

Università degli Studi di Padova



Facoltà di SCIENZE STATISTICHE

Laurea triennale in *STATISTICA E TECNOLOGIE INFORMATICHE*

PROVA FINALE:

**ESPERIENZA DI STAGE IN SIDI SpA:
DATA WAREHOUSE BW E SAP R/3.**

**EXPERIENCE OF STAGE IN SIDI SpA:
DATA WAREHOUSE BW AND SAP R/3.**

Relatore: Ch.mo Prof. SUSI DULLI

Laureando: TOSARELLO ENRICO

n.matr.: 534019 - STI

A.A. 2007/2008

PREFAZIONE

In questo elaborato si vuole descrivere una attività di stage svolta in SIDI SpA.

Esso è suddiviso in due parti: nella prima si vogliono accennare dei concetti teorici utili per capire il significato degli oggetti e strutture utilizzate nell'attività di stage, mentre nella seconda una descrizione del progetto creato nell'esperienza lavorativa in SIDI SpA. I capitoli, dunque, sono così strutturati:

CAPITOLO 1:

In questo capitolo si vuole introdurre il concetto di Data Warehouse e Business Intelligence, strumenti usati per la realizzazione del progetto soffermandosi sulle principali funzionalità e utilità e distinguendo i sistemi OLTP dai OLAP.

CAPITOLO 2:

In questa parte si spiega il sistema SAP R/3 con le relative origini e sottolineando la presenza dell'architettura client-server spiegando quali siano i vantaggi che essa comporta.

CAPITOLO 3:

Nel terzo capitolo ci si sofferma sul sistema SAP BW, spiegando quali sono gli oggetti che esso utilizza e quali di questi sono rilevanti per la reportistica, importante "fase" per il processo decisionale.

CAPITOLO 4:

Svolgimento del progetto svolto in stage spiegando l'infrastruttura creata elencando i vari oggetti che la compongono e i vari passaggi per arrivare a creare il multiprovider utile per le analisi.

Infine vi sono le conclusioni, ringraziamenti e bibliografia.

PREFAZIONE	3
INTRODUZIONE.....	7
CAPITOLO 1 : IL DATAWAREHOUSE.....	10
1.1. Introduzione al Data Warehouse.....	10
1.2. Data Warehouse e Business Intelligence	11
1.3. Tipologie di sistemi: OLTP e OLAP	12
1.4. Analisi Multidimensionale.....	14
1.5. OLAP	16
1.6. Distinzione fra fatti e dimensioni.....	19
1.7. Schema a stella per la memorizzazione dei dati	20
1.8. Schema a fiocco di neve.....	22
CAPITOLO 2 : SAP R/3.....	24
2.1. L'azienda SAP R/3.....	24
2.2. Le origini e l'evoluzione di SAP R/3.....	24
2.3. Business Engineering	25
2.4. SAP: ERP e Data Warehouse.....	26
2.5. Architettura CLIENT-SERVER	27
2.6. Benefici della tecnologia CLIENT-SERVER.....	29
2.7. ABAP/4.....	30
2.8. Struttura modulare.....	31
2.9. Gestione della sicurezza.....	32
2.10. Implementazione: ASAP.....	32
CAPITOLO 3 : SAP BW	35
3.1. Business Information Warehouse SAP BW.....	35
3.2. SAP BW Informational model.....	37
3.3. Estrazione, Trasformazione e Caricamento (ETL)	40
3.4. Operational Data Store (ODS)	44
3.5. Data Warehouse e Data Mart	45
3.6. Le release di SAP BW.....	46
CAPITOLO 4 : SVOLGIMENTO DELL' ATTIVITA'	47
4.1. Obiettivo del progetto	47
4.2. Flusso del modello dati	48
4.3. Contratto.....	49
4.4. Ordini di Acquisto.....	52
4.5. Entrata Merce	55
4.6. Multiprovider APPROVIGIONAMENTI.....	59
4.7. Query Tramite il programma QUERY DESIGNER.....	60
CONCLUSIONI.....	61
BIBLIOGRAFIA	62
RINGRAZIAMENTI.....	64

INTRODUZIONE

In questi ultimi anni si è assistito ad uno straordinario sviluppo delle tecnologie informatiche in campo aziendale. Se solo pochi decenni fa lo sviluppo era legato alle attività operative, oggi l'Information Technology è entrata in modo predominante in tutte quelle operazioni che fanno parte del processo decisionale di un'organizzazione. Tutte le decisioni prese a livello manageriale, dai responsabili delle diverse aree funzionali e non solo, nascono dalla fondamentale necessità di un adeguato livello quantitativo di informazioni. La vita di un'organizzazione e la sua efficacia sono sempre più legate alla creazione e allo sviluppo di un adeguato Sistema Informativo Automatizzato (I.S. Information System).

Tale sistema rappresenta in sostanza un insieme di flussi di informazioni destinati ad essere impiegati per la presa di decisioni a diversi livelli di operatività aziendali. In realtà esso è molto di più, in quanto rappresenta un patrimonio inestimabile e inesplorato per le aziende, e troppo spesso viene usato, o meglio sottoutilizzato e trattato come puro strumento sterile di lavoro routinario. Al contrario, l' I.S. di un'azienda nasconde al suo interno informazioni e conoscenza importantissimi per i processi decisionali della stessa.

Esso si configura come un enorme magazzino che archivia e gestisce i dati provenienti sia dall'interno che dall'esterno dell'azienda e che, se adeguatamente "trattato", è in grado di fornire informazioni e più ancora conoscenza a chi sa interrogarlo.

Per considerare un progetto di Data Warehouse di successo occorre, non solo rendere disponibili informazioni e creare conoscenza, ma anche creare valore aggiunto per il business aziendale; questo valore aggiunto, indipendentemente dal tipo di azienda, può essere così concretizzato:

- Nell'aumento delle quote di mercato (ad esempio alle applicazioni di marketing, gestione complessiva dei clienti, cross-selling);
- Nella riduzione dei costi e relativo aumento dell'efficienza complessiva (ad esempio applicazioni di gestione integrata degli acquisti e delle scorte, controllo della produttività);
- Nel miglioramento della qualità dei servizi erogati (ad esempio applicazioni di customer care, trasparenza e informazioni al cittadino).

Le informazioni sono dati filtrati, sintetizzati e aggregati; esprimono dei concetti, sono il pane quotidiano per coloro che devono prendere decisioni. Se l'organizzazione è in grado di gestirsi, sopra l'informazione si costituisce la "conoscenza", risultato di una gestione razionale e continuativa delle interazioni tra informazioni, azioni passate, esperienza e comprensione. Analisi e condivisione delle informazioni contribuiscono a diffondere a tutti i livelli organizzativi e proprio questo sta alla base della "Business Intelligence" (B.I), concetto coniato nel 1989 da Howard Dresner e diffuso negli anni '90 in poi che mira a trasformare i dati archiviati in un grande repository detto Data Warehouse, in dati utilizzabili, favorendo la distribuzione e condivisione delle informazioni con tutti i dipendenti, funzionari e dirigenti, creando così una conoscenza collettiva all'interno dell'organizzazione.

Il data warehouse diventa dunque uno strumento della B.I per rendere l'azienda intelligente, uno strumento base per diffondere la conoscenza, mezzo per permettere ai responsabili delle decisioni di ottenere risposte rapide ai loro quesiti, arrivando immediatamente alle informazioni di cui hanno bisogno.

Qui nel Veneto non sono molte le aziende che si occupano di Data Warehouse e B.I, ma tra queste c'è sicuramente SIDI S.P.A, azienda nella quale ho svolto la mia attività di stage.

SIDI S.P.A, situata a Perarolo di Vogonza (PD) è all'avanguardia nella consulenza e nell'integrazione di soluzioni gestionali ad alto valore aggiunto e, come Global

System Integrator, è il partner di riferimento di oltre 100 aziende private e pubbliche, cui offre una vasta gamma di servizi mirati per tipologia e mercato.

Per questo SIDI ha creato per ogni area di attività un centro di competenza che raccoglie le esperienze e il know-how necessari per fornire il miglior valore in ogni progetto specifico: SAP, Human Resources, Business Intelligence e Application Management.

Alle aziende che affidano un progetto tecnologico viene offerto:

- un ruolo attivo nella sua realizzazione;
- una risposta qualificata anche alle esigenze più specifiche;
- una rete di partner per affrontare qualsiasi aspetto con un'offerta completa.

SIDI utilizza un sistema ERP per la gestione delle proprie attività, memorizzando le proprie transazioni commerciali in apposite tabelle di un database. Nel progetto faremo riferimento al sistema OLTP come l'insieme di server, basi di dati e moduli di interfaccia che memorizzano le transazioni commerciali. I dati vengono inseriti nel sistema OLTP tramite interfacce client, ovvero programmi che risiedono nei client e si collegano al sistema OLTP per l'esecuzione delle transazioni tramite interfacce WEB. Queste ultime risiedono nel portale aziendale, elemento fondamentale dell'organizzazione commerciale che permette lo svolgimento di diverse attività, operative o semplicemente di analisi, senza vincolo di essere in sede. Il portale diventa un momento fondamentale nel momento in cui viene progettata la parte di reporting dell'infrastruttura di supporto alle decisioni.

CAPITOLO 1 : IL DATAWAREHOUSE

1.1. Introduzione al Data Warehouse

Il concetto di Data Warehouse è sorto dalla convergenza tra il crescente interesse del business ed il progresso della tecnologia nel campo delle basi di dati.

Si tratta di una piattaforma sulla quale vengono archiviati e gestiti dati provenienti dalle diverse aree aziendali; questi dati vengono aggiornati, integrati e consolidati dai sistemi di carattere informazionale per rendere operative tutte le applicazioni di supporto alle decisioni.

Il data warehouse è stato definito storicamente come : “una raccolta di dati integrata, subject - oriented, time variant e non volatile di supporto ai processi decisionali”[In96].

Spiegheremo ora i vari aggettivi con maggior dettaglio.

- Integrato: i dati dopo essere stati estratti da varie sorgenti, devono essere ripuliti, trasformati e modellati in modo tale da essere resi omogenei e gestibili in modo semplice ed efficace. L'integrazione può essere raggiunta con l'omogeneità semantica di tutte le variabili, usando dei metodi di codifica uniformi, utilizzando le stesse unità di misura, ecc.
- Subject-oriented: viene posta l'attenzione sui principali soggetti del business aziendale; i dati vengono organizzati in base alle particolari funzioni o processi aziendali in modo da consentire una visione multidimensionale. Quanto più i dati raccolti saranno mirati al soggetto, tanto più sarà profonda la conoscenza del business aziendale a tutti i livelli e tanto più efficaci saranno le decisioni prese a partire da questa conoscenza.
- Time variant: i dati verranno organizzati in periodi di tempo, consentendo così di fornire una prospettiva storica del business, di analizzare le tendenze passate e di fare previsioni future.

- Non volatile: una volta caricati i dati nel DW, essi sono disponibili solo in lettura e non possono essere modificati. È sempre possibile poi aggiungere nuovi dati alla base già esistente.

Il data warehouse è costituito da una serie di snapshots che rappresentano la situazione delle varie aree da analizzare relative ad un certo periodo di tempo e con un certo grado di granularità, ovvero il livello di dettaglio dei dati. Da notare che la dimensione del tempo ha un ruolo fondamentale e deve essere sempre presente nel momento in cui avviene l'estrazione dei dati.

È importante tenere traccia delle trasformazioni che subiscono i dati nel passaggio di database di produzione al data warehouse, avere quindi delle informazioni che descrivano il significato e la struttura dei dati aziendali; come sono creati, come vi si accede e come possono essere utilizzati. Queste informazioni sono contenute nei cosiddetti **metadati** e assumono un ruolo di fondamentale importanza in tutto il processo di Data Warehousing, dalla sua progettazione fino all'utilizzo per le analisi.

1.2. Data Warehouse e Business Intelligence

Con il data warehouse si utilizza una nuova generazione di software per il supporto decisionale, detta **Business Intelligence**. L'obiettivo è quello di produrre con la massima tempestività le informazioni essenziali e sintetiche in modo da facilitare il processo di decision-making ai manager aziendali.

Da un articolo di Howard Dresner, analista della Gartner Group a cui si deve l'introduzione del termine business intelligence nel 1989, si legge:

“Dal 1996, tramite l'uso di soluzioni di business intelligence come migliore strategia per comprendere il business, il campo di analisi si sposterà definitivamente dagli analisti specializzati a tutti i manager e i professionisti.

Anziché un numero ristretto di analisti che impiegano il 100% del loro tempo analizzando dati, tutti i manager e i professionisti impiegheranno il 10% del loro tempo utilizzando software di business intelligence”.

È importante sottolineare l'importanza dell'integrazione del Data Warehouse con la Business Intelligence. Il data warehouse da solo, infatti, fornisce semplicemente l'infrastruttura per riunire i dati provenienti dai svariati sistemi informativi dell'azienda per immagazzinarli in un formato consistente all'interno di un unico sistema, mentre solo grazie ai tool di business intelligence riesce a dare quei vantaggi strategici appropriati con l'analisi dei dati e con l'estrazione di informazione da questi.

In conclusione possiamo affermare che il data warehouse con l'applicazione dei tool di business intelligence apporta all'azienda tali benefici:

- Aumento del reddito;
- Diminuzione dei costi;
- Aumento della soddisfazione dei clienti e degli impiegati;
- Possibilità di anticipare le richieste del mercato.

1.3. Tipologie di sistemi: OLTP e OLAP

All'interno dell'azienda si trovano, da un lato i sistemi informativi che sovrintendono all'operatività giornaliera dell'azienda (OLTP), dall'altro quelli che soddisfano il bisogno di informazione sull'andamento generale dell'azienda attraverso l'estrazione, rielaborazione e la rappresentazione dei dati (OLAP).

I sistemi OLTP (On Line Transaction Processing) si occupano della registrazione, modifica e visualizzazione dei singoli record in tempo reale a seconda delle esigenze operative delle varie funzioni aziendali. Nei database OLTP i campi che descrivono i dati di interesse sono distribuiti su più tabelle per ragioni di velocità e semplicità di aggiornamento e inserimento. Al loro interno è possibile trovare i dati

storici relativi alla dimensione temporale, ma in genere questi sistemi non conservano dati storici sulle passate gestioni. Molte volte si possono recuperare informazioni di tipo storico attraverso una serie di operazioni di “restore” da analisi di backup.

I sistemi OLAP (On Line Analytical Processing) permettono di operare on-line sui dati del data warehouse e danno la possibilità all’utente di realizzare in maniera del tutto semplice, veloce ed efficace indagini a supporto del proprio processo decisionale. Queste applicazioni lasciano l’utente libero di navigare tra i dati senza costringerlo ad eseguire delle query rigidamente pre-definite come avviene nelle elaborazioni di tipo tradizionale, in cui i problemi sono perfettamente strutturati. In definitiva questi sistemi consentono di:

- supportare le complesse analisi di un processo decisionale;
- analizzare i dati da varie prospettive, ovvero dimensioni;
- agire dinamicamente su enormi quantità di dati.

Le applicazioni OLAP utilizzano modelli multidimensionali di archiviazione dei dati che portano ad una concettualizzazione dei dati molto vicina al modo in cui il manager percepisce la realtà aziendale. Ogni dimensione raccoglie quelli che saranno i criteri in base ai quali si potranno utilizzare i dati nel data warehouse, ad esempio potrà esserci la dimensione geografica oppure quella temporale o di tipologia del prodotto. Queste dimensioni vengono poi rappresentate con una struttura ipercubica con n-lati, la cui intersezione individua una o più misure corrispondenti.

Quindi le differenze tra i due sistemi si possono riassumere in questa tabella:

	OLTP	OLAP
Funzione	gestione giornaliera	supporto alle decisioni
Progettazione	orientata alle applicazioni	orientata al soggetto
Frequenza	giornaliera	sporadica

Dati	recenti, dettagliati	storici, riassuntivi, multidimensionali
Sorgente	singola DB	DB multiple
Uso	ripetitivo	ad hoc
accesso	read / write	read
Flessibilità di accesso	uso di programmi precompilati	generatori di query
Numero di record acceduti	decine	migliaia
Tipo utenti	operatori	manager
Numero utenti	migliaia	centinaia
Tipo db	singola	multiple eterogenee
Performance	alta	bassa
Dimensione DB	100 MG-GB	100 GB-TB

Tabella [DIADE 2000, *Modelli e strutture per il Data Warehousing*]

1.4. Analisi Multidimensionale

Il successo di un data warehouse è strettamente legato alla possibilità che viene data agli utenti di accedere facilmente ai dati sintetici e analitici e di poterli effettivamente analizzare. Il data warehouse deve sapere rispondere alle esigenze dell'utente e comprendere il modo in cui essi potrebbero servirsi dei tool di

business intelligence per analizzare il significato nascosto dei dati e ricavarne i risultati dall'analisi.

Per un'efficace analisi di dati, dunque, è importante tenere in considerazione il modo in cui i manager utilizzano le informazioni per gestire il loro business.

- **I manager pensano in modo multidimensionale**, ad esempio considerano le vendite per semestre per venditore, costi per marca di articoli per regione, ecc. Succede che le informazioni da utilizzare provengano dalle intersezioni di molte variabili come il tempo, fornitore, articoli e via dicendo. Questo processo mentale è ben supportato dall'esplorazione dei dati a livello aggregato, ovvero di informazioni mostrate nel contesto di business. Per venire incontro a questa necessità, i dati possono essere sistemati in strutture multidimensionali e presenti all'utente mediante dei tool che consentono di esplorare le differenti relazioni. Le strutture multidimensionali sono caratterizzate da tre a più dimensioni e questo consente di rappresentare i dati, non più tramite tabelle, ma con cubi multidimensionali, dove ogni dimensione può rappresentare una variabile di business a cui l'utente è interessato. Ad esempio, le dimensioni scelte potrebbero essere le vendite per regione, per articolo, per semestre. Concettualmente, sarebbe possibile ingrandire il cubo aggiungendogli dimensioni, in modo da includere differenti aree geografiche, caratteristiche dei fornitori, ecc. In tal modo l'utente attraverso operazioni di "slice and dice" può ottenere i risultati che sta cercando, scegliere di analizzare più di due dimensioni contemporaneamente, e quindi esaminare l'impatto di variabili diverse sui risultati di business.
- **I manager richiedono l'accesso a dati dettagliati** nei particolari casi in cui si verificano delle anomalie di cui si vogliono trovare le cause, oppure se l'analisi multidimensionale aggregata segnala una tendenza specifica che necessita di un'ulteriore indagine.

Queste due richieste da parte dei manager possono essere effettivamente soddisfatte tramite l'approccio, su cui si basa il data warehouse, di tipo "Analyze-then-Query" realizzabile tramite applicazioni di tipo OLAP e che consiste in un metodo efficace per realizzare i dati del warehouse raggiungendo dei risultati di alto livello. Questo modo sostituisce il più tradizionale "Query first, Analyze later", poiché quest'ultimo tipo di approccio rappresentava limitate capacità di analisi, lunghe attese e una difficile navigazione tra i dati.

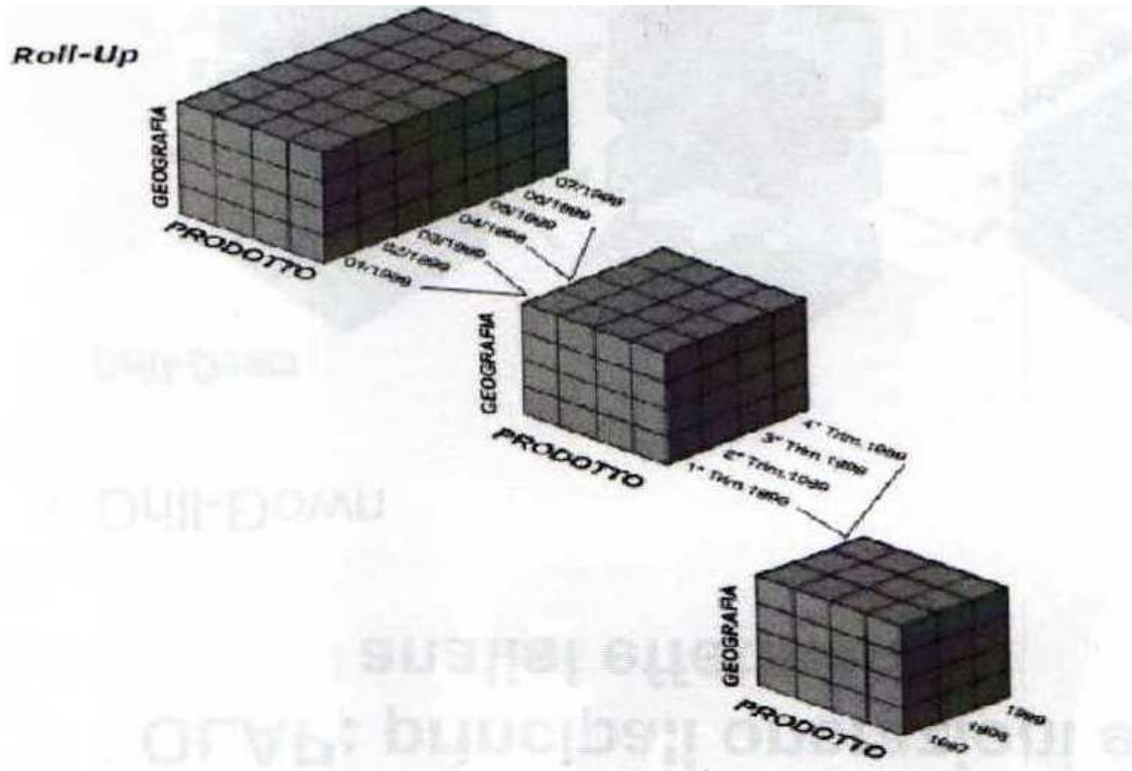
Le applicazioni OLAP estraggono i dati prescelti dai sistemi operazionali per poi costruire delle strutture, contenute in "**sweet spots**", di informazioni selezionate appositamente per agevolare i processi decisionali dei manager. In questo modo le informazioni di maggior interesse a livello manageriale vengono concentrate in un numero ristretto di sweet spots di informazioni così da ridurre tempi di analisi consentendo di incrementare la produzione, ridurre i costi e orientare politiche aziendali maggiormente mirate all'esigenza del cliente.

1.5. OLAP

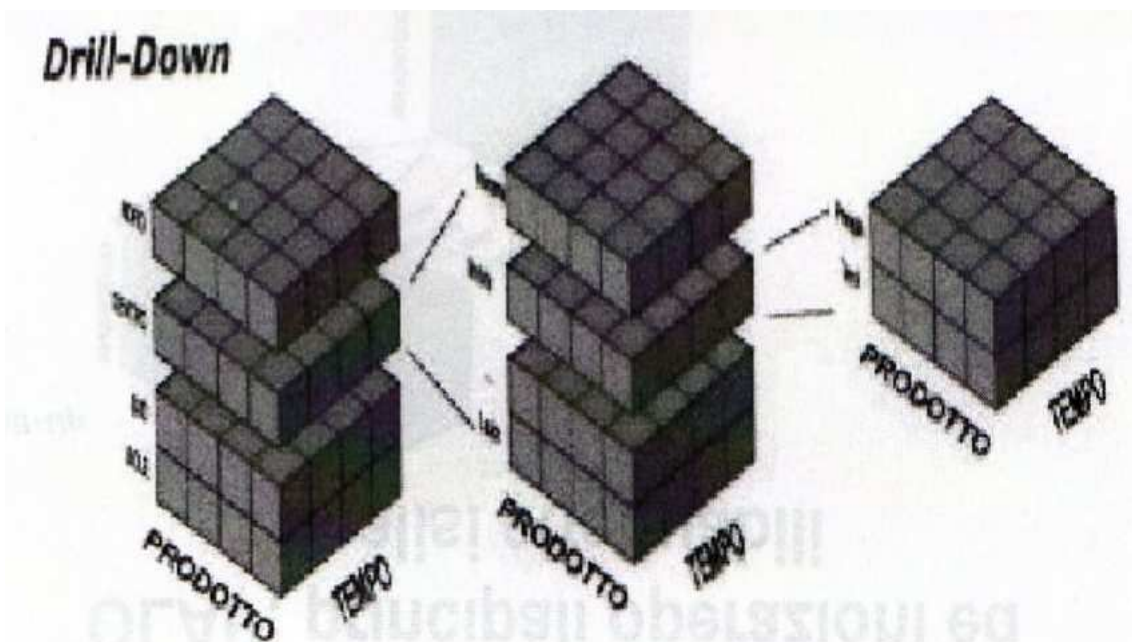
È forse la principale modalità di fruizione delle informazioni contenute in un data warehouse e consente a utenti le cui necessità di analisi non siano facilmente identificabili a priori di analizzare ed esplorare interattivamente i dati sulla base del modello multidimensionale. Una sessione OLAP consiste in pratica in un *percorso di navigazione* che riflette il procedimento di analisi di uno o più fatti di interesse sotto diversi aspetti e diversi livelli di dettaglio.

Ogni passo della sessione di analisi è scandito dall'applicazione di un operatore OLAP che trasforma l'ultima interrogazione formulata in una nuova interrogazione. Gli operatori più comuni sono:

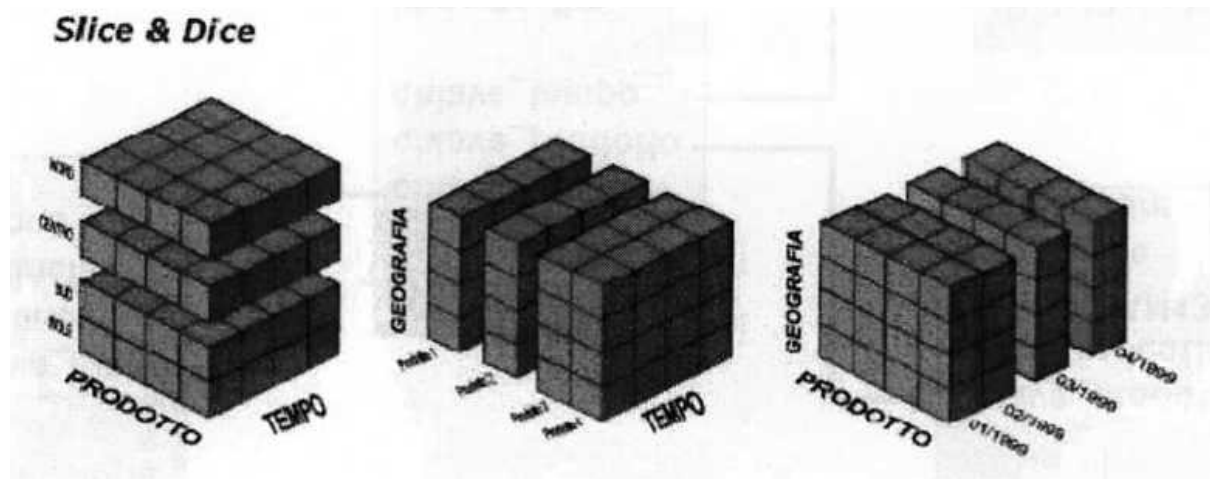
1. Roll-up: letteralmente significa arrotolare o alzare, e induce un aumento nell'aggregazione dei dati eliminando un livello di dettaglio da una gerarchia.



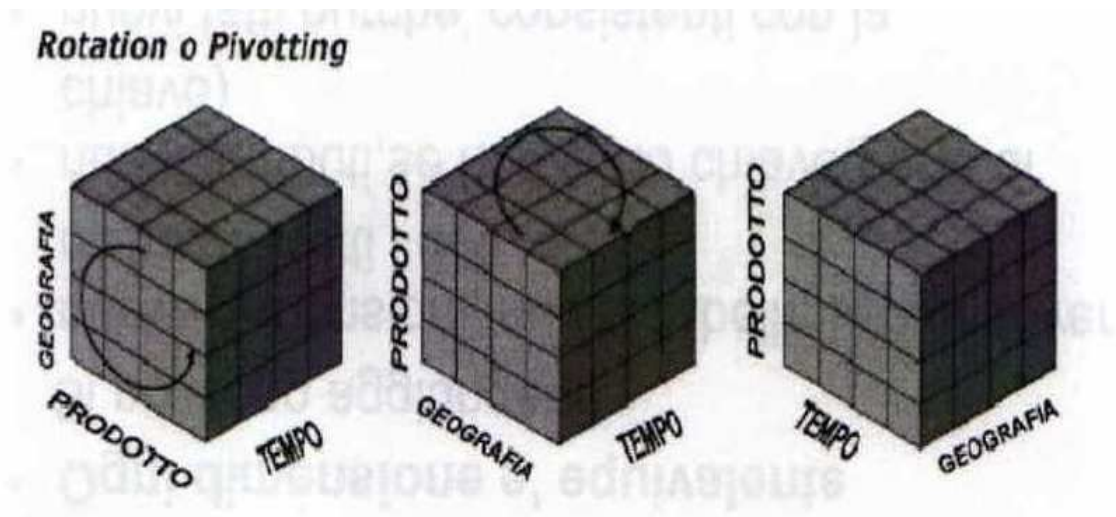
2. Drill-down : è il duale al roll-up; infatti esso diminuisce l'aggregazione dei dati introducendo un ulteriore livello di dettaglio.



3. Slice-and-dice : tagliare a fette, riduce la dimensionalità del cubo fissando un valore per una delle dimensioni, selezione o filtraggio l'operazione che riduce l'insieme dei dati in oggetto di analisi attraverso la formulazione di un criterio di selezione.



4. Rotation : questa operazione consente di scambiare gli assi di visualizzazione delle dimensioni: è possibile spostare una dimensione dall'asse X di un'ipotetica visione cartesiana, all'asse Y o Z.



1.6. Distinzione fra fatti e dimensioni

Definiamo “*FATTI*” le variabili o misure, tipicamente di natura numerica, che sono al centro delle analisi del sistema di Business Intelligence. Sono gli elementi che rispondono alla domanda “di cosa sto parlando”? quando si legge un report o si progettano delle analisi. Esempi di fatti possono essere le vendite, fatturato, il budget, approvvigionamenti oppure la quota di mercato. Naturalmente è importante che il fatto di interesse abbia una evoluzione nel tempo.

Se ad esempio consideriamo le vendite, è importante verificare quali informazioni sono disponibili ed a quale dettaglio: i dati sulle vendite non esistono “di per sè”, ma in un contesto determinato del prodotto venduto, dal negozio dove avviene la vendita, dal momento in cui questa avviene e da qualsiasi altro tipo di informazione relativa all’evento. Queste informazioni vengono raggruppate in “*DIMENSIONI*” che raccolgono quelli che saranno i criteri in base ai quali potremmo analizzare i dati nel data warehouse; ritornando all’esempio delle vendite, potremmo avere una dimensione geografica, una tipologia di prodotto e una temporale (ricordiamo che il tempo è la caratteristica fondamentale del DW).

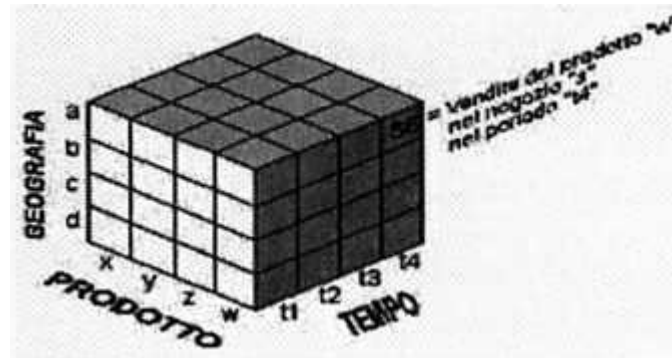
Per riuscire a visualizzare questa struttura prendiamo spunto dal modello multidimensionale, disegnando un cubo che rappresenta il fatto Vendite qualificato dalle dimensioni Geografia, Prodotto e Tempo. Quindi vediamo che ogni cella conterrà il dato delle vendite di un prodotto in un determinato negozio e in un particolare periodo.

In particolare perciò, in un ambiente di data warehouse, è possibile individuare tre tipi di misure:

Misure di flusso : si riferiscono ad un periodo, al cui vengono valutate in modo cumulativo (Numero di prodotti venduti in un giorno, incasso mensile, ecc);

Misure di livello : vengono valutate in particolari istanti di tempo, rappresentano istantanee di specifici aspetti del dominio applicativo (Numero di prodotti in magazzino, numero di degenti ricoverati in un dato giorno, ecc);

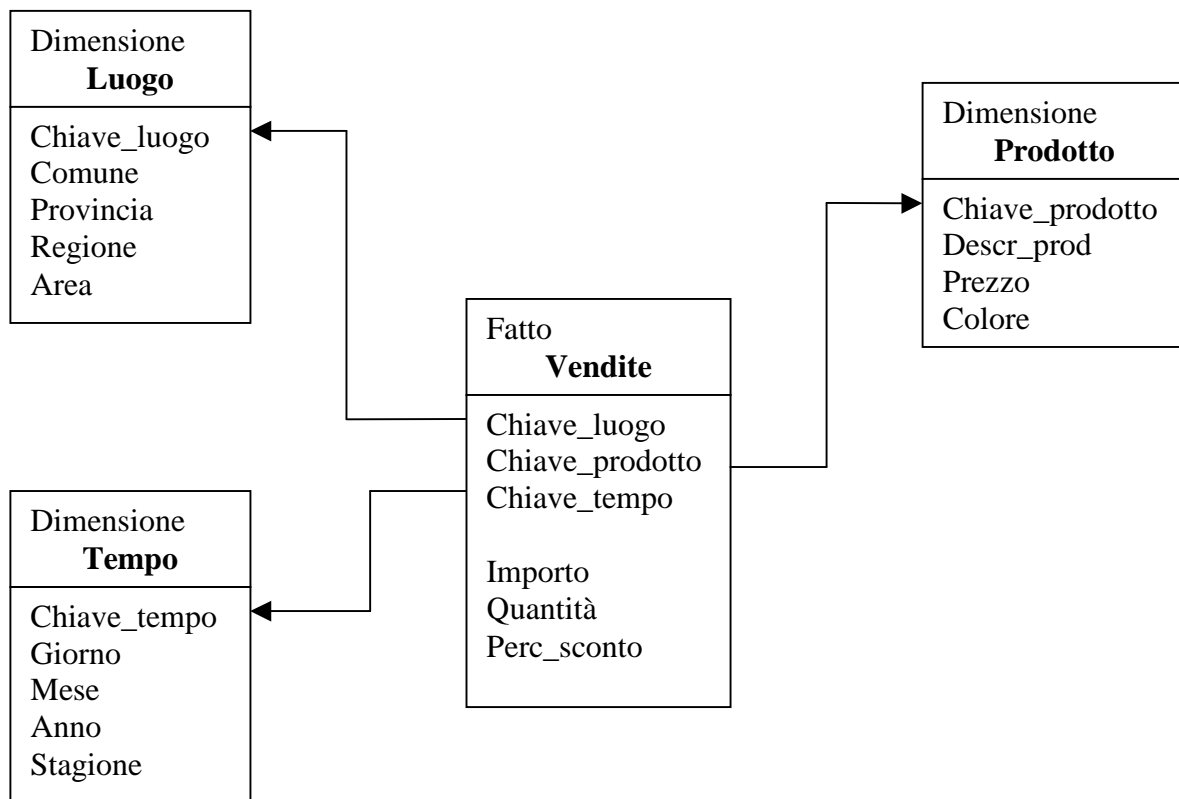
Misure unitarie : vengono valutate in particolari istanti di tempo, ma sono espresse in termini relativi (Prezzo unitario di prodotto, percentuale di sconto, ecc);



1.7. Schema a stella per la memorizzazione dei dati

Lo schema a stella, per la memorizzazione dei dati nel data warehouse, deriva direttamente dal modello concettuale per l'OLAP essendo organizzato in due tipi di tabelle: una tabella dei fatti e una o più tabelle delle dimensioni.

Come riportato dalla figura sottostante, la tabella dei fatti conterrà solo due tipi di elementi: una o più chiavi esterne e delle misure riguardanti il fatto di interesse. Queste misure rappresentano i dati numerici, da sommare lungo una o più dimensioni, dai quali ottenere report e sommatorie.



Modello a stella

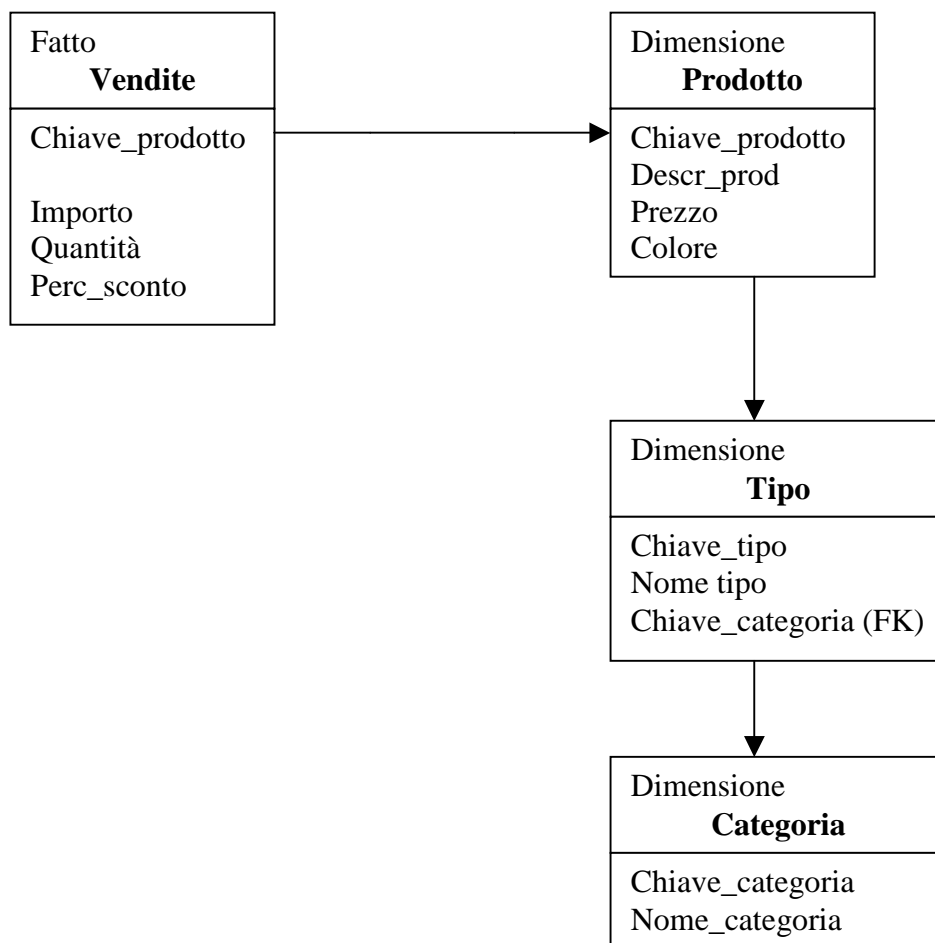
Una tabella delle dimensioni conterrà, invece, una o più chiavi primarie e delle informazioni descrittive sulla dimensione. In queste tabelle sia le informazioni descrittive sia la chiave primaria contenente i dati testuali relativi alla dimensione saranno utilizzati per individuare i criteri di estrazione da applicare alla tabella dei fatti e per fornire informazioni descrittive nei report. Lo schema a stella implica solitamente una relazione uno-a-molti tra le tabelle delle dimensioni e la tabella dei fatti; quindi le tabelle delle dimensioni implementano una memorizzazione denormalizzata degli attributi relativi alla dimensione, mentre la tabella dei fatti rappresenta una vista normalizzata dell'entità dell'oggetto preso in analisi.

1.8. Schema a fiocco di neve

Vi sono situazioni in cui l'uso dello schema a stella non è la scelta ideale. La ragione principale è essenzialmente dovuta a due motivi:

- Lo schema a stella denormalizzato può richiedere troppo spazio su disco;
- Tabelle delle dimensioni troppo grandi possono degradare le prestazioni, andando a cancellare i benefici dell'aggregazione.

La normalizzazione delle dimensioni permette un notevole risparmio di spazio occupato. Un aumento delle prestazioni attraverso l'uso di uno schema a fiocco di neve (o Snow Flakes) è possibile nel caso si voglia memorizzare gli aggregati in differenti tabelle dei fatti. In questo caso un aggregato per categoria di prodotto avrà come dimensione la tabella normalizzata CATEGORIA e non come avveniva nello schema a stella l'intera tabella PRODOTTO.



Schema a fiocco di neve

Supponiamo ad esempio che ci venga posta una query che chieda le vendite totali per categoria di prodotto e di avere a disposizione una tabella aggregata per categoria di prodotto. Supponiamo inoltre che vi siano 100.000 prodotti raggruppabili in 15 categorie.

Nel caso di uno schema a stella viene eseguito un join tra la tabella dei fatti aggregata e l'intera tabella della dimensione dei prodotti (100.000 righe).

Nel caso di uno schema a fiocco di neve, invece, viene eseguito direttamente il join con la tabella normalizzata contenente le informazioni della categoria (15 righe) con un notevole aumento delle prestazioni.

Quest'ultimo modello, però, offre anche degli svantaggi. Il più evidente è la notevole complessità del modello, che è punto di forza dello schema a stella.

CAPITOLO 2 : SAP R/3

2.1. L'azienda SAP R/3

SAP (Sistem Application & Product in Data Processing) è stata fondata nel 1972 in Germania da cinque persone ed è oggi uno dei principali produttori mondiali di software con oltre 19.000 dipendenti. L'azienda, che ha sede a Walldorf, è quotata in borsa ed è presente con proprie strutture nei principali Paesi Europei, nelle Americhe, Asia, Sudafrica e Giappone.

SAP entra in Italia agli inizi degli anni '90 e in questo periodo vengono realizzate le prime implementazioni dei prodotti della società tedesca. Otto anni dopo, nel 1998, grazie al successo riscontrato viene costituita SAP Italia, la filiale italiana che impiega 220 persone nelle sedi di Agrate (MI), Roma, Padova e Bologna.

2.2. Le origini e l'evoluzione di SAP R/3

Inizialmente SAP acquista i diritti di utilizzo di un package di compatibilità finanziaria per poter sviluppare un sistema standard, impiegabile in differenti realtà, in logica realtime. A questo primo pacchetto viene presto affiancato un software per la gestione dei materiali, l'acquisizione delle materie prime, la gestione del magazzino e la fatturazione.

È seguita la realizzazione di un terzo modulo per coprire l'area della gestione degli investimenti e contemporaneamente l'introduzione e la diffusione sul mercato della tecnologia mainframe. A questo punto, nel 1978, i tre pacchetti vengono integrati facendo così nascere il gestionale dalla sigla R/2.

La successiva importante innovazione tecnologica che determina un cambiamento decisivo per SAP è rappresentata dalla nascita della soluzione client server.

Nel 1992 viene rilasciato R/3, la soluzione gestionale standard per le nuove architetture informatiche, che consentono alle imprese di sfruttare le potenzialità e i

vantaggi dell'innovativa logica di elaborazione distributiva. Il sistema R/3 nasce come software aperto, in grado di supportare tutte le piattaforme hardware e i sistemi operativi più diffusi (Microsoft Windows NT, UNIX, AS/400 e OS/390), con un'ampia gamma di moduli specifici per tutte le aree gestionali dell'impresa. Con l'obiettivo di aprire il sistema alle applicazioni esterne, nel 1996 l'azienda introduce la Business Framework Architecture, una nuova architettura che fornisce un ambiente standard e integrato nel quale possono interagire i componenti software di R/3 e quelli sviluppati da terze parti. Grazie alla sua flessibilità e modularità, il pacchetto SAP può essere installato in una molteplicità di settori di mercato e in aziende con diverse dimensioni.

2.3. Business Engineering

Le aziende di tutto il mondo sfruttano l'information technology (IT) per modificare radicalmente il modo di condurre la loro attività. Nel passato, IT veniva usata specialmente per automatizzare le funzioni aziendali già esistenti, ma ora questa può migliorare o cambiare completamente le operazioni di un'azienda.

Questo approccio viene chiamato Business Engineering (BE) e utilizza appunto l'information technology per supportare la riprogettazione dei processi di business in modo da migliorare e accelerare i processi di output ottimizzando quindi i flussi di produzione. Utilizzando la BE per la progettazione di intere catene di processi, le aziende possono meglio integrare tutte le loro attività critiche ed avere dei vantaggi competitivi sul mercato grazie a dei costi più bassi, a tempi di consegna più brevi e maggiore sensibilità nei confronti dei clienti.

La Business Engineering è sorta dal bisogno di sistemi software che possano supportare adeguatamente i sempre più frequenti cambiamenti dei processi di business. Una delle case venditrici di software che ha pensato proprio a questa esigenza di trovare la migliore strategia d'implementazione di nuovi progetti è

stata SAP, società che ha integrato con successo l'information technology con la Business Engineering.

2.4. SAP: ERP e Data Warehouse

SAP R/3 può essere rappresentato come leader dei sistemi ERP (Enterprise Resource Planning), ha al suo interno un insieme di applicazioni in cui sono unificate, sotto un'unica struttura applicativa e su un'unica interfaccia, tutte le divisioni aziendali.

Essendo un sistema ERP, ha la capacità di gestire mediante un unico processo informatico, che viaggia in rete, strutture separate generando grandi quantità di dati integrati. Gli ERP sono sistemi di tipo operativo e transazionale, orientati alle attività di back-office e costituiti da suite di applicazioni che consentono il reengineering, automatizzare, integrare e semplificare i processi aziendali di base.

Tutti questi dati generati però, sono utilissimi dal punto di vista strettamente operativo, devono essere trasformati in informazioni utili per il supporto decisionale del manager, in modo da fornire all'azienda un vantaggio competitivo a lungo termine. È per questo motivo che i dati transazionali (dati di tipo statico organizzati in file sequenziali), provenienti da sistemi ERP devono essere trasformati e integrati verso un sistema di risorse centralizzato, il data warehouse .

Una caratteristica fondamentale del data warehouse è quella della "scalabilità", cioè la capacità di ampliarsi in modo parallelo all'incremento dei dati disponibili. In questo modo i dati andranno raccolti e normalizzati per essere poi elaborati e conservati in prospettiva di un'analisi storica. L'introduzione dei pacchetti ERP ha agevolato dunque la diffusione delle tecniche di data analysis e business intelligence.

Per rispondere alle crescenti esigenze informative aziendali, SAP supporta una forma di multidimensional technology denominata: Infocubi. Essi hanno le caratteristiche dei già accennati cubi multidimensionali, che permettono di

raggruppare gli aspetti di interesse aziendale in dimensioni con le quali, eseguendo un'intersezione, sarà possibile fare analisi sui dati.

L'obiettivo degli infocubi è proprio quello di consentire al manager aziendale di accedere ai dati in modo da rendere facile e veloce l'analisi più di quanto fosse possibile ad un livello di transazione.

2.5. Architettura CLIENT-SERVER

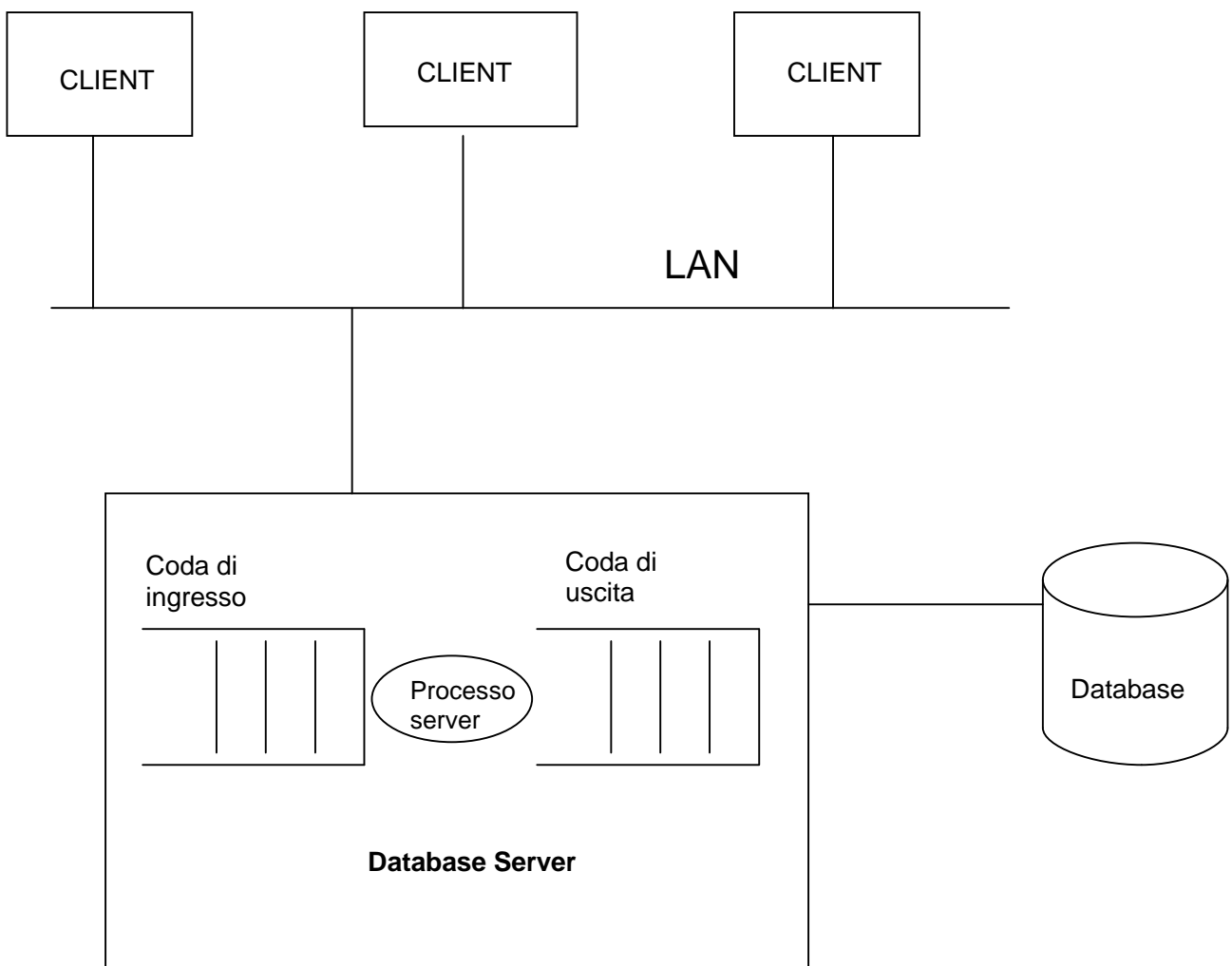
La maggior parte delle applicazioni R/3 sono progettate per un'architettura client-server a tre livelli, collegati da un middleware (software e motore che consentono di collegare i server con i client) che gestisce il processo di distribuzione dei flussi informativi all'interno del sistema e permettere a SAP di creare un ambiente distributivo. Con la tecnologia client-server, il software può essere sia centralizzato in un unico computer sia decentralizzato in molti. Il software client-server è abbastanza flessibile da essere collegato con delle linee telefoniche, sullo stesso computer, utilizzando reti locali o combinazione di queste. Di solito è basato su uno schema a tre livelli che permette al sistema di essere distribuito tra i PC e le workstation. Questo schema permette di fare una separazione funzionale tra i front-end presentation server e i back-end database server, con l'application server operativa nel mezzo.

I tre livelli sono dunque:

- *Database server*: elaboratore centrale che gestisce tutte le funzioni di data base come aggiornamento, cancellazione, modifica e così via. A questo livello il database SAP interagisce con un DBMS (Data Base Management System) che archivia tutti i dati sul disco.
- *Application server*: è collegato al database server e contiene la logica applicativa dei moduli che compongono il sistema R/3. In questo livello dunque i programmi vengono configurati ed eseguiti e i dati elaborati.

Tipicamente l'application server è dedicato ad un gruppo numeroso di utenti, come un reparto. Può contenere software sviluppato all'esterno di R/3 se questo linguaggio è ABAP/4.

- *Presentation server*: sono workstation, personal computer e terminali collegati all'application server utilizzati per la presentazione dei dati dall'utente finale, con interfaccia grafica. Il software include un user-friendly GUI, chiamato SAPGUI (SAP Graphical User Interface), che preleva le richieste dell'utente e le passa all'application server. I sistemi operativi utilizzati per i desktop sono Windows, Macintosh, OSF/Motif, OS/2 Presentation Manager Java.



L'approccio a tre livelli per il sistema client-server è favorito sia dalla robustezza dei computer locali sia dalla potenza computazionale del sistema centrale. Il risultato è un sistema integrato per le procedure e le transazioni aziendali.

2.6. Benefici della tecnologia CLIENT-SERVER

La struttura a tre livelli usata dalla tecnologia client-server crea un ambiente semplice e flessibile che apporta alle compagnie e ai singoli utenti molti benefici quali l'aumento di potere e controllo di ogni aspetto dell'architettura software.

Usando un'architettura client-server il sistema può ottimizzare le risorse informative aziendali, trarre vantaggio dalle nuove tecnologia hardware e adattare le proprie procedure di business per soddisfare esigenze specifiche.

La distribuzione su tre livelli è spesso la migliore per le grandi strutture aziendali decentralizzate. Questa soluzione non solo si applica al presentation level, dove i personal computer possono essere aggiunti semplicemente se il numero degli utenti aumenta, ma si applica anche all'application level: i dedicated computers possono essere aggiunti per le analisi di fondo o per le specifiche aree di applicazione.

Questa applicazione permette, inoltre, di nascondere ai suoi utenti finali il complesso e sofisticato sistema aziendale; i dipendenti possono dare in modo semplice le informazioni che vengono loro richieste dai clienti senza dover conoscere la struttura tecnica del sistema o dover usare difficili comandi di programmazione. In questo modo gli utenti lavorano in modo più veloce e necessitano di minor conoscenza tecnica. Questo sistema permette a molte persone situate in posti molto diversi di vedere e lavorare con le stesse informazioni, condividere files e opinioni, tutto in tempo reale, in modo da facilitare la collaborazione e conseguire gli stessi obiettivi. Per esempio, permette ai manager finanziari situati in differenti posti di conseguire velocemente a dei dati finanziari, prelevando processi e dati di interesse manageriale.

Si tratta di un sistema che utilizza un processo di elaborazione dei dati integrato in tempo reale per il quale l'aggiornamento istantaneo del database poggia sul Repository. Il Repository di R/3 è un database ad oggetti particolarmente specializzato nella definizione, memorizzazione, accesso alle informazioni e interscambio dei dati, dal quale vengono estratte tutte le applicazioni che poi vengono fatte passare all'application server e successivamente, quando necessario, al presentation server.

La flessibilità dell'architettura software client-server permette alle compagnie di salvaguardare i loro investimenti in hardware e in software poichè riescono a fronteggiare adeguatamente le continue innovazioni tecnologiche. Questo software permette inoltre di utilizzare ogni funzione aziendale significativa e percorre tutte le principali piattaforme hardware consentendo alle aziende di poter scegliere l'hardware che meglio si adatta ai bisogni, alle esigenze e alle prestazioni delle singole divisioni e reparti. Questa flessibilità consente di massimizzare la produttività di ogni personal computer, workstation e della classe di macchine mainframe, allevia i problemi quali la distanza fisica, la sicurezza dei dipendenti all'estero, e gli altri costi di investimenti oltre oceano; rende possibile, inoltre, alle compagnie di muoversi velocemente in nuovi mercati interni o esterni.

2.7. ABAP/4

Il sistema SAP R/3 è un pacchetto avanzato di elaborazione dati realizzato in ABAP/4 (Advanced Business Application Programming), un linguaggio di quarta generazione sviluppato interamente dalla società distributrice del software. Le applicazioni utilizzano inoltre l'ABAP/4 Development Workbench, un ambiente di sviluppo che consente sia la semplice predisposizione dei report ad hoc, sia lo sviluppo di moduli specifici in grado di soddisfare funzionalità non supportate dalla versione standard dell'applicativo.

L'interfaccia database tratta l'accesso ai dati usando SAP SQL, che è essenzialmente una libreria dell' SQL standard resa disponibile all'interno dell'ambiente ABAP/4. L'architettura R/3 garantisce che le differenze sintattiche e semantiche delle implementazioni SQL dei vari venditori di database rimangano all'interno di pochi moduli. In questo modo R/3 può supportare tutti i sistemi di database relazionale che si sono dimostrati adeguati. La strutturazione dei dati, infatti, utilizza un sistema di tabelle (matrici che descrivono le relazioni fra insiemi di dati), domini (tipo, formato, lunghezza e valori ammessi per determinati campi delle tabelle) e campi (sottoinsiemi di tabelle), permettendo così di ottimizzare la memorizzazione dei dati e tempi di risposta.

2.8. Struttura modulare

SAP R/3 è un sistema integrato per la gestione dei processi di business e quindi tutte le procedure riferite ad un evento sono correlate tra loro in modo da creare delle catene di attività, di sub-processi e infine di processi abbandonando la logica di divisione del lavoro e specializzazione funzionale.

Le aree gestionali di SAP, che ricoprono particolari settori aziendali, sono strutture a moduli, i quali sono vendibili anche singolarmente ma non esauriscono al proprio interno i processi a supporto dei quali sono stati sviluppati. Questa modularità permette alle aziende di scegliere quali moduli installare ed implementare sfruttando le funzionalità offerte dai singoli moduli e allo stesso tempo riducendo i costi di hardware, di implementazione e di formazione. Le implementazioni parziali del prodotto devono essere opportunamente interfacciate con sistemi legacy.

2.9. Gestione della sicurezza

In un sistema complesso come R/3, è particolarmente importante proteggere i dati in modo adeguato attraverso l'utilizzo di sistemi di autorizzazione.

Tecnicamente, nell'ambito del record anagrafico di ogni singolo utente viene definito nel dettaglio l'accesso a:

1. transazioni a video;
2. elementi organizzativi del sistema;
3. tabelle e oggetti di dati.

Ogni utente ha un suo profilo che raggruppa l'ampio insieme di autorizzazioni gestite dal sistema di autorizzazioni SAP. Tutte le autorizzazioni utente vengono gestite centralmente e ogni modifica ha effetto immediato per tutti i collaboratori.

Per qualsiasi tipo di problema SAP dispone di un help online, richiamabile in qualsiasi momento.

2.10. Implementazione: ASAP

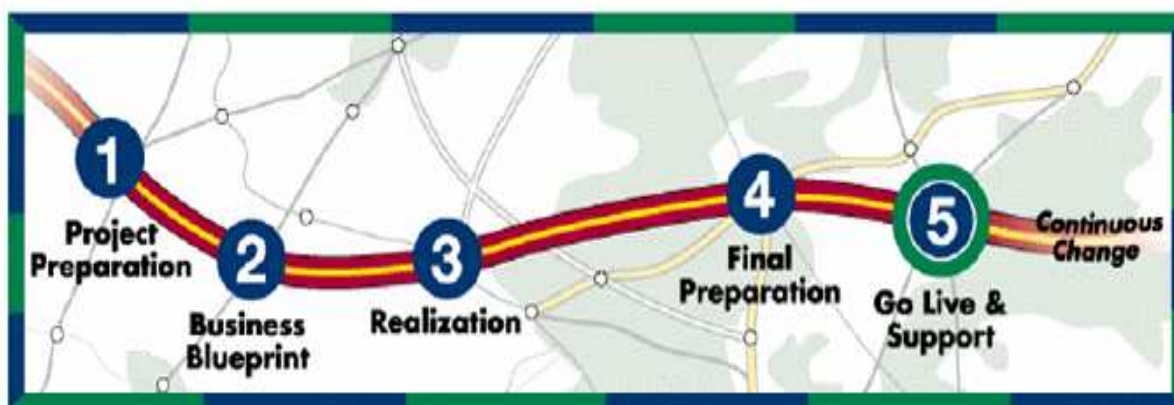
ASAP (Accelerated SAP) è la metodologia di implementazione del package SAP R/3 sviluppata da SAP e dai suoi consulenti. Essa si compone di cinque fasi principali:

- 1 - *Preparazione*: durante questa fase vengono definiti gli obiettivi del progetto sia in termini informativi che tecnologici. Dopo aver individuato la struttura delle funzioni e dei processi aziendali interessati dall'intervento, si individuano i relativi fabbisogni informativi necessari al raggiungimento degli obiettivi definiti. È importante poi identificare, dal punto di vista tecnologico, le caratteristiche dell'infrastruttura che l'azienda intende adottare a supporto del nuovo applicativo.

- 2 - *Blueprint*: in questa fase si effettua l'analisi funzionale; viene condotta un'analisi del sistema R/3 in modo da conoscere bene le caratteristiche operative e derivarne un modello di configurazione capace di supportare le esigenze procedurali e funzionali del cliente. Viene inoltre predisposto il sistema di sviluppo hardware e software, si procede successivamente al disegno dell'architettura di rete necessaria a supportare l'utilizzo di R/3 all'interno dell'organizzazione ed eventualmente verso l'esterno (Es: internet).
- 3 - *Realizzazione*: in questa fase si esegue la configurazione del prodotto seguendo le esigenze specificate dall'utente; essa può essere scomposta in tre tipologie di attività:
- Realizzazione del prototipo, in cui si definisce un prototipo funzionale del nuovo sistema informativo e i processi che questo supporta.
 - Parametrizzazione, in cui vengono impostati tutti i parametri di R/3 che consentono l'utilizzo del sistema secondo le modalità definite in fase di disegno concettuale e validate con il prototipo.
 - Integrazione / personalizzazioni / interfacce / conversioni, in cui viene analizzato l'impatto delle integrazioni, delle personalizzazioni ritenute inevitabili e delle interfacce individuate nella fase di definizione del prototipo specifico; si definiscono quali requisiti dei report che il sistema deve generare e come effettuare le conversioni dei dati ritenute necessarie.
- 4 - *Passaggio in produzione*: a questa fase si attribuisce l'implementazione ovvero l'avviamento del sistema. In essa vi si compongono le seguenti attività:
- Installazione del sistema;
 - Formazione degli utenti;
 - Collaudo.

5. – *Presidio*: In quest'ultima fase si racchiudono le seguenti attività:

- Messa in esercizio, e quindi avviamento del sistema;
- Assistenza nelle fasi di “chiusura” delle procedure;
- Misurazione e valutazione delle performance di sistema;
- Dove necessario, effettuazione di interventi correttivi.



CAPITOLO 3 : SAP BW

3.1. Business Information Warehouse SAP BW

In un mercato caratterizzato da una forte competizione e rapidi cambiamenti, nasce l'esigenza di prendere decisioni in maniera rapida e sulla base delle analisi di fatti opportunamente interpretati. La soluzione a questa esigenza può identificarsi con l'implementazione di SAP BW, il quale rappresenta il fulcro della Business Intelligence nella nuova suite di prodotti SAP, mirata a gestire la più completa integrazione tra componenti e processi per il supporto alle decisioni.

L'utilizzo di tale strumento permette, attraverso l'analisi dei dati, la diffusione dell'informazione sull'andamento del business tra "decision marker" nei tempi e modi necessari per rispondere in maniera competitiva alle sfide di mercato.

Per fare questo, il sistema BW, estrae i dati dal sistema R/3, o da altre sorgenti, e li mantiene in un data warehouse dove sono organizzati e gestiti per mezzo di un'interfaccia grafica garantendo la massima accessibilità e facilità di gestione.

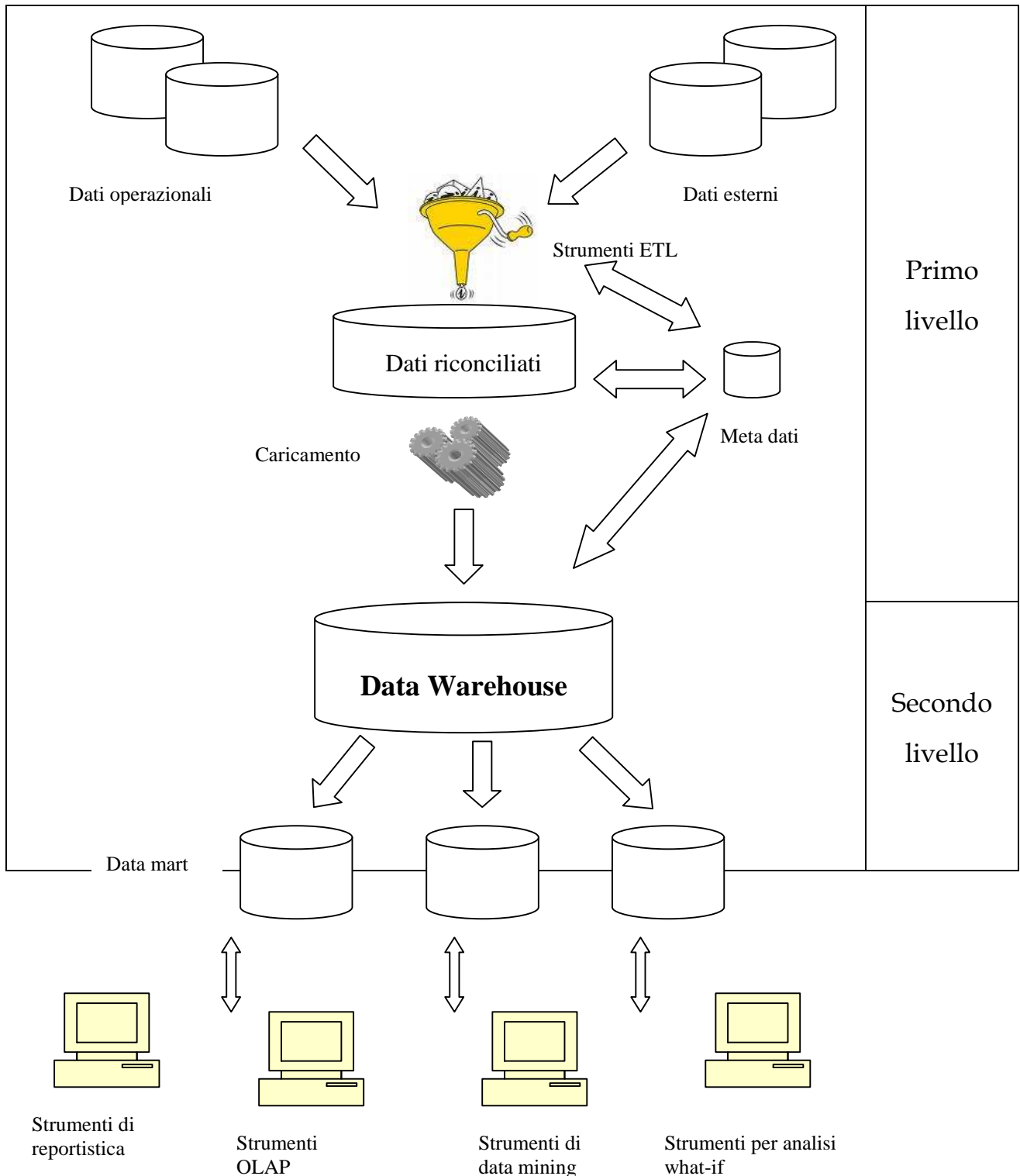
Nonostante BW sia un ambiente separato da quello operativo, con una base di informazioni dedicata, esso è perfettamente integrato con il sistema SAP R/3 rispettando i vincoli di sicurezza e di integrità permettendo un accesso intelligente e guidato non solo ai dati, ma soprattutto ai suoi metadati.

Per le piccole medie imprese SAP ha predisposto una serie di regole e task preconfigurati, i *Business Content*, che consistono in un numero pari a 100 infocubi predefiniti, ossia visioni multidimensionali specifiche per determinare analisi, completi di query e report preimpostati. Questa soluzione fornisce delle soluzioni per le più frequenti esigenze di business, senza affrontare i tempi e i costi di analisi e sviluppo di una soluzione personalizzata.

L'architettura di SAP BW può essere schematizzata in tre livelli:

- Nel primo livello avviene l'estrazione dei dati tramite sistemi OLTP;

- Nel secondo livello troviamo il Business Information Warehouse Server dove avviene la gestione e l'elaborazione dei dati grazie a sistemi OLAP;
- Nel terzo livello ha invece le applicazioni di frontend col Business Explorer che consente la visualizzazione dei dati su fogli excel.



	Terzo livello
--	------------------

Per soddisfare le diverse necessità informative, SAP BW si serve di tre strutture concettuali contenute nel server: un operational data store (ODS), un data warehouse, e dei modelli multidimensionali, che insieme vanno a costituire il Sap BW informational model.

L' ODS è anche detto *livello dei dati riconciliati* che materializza i dati operazionali ottenuti a valle del processo di integrazione e ripulitura dei dati sorgente: quindi dati integrati, consistenti, corretti, volatili e dettagliati. Come illustrato dalla figura sovrastante, il DW viene allora alimentato non più direttamente dalle sorgenti, ma piuttosto dai dati riconciliati.

Il vantaggio principale di avere i dati riconciliati è che esso crea un modello di dati comune e di riferimento per l'intera azienda, introducendo al contempo una separazione netta tra le problematiche legate all'estrazione e integrazione dei dati dalle sorgenti e quelle inerenti all'alimentazione del DW. D'altro canto, i dati riconciliati introducono un'ulteriore ridondanza rispetto ai dati operazionali del sorgente.

3.2. SAP BW Informational model

Il SAP BW informational model si basa principalmente sulle strutture cubiche chiamate InfoObject, le quali contengono soggetti di business e quindi dati relativi ai clienti, ordini di vendita, ordini di acquisto, profitti, ecc.

Essi poi contengono i metadati che descrivono i dati contenuti negli stessi InfoObject, dando informazioni quali l'origine, la storia e le proprietà tecniche dei dati stessi. I metadati hanno il compito di eliminare le diversità semantiche presenti nei dati che attraversano il sistema e le varie unità strutturali, assicurando la loro consistenza e attendibilità. Hanno inoltre il ruolo molto importante di trasformare i dati in informazioni e di spiegare come i vari elementi sono legati tra loro.

Gli elementi chiave di SAP BW informational model sono:

- Data Sources;
- Persistent Staging Area (PSA);
- Info Sources;
- Oggetti ODS;
- Info Cubi.

Data Sources

Contiene le descrizioni dei dati di origine e le trasferisce al SAP BW attraverso una struttura piatta a due dimensioni (Tabella).

PSA

È un insieme di database tabellari trasparenti che costituiscono l'area di deposito iniziale e temporaneo per i dati; qui i dati richiesti e non modificati rispetto al sistema sorgente vengono salvati secondo la modalità di struttura definita dal Data Source.

Info Sources

Sono raggruppamenti di Info Object che vengono associati assieme logicamente da un punto di vista aziendale. Possono contenere dati transazionali, generati da processi OLTP, master data, che possono essere indirizzi o strutture organizzative che qualificano il dato come tale in qualsiasi momento e che solitamente rimangono immutati nel tempo.

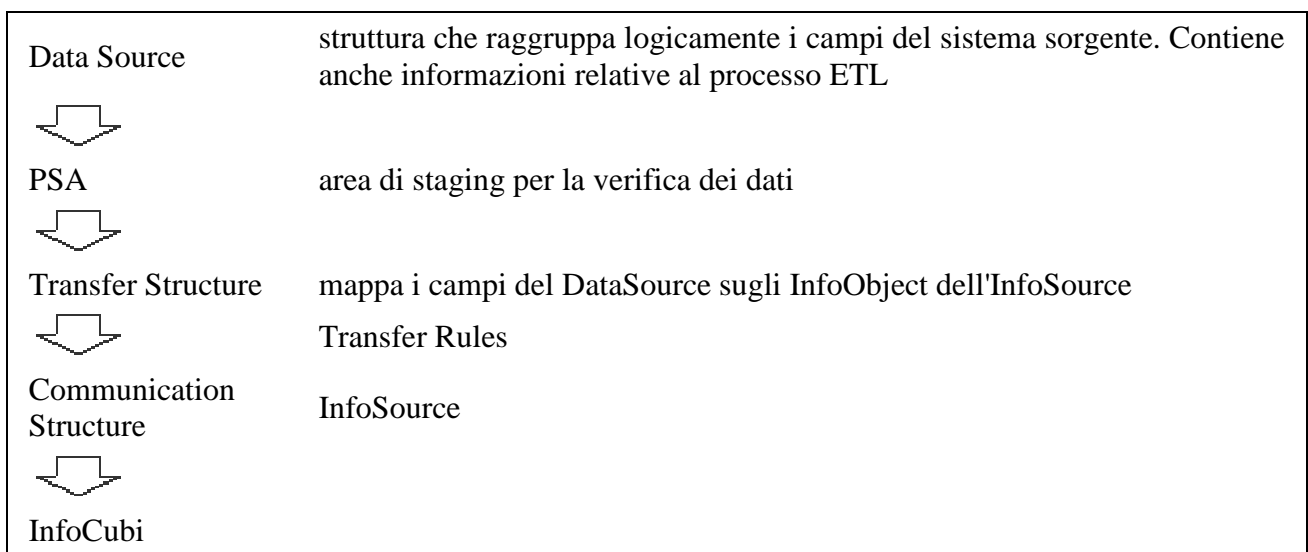
Oggetti ODS

Un oggetto ODS descrive un insieme consolidato di dati provenienti da uno o più Info Sources, immagazzinandoli in tabelle (quindi database a due dimensioni). Alcuni ODS vengono aggiornati negli Info Cubi, altri tramite un aggiornamento delta. I dati contenuti nei vari oggetti ODS possono essere analizzati e visualizzati allo stesso modo di quelli contenuti negli Info Cubi tramite Business Explorer, che è un tool di business intelligence di mysap BI.

Info Cubes

Sono contenitori che organizzano i dati in una struttura multidimensionale, dove ogni dimensione rappresenta un aspetto di interesse aziendale. Si può accedere per avere la reportistica e l'analisi OLAP grazie al Business Explorer.

Flusso di caricamento dei dati negli InfoCubi



La figura sovrastante rappresenta un possibile flusso di dati in SAP BW; possono verificarsi altre situazioni, come ad esempio, che i dati vadano direttamente all'InfoCubo (nella release 7.0 tramite la transfer rules accade proprio questo), o molti Data Sources vengano assegnati ad un singolo Info Sources.

Le regole di trasferimento (Transfer rules) mappano i campi del Data Sources negli Info Object che compongono l'oggetto finale (Info Sources, OSD, Info Cubes, ecc) precisando anche quali dati verranno trasferiti, mentre le regole di aggiornamento (Update rules) trattano il flusso successivo di dati che va dagli InfoSources agli oggetti ODS e agli InfoCubes.

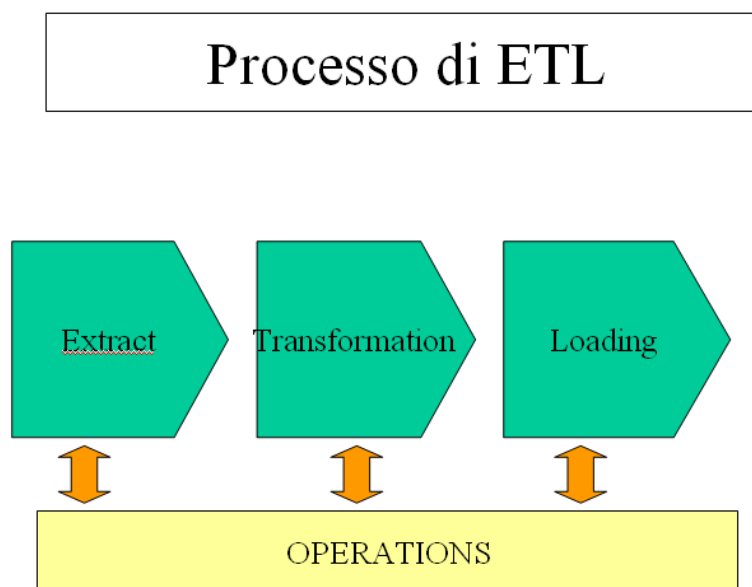
In molti casi, i dati contenuti nella PSA hanno un insieme di metadati incompleti; questi infatti sono aggiunti man mano che gli InfoObject vengono creati e vanno a creare l'Info Sources. Durante il processo che va dal Data Source all' Info Source i dati acquistano valore poiché vengono trasformati in informazione di interesse aziendale. SAP BW offre una ricca libreria di regole di trasformazioni che applicano la logica di business ai dati attraverso i quali la conversione temporale, operazioni su stringhe, aggregazioni, ecc.

L'accesso in SAP BW avviene grazie a:

1. **Info Providers**, che rappresentano gli oggetti che si possono analizzare in SAP BW. Essi possono essere di due tipi: uno che contiene fisicamente i dati, come ad esempio InfoObject, InfoCubi e ODS, e l'altro fornisce un modo di vedere i dati (ad esempio InfoSet, Remote Cubes).
2. **Multi Providers**, sono usati per combinare i dati provenienti da vari oggetti. Il MultiProvider non contiene dati, ma fornisce l'accesso ai dati attraverso vari Info Providers e li rende disponibili per report e analisi.

3.3. Estrazione, Trasformazione e Caricamento (ETL)

Capita spesso che i dati siano frammentati e provenienti da svariati database e applicazioni; questi, per rispondere alle diverse richieste informative aziendali, vengono integrati, standardizzati, sincronizzati e arricchiti grazie al processo ETL. ETL è l'acronimo di Extraction, Transformation, Loading e si riferisce alle fasi di estrazione, trasformazione e caricamento dei dati nel DW. In particolare Kimball [1] definisce in dettaglio il processo di ETL aggiungendo le fasi di Cleaning e Conforming relative alla staging area.



ESTRAZIONE:

L'estrazione è l'operazione di trasferimento dei dati dal sistema sorgente all'area di staging del data warehouse. Durante questa fase vengono trasferiti un insieme di dati sulla base dei filtri predefiniti (*estrazione statica*) o a seconda delle modifiche avvenute nel sistema sorgente (*estrazione incrementale*).

TRASFORMAZIONE:

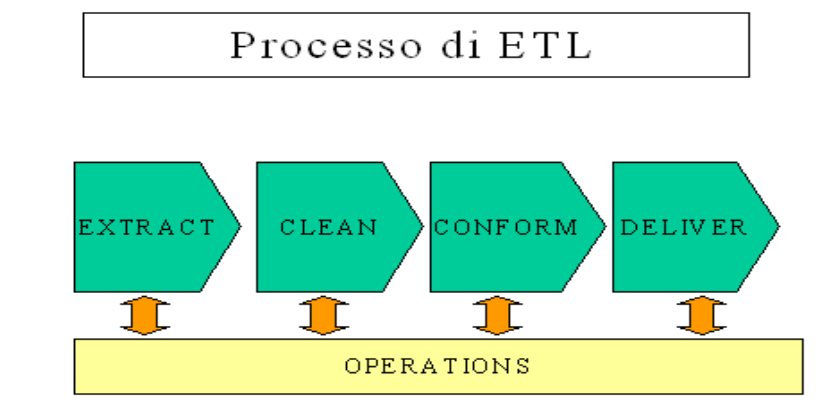
L'obiettivo della fase di trasformazione è quello di prelevare i dati ripuliti nell'area di staging e consolidarli nel livello riconciliato. L'idea è quella di portare informazioni simili provenienti da sorgenti eterogenee ad un modello definito nel

data warehouse. In questa fase, quindi, diverse classificazioni di prodotto possono essere normalizzate in un'unica classificazione, dati presenti in diverse unità di misura possono essere convertiti in una sola, e così via.

CARICAMENTO:

In quest'ultima fase i dati vengono trasferiti nella destinazione, il data warehouse. I dati vengono trasferiti secondo due modalità : *Refresh* (trasferimento integrale) o *update* (trasferimento incrementale).

Facendo riferimento al Kimball[1], il processo di ETL può essere schematizzato dalla seguente figura



dove:

PULITURA:

In questa fase, i dati che risiedono nell'area di staging vengono trattati in modo da migliorare la qualità, con l'integrazione di dati duplicati, la correzione di inconsistenze o dati mancanti, ed altri errori che rendono di difficile interpretazione le informazioni.

CONFORMITÀ:

Questa fase fa in modo che quando si crea un Data Mart successivo ad uno precedentemente creato, esso sia conforme. Conforme significa che abbiano almeno

una dimensione in comune (la stessa); dal punto di vista architettuale è richiesta l'Enterprise Bus Architecture.

DELIVERY :

È la fase finale del processo di ETL, dove i dati puliti e conformati vengono modellati in strutture dimensionali. A queste effettivamente accedono sia gli utenti finali sia le applicazioni.

A seconda della modalità di estrazione, si possono individuare dati estratti dall'application level e al database o file level.

All'application Level i dati vengono estratti come oggetti di business; questo è il metodo più veloce e meno complicato, specialmente quando l'estrazione coinvolge molte applicazioni. Infatti un oggetto di business può essere rappresentato da una serie di tabelle legate tra loro e l'applicazione già rappresenta la relazione tra le tabelle. L'estrazione all' application level può avvenire tramite due componenti:

- Componenti SAP : l'integrazione tra SAP BW e le altre componenti di SAP si basa sull' *Application Programming Interface (API)*, che rende possibile una replicazione completa ed affidabile dei dati grazie a degli estrattori che incapsulano l'applicazione logica. Un estrattore è dunque un programma che riempie la struttura di un Data Source con dati provenienti da SAP.
- Componenti non SAP : per questo tipo di estrazione SAP BW si serve di un'interfaccia chiamata Staging Business Application Programming Interface (BAPI).

L'estrazione **al database o file level** consiste nel prelevare tutti i dati e i corrispondenti metadati direttamente dalle tabelle di interesse e solitamente risulta più complesso ed elaborioso.

Per estrarre i dati da questo livello, esistono i seguenti modi:

* Tramite *DB Connect* per estrarre i dati direttamente dalle tabelle del RDBMS; provvede anche all'inserimento dei metadati replicando le tabelle dei metadati nel metadata repository di SAP BW.

* Tramite *Flat File Transfer* per trasferire in SAP BW dati e metadati provenienti da file piatti (Flat File) come ad esempi CSV, ASCII o XLS.

* Tramite il *Formato XML* (Extensible Markup Language), che è un linguaggio standard di alto valore per i documenti contenenti informazioni strutturali.

Essi sono poi resi disponibili a SAP BW in base alle definizioni date per i dati d'origine nei vari Data Source. I dati vengono poi contenuti fisicamente nella PSA, essi provengano da qualsiasi Data Source o Source System.

I dati vengono trasferiti a BW negli *InfoPackage*, che determinano quali dati definiti in un Data Source potrebbero essere richiesti da un source system. Un InfoPackage può richiedere dati transazionali e master data usando dei parametri specifici, come ad esempio il controllo di un certo ordine in data di creazione 01/01/2004.

In questo modo gli InfoPackage possono descrivere dei sottoinsiemi target di dati descritti nel Data Source. L' *Administrator Workbench* di SAP BW si occupa di schedulazione e monitoraggio dei trasferimenti degli InfoPackage; esso inoltre importa ed esporta i file XML da SAP BW.

3.4. Operational Data Store (ODS)

Un data warehouse ottimizza le query sulla base di un gran numero di dati storici e granulati in modo da fare da supporto alle decisioni aziendali strategiche, piuttosto che a quelle di tipo operativo. Un ODS è strutturato per facilitare le query che si basano su quella ristretta quantità di dati granulari che vengono frequentemente

aggiornati. Contiene infatti dati dettagliati che fanno da supporto tattico alle decisioni aziendali che devono essere prese giornalmente.

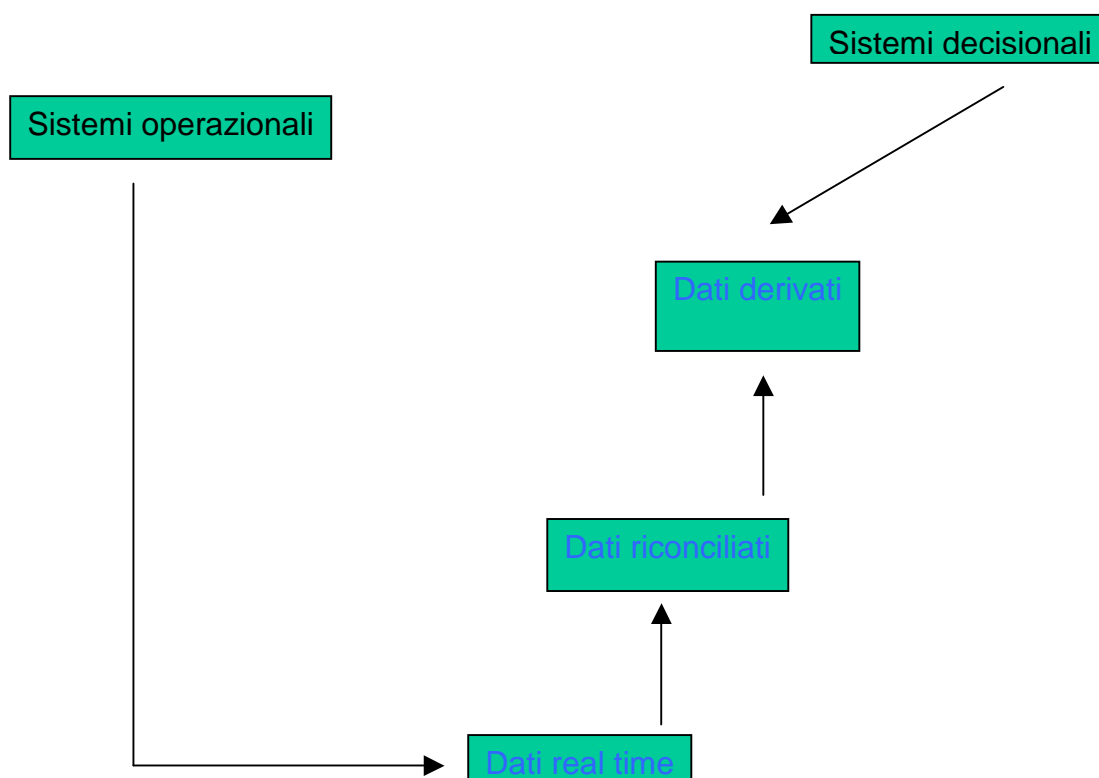
L'ODS è anche un ambiente di informazione continuamente aggiornato di supporto ai reporting operazionali, interagendo coi sistemi transazionali, il datawarehouse o con applicazioni analitiche.

3.5. Data Warehouse e Data Mart

Mentre un Datawarehouse è un “magazzino” centralizzato di tutti i dati importanti riguardanti quegli aspetti a cui gli utenti sono interessati, un Data Mart contiene i dati che servono ad un'area funzionale o aziendale decentralizzata.

L'utilizzo di un data mart ha comunque i suoi pro e contro che vanno valutati al momento della sua implementazione. Questa, infatti se botton-up, è molto più veloce e meno costosa rispetto all'implementazione di un data warehouse (top-down), costando a volte l'80% in meno di un intero data warehouse. Questi vantaggi però si annullano man mano che il data mart prolifera complicando l'integrazione dei dati, aumentando il numero di dati inconstistenti e richiedendo più regole di business.

In definitiva, non esiste una tipologia di implementazione ottimale di data warehouse, ma questa dipende dall'azienda, dalla sua struttura, dai suoi bisogni di informazioni, dai suoi piani e strategie. In questo modo SAP BW si presenta come un approccio flessibile alla tipologia di data warehouse.



3.6. Le release di SAP BW

La sequenza delle versioni di SAP BW è molto importante poiché rappresentano la significativa evoluzione dell'architettura che è avvenuta da un prodotto all'altro nel giro di poco tempo. In breve esse sono:

- BW 1.2b - introduzione agli InfoCubi e alla Business Content (soluzioni preconfigurate per le più frequenti esigenze di business);
- BW 2.0 - introduzione agli ODS e all'interfaccia internet mySAP.com;
- BW 2.1 - introduzione alle componenti di analisi;
- BW 3.0 - ulteriore accrescimento dell'ODS fino a diventare un data warehouse, con la creazione di applicazioni di analisi, società, e così via;
- BW 7.0 - ultima release in commercio.

CAPITOLO 4 : SVOLGIMENTO DELL'ATTIVITA'

4.1. Obiettivo del progetto

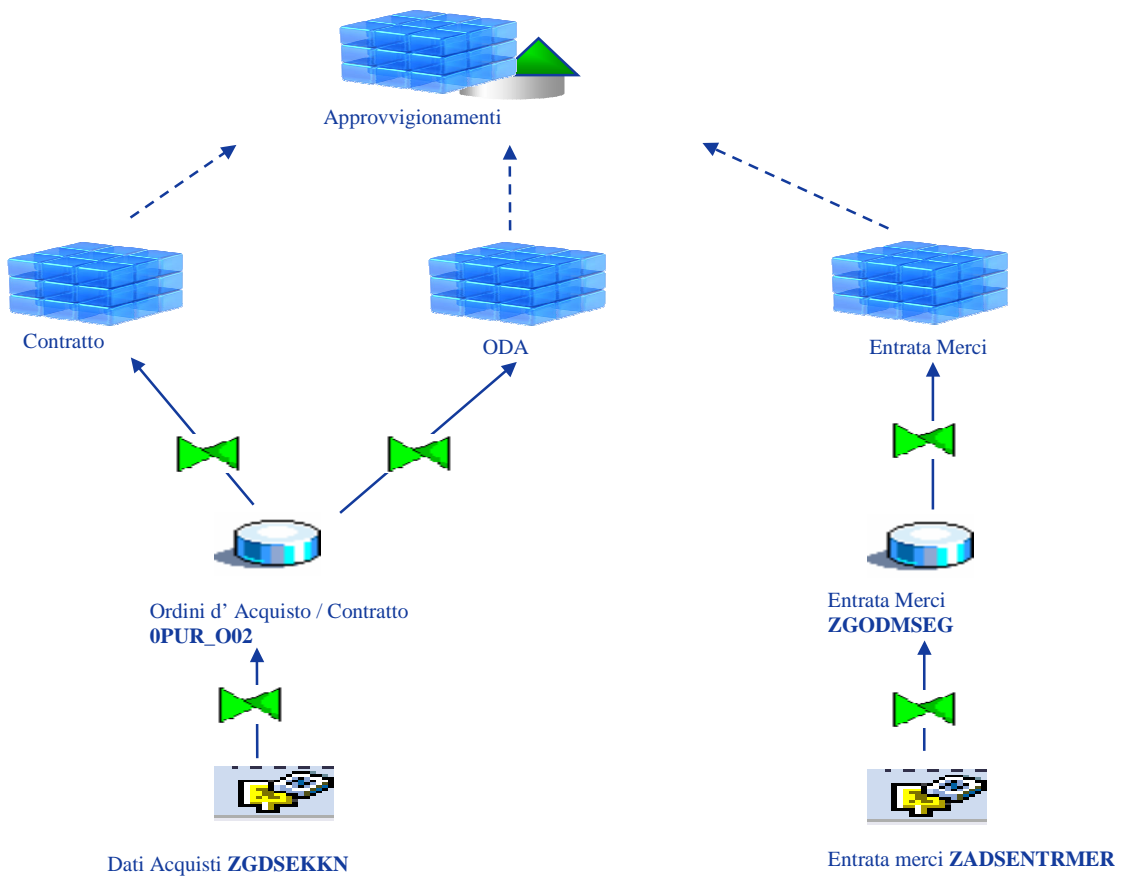
Prendendo come base di riferimento l'implementazione dell'attuale DataWarehouse di una società elettrica, questa fase di progetto costituisce il secondo passo per portare in esercizio il nuovo DataWarehouse, basato sulla piattaforma Sap Netweaver.

L'obiettivo consiste in una nuova soluzione proposta da SIDI per rispondere all'esigenza di una dettagliata reportistica di analisi per l'Area Approvvigionamenti.

Le esigenze sono state rilevate durante una serie di incontri con i referenti della società elettrica, individuati come responsabili funzionali dei processi, e da SIDI - durante le fasi di colloquio con gli utenti - sulla base di pregresse esperienze sulle tematiche in oggetto.

Si prefigge quindi di ripercorrere tutto il processo logico che ha portato ad individuare la soluzione finale in modo da poter analizzare con criticità le condizioni che hanno guidato le principali scelte di implementazione.

4.2. Flusso del modello dati



Come si può vedere dal flusso, il Multi Provider Approvvigionamenti farà da vista a tre cubi:

1. Contratto;
2. Ordini d' acquisto (ODA);
3. Entrata merci.

Vediamo dunque come sono stati implementati tali cubi.

4.3. Contratto

L'Infocube di analisi CONTRATTO prende i dati dal livello riconciliato sottostante e precisamente dall'ODS "Ordini d'Acquisto / Contratto" - 0PUR_O02. Tale ODS è alimentato da Sap R/3 tramite un estrattore custom ZGDSEKKN. L'estrazione è stata effettuata tramite una view chiamata ZGBWEKKN (in essa è eseguito una join tra le tabelle EKKO, EKPO, EKKN). Il nostro estrattore, quindi, prenderà i dati dalle seguenti tabelle:

- EKKO (Testata documento Acquisti),
- EKPO (Posizioni documento Acquisti)
- EKKN (Contabilizzazione nel documento acquisti).

I dati che vengono portati sul cubo sono scremati come segue:

Sono estratti solo i documenti con Categoria Doc. Acquisti = "Contratto" (BSTYP = K). Sono stati esclusi i Tipo Contratto uguali a "MK" e "WK" e i contratti che hanno tutte le posizioni cancellate (questi non saranno proprio portati sull'ODS).

Per prima cosa si è creato il nostro estrattore custom in R/3 (ZGDSEKKN) e replicato tale Data Sources nel sistema BW riportandoci così i campi che si desidera alimentare.

Una volta replicato il DTS si è proceduto con una trasformazione per mappare i campi del data sources con quelli del livello di riconciliazione ODS chiamato 0PUR_O02:

Data Source Impegni Fin. ODA (ZGDSEKKN)				Info regole: Gruppo standard						
Pos.	Ch	Campo	Descr.	Tip.	Nome reg.	Pos.	Ch	InfoObject	Ico	Descr.
1		EBELN	Doc. acquisti	=	ZGIONRODA	1		ZGIONRODA		Ordine di Acquisto
2		EBELP	Posizione	=	ZGIOPSODA	2		ZGIOPSODA		Posizione nel documento ordine di acquisto
3		ZEKKN	Contab. progr.	=	0BBP_ACC_NO	3		0BBP_ACC_NO		N. progressivo della contabilizzazione
4		SAKTO	Conto Co.Ge.	×	0DOC_ITEMS	4		0DOC_ITEMS		N. di posizioni documento
5		GSBER	Sett. cont.	×	0PDVIEYCR	5		0PDVIEYCR		Cons. troppo in anticipo risp. data consegna rich.
6		KOSTL	CdC	×	0PDVVEYCR	6		0PDVVEYCR		Valore cons. troppo in anticipo sulla data cons. rich.
7		ANLN1	Cespite	×	0PDVILECR	7		0PDVILECR		Cons. troppo in ritardo rispetto data cons. rich.
8		ANLN2	N. second.	×	0PDVVLECR	8		0PDVVLECR		Valore cons. troppo in ritardo sulla data cons. rich.
9		AUFNR	Ordine	×	0PDVIELCR	9		0PDVIELCR		Cons. troppo in anticipo/in ritardo rispetto data con
10		KOKRS	Contr. area	×	0PDVVELCR	10		0PDVVELCR		Valore cons. troppo in ant./rit. sulla data cons. rich
11		PRCTR	Profit center	×	0PDVIOVER	11		0PDVIOVER		Consegna superiore alla richiesta
12		PS_PSP_PNR	Elem. WBS	×	0PDVVOVER	12		0PDVVOVER		Valore consegnato superiore alla richiesta
13		NPLNR	Network	×	0PDVIUNDR	13		0PDVIUNDR		Consegna inferiore alla richiesta
14		AUFPL	N. ciclo oper.	×	0PDVVUNDR	14		0PDVVUNDR		Valore consegnato inferiore alla richiesta
15		APLZL	Contatore	×	0PDVIEYSC	15		0PDVIEYSC		Cons. troppo in anticipo risp. data schedul.
16		MENGE	Quantità	×	0PDVVEYSC	16		0PDVVEYSC		Valore cons. troppo in anticipo sulla data schedul.

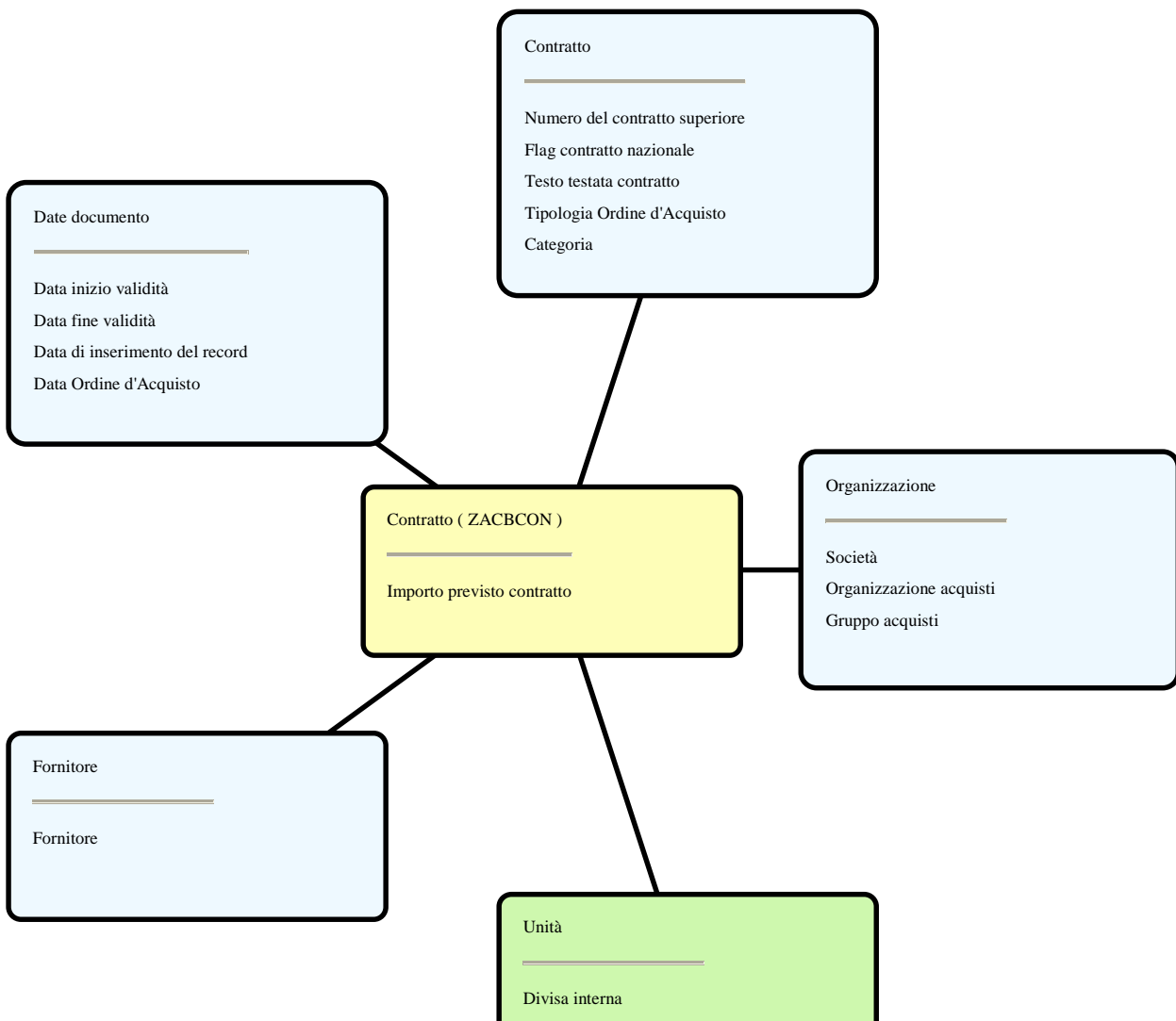
Una volta mappati i campi, si è proceduto con il trasferimento dei dati nella Staging Area (PSA) tramite un infopackage; successivamente i dati sono stati trasferiti da tale area alla destinazione che è l'ODS tramite un DTP applicando un filtro per escludere i tipo contratto uguali a "MK" e "WK" e i contratti che hanno tutte le posizioni cancellate. Ora che nell'ODS contratti sono presenti i dati "puliti" con il

processo di ETL si è potuto applicare una seconda transfert che porta i dati dall' ODS → CUBO; viene creato, dunque, un secondo “livello” in quanto il cubo offre una maggior aggregazione e massima efficienza di reporting. Nell' ods invece c'è un maggior dettaglio ma meno efficiente il reporting (ecco il perché di più livelli). Per verificare che effettivamente i dati coincidano, si è fatta una validazione dei dati creando due query tramite l'uso di Query Designer (SAP BW 3.x) :

- 1- la prima query prende i dati dal cubo;
- 2- la seconda query prende i dati dall'ODS.

L'esito delle query è possibile visualizzarlo tramite excel, sfruttando Analyzer (presente anch'esso nella suite di SAP) e così, affiancando i due risultati, si è voluto verificare se effettivamente i dati presenti nel cubo coincidano con quelli dell'ODS. Ma la vera validazione è stata, però, quella di confrontare i dati ottenuti nel cubo, con quelli presenti nel sistema operativoale R/3.

In seguito è riportato lo schema a stella del cubo CONTRATTO:



4.4. Ordini di Acquisto

L'Infocube di analisi ORDINI D'ACQUISTO prende i dati dal livello sottostante di ODS e precisamente dall'ODS "Ordini d'Acquisto / Contratto" - 0PUR_O02. L'ODS, che è lo stesso visto per i contratti, è alimentato da Sap R/3 tramite l'estrattore custom ZGDSEKKN. L'estrazione è stata effettuata sempre tramite una view chiamata ZGBWEKKN che è la stessa dei contratti. Il nostro estrattore, quindi, prende i dati dalle seguenti tabelle:

- EKKO (Testata documento Acquisti),
- EKPO (Posizioni documento Acquisti)
- EKKN (Contabilizzazione nel documento acquisti).

Sono estratti solo i documenti con Categoria Doc. Acquisti = "Ordini d'Acquisto" (BSTYP = F). Inoltre sono analizzati solo gli OdA con Tipo Documento Acquisti uguale a:

- APP** OdA chiuso Appalto,
- SER** OdA chiuso Servizi,
- FOR** OdA chiuso Fornitura,

Infine sono portati solo gli OdA con Flag Incompleto uguale a blank (sono gli ordini di acquisto conclusi).

Come sempre si è creato il nostro estrattore nel sistema R/3 per poi essere replicato in BW. Una volta presente nel sistema OLAP è stata creata la trasformazione che collega i campi del data sources a quelli dell'ODS degli ordini di acquisto.

15	APLZL	Contatore	=	ZGKIOWR	42	ZGKIOWRT	Importo CQ di Riferimento
16	MENGE	Quantità	=	DCURRENCY	43	DCURRENCY	Chiave divisa
17	NETWR	Valore netto	=	DGL_ACCOUNT	44	DGL_ACCOUNT	Conto Co.Ge.
18	LOEKZ	Cd. cancell.	=	DASSET_MAIN	45	DASSET_MAIN	Numero principale cespite
19	BEDAT	Data documento	=	DASSET	46	DASSET	N. secondario cespite
20	AEDAT	Data modifica	=	DCOORDER	47	DCOORDER	Numero ordine
21	BUKRS	Società	=	DCO_AREA	48	DCO_AREA	Controlling area
22	WERKS	Divisione	=	DPROFIT_CTR	49	DPROFIT_CTR	Profit center
23	MATKL	Gruppo merci	=	ZGIOPSPNR	50	ZGIOPSPNR	Chiave tabella WBSEL
24	KONNR	Contratto	=	DNETWORK	51	DNETWORK	Network
25	LIFNR	Fornitore	=	DACTIVITY	52	DACTIVITY	Operazione network
26	BSART	Tipo documento	=	DCHRT_ACCTS	53	DCHRT_ACCTS	Piano dei conti
27	BSTYP	Categoria doc.	=	DACTY_ELEMT	54	DACTY_ELEMT	Elemento operazione network
28	EKORG	Org. acquisti	=	ZGIOORIG	55	ZGIOORIG	Codice Oggetto di Contabilizzazione
29	EKGRP	Gruppo acquisti	=	DORIG_PCA	56	DORIG_PCA	Tipo oggetto originale in contabilità per profit center
30	MEINS	UM ORDACQ	=	DQUANTITY	57	DQUANTITY	Quantità
31	WAERS	Divisa	=	DUNIT	58	DUNIT	Unità di misura
32	KNTPP	Tp.contab.	=	DAMOUNT	59	DAMOUNT	Importo
33	KTRMB	Doc. contratto					

Una volta mappati i campi si sono trasferiti i dati prima nella PSA tramite un infopackage e in seguito nell'ODS tramite DTP.

Una volta che i dati hanno raggiunto il livello riconciliato, sono stati caricati anche nel cubo ordini d'acquisto per permettere una maggior efficienza nell'analisi di reporting.

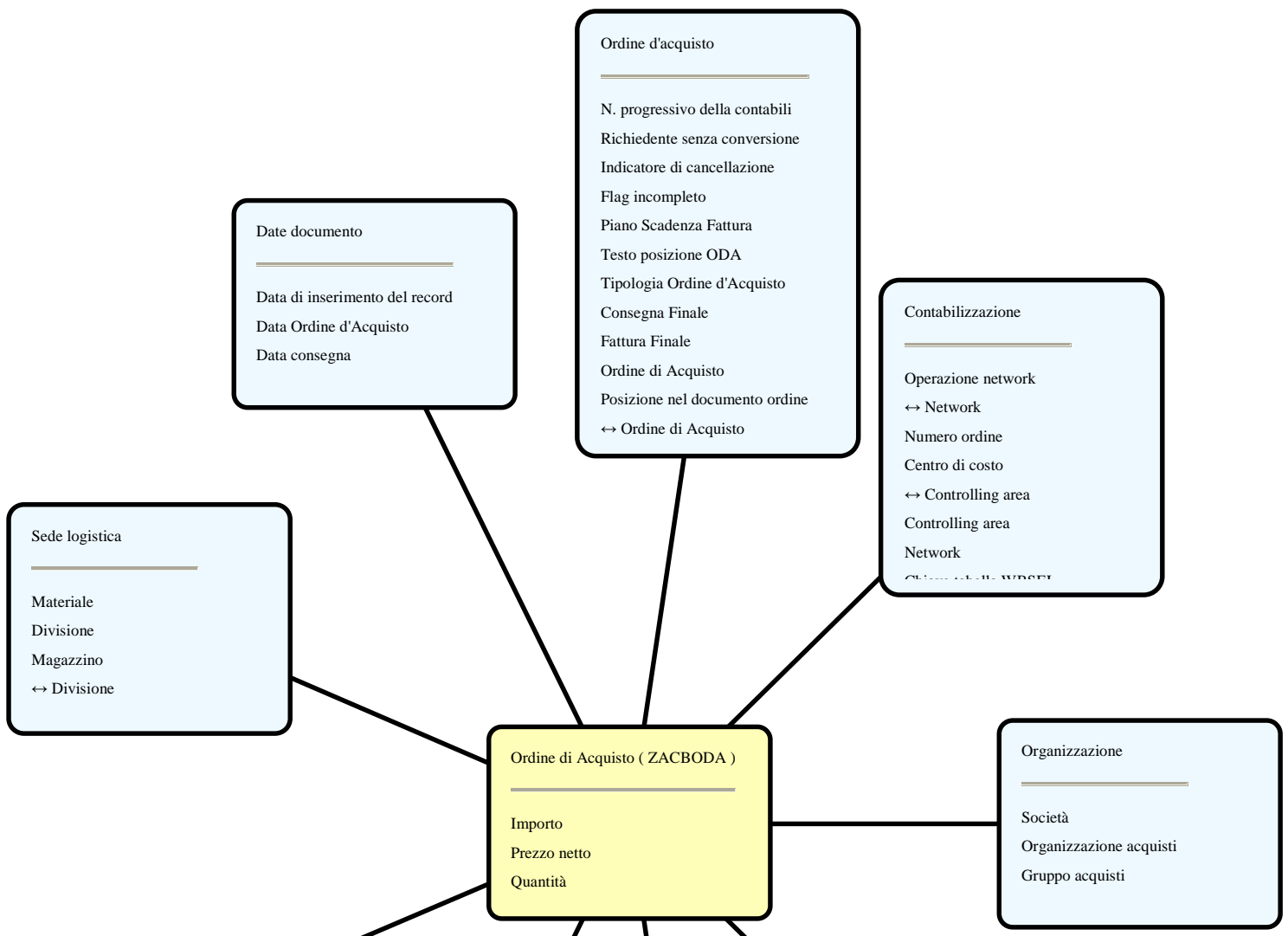
Per verificare che effettivamente i dati presenti in BW siano effettivamente esatti, si è fatta una validazione dei dati creando due query tramite l'uso di Query Designer (SAP BW 3.x) :

1. la prima query prende i dati dal cubo;
2. la seconda query prende i dati dall'ODS.

L'esito delle query è possibile visualizzarlo tramite excel, sfruttando Analyzer e così, facendo un match tra i due risultati, si è voluto verificare se effettivamente i dati presenti nel cubo coincidano con quelli dell'ODS.

In seguito si è fatta la vera validazione confrontando i dati del cubo con i dati presenti in R/3.

Lo schema a stella del cubo ORDINI D'ACQUISTO:



4.5. Entrata Merce

L'Infocube di analisi ENTRATA MERCE (EM) il cui nome tecnico è ZACBENTM, prende i dati dal livello riconciliato ZGODMSEG "Entrata Merce". Tale ODS è alimentato da Sap R/3 tramite un estrattore custom ZADSENTRMER. L'estrazione è effettuata tramite una Function Module chiamata ZAFM_DSENTRMER. All'interno di questa function c'è una chiamata alla classe ZACL_ZADSENTRMER che serve per uniformare le nostre implementazioni. Il nostro estrattore prende i dati dalle seguenti tabelle:

- MKPF (Testata doc. materiale);
- MSEG (Segmento doc. materiale).

I campi estratti da questa Function Module vanno a valorizzare i campi della struttura estrattiva del DTS chiamata ZASTDSENTRMER.

I dati da portare sul cubo sono scremati come segue:

Sono estratte solo le EM che hanno un riferimento all'Ordine di Acquisto (OdA) e sono da escludere le EM relative ad Ordini di Trasferimento (range inizia per 5).

Una volta creato il nostro estrattore, abbiamo replicato tale Data Sources nel sistema BW riportandoci così i campi che vogliamo alimentare.

Una volta replicato il DTS si è proceduto con una trasformazione per mappare i campi del data sources con quelli del livello di riconciliazione ODS chiamato ZGODMSEG.

Entrata Merce (ZADSENTRMER)			Gruppo regole: Gruppo standard						
Pos.	Ch	Campo	Descr.	Gruppo regole: Gruppo standard D2 S	Pos.	Ch	InfoObject	Ico	Descr.
1		MBLNR	Doc. mat.	=		1	OMAT_DOC	📄	Documento materiale
2		MJAHR	Es. doc. mat.	=		2	OMAT_ITEM	📄	Posizione documento materiale
3		ZEILE	Posizione	=		3	ZAIOSDOC	📄	Esercizio del Documento
4		BUKRS	Società	=		4	OLOC_CURRCY	🌐	Divisa interna
5		EKGRP	Gruppo acquisti	=		6	OCOMP_CODE	🏢	Società
6		KOSTL	CdC	=		7	OPUR_GROUP	🏢	Gruppo acquisti
7		AUFNR	Ordine	🏢		8	OPURCH_ORG	🏢	Organizzazione acquisti
8		PS_PSP_PNR	Elem. WBS	=		9	OCOSTCENTER	🏢	Centro di costo
9		VORNR	Operazione	=		10	OCOORDER	🏢	Numero ordine
10		NPLNR	Network	=		11	ZGIOPSPNR	🏢	Chiave tabella WBSEL
11		LFBNR	Doc. di rif.	=		12	OACTIVITY	🏢	Operazione network
12		XBLNR	Riferimento	=		13	ONETWORK	🏢	Network
13		EBELN	Doc. acquisti	=		14	OCO_AREA	🏢	Controlling area
14		EBELP	Posizione	=		15	OPU_RDOCNR	🏢	N. di riferimento di un document
15		RI DAT	Data documento	=		16	ZAI0XBLNR	🏢	Bolla di Consegna

Ora che i campi sono stati mappati, si è proceduto con il trasferimento dei dati nella Staging Area (PSA) tramite un infopackage; successivamente i dati saranno trasferiti da tale area alla destinazione che è l'ODS tramite un DTP. In quest' ultima transfert è stato applicato un filtro escludendo tutti i documenti acquisiti che iniziano con 5 (il cliente richiede di non estrarre ordini di trasferimento, per convenzione sono gli ordini che iniziano per 5).

Ora che nell'ODS entrata merce sono presenti i dati "puliti" con il processo di ETL possiamo applicare una seconda transfert che porta i dati relativi all' entrata merce dall'ODS al nostro cubo ENTRATA MERCE (viene creato, dunque, un secondo "livello" in quanto il cubo offre una maggior aggregazione e massima efficienza di reporting. Nell' ods invece c'è un maggior dettaglio ma meno efficiente il reporting (ecco il perché di più livelli).

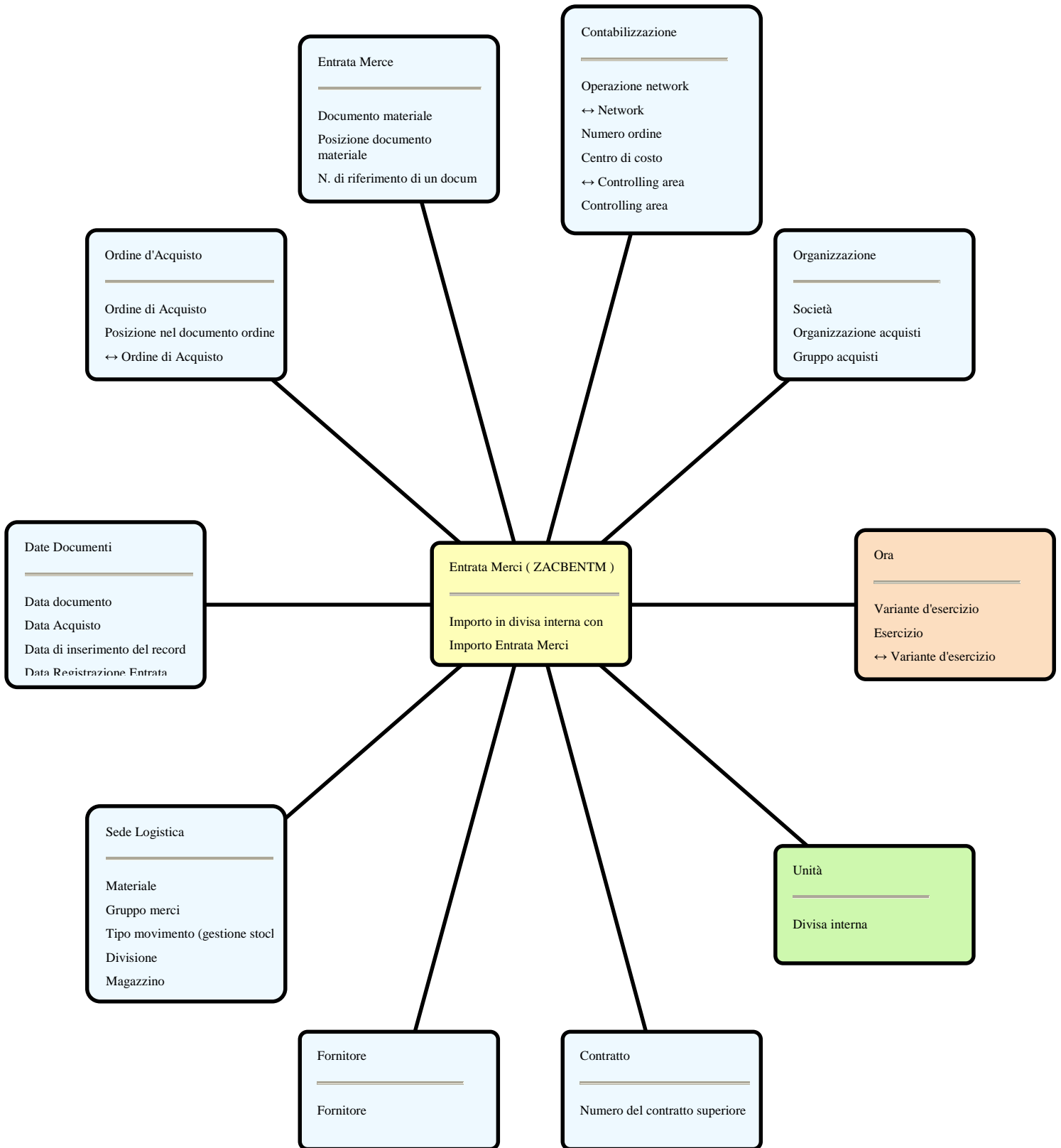
Per verificare che effettivamente i dati coincidano, si è fatta una validazione dei dati creando due query tramite l'uso di Query Designer (SAP BW 3.x) :

- 1 - la prima query prende i dati dal cubo;
- 2 - la seconda query prende i dati dall'ODS.

L'esito delle query è possibile visualizzarlo tramite excel, sfruttando Analyzer e così, affiancando i due risultati, si è voluto verificare se effettivamente i dati presenti nel cubo coincidano con quelli dell'ODS.

La vera validazione è stata, però, quella di confrontare i dati ottenuti nel cubo, con quelli presenti nel sistema operativo R/3.

Lo schema a stella dell'Infocube ZACBENTM è il seguente:



4.6. Multiprovider APPROVIGIONAMENTI

Il nostro oggetto finale per la reportistica e analisi per decisioni è dunque una vista che fa riferimento ai tre cubi precedentemente realizzati.

Le viste non sono altro che relazioni derivate (relazione il cui contenuto è funzione del contenuto di altre relazioni), tabelle di tabelle, definite per mezzo di funzioni non memorizzate nella base dati ma utilizzabili nelle query come se lo fossero. Differenti sono le viste materializzate (utilizzate nel nostro caso) che sono relazioni effettivamente memorizzate nella base dati. I vantaggi di queste ultime sono i seguenti:

- le viste materializzate sono sempre disponibili per le query mentre le virtuali (tipiche dei sistemi relazionali) devono essere ricalcolate ad ogni query;
- le materializzate devono essere consistenti con le relazioni da cui derivano, le virtuali non hanno problemi di allineamento;
- se gli aggiornamenti sono rari rispetto alle interrogazioni, o la vista è complessa, le viste materializzate offrono un sistema più veloce di risposta alle query;

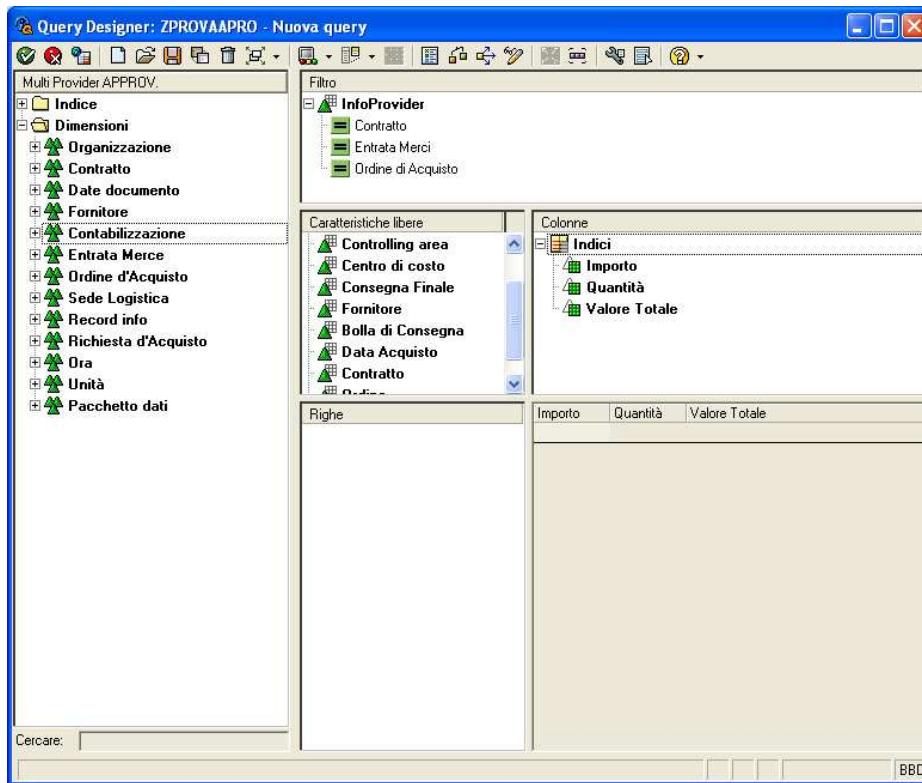
Normalmente non si può influire sulla materializzazione delle viste perché il sistema nasce implementato in modo relazionale o multidimensionale. Il sistema SAP da noi utilizzato, è un sistema multidimensionale e pertanto dispone di viste materializzate.

Riportiamo qui in seguito uno screen-shot relativo ad una query creata con le viste materializzate, ricordando ancora che i dati non sono significativi perché presi in sistema di test.

4.7. Query Tramite il programma QUERY DESIGNER

Esempio di query create utilizzando il multiprovider APPROV.

Screen-shoot che indica il programma Query Designer per la creazione delle query:



Screen-shoot che indica un esempio di query elaborata tramite excel:

	A	B	C	D	E	F	G	H	I
15	InfoProvider	ZACBENTM, ZACBCON, ZACBODA							
16									
17	Ordine	Contratto	Fornitore	Consegna Finale	Importo	Quantità	Valore Totale		
18	600000003	#	100009	X	199.636,47 EUR				
19	600000005	#	100009	X	84.670,00 EUR				
20	600000012	#	100009	X	17.400,00 EUR				
21	600000017	#	100009	X	14.580,00 EUR				
22	600000081	#	100009	X	4.333.193,54 EUR				
23	600000275	#	100009	X	337.861,82 EUR				
24	600000414	#	100070	X	*				
25	600000447	#	100009	X	103.773,90 EUR				
26	700000951	#	100009	X	250,00 EUR				
27	700000938	#	100009	X	12,39 EUR				
28	700000939	#	100009	X	7,82 EUR				
29	700000985	#	100009	X	8,68 EUR				
30	21000000017	#	100009	X	72.900,00 EUR				
31	300000000476	#	100009	X	200,00 EUR				
32	310000000000	#	100009	X	31.000,00 EUR				
33				#	100,00 EUR				
34	310000000010	#	100005	X	77,00 EUR				
35				500001	X	154,00 EUR			
36	310000000022	#	100009	X	0,00 EUR				
37	310000000023	#	100009	X	695,64 EUR				
38	310000000024	#	100009	X	329,60 EUR				
39	310000000049	#	100009	X	703.500,00 EUR				
40	310000000053	600000133	500001	X	241,00 EUR				
41	310000000057	#	100008	X	220.500,00 EUR				
42	310000000061	#	100009	X	743.179,36 EUR				
43	310000001998	600000136	400001	X	555,00 EUR				

CONCLUSIONI

L'esigenza di un supporto alle decisioni è diventata oggi un'attività il cui ritorno dell'investimento è riconosciuto, apprezzato e richiesto.

Per questo il management "sponsorizza" la creazione di un'infrastruttura per il supporto alle decisioni, in fase finale come un insieme di report consultabili via web (ma non solo). I risultati delle query ottenuti sono stati calcolati con dati di sviluppo (ambiente di test), per questo non altamente significativi.

L'esperienza di stage mi ha aiutato molto a comprendere meglio i concetti studiati nel corso universitario di SISTEMI INFORMATIVI AZIENDALI, integrando così la pratica alla teoria. Ho scoperto nel sistema SAP uno strumento potentissimo per il supporto decisionale, rendendo possibili le informazioni necessarie al management in modo rapido e significativo, ricordando che proprio la velocità di possedere delle informazioni al momento giusto può comportare un notevole vantaggio competitivo.

In questa esperienza, le maggiori difficoltà rilevate sono state in fase di progettazione della infrastruttura finale e nella "realizzazione" della stessa.

Spesso è capitato che il cliente finale si sia reso conto che la soluzione ottimale fosse differente da quella programmata e questo ha comportato, nel migliore dei casi, delle modifiche agli oggetti finali e nel peggiore la cancellazione e la creazione di nuovi oggetti. Ne segue, come è noto a chi lavora in questo settore, che la fase realizzativa di un progetto prevede un arco temporale che va da un paio di mesi fino ad arrivare a un anno, anno e mezzo. La mia attività di stage si è conclusa ma non si è concluso totalmente il progetto che è destinato ad ulteriori modifiche e si protrarrà fino a fine ottobre; rimane il fatto che sono totalmente soddisfatto dell'esperienza maturata in un arco temporale relativamente breve.

BIBLIOGRAFIA

- *Data Warehouse Teoria e pratica della progettazione*, M. Golfarelli - S. Rizzi, McGraw-Hill, 2002;
- *Data warehouse Toolkit*, R. Kimball - M. Ross, HOELPI, 2003;
- *Data Warehouse teoria ed esercizi*, S. Dulli - S. Furini - E. Peron, Progetto, 2008;
- *Modelli e strutture per il Data Warehousing*, S. Dulli - V. Favero, Diade, 2000;
- *Processi aziendali e sistemi informativi*, Bracchi - Motta, Franco Angeli, 1997.
- Altri appunti del corso di Sistemi Informativi Aziendali tenuto dalla professoressa S. Dulli.

RINGRAZIAMENTI

Il primo ringraziamento va di diritto alla professoressa Dulli per la sua volontà di tenermi in tesi e la sua pazienza per avermi accolto nel suo studio per chiarimenti e aiuti anche fino alle otto e mezza di sera.

Il secondo ringraziamento va alla mia famiglia per avermi sostenuto e creduto sempre nelle mie capacità e sono contento di averli resi orgogliosi di me.

Un ringraziamento speciale va alla mia ragazza Giulia per non aver mai smesso, e sottolineo mai, di ripetermi di studiare nei tre anni passati all'università e di scrivere la tesi negli ultimi due mesi. La ringrazio perché mi è sempre stata vicino e aiutato a superare tutte le mille e una preoccupazioni pre-laurea.

Un grazie anche agli amici che mi sono stati molto vicino in quest'ultimo periodo di nervosismo portandomi di continuo serenità e allegria.

Infine un ringraziamento a SIDI SpA per avermi tenuto in stage permettendomi di cominciare una formazione lavorativa nel campo nel quale vorrei specializzarmi e Massimiliano (tutor aziendale) per essere sempre disponibile a rispondere a tutte le mie domande spesso anche troppo frequenti ed aiutandomi a crescere professionalmente.

Grazie di cuore a tutti voi.

ENRICO.