_pelanggan	Copy
notelp	varchar	15		phone	varchar	15	tm_pelanggan	Copy
nama	varchar	40		Pre_name	varchar	40	tm_pelanggan	Copy
alamat	varchar	40		Addr	varchar	40	tm_pelanggan	Copy
kota	varchar	15		city	varchar	15	tm_pelanggan	Copy
paket	varchar	15		package	varchar	15	tm_pelanggan	Copy

b. Metadata Tabel Dimensi Produk :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_produk
Deskripsi Tabel : Tabel dim_produk berisi data produk yang dipasarkan pada transaksi pemasaran telekomunikasi.

Tabel 4.7 Metadata Tabel Dimensi Produk

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_produk	Int	4	Surrogate Key	-	-	-	-	Create
cprod	varchar	10		abrv_prod	varchar	10	tm_produk	Copy
nama_produk	varchar	30		Prod_brand	varchar	30	tm_produk	Copy

c. Metadata Tabel Dimensi Lokasi :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_lokasi
Deskripsi Tabel : Tabel dim_lokasi berisi data lokasi untuk daerah pelayanan Telekomunikasi pada transaksi pemasaran DTT dan DTB

Tabel 4.8 Metadata Tabel Dimensi Lokasi

Field	Tipe	Uku-ran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_lokasi	Int	4	Surrogate Key	-	-	-	-	Create
kd_lokasi	varchar	10		c_datel	varchar	10	tm_lokasi	Copy
nama_lokasi	varchar	30		desc_datel	varchar	30	tm_lokasi	Copy

d. Metadata Tabel Dimensi Harga :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_harga
Deskripsi Tabel : Tabel dim_harga berisi data harga produk Telkom yang Baik pasang baru, abonemen pada transaksi pemasaran.

Tabel 4.9 Metadata Tabel Dimensi Harga

Field	Tipe	Uku-ran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_harga	Int	4	Surrogate Key	-	-	-	-	Create
kd_harga	varchar	10		c_price	varchar	10	tm_price	Copy
nama_harga	varchar	30		desc_price	varchar	30	tm_price	Copy
harga	varchar	15		price	varchar	15	tm_price	Copy

e. Metadata Tabel Dimensi LineCat :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_linecat
Deskripsi Tabel : Tabel dim_linecat berisi informasi katagori pelanggan Line Catagory

Tabel 4.10 Metadata Tabel Dimensi LineCat

Field	Tipe	Uku-ran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_linecat	Int	4	Surrogate Key	-	-	-	-	Create
kd_linecat	varchar	10		ccat	varchar	10	P_Catagorie	Copy
nama_linecat	varchar	15		desc_ccat	varchar	30	P_Catagorie	Copy

f. Metadata Tabel Dimensi Waktu :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_waktu
Deskripsi Tabel : Tabel dim_waktu berisi informasi mengenai data waktu yang Tersimpan pada transaksi pemasaran telekomunikasi.

Tabel 4.11 Metadata Tabel Dimensi Waktu

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_waktu	Int	4	Surrogate Key	-	-	-	-	Create
tahun	Int	4		year	DateTime	4	tm_time	Copy
bulan	Int	2		month	DateTime	4	tm_time	Copy
tanggal	Int	2		date	DateTime	4	tm_time	Copy

g. Metadata Tabel Dimensi NMS :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_nms
Deskripsi Tabel : Tabel dim_nms berisi informasi tentang data Potensi Infrastruktur telekomunikasi yang ada di Telkom

Tabel 4.12 Metadata Tabel Dimensi NMS

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_nms	Int	4	Surrogate Key	-	-	-	-	Create
kd_nms	varchar	10		kd_nms	varchar	5	File text	Copy
nama_infra	varchar	15		nama_infra	varchar	15	File text	Copy
dapot_infra	varchar	15		dapot_infra	varchar	15	File text	Copy

h. Metadata Tabel Dimensi Kompetisi :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_kompetisi
Deskripsi Tabel : Tabel dim_kompetisi berisi informasi tentang nama kompetitor beserta data potensinya.

Tabel 4.13 Metadata Tabel Dimensi Kompetisi

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_komp	Int	4	Surrogate Key	-	-	-	-	Create
kd_komp	varchar	10		kd_Komp	varchar	10	File Excel	Copy
nama_komp	varchar	15		nama_Komp	varchar	15	File Excel	Copy
lis	varchar	15		lis	varchar	15	File Excel	Copy
produk	varchar	15		produk	varchar	15	File Excel	Copy
sales	varchar	15		sales	varchar	15	File Excel	Copy
churn	varchar	15		churn	varchar	15	File Excel	Copy

i. Metadata Tabel Dimensi Gamas :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : dim_gamas
Deskripsi Tabel : Tabel dim_gamas berisi informasi gangguan yang bersifat masal karena adanya teknik maupun non teknik.

Tabel 4.14 Metadata Tabel Dimensi Gamas

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
sk_gamas	Int	4	Surrogate Key	-	-	-	-	Create
kd_gamas	varchar	10		kd_gamas	varchar	5	File Excel	Copy
nama_gamas	varchar	15		nama_gamas	varchar	15	File Excel	Copy
lok_gamas	varchar	15		lok_gamas	varchar	15	File Excel	Copy

j. Metadata Tabel Fakta LIS Pelanggan :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : fakta_lis_pelanggan
Deskripsi Tabel : Tabel fakta_lis_pelanggan berisi informasi jumlah pelanggan di seluruh wilayah telekomunikasi

Tabel 4.15 Metadata Tabel Fakta LIS Pelanggan

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
kode_pasang baru	Var char	15	Kode unik pelanggan	pasang baru	Var char	15	Trans_pasang baru	Transform
sk_pelanggan	Int	4	Surrogate Key	sk_pelanggan	Int	4	dim_pelanggan	Copy
sk_produk	Int	4	Surrogate Key	sk_produk	Int	4	dim_produk	Copy
sk_lokasi	Int	4	Surrogate Key	sk_lokasi	Int	4	dim_lokasi	Copy
sk_linecat	Int	4	Surrogate Key	sk_linecat	Int	4	dim_linecat	Copy
sk_waktu	Int	4	Surrogate Key	sk_waktu	Int	4	dim_waktu	Copy
Jumlah_Pelanggan	Int	15						Calculate

k. Metadata Tabel Fakta Pendapatan :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : fakta_pendapatan
Deskripsi Tabel : Tabel fakta_pendapatan berisi informasi tentang jumlah pendapatan pasang baru dan pemakaian telekomunikasi

Tabel 4.16 Metadata Tabel Fakta Pendapatan

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
kode_billing	Var char	20	Kode unik billing	billing	Var char	20	Trans_billing	Transform
sk_pelanggan	Int	4	Surrogate Key	sk_pelanggan	Int	4	dim_pelanggan	Copy
sk_produk	Int	4	Surrogate Key	sk_produk	Int	4	dim_produk	Copy
sk_lokasi	Int	4	Surrogate Key	sk_lokasi	Int	4	dim_lokasi	Copy
sk_linecat	Int	4	Surrogate Key	sk_linecat	Int	4	dim_linecat	Copy
sk_harga	Int	4	Surrogate Key	Sk_harga	Int	4	dim_harga	Copy
Jumlah_Pendapatan	Int	15						Calculate

l. Metadata Tabel Fakta Potensi :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : fakta_potensi
Deskripsi Tabel : Tabel fakta_potensi berisi informasi tentang jumlah data potensi infrastruktur telekomunikasi

Tabel 4.17 Metadata Tabel Fakta Potensi

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
Kode_infrastruktur	Var char	15	Kode unik infrastruk tur	infrastruktur	Var char	15	Trans_infrastruktur	Transform
sk_pelanggan	Int	4	Surrogate Key	sk_pelanggan	Int	4	dim_pelanggan	Copy
sk_produk	Int	4	Surrogate Key	sk_produk	Int	4	dim_produk	Copy
sk_lokasi	Int	4	Surrogate Key	sk_lokasi	Int	4	dim_lokasi	Copy
sk_nms	Int	4	Surrogate Key	sk_nms	Int	4	dim_nms	Copy
Jumlah_Potensi	Int	15						Calculate

m. Metadata Tabel Fakta Eksternal :

DMBS : MySQL
Nama Database : WARROOM
Nama Tabel : fakta_eksternal
Deskripsi Tabel : Tabel fakta_eksternal berisi informasi tentang potensi kompetitor dan data Gamas

Tabel 4.18 Metadata Tabel Fakta Eksternal

Field	Tipe	Ukuran	Ket	Sumber Data				Trans formasi
				Field	Tipe	Ukuran	Tabel	
Kode_eksternal	Var char	15	Kode unik eksternal	eksternal	Var char	15	Trans_eksternal	Transform
sk_pelanggan	Int	4	Surrogate Key	sk_pelanggan	Int	4	dim_pelanggan	Copy
sk_produk	Int	4	Surrogate Key	sk_produk	Int	4	dim_produk	Copy
sk_lokasi	Int	4	Surrogate Key	sk_lokasi	Int	4	dim_lokasi	Copy
sk_komp	Int	4	Surrogate Key	sk_komp	Int	4	dim_kompetisi	Copy
sk_gamas	Int	4	Surrogate Key	sk_gamas	Int	4	dim_gamas	Copy
Jumlah_eksternal	Int	15						Calculate

4.1.3.2 Kebutuhan Kapasitas Media Penyimpanan

Kebutuhan kapasitas media penyimpanan meliputi kapasitas media penyimpan untuk tabel dimensi dan tabel fakta, dengan repetisi ekstrak data sesuai target laporan dan kegiatan *management review* dilakukan satu minggu satu kali, sehingga jumlah untuk 5 tahun menjadi 4x12x5=240 kali proses penyimpanan.

Dengan menggunakan perhitungan dari SQL Server 2008 *Book Online* pada bagian "*Estimating the Size Of Database"* yaitu $Rn = R * (n + (1+i)^n)$, dimana R (jumlah record), n (tahun), i (prosentase pertumbuhan record per tahun), antara lain meliputi :

a. Num_Row = Jumlah Baris/Jumlah Record
b. Num_ Col = Menentukan Jumlah kolom dalam tabel
c. Fixed _Data _Size = Jumlah byte yang dibutuhkan oleh semua kolom.
d. Null_Bitmap = Bit status null kolom = 2 + ((Num_Col + 7)/8
e. Row_Size = Fixed _Data_Size + Null_Bitmap + 4
f. Row_Per_Page = 8096/ (Row_Size + 2)
g. Num_Of_Page = Num_Row / Row_Per_Page
h. Num_Of_Byte = 8192 * Num_Of_Page

Maka hasil analisis perhitungan kapasitas media penyimpanan yang perlu disiapkan untuk menampung data, yang disimpan pada tabel dimensi dan tabel fakta, dengan durasi penyimpanan sampai dengan 5 (lima) tahun dapat dijelaskan pada tabel 4.19 dan 4.20.

a) Analisis Kapasitas Media Penyimpanan Tabel Dimensi

Tabel 4.19 Analisis Kapasitas Media Penyimpanan Tabel Dimensi

Tabel Dimensi	Num_Row	Num_Col	Fixed_Data Size	Null_Bitmap	Row_Size	Rows_Per_ Page	Num_Of_ Pages	Num_Of_ Byte	Rep/Mg	Total
Pelanggan	9.234,66	7	94	3,75	101,75	78,03	118,34	969.457	240	232.669.741
Produk	2.848,99	3	29	3,25	36,25	211,66	13,46	110.266	240	26.463.872
Lokasi	2.848,99	3	29	3,25	36,25	211,66	13,46	110.266	240	26.463.872
Harga	2.848,99	3	29	3,25	36,25	211,66	13,46	110.266	240	26.463.872
Line Cat	2.848,99	3	29	3,25	36,25	211,66	13,46	110.266	240	26.463.872
Waktu	1.571,86	4	16	3,375	23,375	319,05	4,93	40.359	240	9.686.123
NMS	2.848,99	3	29	3,25	36,25	211,66	13,46	110.266	240	26.463.872
Kompetisi	8.743,46	7	89	3,75	96,75	81,98	106,65	873.655	240	209.677.127
Gamas	4.322,61	4	44	3,375	51,375	151,68	28,50	233.455	240	56.029.212

b) Analisis Kapasitas Media Penyimpanan Tabel Fakta

Tabel 4.20 Analisis Kapasitas Media Penyimpanan Tabel Fakta

Tabel Fakta	Num_Row	Num_Col	Fixed_Data Size	Null_Bitmap	Row_Size	Rows_Per_ Page	Num_Of_ Pages	Num_Of_ Byte	Rep/Mg	Total
Lis Pelanggan	3.831,40	7	39	3,75	46,75	166,07	23,07	188.996	240	45.358.969
Pendapatan	3.831,40	7	39	3,75	46,75	166,07	23,07	188.996	240	45.358.969
Potensi	3.831,40	7	39	3,75	46,75	166,07	23,07	188.996	240	45.358.969
Eksternal	4.224,37	8	43	3,875	50,875	153,12	27,59	226.012	240	54.242.885

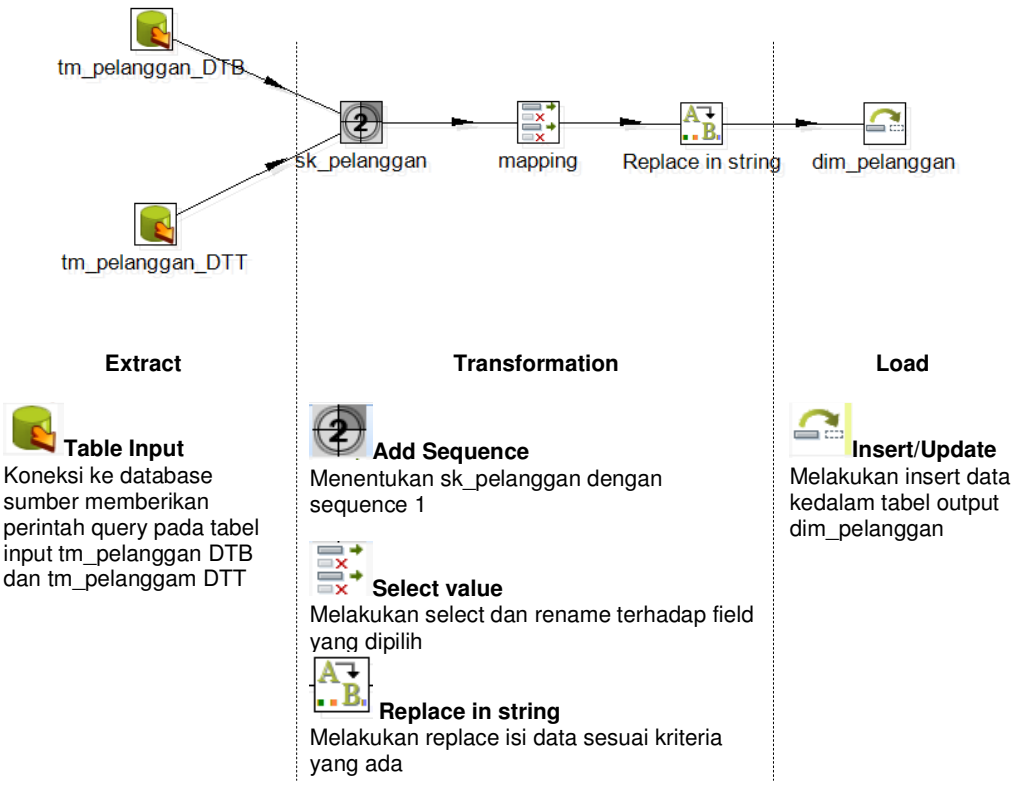
Berdasarkan hasil analisis kapasitas media penyimpanan data yang dibutuhkan pada perancangan *data warehouse* tersebut, maka dapat diperkirakan spesifikasi teknis dan kebutuhan besaran kapasitas *hard disk* minimum yang perlu disiapkan dan kemungkinan adanya perubahan atau lonjakan data selama kurun waktu 5 (lima) tahun tersebut.

4.1.4 *ETL Design*

Pada tahap ini membahas tentang menggabungkan data dari berbagai sumber internal dan eksternal dengan proses *Extraction Transformation Load* (ETL), yaitu proses *query* dari sumber data, lalu dilakukan perubahan, pembersihan menjadi data yang konsisten, data yang sudah di transformasi selanjutnya disimpan ke dalam tabel dimensi *data warehouse*. *Tools* yang digunakan untuk proses ETL ini adalah *Pentaho Data Integration* (PDI). Pada tahap ETL *design & development* ini terbagi atas dua tahap, yakni *dimension table staging* dan *fact table staging.*

4.1.4.1 *Dimensional Table Staging*

Secara umum proses ETL untuk data *staging* pada perancangan tabel dimensi dilakukan sesuai kaidah *kettle Spoon Stable release 4.2.0 Pentaho Data Integration.* Pentaho memberikan gambaran aliran proses untuk menjelaskan tahapan proses ETL-nya, sebagai berikut:

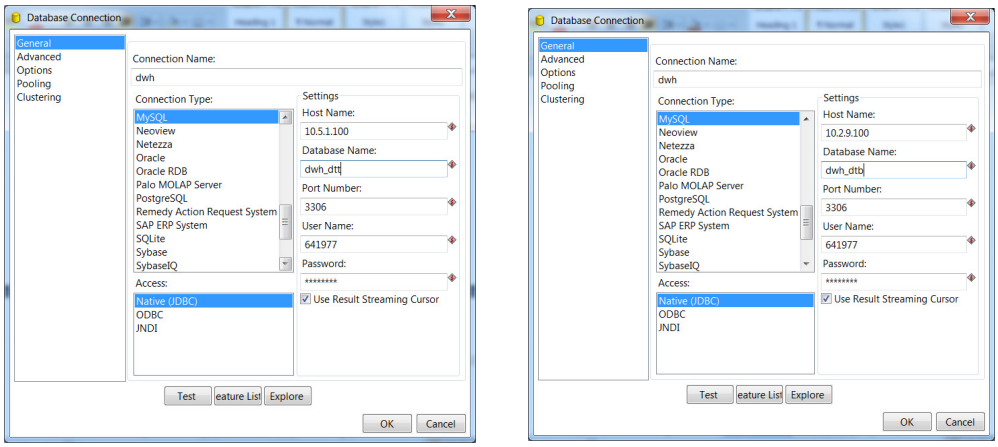


Gambar 4.2 Proses ETL Dimensi Pelanggan

Pada gambar 4.2 diatas menjelaskan proses mulai dari *Extract, Transform* sampai dengan *Loading* untuk membentuk tabel dimensi pelanggan. Prosesnya dibagi menjadi 3 tahap, dan tiap tahap perlu dilakukan pemasukan data sesuai dengan konektifitas ke *database Server*.

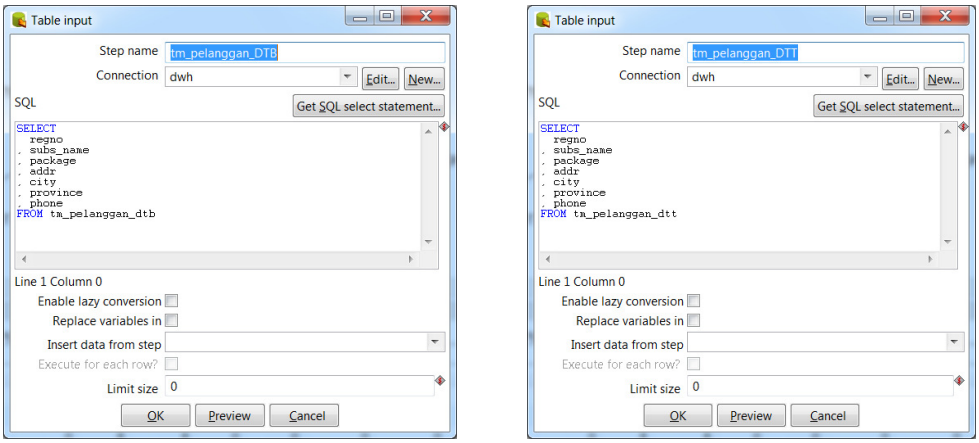
a) Proses *Extract* pada dimensi pelanggan :

Pada tahap ini dilakukan koneksi ke *database* sumber telkomdwh pada tabel master_pelanggan, misalnya untuk data pelanggan ini didapat dari 2 (dua) sumber, tabel master pelanggan DTB (IP address 10.2.9.100) dan dari tabel master pelanggan DTT (IP address 10.5.1.100).



Gambar 4.3 Proses koneksi ke database sumber

Melakukan query pengambilan data dari data sumber pada tabel tm_pelanggan_DTT dan tm_pelanggan _DTB.



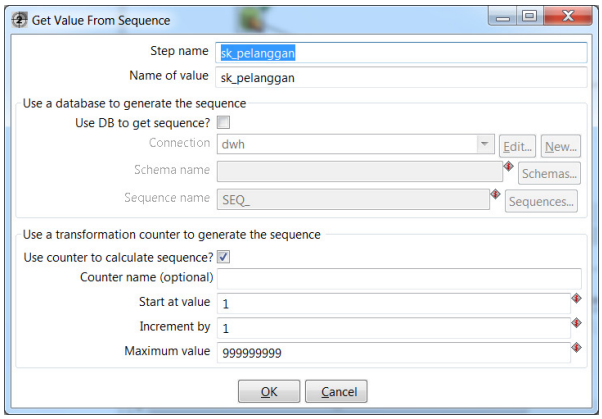
Gambar 4.4 Proses Ekstrak Dimensi Pelanggan

b) Proses *Transform* pada dimensi pelanggan :

Pada tahapan ini terdiri dari proses *add sequence, mapping* dan proses *Replace in string*.

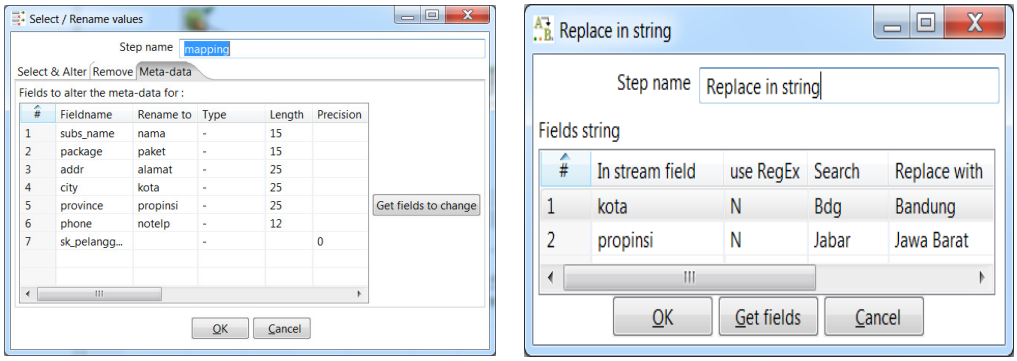
Tahap *Add Sequence* untuk pembuatan *surrogate key* pada dim_pelanggan berupa nilai urutan angka yang dimulai dari satu. Penggunaan *surrogate key*

ini didasarkan pada pemilihan terhadap tipe SCD-2 untuk penyimpanan data dengan cara menambahkan *record* baru, namun apabila kita menggunakan SCD tipe-1 maka kita dapat menggunakan *Primary Key*.



Gambar 4.5 Membuat *Surrogate Key* Dimensi Pelanggan

Tahap *mapping* melakukan *Select Value* untuk menentukan *select* dan *rename* data sesuai field pada tabel dim_pelanggan.



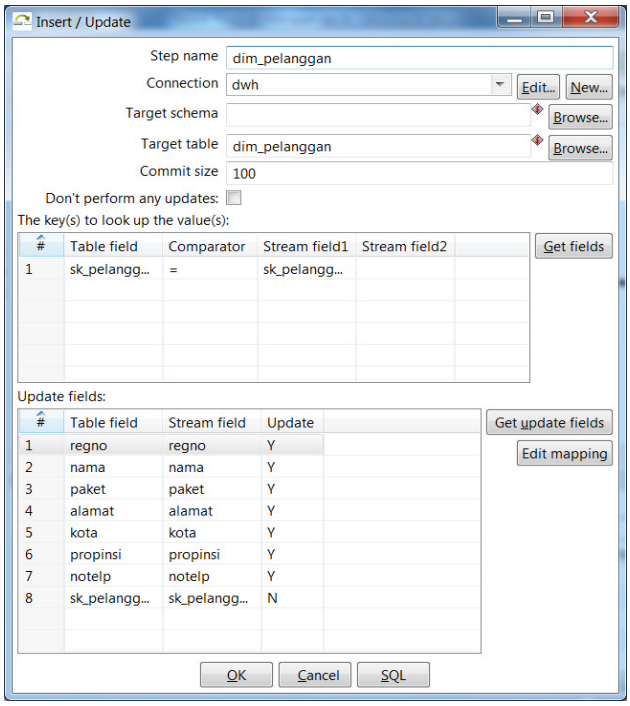
Gambar 4.6 *Transform* Dimensi Pelanggan

Tahap *Replace in string* untuk melakukan penyesuaian data dan pembersihan data tujuannya agar data yang telah diekstrak dapat dilakukan perubahan untuk disesuaikan dengan format yang telah disepakati antara lain nama *field*, panjang, tipe sampai dengan keseragaman isi data,

misalnya nama kota dari sumber data tertulis Bandung, Bdg, Bandoeng maka perlu disamakan penulisannya menjadi Bandung. Contoh lain nama propinsi dari sumber data tertulis Jawa Barat, Jabar maka perlu disamakan menjadi Jawa Barat.

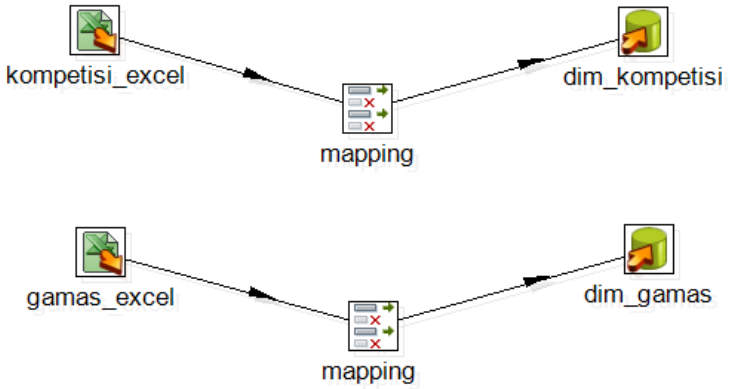
c) Proses *Loading* pada dimensi pelanggan :

Data-data tersebut selanjutnya di-*load* dengan melakukan *insert* data kedalam tabel dimensi dim_pelanggan, karena model penyimpanan datanya kita mengacu pada SCD type-2



Gambar 4.7 Proses Insert Data Dimensi Pelanggan

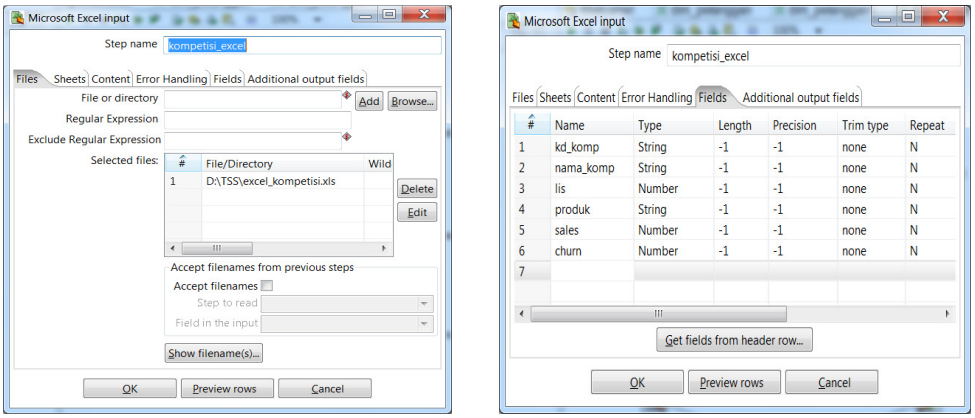
Proses ETL untuk membentuk tabel dimensi lainnya yang berasal dari data eksternal, dapat berupa data text hasil *dump* data atau data excel yang berasal dari himpunan data kompetitor, Proses ETL dari data sumber yang berupa file dengan format *spread sheet* misalnya *Microsoft Excel 97-2003 workbook*. Ilustrasi dari proses tersebut dapat dilihat pada gambar 4.8.



Gambar 4.8 Proses ETL Data *Spread sheet*

a) Proses *Extract* pada dimensi kompetisi

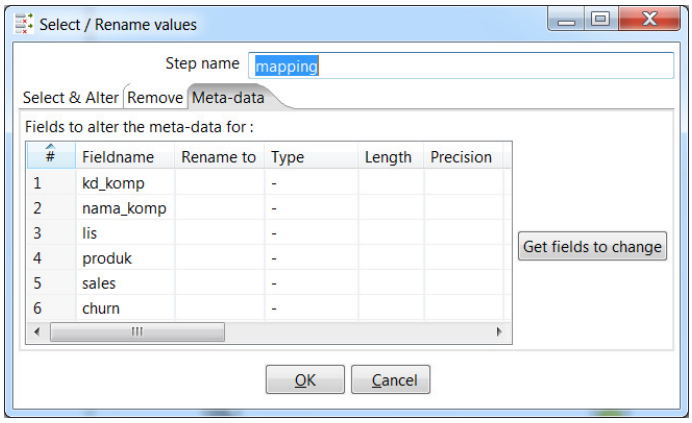
Pada Tahap ini dilakukan pengambilan data dari sumber data eksternal berupa file dengan format Excel yang berisi data market perusahaan kompetitor yang dihimpun dari berbagai sumber informasi, misalnya media massa, agen data dan lain-lain. Dari input file tersebut selanjutnya dilakukan penentuan kolom data Excel yang akan di Ekstrak sesuai dengan nama field tabel dimensi dim_kompetisi.



Gambar 4.9 Proses *Extract* Data Excel

b) Proses *Transform* pada dimensi kompetisi

Melakukan *mapping* dengan menambahkan *Select Values* untuk menentukan metadata ke tabel dim_kompetisi dan melakukan *rename* sesuai *field* tabel dimensi dim_kompetisi.

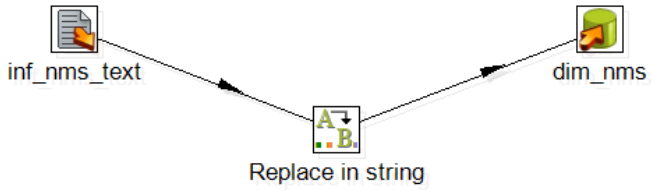


Gambar 4.10 Proses *Transform* Data Excel

c) Proses *Load* pada dimensi kompetisi

Data-data tersebut selanjutnya di *load* ke tabel output tabel dimensi dim_kompetisi.

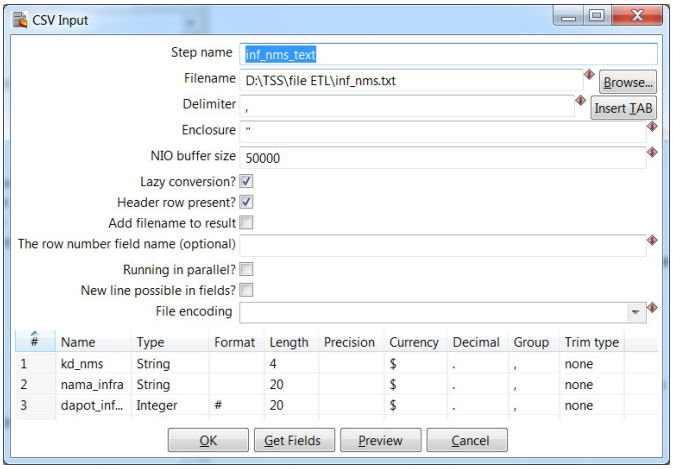
Proses ETL dari data sumber NMS berformat CSV (text)



Gambar 4.11 Proses ETL Data Text

a) Proses *Ekstrak* pada dimensi nms

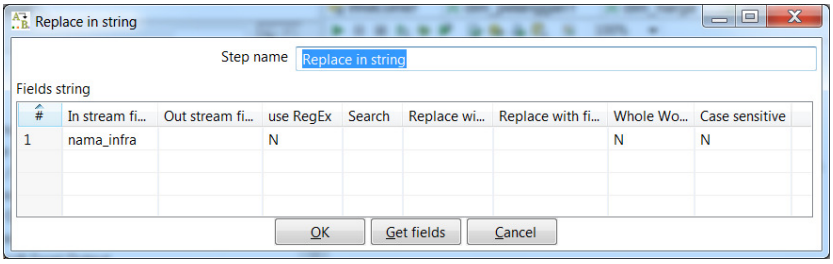
Pada tahap ini dilakukan pengambilan data dari sumber data berupa file text yang berasal dari output system NMS, karena data text tidak mempunyai header, maka perlu membentuk *field* dengan memanfaatkan *delimiter* berbentuk koma, spasi atau lainnya.



Gambar 4.12 Proses *Extract* untuk CSV input

b) Proses *Transform* pada dimensi nms

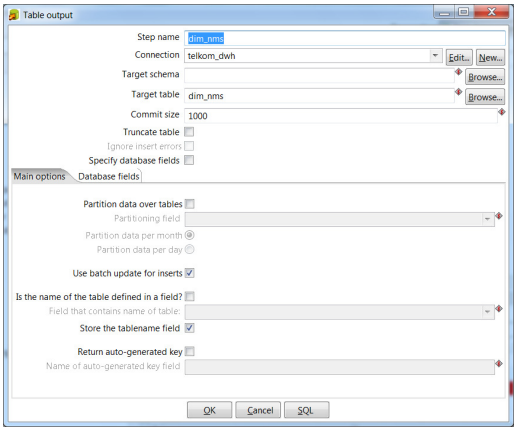
Menambahkan *Replace in string* apabila ada data yang ingin diganti agar data menjadi konsisten.



Gambar 4.13 Proses *Transform* untuk CSV input

c) Proses *Load* pada nms

Data selanjutnya data yang sudah melalui proses *ekstract* dan *transform* tersebut, maka selanjutnya data akan di *loa*d kedalam tabel dimensi nms.



Gambar 4.14 Proses *Load* untuk CSV input

Secara lengkap tabel dimensi lain yang dibentuk dari data database sumber antara lain :

a) Table Dimensi Pelanggan

b) Tabel Dimensi Produk

c) Tabel Dimensi Lokasi

d) Tabel Dimnesi Line Category

e) Tabel Dimensi Waktu

f) Tabel Dimensi NMS

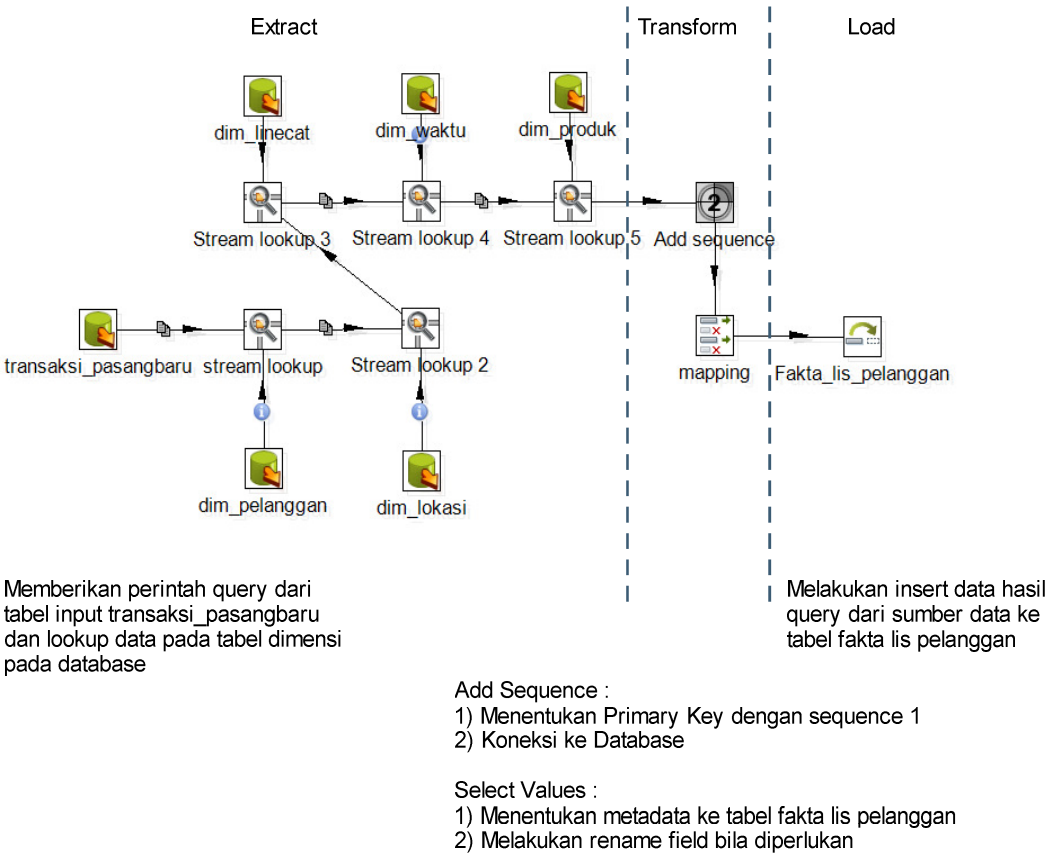
g) Tabel Dimensi Kompetitor

Prosedur detail dari proses ETL (*Ekstrak Transform Load*) masing-masing tabel dimensi dapat dilihat pada lampiran.

4.1.4.2 *Fact Table Staging*

Tahap selanjutnya adalah table fakta, antara lain tabel fakta LIS Pelanggan, Pendapatan, Potensi, Eksternal yang sesuai *dimensional modeling* dibentuk dari

tabel-tabel dimensi dengan skema bintang (*star schema*). Pada gambar 4.10 adalah salah satu proses ETL untuk membentuk tabel fakta yang menjelaskan proses yang dimulai dari *Extract, Transform* sampai *Loading* terhadap tabel data transaksional dan tabel dimensi sehingga membentuk tabel fakta lis pelanggan ke *data warehouse.*



Gambar 4.15 Proses ETL Fakta LIS Pelanggan

a) Proses *Extract* dalam fakta_lis_pelanggan :

Mengambil data transaksi pasang baru dan dim_pelanggan_sk_pelanggan, dim_produk_sk_ produk, dim_lokasil_sk_lokasi, dim_waktu_sk_waktu dari tabel dim_pelanggan, dim_produk, dim_lokasi, dim_linecat, dim_waktu.

b) Proses *Transform* dalam fakta lis pelanggan :

Membuat *Surrogate Key* untuk fakta_lis_pelanggan.

c) Proses *Loading* dalam fakta_lis_pelanggan :

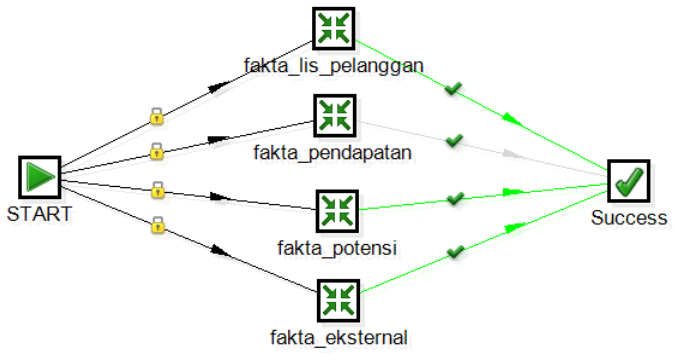
Data-data tersebut selanjutnya diload kedalam tabel fakta yaitu fakta_lis_pelanggan.

Tabel fakta yang dibentuk dari tabel dimensi yang lain adalah Tabel Fakta Pendapatan, Tabel Fakta Potensi, Tabel Fakta Eksternal. Prosedur detail proses ETL masing-masing tabel fakta dapat dilihat pada lampiran.

4.1.4.3 Penjadwalan Proses Otomatisasi ETL

Setelah proses transformasi data menjadi bentuk tabel dimensi atau bentuk tabel fakta. Semua transformasi tersebut masih dicontohkan untuk dijalankan secara manual. Untuk membuat transformasi menjadi otomatis Pentaho menamakan otomatisasi ETL dengan *job*. Fungsinya antara lain menentukan *Desain job* dan melakukan *Setting job scheduling*. Berikut adalah langkah-langkah melakukan penjadwalan ETL :

a) Menentukan *desain job* pada proses transformasi yang akan dijadwalkan ETL-nya dengan Pentaho *Data Integration*.

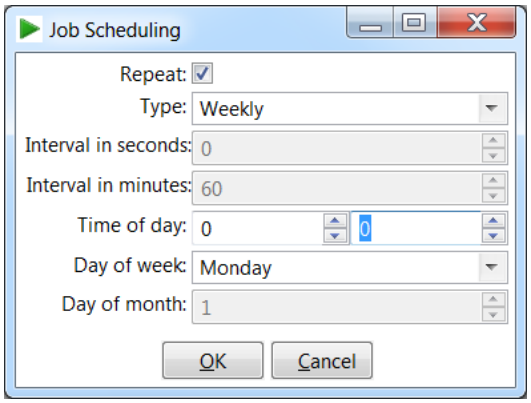


Gambar 4.16 Proses Otomatisasi ETL

Proses desain job pada gambar 4.16 menunjukkan ada empat tabel fakta yang akan menerima insert data dari sumber data tabel transaksional dan tabel dimensi secara otomatis.

b) Melakukan *setting Job Scheduling*

Pada gambar 4.17 adalah format *setting job scheduling* yang dapat diatur pada menu *Pentaho Data Integration* (PDI) untuk membuat eksekusi proses ETL secara otomatis, contoh diatas *job scheduling* diatur secara berulang setiap hari Senin pada jam 00.00.



Gambar 4.17 *Setting Job Scheduling*

Data operasional yang terdiri dari data akuisisi pelanggan yang berasal dari 61 (enam puluh satu) lokasi Wilayah Telekomunikasi (Witel) secara periodik dilaporkan oleh manajemen Witel ke Tingkat Divisi melalui Sistem Informasi sesuai divisi masing-masing yaitu Divisi Telkom Barat (DTB) dan Divisi Telkom Timur (DTT). Data tersebut beserta Sistem Informasi lainnya (TeNOSS, TIBSS, TICARES dan TREMS) diintegrasikan kedalam *Data Warehouse* melalui proses ETL.

Daftar lokasi kantor Telkom mulai dari *Corporate*, Divisi dan Wilayah Telekomunikasi dapat dilihat pada tabel 4.21.

Tabel 4.21 Daftar Lokasi Kantor Telkom

NO	CORPORATE
1	BANDUNG

NO	DIVISI
1	DIVISI TELKOM BARAT (DTB-JAKARTA)
2	DIVISI TELKOM TIMUR (DTT-SURABAYA)

DIVISI TELKOM BARAT

NO	WILAYAH TELEKOMUNIKASI (WITEL)
1	NAD (ACEH)
2	SUMUT TIMUR (PEMATANG SIANTAR)
3	SUMUT BARAT (MEDAN)
4	RIAU KEPULAUAN (BATAM)
5	RIAU DARATAN (PEKAN BARU)
6	SUMATERA BARAT (PADANG)
7	JAMBI
8	BENGKULU
9	BANGKA BELITUNG (PANGKAL PINANG)
10	SUMATERA SELATAN (PALEMBANG)
11	LAMPUNG (BANDAR LAMPUNG)
12	BANTEN BARAT (SERANG)
13	BANTEN TIMUR (TANGERANG)
14	JAKARTA BARAT
15	JAKARTA PUSAT
16	JAKARTA SELATAN
17	JAKARTA UTARA
18	JAKARTATIMUR
19	JABAR BARAT UTARA (BEKASI)
20	JABAR UTARA (KARAWANG)
21	JABAR BARAT (BOGOR)
22	JABAR SELATAN (SUKABUMI)
23	JABAR TENGAH (BANDUNG)
24	JABAR TIMUR (CIREBON)
25	JABAR TIMSEL (TASIK MALAYA)

DIVISI TELKOM TIMUR

NO	WILAYAH TELEKOMUNIKASI (WITEL)
1	JATENG BARAT UTARA (PEKALONGAN)
2	JATENG BARAT SELATAN (PURWOKERTO)
3	JATENG UTARA (SEMARANG)
4	JATENG TIMUR UTARA (KUDUS)
5	JATENG TENGAH (SALATIGA)
6	JATENG SELATAN (MAGELANG)
7	DI YOGYAKARTA
8	JATENG TIMUR SELATAN (SOLO)
9	JATIM TENGAH (KEDIRI)
10	JATIM BARAT (MADIUN)
11	JATIM SELATAN (MALANG)
12	JATIM UTARA (GRESIK)
13	JATIM SURAMADU (SURABAYA)
14	JATIM TENGAH TIMUR (SIDOARJO)
15	JATIM TIMUR (JEMBER)
16	JATIM SELATAN TIMUR (PASURUAN)
17	KALTIM TENGAH (SAMARINDA)
18	KALTIM UTARA (TARAKAN)
19	KALTIM SELATAN (BALIKPAPAN)
20	KALBAR (PONTIANAK)
21	KALTENG (PALANGKARAYA)
22	KALSEL (BANJARMASIN)
23	BALI SELATAN (DENPASAR)
24	BALI UTARA (SINGARAJA)
25	NTB (MATARAM)
26	NTT (KUPANG)
27	SULSEL (MAKASAR)
28	SULSES BARAT (PARE-PARE)
29	SULTENG (PALU)
30	SULTRA (KENDARI)
31	SULUT SELATAN (GORONTALO)
32	SULUT (MANADO)
33	MALUKU UTARA (TERNATE)
34	MALUKU (AMBON)
35	PAPUA BARAT (SORONG)
36	PAPUA (JAYAPURA)

4.1.5 Solusi Terhadap Problem *Transformasi* Data

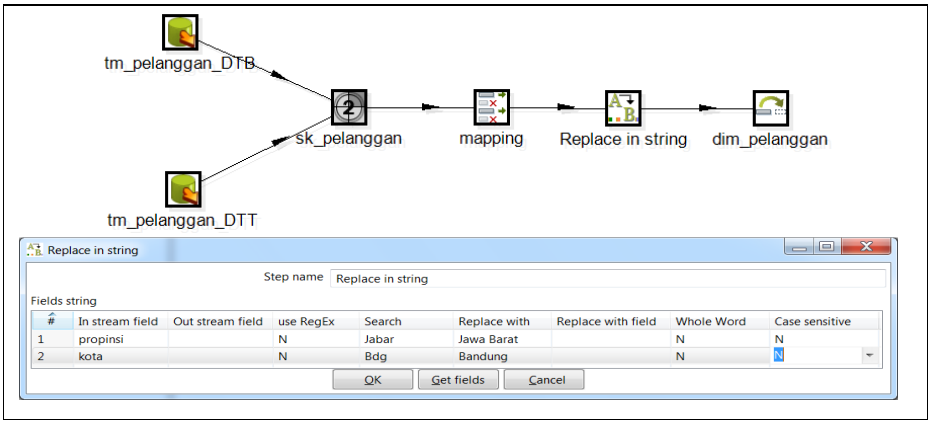
Dari beberapa permasalahan yang sudah diungkapkan pada bab sebelumnya, maka ada berberapa cara yang digunakan sebagai solusi terhadap problem transformasi data dari permasalahan Redundansi data, inkonsistensi data, isolasi data, *Backup* data, *Security* data, antara lain :

4.1.5.1 Redundansi Data

Untuk mengatasi adanya Redundansi data sehubungan adanya data yang sama akibat dari pengulangan penulisan atau keberadaan data dari file-file atau program aplikasi yang disusun oleh pengembang yang berbeda-beda, yaitu dengan cara menerapkan sistem teknologi basis data (*database*) pada *Data Warehouse*, dimana dalam proses ETL, ekstraksi data diambil dari sumber data dimana sumber datanya mempunyai kompetensi untuk memberikan data tersebut. Sebagai contoh kebutuhan data pelanggan antara lain nama, alamat dan nomor telepon bisa didapatkan dari sumber data Customer dan sumber data Billing. Sehingga data yang sama cukup dipilih satu sumber, dengan dasar pemilihan data yang dinilai lebih kompeten dan dipercaya mengeluarkan data pelanggan, dalam hal ini bisa dipilih ekstraksi dari data Customer.

4.1.5.2 Inkonsistensi Data

Untuk mengatasi adanya inkonsistensi data, yaitu sama halnya dengan redundansi data namun merupakan data yang berbeda pada tempat atau file yang sama. Misalnya kemungkinan yang terjadi adalah perbedaan penulisan nama kota Bandung dan Bdg, nama propinsi Jawa barat dan Jabar. Permasalahan ini dapat dilakukan perbaikan pada tahap Transformasi data ETL *Replace in String*, yang dapat digambarkan pada Gambar 4.18.



Gambar 4.18 Step *Replace in String* untuk *Cleaning Data*

4.1.5.3 Isolasi Data

Untuk mengatasi Isolasi data, jika data tersebar dalam beberapa file atau table dalam bentuk format yang tidak sama, maka data harus dibuat pada satu format yang sama. Format yang sama ini akan mempermudah bagi programmer untuk membuat akses data pada program aplikasinya.

4.1.5.4 *Backup* Data

Backup dilakukan dengan membuat salinan data untuk mengantisipasi apabila *data warehouse* mengalami masalah. Salinan data dapat disimpan ke dalam *harddisk portable* tujuannya agar penyimpanan terpisah dari aslinya. Penggunaan *harddisk portable* ini dengan pertimbangan lebih mudah digunakan, praktis dan lebih ekonomis.

4.1.5.5 *Security* Data

Security data atau keamanan data menjadi perhatian khusus, data yang sudah disimpan harus dapat dijaga dari ancaman permasalahan internal ataupun ancaman permasalahan eksternal.

1) Permasalahan Internal

Untuk menjaga keamanan data dari permasalahan internal misalnya gangguan perangkat atau *hardware error* disisi *server data warehouse* maka secara berkesinambungan selalu dilakukan *preventif maintenance*, sehingga dapat sedini mungkin dapat diketahui performansi *hardware*-nya, selain itu dapat menerapkan sistem redundansi server yaitu apabila *main server* mengalami *trouble*, maka secara otomatis dapat diambil alih oleh *backup server,* konsekuensinya adalah perlu diaplikasikannya sistem sinkronisasi data. Namun apabila *trouble*-nya disisi *server* sumber misalnya terjadi *downtime*, maka sebelum waktu pelaksanaan otomatis job ETL oleh Pentaho, dapat dilakukan proses pengecekan kesiapan perangkat sumbernya bisa dengan menambahkan *auto command* dengan *test connection,* apabila ditemukan kondisi abnormal tersebut sedini mungkin ditransformasikan menjadi informasi *early warning* yang segera dapat disampaikan ke pemilik otorisasi *server* sumber.

2) Permasalahan Eksternal

Sedangkan untuk menjaga keamanan data dari pihak-pihak lain yang tidak bertanggung jawab misalnya *hacker* dan *cracker*, terdapat beberapa teknik pengamanan data untuk mengantisipasi hal tersebut antara lain :

(1) Dengan memberlakukan autentikasi dan otorisasi. Autentikasi dilakukan dengan cara setiap user yang ingin mengakses data harus melakukan *login* terlebih dahulu. *Login* dilakukan dengan cara memasukkan *username* dan *password* yang terenkripsi, Autentikasi ini sesuai dengan *database* operasional yang mengacu pada tabel staff. Sedangkan *otorisasi* dilakukan dengan membatasi hak akses pengguna terhadap data dan secara otomatis

melakukan permintaan kepada user secara berkala untuk mengganti *password*-nya, dengan kombinasi *password* huruf, angka dan huruf kapital.

(2)Pengamanan jaringan dengan menggunakan *internet Firewall*, yaitu sistem *proxy* dan sistem *filtering*. sistem ini akan mengidentifikasi data-data dalam suatu jaringan agar tidak dapat diakses oleh pihak lain dari luar. Firewall akan mengontrol, mengatur, mengendalikan siapa-siapa saja yang dapat mengakses di jaringan itu, seperti sistem *clustering* karena sifatnya intranet. Tipe sistem proxy, memberi kebebasan pada user yang terkoneksi dalam jaringan untuk saling bertukar data, mengakses data serta melakukan perintah lainnya, tetapi membatasi pengguna luar untuk bergabung dalam jaringan. Tipe Proxy membaca alamat IP masing-masing client yang sudah legal terhubung agar bisa berkoneksi, jika ada perangkat lain dengan IP yang belum terverifikasi maka akses tidak diberikan. Sedangkan Tipe *Filtering*, pengamanan tipe ini juga akan mengontrol keluar masuknya data yang beredar antar perangkat, baik yang di dalam ataupun yang di luar jaringan. Data yang terdapat pada jaringan di filter terlebih dahulu keamanannya, apabila ada unsur data yang tidak dikenal dan membahayakan sebagai virus. Jadi hanya data-data tertentu saja yang diizinkan untuk beredar pada jaringan dengan pengamanan tipe *filtering*.

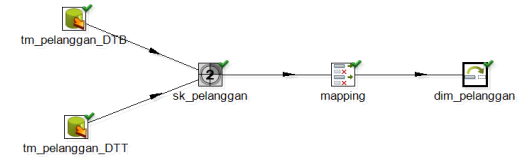
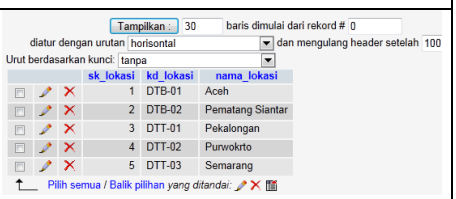
(3)Menggunakan sistem Kriptografi, yaitu sistem pengamanan yang mengutamakan pengolahan data menjadi bentuk yang rumit, dan hanya si pemilik yang mengetahui maksud dan tujuannya. Kriptografi dikenal juga dengan seni atau ilmu merahasiakan suatu informasi. Informasi diolah sedemikian rupa agar tidak terbaca oleh orang lain.

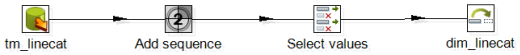
(4)Monitoring Jaringan, yaitu dengan memanfaatkan aplikasi yang dapat melihat lalu lintas data atau kegiatan dalam suatu jaringan, serta dapat menangkap informasi dan mengontrol segala kegiatan yang terjadi pada jaringan.

4.1.6 Hasil Proses ETL

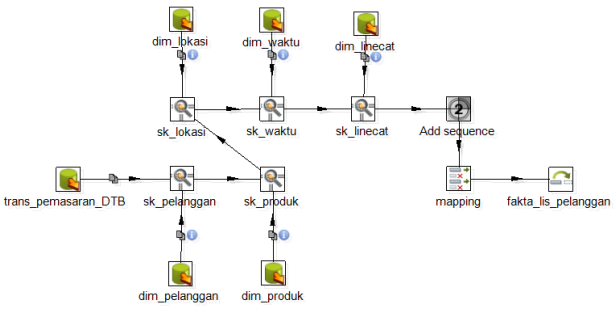
Berikut ini adalah contoh hasil proses ETL yang dilakukan dengan *tools Pentaho Data Integration.*

Tabel 4.22 Tabel Dimensi hasil proses ETL

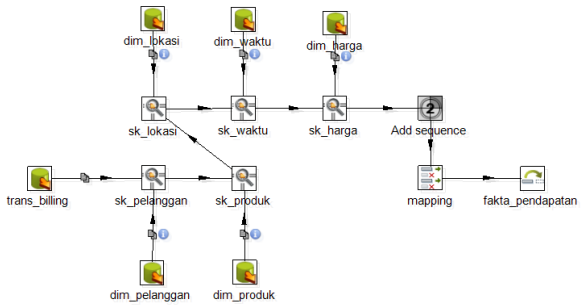
Tabel	Input	Output
dim_pelanggan		
		
dim_harga		
		
dim_lokasi		
		

dim_linecat	  
dim_produk	  

Tabel 4.23 Tabel Fakta hasil proses ETL

Tabel	Output
fakta_lis_pelanggan	 

fakta_pendapatan



Tampilkan : 30 baris dimulai dari rekord # 0

diatur dengan urutan horisontal dan mengulang header setelah 100 sel.

Urut berdasarkan kunci: tanpa

sk_pelanggan	sk_lokasi	sk_produk	sk_waktu	sk_linecat	nama_produk	nama_lokasi	tanggal	Jml_pendapatan
1	1	2	1	1	Telepon	Aceh	01032014	1000000
2	1	1	2	1	Speedy	Aceh	01032014	1000000
3	1	2	3	1	UseeTV	Aceh	01032014	1000000
4	2	2	1	1	Telepon	Pematang Siantar	01032014	1000000

Pilih semua / Balik pilihan *yang ditandai:*

4.2 *Application Track*

Pada diagram *Life Cycle Kimball* pembahasan menitik beratkan pada perancangan penggunaan aplikasi *Web Service* pada sisi *end user application.* Tahapan ini berhubungan dengan *end user application*, dimulai dengan menetukan spesifikasi dari komponen-komponen pendukung teknologi *Web Service* sampai dengan merancang aplikasi *web service* tersebut.

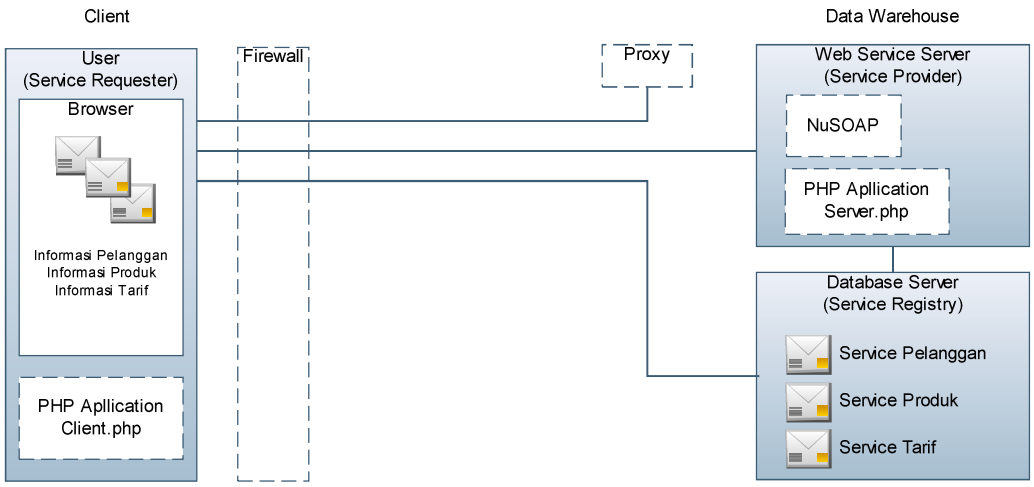
4.2.1 Perancangan *Web Service*

Perancangan teknologi *web service* ini terdiri dari beberapa komponen pendukung antara lain *user, web service server dan database server*. Pada gambar 4.19 menunjukkan rancangan *web service*, yang terdiri dari :

1) 3 (tiga) *User* yang bertindak sebagai *service requester* untuk meminta layanan data informasi pelanggan, produk dan tarif. Pada sisi *user* ini diperlukan *script* yang dicontohkan dengan bahasa pemrograman PHP dengan client.php.

2) *Web service serve*r sebagai *Service Provider*, merupakan server yang digunakan sebagai penyedia data. Pada sisi *provider* diperlukan *script* yang dicontohkan bahasa pemrograman PHP dengan server.php.

3) *Database server* sebagai *Service Registry* dalam prakteknya server ini menjadi satu dengan *web service server.* Fungsinya adalah sebagai tempat penyimpanan database yang telah didaftarkan pada *provider*.



Gambar 4.19 Perancangan Arsitektur Aplikasi *Web Service*

4.2.2 Spesifikasi *Web Service*

Aplikasi *end user* ini memanfaatkan data yang sudah diintegrasikan kedalam *data warehouse*, untuk pembahasan tersebut peneliti lebih menitik beratkan pada pembahasan aplikasi *web service*, antara lain tentang spesifikasi *web service* dalam bentuk komunikasi, template pengiriman data dan parameter yang dipergunakan untuk interaksi antara *data warehouse* dan aplikasi pengguna data.

Spesifikasi *web service* yang dirancang sebagai sistem layanan untuk melayani aplikasi pengguna menggunakan komponen-komponen pendukung berdasarkan teknologi terbuka antara lain *Simple Object Access Protocol* (SOAP),

eXtensible Markup Language (XML), Universal Description Discovery Integration (UDDI) dan *Web Servicess Description Language* (WSDL).

Dalam perancangan *web service* ini dicontohkan pada sebuah server berbasis Windows-7 yang telah dinstalasi program XAMPP (X Apache MySQL PHP dan Perl) versi 1.7.3, yang berisi software *web server apache* yang di dalamnya sudah terdapat database seperti MySQL (*Structured Query Language*) dan PHP (*PHP HyperText Preprocessor). Database server* menggunakan MySQL, karena MySQL merupakan salah satu perangkat lunak basis data yang sangat cepat, *multi-thread*, *multi-user*, dan server basis data SQL yang kuat dan bersifat *open source*. Dalam perancangan *web service* mengunakan komponen NuSOAP yang merupakan *library* yang digunakan *web service* berbasis SOAP yang ditulis dengan menggunakan bahasa PHP.

4.2.3 Daftar *Service*

Berikut ini merupakan daftar *service* yang akan digunakan untuk melayani kebutuhan data dari user atau *Service Requester.*

Tabel 4.24 Daftar *Service*

Nama Service	Parameter	Deskripsi
Pelanggan	- Notelp - Nama - Alamat - Paket	Service untuk menampilkan informasi data terbaru seluruh pelanggan yang terdaftar
Produk	- Kode produk - Nama produk	Service untuk menampilkan informasi seluruh produk yang dipasarkan
Tarif	- Produk - Tarif	Service untuk menampilkan informasi tariff, berdasarkan katagori pelanggan, paket, abonemen dan pemakaian

4.2.4 Analisis Proses Transaksi

Berdasarkan hasil dari identifikasi *service*, berikut akan dijelaskan proses yang akan terjadi dari sistem *web service* yang akan dirancang agar dapat menjamin komunikasi dan keamanan data dari satu aplikasi ke aplikasi yang lain.

1) Proses pada *server web service* (*provider entity*)

   Di dalam proses pada *provider entity* atau penyedia layanan adalah sebagai langkah awal yaitu menganalisis informasi apa yang akan ditampilkan sesuai dengan kebutuhan pengguna informasi. Pada penelitian ini akan mengembangkan aplikasi *web service* dengan mengunakan PHP dan aplikasi *web service* akan mengunakan komponen pendukung yaitu NuSOAP. Mengimplementasikan rancangan informasi yang akan ditampilkan yang berupa *source code* dengan mengunakan bahasa PHP. Layanan *web service* siap untuk di akses oleh *client* manapun.

2) Proses pada *client web service* (*requester entity*)

   Pada proses di *client web service* atau *requester entity* sebagai langkah awal yaitu mengidentifikasi dan menganalisis data atau informasi yang disediakan oleh *server* yang selanjutnya akan ditampikan pada sistem informasi *client*. Mengimplementasikan rancangan informasi yang akan ditampilkan yang berupa *source code* dengan mengunakan bahasa PHP atau mengunakan bahasa pemrograman sesuai dengan sistem informasi *client*. Selanjutnya, proses pengintegrasian layanan *web service* dengan sistem informasi *client* atau *requester* untuk dapat mengunakan layanan *web service* dari *server* atau *provider entity*, sehingga informasi dapat ditampilkan dengan baik.

4.2.5 Aplikasi *Web Service*

Aplikasi web *service* yang digunakan dalam penelitian ini menggunakan NuSOAP *toolkit*. Prosedur yang perlu dilakukan dalam mengimplementasikan *web service* dengan NuSOAP tersebut adalah sebagai berikut :

1) Melakukan ekstrak file NuSOAP, hasilnya berupa folder bernama /lib yang berisi kumpulan *class* dan meletakkan pada folder htdocs pada XAMPP.
2) Menyiapkan data yang dapat diakses oleh *client*, dalam kasus ini berupa tabel pelanggan, tabel produk dan tabel tarif pada *database* MySQL.
3) Menyiapkan script PHP *web service* di sisi *server* yang berfungsi untuk menangkap *request* dari *client.*

Pada tabel 4.25 adalah contoh kasus pencarian data pelanggan, dimana antara *Server* dengan *Client* saling berkomunikasi melalui NuSOAP *Web Service*.

Tabel 4.25 *Web Service* sisi *server*

Sisi *Server*	*Script* dengan PHP
Koneksi library NuSOAP	`require_once('lib/nusoap.php');`
Membuat object dari classsoap_server	`$server = new soap_server;`
Registrasi method	`$server->register('search');`
Koneksi ke database	`mysql_connect('dbhost', 'dbuser', 'dbpass');` `mysql_select_db('dbname');`
Query pencarian data pelanggan	`$query = "SELECT * FROM fact_lis_pelanggan WHERE notelp = '$key' OR nama LIKE '%$key%' OR alamat LIKE '%$key%'";`
Menyiman data hasil dalam bentuk array	`$result[] = array('notelp' => $data['notelp'], 'nama' => $data['nama'],'alamat' => $data['alamat']);`
Mereturn array hasil pencarian	`return $result;`
Mengecek Variable HTTP_RAW_POST_DATA	`$HTTP_RAW_POST_DATA = isset($HTTP_RAW_POST_DATA) ? $HTTP_RAW_POST_DATA : '';` `$server->service($HTTP_RAW_POST_DATA);`

Tabel 4.26 *Web Service* sisi *client*

Sisi *Client*	*Script* dengan PHP
Membuat *keyword* pencarian	`Keyword Pencarian <input type="text" name="key"> <input type="submit" name="submit" value="Search">`
Mengarahkan ke Komputer Server	`$client = new nusoap_client('http://…/server.php');`
Proses *call method* sesuai parameter keyword	`$result = $client->call('search', array('key' => $key));`
Menampilkan hasil	`if (is_array($result)) { echo "<tr><th>NOTELP</th><th>NAMA</th><th>ALAMAT</th></tr>"; foreach($result as $data){ echo "<tr><td>".$data['notelp']."</td><td>".$data['nama']."</td><td>".$data['alamat']."</td></tr>";`

Struktur XML dalam SOAP selama proses transaksi antar sistem baik dalam proses *call/request* maupun *respons* dalam *web service* dapat digambarkan sebagai berikut :

Tabel 4.27 Struktur XML *Request*

Request : POST /nusoap/server.php HTTP/1.0 Host :localhost User-Agent: NuSOAP/0.7.2/(1.95) Content-Type: text/xml; charset=ISO-8859-1 SOAPAction:”” Content-Length: 913 <?xml version=”1.0” encoding=”ISO-8859-1”?> <SOAP-ENV:Envelope SOAP-ENV:encodingStyle= ”http://schemas.xmlsoap.org/soap/encoding/” xmlns:SOAP-ENV=”http://schemas.xmlsoap.org/soap/envelope/” xmlns:xsd=”http://www.w3.org/2001/XMLSchema xmlns:xsi=”http://www.w3.org/2001/XMLSchema-instance” xmlns:SOAP-ENC=”http://schemas.xmlsoap.org/soap/encoding/”> <SOAP-ENV:Body> <ns8591:Data_Pelanggan xmlns:ns8591=”http://tempuri.org> <notelp xsi:type=”xsd:string”>022210000</notelp> </ ns8591:Data_Pelanggan> </SOAP-ENV:Body> </SOAP-ENV:Envelope>	SOAP-ENV:Envelope SOAP-ENV:Header SOAP-ENV:Body

Tabel 4.28 Struktur XML *Respons*

Respons :	
HTTP/1/1 200 OK Date: Wed, 21 Mei 2014 10:16:22 GMT Server: Apache/1.3.23(Win32) X-powered-By: PHP/5.2.17 X-SOAP-Server:NuSOAP/0.9.5 (1.123) Content-Length: 653 Connection: close Conent-Type: text/xml; Charset=ISO-8859-1	
<?xml version=”1.0” encoding=”ISO-8859-1”?> <SOAP-ENV:Envelope SOAP-ENV:encodingStyle=”http://schemas.xmlsoap.org/soap/encoding/” xmlns:SOAP-ENV=”http://schemas.xmlsoap.org/soap/envelope/” xmlns:xsd=”http://www.w3.org/2001/XMLSchema xmlns:xsi=”http://www.w3.org/2001/XMLSchema-instance” xmlns:SOAP-ENC=”http://schemas.xmlsoap.org/soap/encoding/”> <SOAP-ENV:Body> <ns1:Data_Pelanggan xmlns:ns1=”http://tempuri.org> <return xsi=type=”xsd:string”> Budi</return> </ ns1:Data_Pelanggan> </SOAP-ENV:Body> </SOAP-ENV:Envelope>	SOAP-ENV:Envelope SOAP-ENV:Header SOAP-ENV:Body

Source code secara lengkap yang ada pada sisi *server* dan *source code* yang ada di sisi *Client,* dapat dilihat pada lampiran.

4.3 *Technology Track*

Setelah proses perancangan data dan perancangan aplikasi sudah disiapkan maka tahap selanjutnya adalah perancangan teknologi berupa spesifikasi *hardware & software* pada *War Room* Telkom dan desain arsitektur teknologi integrasi data secara keseluruhan.

4.3.1 Kebutuhan Perangkat

Untuk keperluan perancangan arsitektur teknologi tentunya memerlukan data teknis terhadap kebutuhan-kebutuhan *hardware* dan *software.* Dalam perancangan ini sebagian besar memanfaatkan *hardware* eksisting mulai dari server

sampai dengan infrastruktur koneksi internet, sedangkan untuk *software* menggunakan *software* yang umum dan bersifat *open source*.

4.3.1.1 *Hardware*

Perangkat keras (*hardware*) berupa sebuah komputer yang dipergunakan sebagai *server, client* dan perangkat untuk interkoneksi ke jaringan internet, dengan menggunakan spesifikasi sebagai berikut :

1) Spesifikasi *Server* :
   - Server : IBM Blade Center HS22
   - Form factor : Single wide (30 mm)
   - Processor : Choice of two intel Xeon 5600 series up to 3.60 GHz
   - Number of processor : (std/max) 1/2
   - Memory : DDR-3 VLP DIMM slots
   - Expansion slots : one CIOv slot and one CFFh slot Hard disk 500GB
   - Disk bays : two hot-swap bay supporting SAS HDDs or solid state drives
   - Max internal storage : Up to 1.0 TB total internal storage
   - Network interface : model Broadcom 5709S on board NIC with dual Gigabit Ethernet ports with TOE.
2) Spesifikasi *Client* :
   - Intel Core i5
   - Memory 4 GB RAM
   - Hard disk 500 GB
   - Monitor 14” LCD
3) Modem ADSL untuk koneksi internet dan TP-Link 24 Port untuk mendukung jaringan / Local Area Network (LAN).

4.3.1.2 *Software*

Perangkat lunak (*software*) yang terdiri dari *web server*, *database server* dan aplikasi *tools* pendukung, sebagai berikut :

1) Web Server XAMPP versi 1.7.3

   XAMPP merupakan tool yang menyediakan perangkat lunak dalam satu buah paket, yang terdiri dari :

   - Apache 2.2.14 (IPv6 enabled) + Open SSL 0.9.81
   - phpMyadmin 3.2.4
   - PHP 5.3.1
   - Perl 5.10.1
   - Filezilla FTP server 0.9.33
   - Mercury Mail Transport system 4.72

2) My SQL
   - Server: localhost via TCP/IP
   - Versi Server: 5.1.30-community
   - Protocol version: 10
   - Charset MySQL: UTF-8 Unicode (utf8)

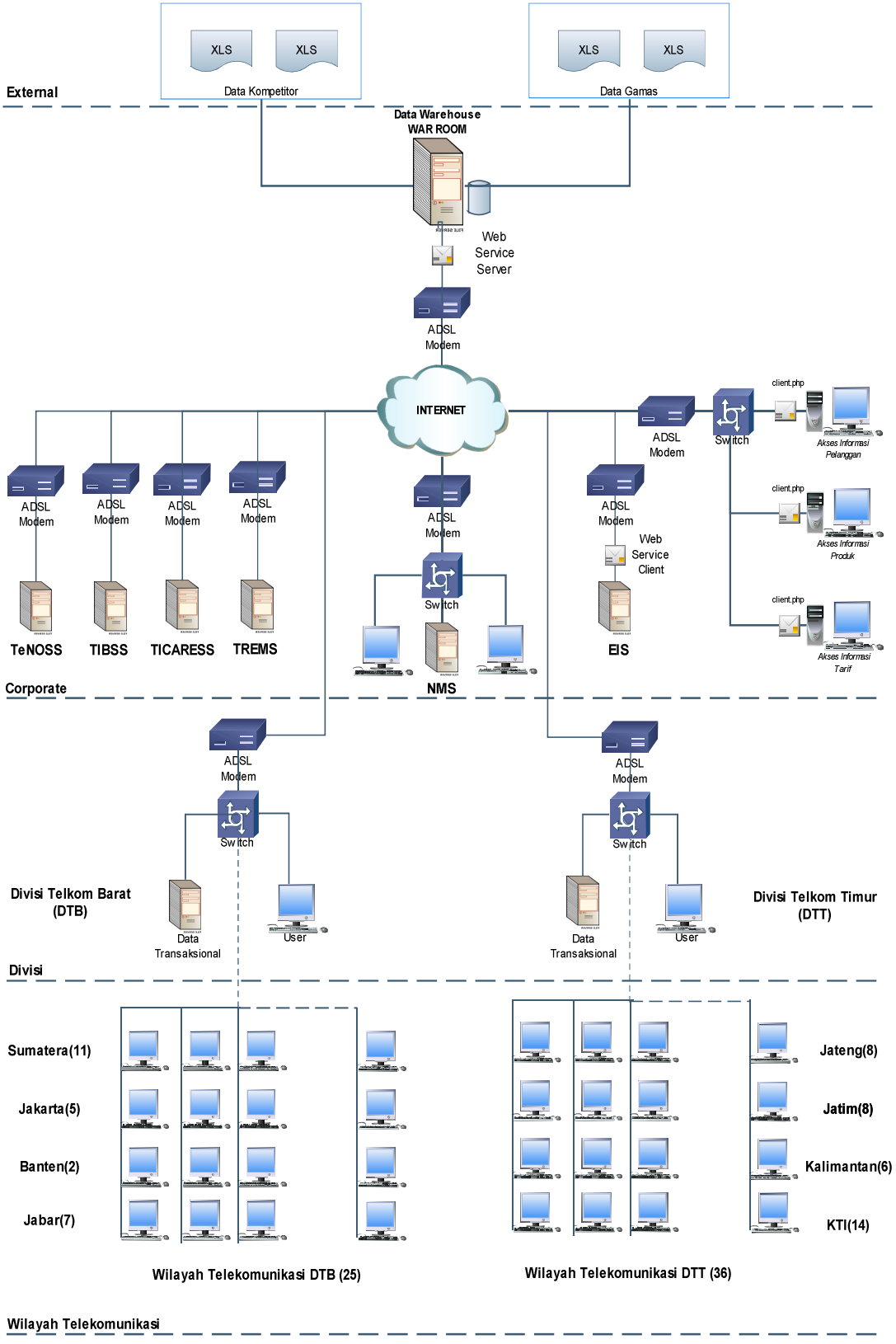
3) Aplikasi tools pendukung
   - *Pentaho Data Integration* Stable release versioan 4.2.0, yaitu *versi community* yang bersifat *open source* yang digunakan untuk proses *Extract Transform Load*.
   - MySQL workbench 5.2.47 CE, yang digunakan untuk merancang diagram skema database yang terkoneksi dengan database MySQL.
   - Tools pendukung lainnya.

4.3.2 Arsitektur Teknologi Integrasi Data

Pendekatan teknologi mencakup rancangan arsitektur teknologi dan keperluan perangkat baik sisi *hardware* maupun *software* yang diperlukan untuk keperluan integrasi data. Perancangan arsitektur teknologi untuk menggambarkan proses integrasi data yang dilakukan untuk keperluan data pada unit *War Room* Telkom terdiri dari 3 (tiga) bagian yaitu :

1) Sumber data.
   (1) Data Internal *Corporate* (TeNOSS, TREMS, TIBS dan TICARES) dan data Divisi (DTB dan DTT) yang berasal dari 61 (enam puluh satu) lokasi wilayah telekomunikasi berupa data tabel dari database sumber yang dipilih.
   (2) Data Eksternal Kompetitor dan Gangguan massal, data ini tidak berbentuk tabel dari database, namun berbentuk file *spread sheet* dan data *text*.
2) *Data warehouse War Room* Telkom
   Data dilakukan proses *Extract Transform Load* (ETL) dengan bantuan *tools Pentaho Data Integration* untuk melakukan proses pembersihan dan penyeragaman data sesuai dengan kaidah yang telah ditentukan, selanjutnya di *load* kedalam media penyimpanan *data warehouse*.
3) *Web service*
   Data yang telah terintegrasi dapat dimanfaatkan untuk operasional *War Room* Telkom dengan menyediakan beberapa layanan informasi Pelanggan, Informasi Produk dan Informasi Tarif dengan menempatkan aplikasi *web service* pada sisi *Client* dan sisi *Server*.

Pada gambar 4.20 dapat dijelaskan arsitektur teknologi yang memperlihatkan keterkaitan antara ketiga bagian integrasi data tersebut.



Gambar 4.20 Perancangan Arsitektur Teknologi Integrasi Data

Proses mengekstrak data dari berbagai sumber yang berhubungan dengan keperluan unit *War Room* Telkom, antara lain data pemasaran dari sistem aplikasi data transaksional DTB dan DTT, data *billing*, data pelanggan data produk, data potensi infrastruktur serta data-data eksternal yang diperlukan untuk analisa kompetitor dan gangguan masal sebagai *early warning*. Kemudian data tersebut diambil dan dilakukan transformasi data agar data berubah menjadi bentuk yang terintegrasi dan konsisten dan akhirnya data di *load* kedalam *Data Warehouse* yang terpusat dan independen.

Arsitektur yang terpusat di *Corporate* pada didasarkan pada hal-hal sebagai berikut :

1) Seluruh data operasional dari kinerja masing-masing unit bisnis dilaporkan sampai tingkat *Corporate*.
2) Pengembangan rancangan *data warehouse* terpusat relatif lebih mudah dan praktis daripada *data warehouse* yang terdistribusi.
3) Data yang dibutuhkan unit *War Room* Telkom dari masing-masing data operasional akan tersimpan didalam satu *server.*
4) Mudah dalam menjaga konsistensi data terhadap segala perubahan karena proses *Extraction Transformation Load* dilakukan dalam satu proses yang terpusat.
5) Mudah dalam melakukan pengoperasian dan pemeliharaan baik daris sisi *hardware* maupun *software*.