Data Warehousing & Business Intelligence

ORACLE

PartnerNetwork
Certified Specialist

Oracle Business
Intelligence
Foundation

Multicom d.o.o.
Vladimira Preloga 11
10000 Zagreb

multicom@multicom.hr
www.multicom.hr

Table of Contents

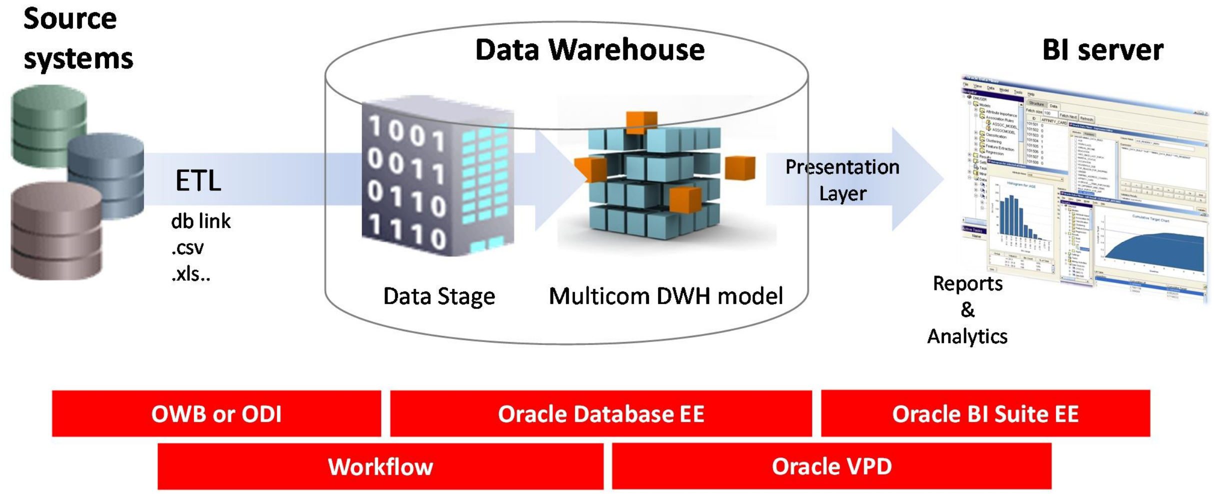
1 General

The prime goal of DWH technology is to collect, integrate, clean and store data in consistent manner to the database system, usually called “History Store”. Due to large quantities of data, the organizations often find that use of thematic databases cointaing just part of the history store leads leads to cost savings (infrastructure costs) and increased availability and move-ability of the data (user can analyze the data on his laptop or department server without connection to history store). Such thematic databases are called “Data Marts”.

Once data exists in history store or data marts, it makes no difference, the user can leverage the value of the data in one or more styles available today:

- Enterprise reporting – the classical tabular, printable report, generated on demand or prepared before being requested.
- Online analytical processing analysis – the fast browsing and querying through the large quantities of the pre-computed aggregated data organized in a conceptual cube.
- Ad Hoc Query Analysis – by writing database queries a user can explore the data in a way not foreseen before creation of reports and cubes.
- Statistical Analysis and Data Mining – using mathematical and statistical models built into data mining tools combined with data in history store and data marts, the user tries to find unknown data patterns and relationships.
- Alerting and Report delivery – the ability of the system to inform the user that requested data becomes available or that there is significant event occurred.
- Corporate Performance Management – monitoring the performance indicators, those that show organization performance (or health) in the past, in the present and how far the organization is from its future targets. Technologically, the CPM is done through visually appealing and easy to understand dashboards and scorecards views (or screens).

Multicom offers an integrated data warehouse (DWH) and Business Intelligence (BI) solution based on Oracle DWH & BI technologies.



Proposed solution is based on Oracle products for several reasons:

✓ Integration – Oracle BI and Oracle database are highly integrated. All the processes (BI queries, data mining, etc.) are executed in database without the need for the extraction.

- ✓ Security – The same security rules are applied in BI and in database. Once the data is in database it is encrypted, there is no need for the export to other systems for BI purposes (cubes, data marts, etc.)
- ✓ Resources optimization – BI processes use database engines. No need for extra servers and storage and data extraction for BI purposes (like with other tools).

The proposed solution uses only proven technology from Oracle, the leading vendor in data warehousing and business intelligence technology. It meets the following goals:

- **Longevity and ability to evolve**:
  - Achieved by proven and sound architecture – by separating concerns across layers of DWH system (Information integration, history store, information delivery and metadata aspect), flexibility and ability to evolve is achieved. Every layer can be changed and evolved separately from others, thus saving money and time. The central meta data repository allows to change some layer, like information delivery (BI), with one from different vendor, but saving large portion of already defined objects like reports, queries, KPIs, etc.
  - Utilizing technology according to the real needs but considering the future trends – the sound future trends like Service-oriented architecture (SOA) enabled BI, Master Data Management, Integrated search, Real Time Performance Management, Text mining, Integrated Search and Master Metadata repository are taken into account when preparing the proposed solution.
- **Credibility**:
  - Integrate data from relevant OLTP & legacy systems – different technologies used to build operating systems in enterprise environment require the flexible tool for data extraction. The proposed solution tool for information integration is able to connect with more than 40 various databases, middleware and transactional systems.
  - Ensure that only clean data enters the system – the proposed solution is equipped with Data Quality technology with some of the most advanced algorithms for eliminating data duplication & performing data enrichment.
- **Usability**:
  - Simple to use end user tools – the proposed information delivery tools are built as Web applications and technology used there are invented to mimic how a person uses a paper document. Built in drag & drop, links, online help, integrated search of objects stored in history store and meta data repository allows the end user:
    - To easily find the data – searching for reports & objects
    - To understand the data – various description of data objects, reports and models
    - Find out the quality of the data – how fresh is the data shown on the report, when was object updated
    - To do data lineage – to find out where from the data was collected
  - Simple access to the data – by just using Internet browser and single sign on concept, the user can access all data stored in the history store or data marts.
- **Security**:
  - Clear policies how system is used and by who – to protect the systems most valuable asset, the information, from unauthorized persons, comes from ability of database technology to constraint access on the row/table/scheme level. The information delivery technology allows setting constraints on almost every definable object like report, dashboard, scorecard, etc. or even on just part of the single view. The policies are enforced by using known security techniques:
    - Authentication: validation of user identification through user ID and password credentials.
    - Authorization: defining what information users can access
    - Auditing: logging what users did with the information.
    - Encryption: protecting information from being read as it is transferred over a network or stored in a data set, table, server or application.

- **Performance**
    - Scalability – the proposed solution can be scaled well beyond today or future requirements. Such ability comes from the fact that software components can be deployed on the computer clusters for data collection and web farms for information delivery.
    - Parallelism – the ability to launch processing work items in parallel instead sequentially, reduces the time to finish for the whole work. Having many workflows (jobs) running in parallel allows to have fresh data with very little delay from the time the data is created in the enterprise operating systems.

2 Solution Description

The architecture of DWH/BI system is commonly divided into following areas:

- Information integration services – also referred to as ETL (extract, transfer and load), set of components tools and interfaces for extraction of data from source system, assuring consistency and quality and loading data into history database.
- DWH (History area) – the database that stores operational data for long period of time
- Information delivery – the information delivery services extract information from the data (stored in DWH) provides these to users in various forms.

This proposed system is independent of the HW technology in sense that it is deployable on various processors and operating systems and additionally, on single machine or large computer grids and web farms.

2.1 General overview of the Data Warehouse environment

In this chapter we describe data warehousing concepts and possible architectures. In the following Picture, we depict an example enterprise data warehouse, where the arrows show the data flow among components.

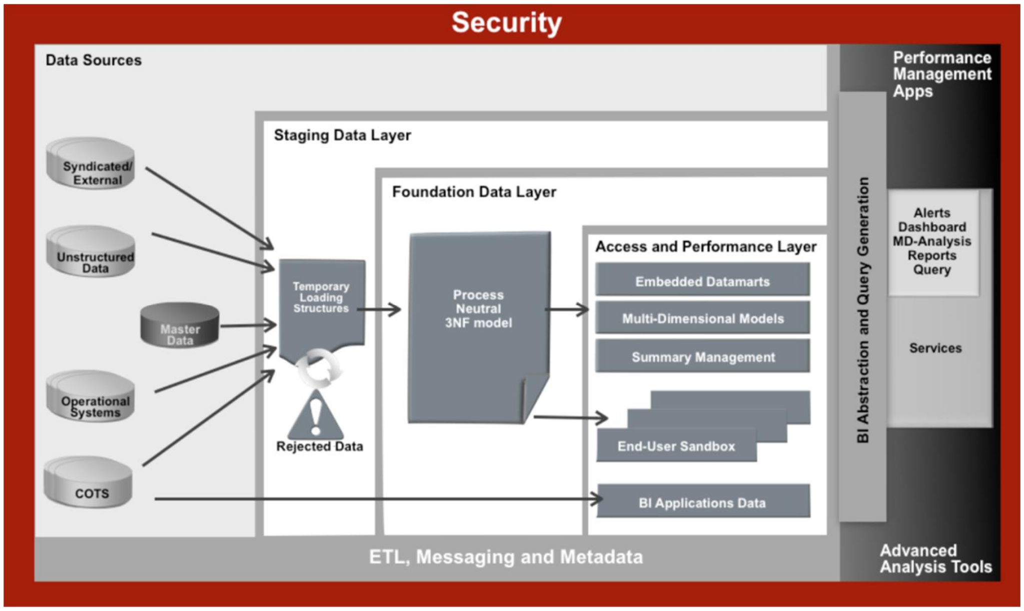


Figure Enterprise data warehouse environment

The DWH components differ not only by content of data but also by the way they store the data and by whom it can be accessed.

- *Staging Data Layer:* For handling data extracted from source systems. There can be data transformations at this point and/or as the data is loaded into the data warehouse. The structure of the staging area depends on the approach and tools used for the extract, transform, and load (ETL) processes. The data model design affects not only performance, but also scalability and ability to process new data without recreating the entire model.
- *Foundation Data Layer*: This is the area, also called the system of record (SOR) that contains the history data in 3NF and is typically not accessed for query and analysis. Use it for populating the summary area, analytical areas, and the dependent data marts.
- *Access and Performance Layer*: Contains multidimensional (MD) structures, such as the star schema (also referred to as Subject Areas), snowflakes, or multi-star schemas, constructed for high performance data analysis.

- *Embedded Data marts and Summary Management:* This area contains aggregations. Structures are usually derived from the data warehouse where one or more attributes are at the higher grain (less detail) than in the data warehouse. These are constructed for high performance data analysis where low level detail is not required.

2.2 ETL architecture and processes

One of the key processes in the DWH system is integration with operational environment. The data is gathered from environment into DWH through integration type called ETL. The data is extracted from the source systems, transformed, quality assured and finally transferred to the history store for long time storage.

Abbreviated term, the ETL stands for:

- Extraction – get the data from data source system
- Transformation – manipulate the data until it fits the business needs and history store data model
- Load – persist data into target database (data warehouse, data mart)

The ETL Layer extracts, transforms, and loads validated source system data into DW using Multicom ETL Framework. The key components of Multicom ETL Framework are:

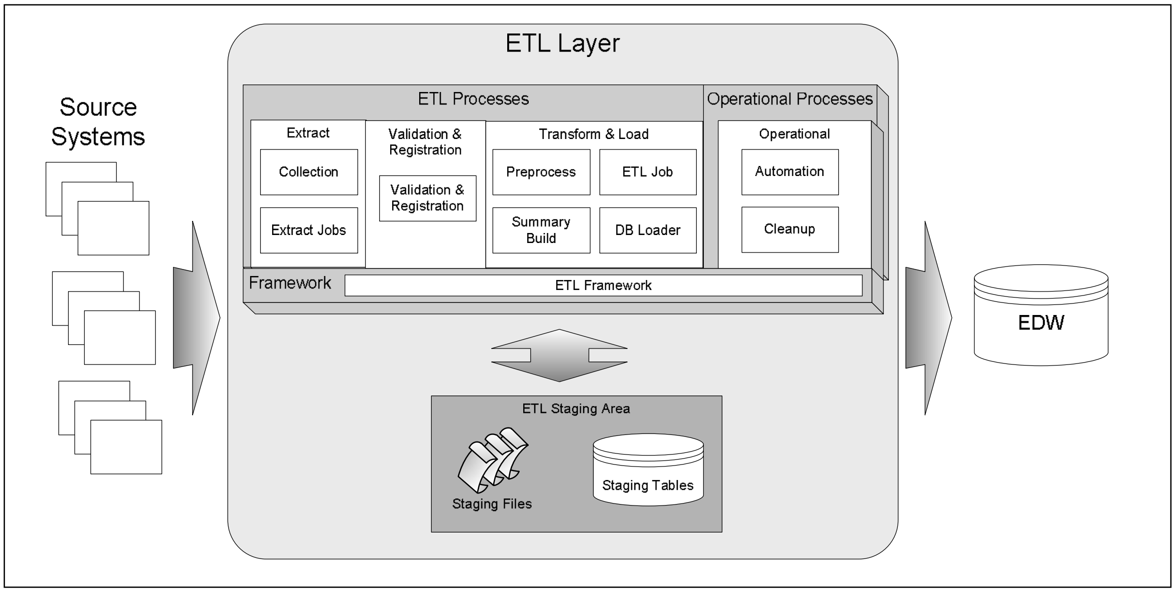


Figure ETL Architecture

The ETL Layer consists of the following modules:

- **ETL Processes** – contains the data movement pieces of the architecture. These modules perform the main functions of the ETL Layer in extracting data from the source systems, validating input data, transforming the source data according to the business rules and loading the EDW. The ETL processing will be performed through Oracle Warehouse Builder mappings.
- **Operational Processes (Workflow)** – contains the operational maintenance pieces of the ETL Layer that supports the scheduling and automation of ETL jobs and cleanup of the files in the staging area. The Automation will be performed through Oracle DWH workflow processes.
- **Framework** – contains the utilities and scripts for common functions of logging, metadata capture, process control, and automation utilities used by components in the ETL Layer modules – ETL Processes and Operational Processes.

It is important to notice that Multicom ETL loading framework is independent of ETL tool being used. The basic of the framework is a process execution hierarchy (metadata) and log table with process execution details. The log table is the main criteria for process execution skipping or restarting. Process transitions can be marked as critical (halting the overall execution) or non-critical.

Keeping the basic framework independent of ETL tool makes later tool upgrades or even switches to other tools, risk free. Hybrid combinations are also possible where different parts of the overall DWH refresh process are implemented by using different ETL tool.

Multicom's standardized ETL framework assures:

- Standard mapping design principles (pre and post mapping procedures);
- Naming conventions;
- Additional attributes ensuring data lineage and quality, metadata, ETL consistency (source table, destination table, business keys/IDs, extraction date, extraction run_id, etc.);
- Ability to run standard set with single command via graphical interface or command line. Passing of parameters is also supported.
- Ability to restart processes after failure with no processing and performance overhead;
- Double processing protection and process skip protection;
- Data and process dependency check and logging;
- Visual process execution state monitoring;
- API for programmatically starting ETL processes (push method);
- Staging area maintenance;
- Ability to run ETL scheduled processes (sub-processes) more than once per day;
- Ability to run ETL processes manually;
- Ability to extract and load only selected parts of data sources;
- Ability to monitor and log all events and operations together with performance data. Runtime repository stores complete metadata regarding process and mapping execution (start time, end time, records selected, records inserted, records merged etc.) together with error details.
- Data quality check combined with notification activities.

The user-friendly interface offers graphical access to the most common features that a dependency engine supports. The user can design the complete process including email notifications etc. Code generation for Process Flow definitions consists of industry standard XML Process Definition Language (XPDL). Process flow definitions can contain a multitude of activities, including mappings, transformations, external processes and file-based activities such as FTP or “whether file exists”.

2.2.1 Data Quality

The ETL process includes one additional discipline, the Data Quality Management (DQM). The DQM goal is to have integrated and correct data stored in the history store. The DQM deals with following activities:

- Data Profiling - analyzing operational data for error patterns and generating Meta data about tables (e.g. potential primary key, referential integrity issues, data scheme in 1st or 2nd normal form, etc.)
- Data cleansing – Parsing attributes and standardizing data formats (e.g. dates, sex, product codes, addresses, names, etc.)
- Data enrichment - supplementing of the data set by adding missing or new information (e.g. postal codes, product names, full addresses)
- Data integration - merge identical data sets from potentially different data sources (e.g. when there are two or more customer records, are they pointing to the same customer?).

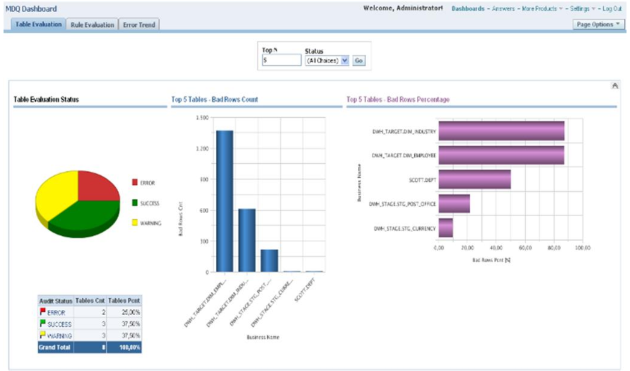
Multicom Data Quality (MDQ) is an application for data quality assurance in information systems developed on top of Oracle Database. It enables evaluation and monitoring of all kinds of IT business rules. Application comprises the following components:

Å Metadata repository (data models, business rules and data quality logs)

Å Graphical user interface for metadata entry and operational reports view

Å High-performance PL/SQL engine for data quality evaluation and logging

Å BI graphical interface with predefined reports for monitoring the data quality logs



Top N: 5
Status: (All Choices)

Audit Status	Tables Cnt	Tables Pcnt
ERROR	2	25,00%
SUCCESS	3	37,50%
WARNING	3	37,50%
Grand Total	8	100,00%

2.3 DWH

Oracle Database is the repository for the proposed Data Warehouse data model. Oracle Database 11gR2 is the proven release of the leading relational database for business intelligence and data warehousing. Oracle is most often chosen for data warehousing because of its success in satisfying the core requirements for data warehousing: performance, scalability, and manageability. Since data warehouses will store larger volumes of data, support more users, and require faster performance, these core requirements remain key factors in the successful implementation of data warehouses.

Oracle is an industry leader in performance, which makes it very suitable for the Data Warehousing area, among others. It boasts a powerful array of performance features including:

- partitioning
- bitmap indexes
- table and index compression
- materialized views
- advanced query optimization
- virtual private database

Scalability is also a deciding factor, with Oracle being the only database vendor which implements the grid processing concept.

The DW components differ not only by content of data but also by the way they store the data and by whom it can be accessed.

- *Staging area:* For handling data extracted from source systems. There can be data transformations at this point and/or as the data is loaded into the data warehouse. The structure of the staging area depends on the approach and tools used for the extract, transform, and load (ETL) processes. The data model design affects not only performance, but also scalability and ability to process new data without recreating the entire model.
- *Data warehouse*: This is the area, also called the system of record (SOR), that contains the history data in 3NF and is typically not accessed for query and analysis. Use it for populating the summary area, analytical areas, and the dependent data marts.
- *Summary area:* This area contains aggregations. Structures are usually derived from the data warehouse where one or more attributes are at the higher grain (less detail) than in the data warehouse. These are constructed for high performance data analysis where low level detail is not required.
- *Analytical area*: Contains multidimensional (MD) structures, (also referred to as Subject Areas), constructed for high performance data analysis.

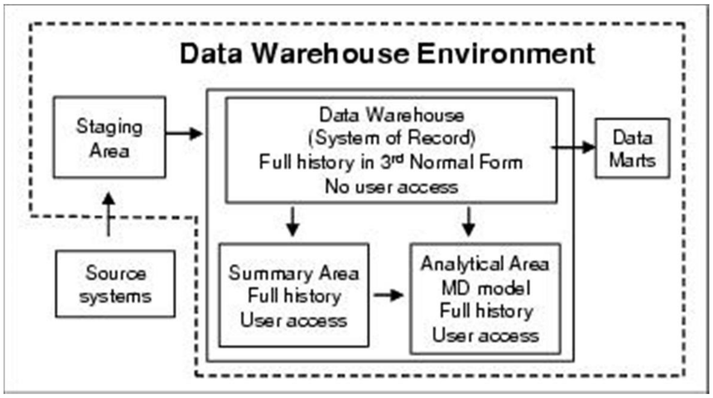
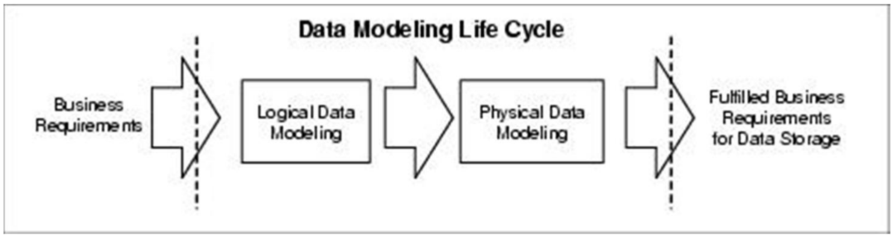

Data Warehouse Environment
Staging Area
Data Warehouse (System of Record) Full history in 3rd Normal Form No user access
Data Marts
Source systems
Summary Area Full history User access
Analytical Area MD model Full history User access

Figure Enterprise data warehouse environment

2.3.1 Data Warehouse Model

Based on experience, Multicom has developed own models based on best practices in the telecom, utility an financial industries.
The goal of the data modeling life cycle is primarily the creation of a storage area for the business data.
By identifying the exact required data, data processing activity will produce the exact required information and avoid producing unnecessary information, as depicted in the following figure:



A generic data modeling life cycle

In the context of data warehousing, data modeling is a critical activity. Overview of the life cycle components:

- Logical data modeling - This component defines a network of entities and relationships representing the business information structures and rules. The logical data warehouse model contains all the business entities for specific business subject areas, relationships between the entities, and attributes of each entity comprising the target model. Parameters such as fact table names, fact table columns, dimension names, dimension columns, business definition for the data fields, level of granularity, record creation, etc.
- Physical data modeling - This component maps the logical data model to the target database management system (DBMS) in a manner that meets the system performance and storage volume requirements.

Data model components:

	Stage	Target	Data Marts
Reason	Gateway into analytics environment from multiple sources	‘Working store’ for analytics applications and engines	Organizing analytics results for reporting
Organization	ÅDenormalized Flat extract from Source by product	ÅDenormalized Structure tuned for app performance	ÅDenormalized Star, OLAP cubes
Subject areas	By source ‘business process’ – eg. Withdrawals, deposits etc.	Combination of common areas (customer, GL, account) and app-specific(ALM, OR, MR)	App-specific – eg. ALM, Profitability BI, Basel II reports etc.
Updates	ÅNo/low updates	ÅFrequent as a result of application processing	ÅFrequency depends on reporting needs
Processing	ÅDe-duplication Data Cleansing Dimensional Conformance Reconciliation	ÅEngine Specific processing (CFE, allocations, Monte-carlo for risk apps)	ÅAggregations Drill throughs

2.3.2 Metadata and Data Dictionary

The generic metadata repository should be capable of sourcing, sharing, storing, and reconciling four groups of metadata:

- **Business metadata** - business rules, definitions, business domains, terms, glossaries, algorithms and lineage information using business language (for business users).
- **Technical metadata -** defines source and target systems, and their table and fields structures and attributes, derivations and dependencies (for specific users of Business Intelligence, OLAP, ETL, Data Profiling, and ER Modelling tools).
- **Operational metadata** – the data about operational application execution (events) and their frequency, record count, component-by-component analysis, and other granular statistics.
- **Project metadata -** documents and audits development efforts, assigns stewards, and handles change management (for operations, stewards, tool users, and management).

Multicoms ETL framework includes following functionality:

- Lineage and impact analysis - Includes interactive analysis available in the Design Centre
- Change propagation - Includes automatic propagation of property changes to impacted objects
- Extensibility - Includes project based and public based user-defined objects, user-defined associations, and user-defined modules. Includes creating icon sets and assigning custom icons to objects
- Advanced configuration management – Includes project based configuration management, handles transition between test&dev to production environment.

2.3.3 Security

BI-DWH platforms are typically secured against unauthorized access to data, reports, and other functionality. In this particular project security is of highest importance since data is sensitive and of high business value and must under no circumstances be shared with different report consumers

Data security is performed through the following mechanisms, according to the defined needs and requirements:

- Centralized user authentication/authorization (LDAP)

- Multiple levels of user privileges and access to data and BI-DWH components (administrators, staging area, presentation model, procedures and processes)
- HTTPS protocol for access to sensitive web application content
- Database-level user authorization
- Application-level, module-level and report-level user authorization
- Segmentation of user privileges (department/position)

The use of **Oracle Virtual Private Database (VPD)** mechanism ensures that:

- It is possible to create arbitrarily complex security rules
- Security rules are created and maintained in a single location
- Since the rules are being evaluated and applied at a low level (data access layer), it is not possible to subvert or bypass them by using a different channel or application
- Multiple users can e.g. use the same report, but each user will be able to see only rows and columns that correspond to his access privileges

Additional security measures (physical security, network security, remote access, password policy, auditing, monitoring and alerting, anti-virus protection) could be implemented in accordance with corporate standards of the organization.

2.4 Information Delivery Services (End User BI Tools)

The data stored in history store or data marts have to be made available to the business users in simple to use and intuitive way. The access to the data has to be controlled and secured but in a way that does not impact usability and user acceptance. As a fact, the business user are not using the data in the system in the same way, some feel most comfortable with printable, tabular report while others prefer to use multidimensional report models and dashboards. Based on business user needs and preferences there are five styles of BI use:

- Enterprise reporting
- Ad Hoc Query analysis
- Performance management
- OLAP analysis
- Alert notifications

When an enterprise wishes to distribute standard operational reports or financial reports to all stakeholders in the organization, enterprise Reporting is used. Since the 1950s, corporations have found clear returns on their investment in operational and financial reporting. Consequently, Enterprise Reporting is the most widespread Style of BI – ranging from its earliest adoption as mainframe green-and-white banded paper reports to today’s web-based reports. An enterprise report is a tabular report with optional graphical elements like charts, images (geographical maps) and controls for filtering and sorting. The reports can be simple and static (table view) and very complex with input parameters, cross tabbed, linked to more detailed reports, etc.

Ad Hoc Query and Analysis is the Style of BI that enables a user to construct a query (SQL, MDX, XML) and execute it against database system. Although the ad-hoc query allows retrieving and data in any combination imaginable, it is not widely used, except by the report developers.

Generally, the performance management is a set of processes for measuring the business performance (health) of organization. The performance of the organization or a single business process is measured with values called metrics or indicators. There are four types of indicators:

- Key performance indicators – a way to measure progress toward organization strategy goals
- Performance indicators – a way to measure progress toward department/process goals
- Key result indicators – overview of the past performances for upper management
- Result indicators – overview of the past performances for department/process staff

The technological part of performance management in context of BI are the tools that allows retrieving, formatting and visualizing metrics and their goals in graphically rich single screen view. There are two distinct ways to create a performance management view:

- Scorecard – the view on the key performance indicators (metrics + predefined goals) intended for the high executives.
- Dashboard – the view on the metrics that allows monitoring of business process(s). The refresh of metrics should be as close to real time as possible since they tend to show the current health of business process.

The coined term OLAP (Online analytical processing) refers to how database system stores data and client tools that allow search operations on the same data. There are three types of data storing strategies ROLAP (Relational OLAP), MOLAP (Multidimensional OLAP) and HOLAP (Hybrid OLAP) but they all give the business user ability to perform following generic queries:

- Drill-down: raising the degree of detail of data in a view by transferring to a lower aggregation stage (e.g. in the time dimension, transferring from 'year' as the aggregation stage to the ‘month’ stage).
- Roll-up: reducing the degree of detail of data in a view by transferring to a higher aggregation stage (e.g. for the time dimension, transferring from the 'day' aggregation stage to the 'year' stage via the 'month' stage). This is an opposite operation than drill-down.
- Slice-and-dice: navigation in a multidimensional data space by focusing on specific aspects, for example distribution of sales for a certain product over various regions and time periods (e.g. by applying various data filters).
- Drill-through (or reach-across): when a user wishes to see detailed data, beyond those stored in the cube, the tool retrieves the data from the history store. (e.g. when user drills down to day level and wants to go to transaction level).

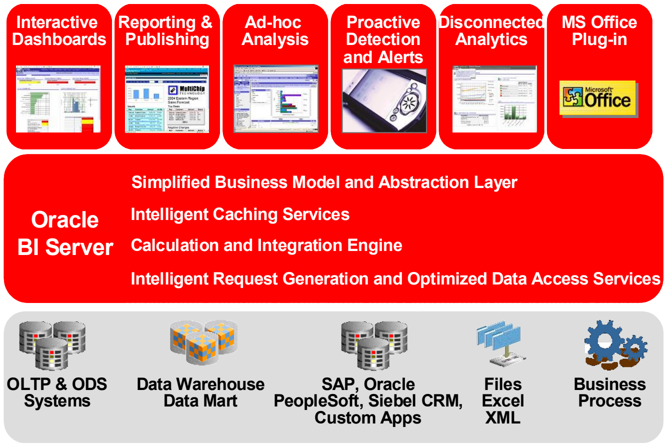
The OLAP is popular way to analyze vast amounts of aggregated data since the query takes just few seconds or less to execute. Additionally, the intuitive browsing through hierarchies and combing different data dimensions with measures makes this kind of analytics suitable for most users.

There are various events that occur in the enterprise every day and the data trail for some of them remains stored in operational systems and later in data warehouse. As a consequence, the event data will be displayed in one of the BI outputs (reports, dashboards, etc.) which assume the proactive approach from business user (he has to request information). Since it is not really possible for the business users to be proactive whole day (24 hours) the BI tools come equipped with alert notification functionality. The alert notification is ability of the BI platform to send e-mail (or SMS/MMS) when important event occurs. The logic behind, is a database query and threshold value or simply a rule. When rule is applied successfully on the data the predefined message with possible attachments is sent to concerned business users.

2.4.1 End User Tools / BI Applications

The proposed solution business intelligence platform with above described functionality and many more is the Oracle Business Intelligence Enterprise Edition tool stack. Built directly on top of database drivers is the Oracle BI Server as semantic layer. The BI server contains model of all objects definitions used in end user application, from tables, columns, queries up to metrics and goals. This, semantic layer enables the end user and BI developer to create reports, dashboards and scorecards without deep, detailed knowledge of database scheme or how to write complex data retrieval queries

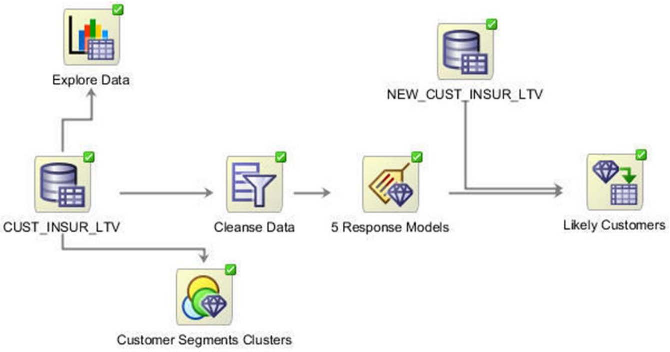
The Oracle Business Intelligence Suite Enterprise Edition Plus is a member of Oracle BI offering with the enterprise components as displayed on picture.


Interactive Dashboards
Reporting & Publishing
Ad-hoc Analysis
Proactive Detection and Alerts
Disconnected Analytics
MS Office Plug-in
Oracle BI Server
Simplified Business Model and Abstraction Layer
Intelligent Caching Services
Calculation and Integration Engine
Intelligent Request Generation and Optimized Data Access Services
OLTP & ODS Systems
Data Warehouse Data Mart
SAP, Oracle PeopleSoft, Siebel CRM, Custom Apps
Files Excel XML
Business Process

Oracle Business Intelligence Suite delivers a unified, integrated BI infrastructure featuring a comprehensive set of products that are available today and span query and analysis, enterprise reporting, mobile analytics, dashboards and portal technology, integration with Microsoft Office and Excel, intelligent workflow, real-time alerting, Business Activity Monitoring (BAM), and more.
Besides providing full spectrum of BI functionality, OBI EE platform is based on modern Web Service Oriented Architecture (SOA), which ensures real next generation business intelligence.

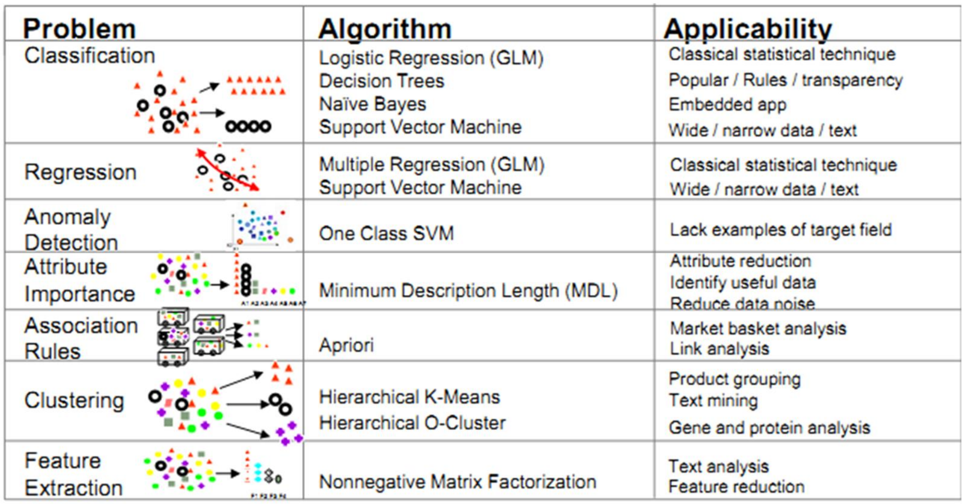
2.4.2 Oracle Advanced Analytics

Oracle Advanced Analytics is a separately licensed extension of the Oracle Database that enhances it into a comprehensive advanced analytics platform through two major components: Oracle Data Mining and Oracle R Enterprise. With Oracle Advanced Analytics, customers have a comprehensive platform for real-time analytics that delivers insight into key business subjects such as churn prediction, product recommendations and fraud alerting.



Its key features are:

- Powerful and scalable architecture for performing in-database predictive analytics, data mining and statistics
- Easy to use: SQL Developer/ODM workflow GUI or any R GUI work directly on database tables/views
- Integration with open source R algorithms
- Functionality accessible via SQL, PL/SQL, R and database APIs

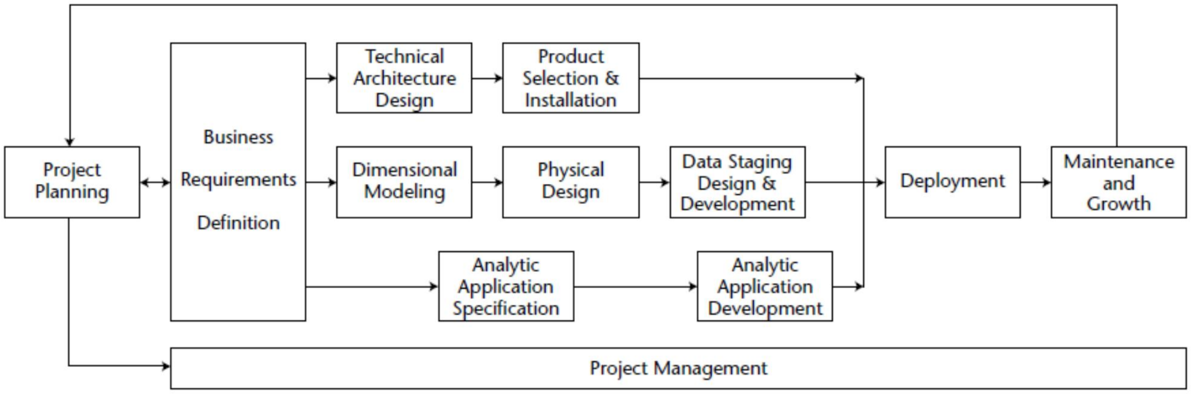


Problem	Algorithm	Applicability
Classification	Logistic Regression (GLM) Decision Trees Naïve Bayes Support Vector Machine	Classical statistical technique Popular / Rules / transparency Embedded app Wide / narrow data / text
Regression	Multiple Regression (GLM) Support Vector Machine	Classical statistical technique Wide / narrow data / text
Anomaly Detection	One Class SVM	Lack examples of target field
Attribute Importance	Minimum Description Length (MDL)	Attribute reduction Identify useful data Reduce data noise
Association Rules	Apriori	Market basket analysis Link analysis
Clustering	Hierarchical K-Means Hierarchical O-Cluster	Product grouping Text mining Gene and protein analysis
Feature Extraction	Nonnegative Matrix Factorization	Text analysis Feature reduction

Standard data mining algorithms supported by Oracle Data Miner

3 Metodology

Multicom understands that to successfully implement business applications today, organizations must embrace a proven, structured method to guide the implementation, manage risk, and avoid missteps. The method must be flexible enough for the implementation effort to be tailored to the specific and unique needs of the organization.
Business dimensional lifecycle diagram:


Project Planning
Business Requirements Definition
Technical Architecture Design
Product Selection & Installation
Dimensional Modeling
Physical Design
Data Staging Design & Development
Analytic Application Specification
Analytic Application Development
Deployment
Maintenance and Growth
Project Management

Our proposed project methodology is aligned with the Oracle DWMft method (Data Warehousing Method FastTrack), PMI „PMBook 2004“, R. Kimball methodologies and is adjusted based on our experience from previous projects.

Multicom experts have years of experience in implementing this technology in large scale implementations. We consider our strength to know exactly how the of-the-shelf products behave and how to use them in an optimal way. Over years we have developed own frameworks using this technology to offer clients specific solutions for common problems in DWH environments.