Fall 2011

Data Warehouse and Business Intelligence: Comparative Analysis of Olap tools

Mahesh Kumar Bhetwal
Regis University

Regis University
College for Professional Studies Graduate Programs
Final Project/Thesis

Disclaimer

DATA WAREHOUSE AND BUSINESS INTELLIGENCE: COMPARATIVE ANALYSIS OF

OLAP TOOLS.

A THESIS

SUBMITTED ON 15 NOVEMBER, 2011

TO THE DEPARTMENT OF COMPUTER SCIENCE

OF THE SCHOOL OF COMPUTER & INFORMATION SCIENCES

OF REGIS UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF MASTER OF SCIENCE IN
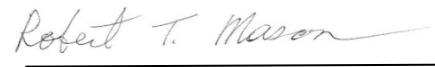
DATABASE TECHNOLOGY

BY

_____

Mahesh Kumar Bhetwal

APPROVALS

_____

Dr. Bob Mason, Thesis Advisor

_____

Dr. Richard L. Blumenthal

_____

Nancy Birkenheuer

Abstract

Data Warehouse applications are designed basically to provide the business communities with accurate and consolidated information. The objective of Data Warehousing applications are not just for collecting data and reporting, but rather for analyzing, it requires technical and business expertise tools.

To achieve business intelligence it requires proper tools to be selected. The most commonly used Business intelligence (BI) technologies are Online Analytical Processing (OLAP) and Reporting tools for analyzing the data and to make tactical decision for the better performance of the organization, and more over to provide quick and fast access to end user request.

This study will review data warehouse environment and architecture, business intelligence concepts, OLAP and the related theories involved on it. As well as the concept of data warehouse and OLAP, this study will also present comparative analysis of commonly used OLAP tools in Organization.

Acknowledgements

I am heartily thankful to my advisor, Dr. Bob Mason, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I would also thank to my friend, Mr. Mohan Uprety, for his support to collect and find the reference materials for this research. I would also like to thank to Dr. Michael J. Garcia, Lead Faculty of Humanities and Social Sciences, at Regis University who edited the paper and provided me valuable feedback that help me out to improve my research paper. I like to thank my wife, Anita Bhetwal, for her incredible support, suggestions, and feedback to accomplish this project.

Lastly, I offer my regards and blessing to all of those who supported me in any respect during the completion of the project.

Table of Contents

List of Figures

List of Tables

Chapter – 1 Introduction

A data warehouse is a relational database that is designed for data analysis rather than for transaction processing. It usually contains historical data derived from transactional data, but it can include data from other sources. It has a separate analysis workload from a transaction workload and enables an organization to consolidate data from several sources. It is not just a single repository system, but rather enables decision support systems and knowledge based applications to support decision making and long term business plans. Data warehouse applications are designed basically to provide the business communities with accurate and consolidated information. The objective of Data Warehousing applications are not just for collecting data and reporting, but rather for analyzing, it requires technical and business expertise tools.

Project Objectives

Business intelligence (BI) mainly refers to computer-based techniques used in identifying, extracting, and analyzing business data, such as sales revenue by products and/or departments, or by associated costs and incomes. Business users and managers have difficulties in choosing appropriate BI tool because various kinds of BI tools are available in the marketplace. The purpose of this study was to help business users and managers to choose appropriate BI tools for their organization.

This OLAP review motivated the researcher to explore reporting functionality of different OLAP tools and to review available literature about data warehousing, OLAP, BI, and OLAP tools to help business managers in choosing appropriate tools for data reporting and analyzing. This study will review data warehouse environment and architecture, business intelligence concepts, OLAP and the related theories involved on it. As well as the concept of

data warehouse and OLAP, this study will also present comparative analysis of commonly used OLAP tools in Organization. This research will address the research question: What is the best available tool for Business Intelligence based on the evaluation criteria of a particular organization?

Project Scope

The scope of this study was to describe the work carried out to achieve the research aim, this report discussed the Data Warehousing environment and its architecture. Four different OLAP tools were compared. Due to the limited availability and access of software and database, this study focused on two main OLAP tools, Cognos & Business Objects, to analyze and compare them.

Chapter 2 – Review of Literature and Research

The main objective of a data warehouse is to help with the analysis of the data. The data warehouse provides access to different environments of an enterprise data. For instance, when we build a data warehouse about a company's accounting data then the data warehouse concentrates on a single subject (i.e. accounts of that company).

In the early 1990's, data warehousing was proposed as a general purpose solution to the problem of satisfying organizational management information needs (Moody & Kortink, 2000). Data warehouses are databases devoted to analytical processing. They are used to support decision-making activities in most modern business settings, when complex data sets have to be studied and analyzed (Bonofati et al, 2001). Inmon (1995, p.1) defined a data warehouse as "a subject oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process." Since then, data warehouses have become very large and single subject data marts have proliferated. Kimball (1996, p.5) defined a data warehouse as

"a copy of transaction data specifically structured for query and analysis.". A data warehouse

contains massive amounts of highly detailed, time-series data used for decision support.

Specialized software extracts data from operational databases, then summarizes, reconciles, and

manipulates it for business use (Breslin, 2004). A data warehouse is generally understood as an

integrated and time-varying collection of data primarily used in strategic decision making by

means of online analytical processing (OLAP) techniques. An organization must choose a set of

data warehouse design and maintenance tools from among scores of software tools commercially

available in market. Mimno (1997) stated that selecting data warehousing product is complex

because there are not any single vendor that provides a fully integrated set of products that

support all components of a data warehouse. As a result, data warehouses are typically built

using products from multiple vendors. It is difficult to decide on the combination of products that

are the best for an organization.

Data Warehouse Architecture

A data warehouse is a relational database that is designed for query and analysis rather

than for transaction processing. Data warehouse usually contains historical data derived from

transaction data, but it can include data from other sources. A data warehouse separates analysis

workload from transaction workload and thus enables an organization to consolidate data from

several sources (Oracle, 2002). In addition to a relational database, a data warehouse

environment includes an extraction, transformation, and loading (ETL) solution, an online

analytical processing (OLAP) engine, client analysis tools, and other applications that manage

the process of gathering data and delivering it to business users. Furthermore, the data in the

warehouse must be integrated and consistent. Data integration is a central problem in the design

of Data Warehouses and Decision Support Systems. When data passes from the different sources

to the Data Warehouse, possible inconsistencies and redundancies should be resolved, so that the

warehouse is able to provide an integrated and reconciled view of data of the organization. Data warehouse architecture is a description of the elements and services of the warehouse, with details showing how the components will fit together and how the system will grow over time. There is always architecture, either ad hoc or planned, but experience shows that planned architectures have a better chance of succeeding (Hadley, 2002). Zachman recommends a reasonable compromise consisting of a four-layer process: business requirements, technical architecture, standards, and products.

Data warehouses and their architectures vary depending upon the specifics of an organization's situation. Three common architectures are:

- Basic Data Warehouse Architecture

- Data warehouse Architecture with a Staging Area

- Data Warehouse Architecture with a Staging Area and Data Marts.

Figure 1 shows a simple architecture for data warehouse. End users directly access data derived from several source systems through the data warehouse.
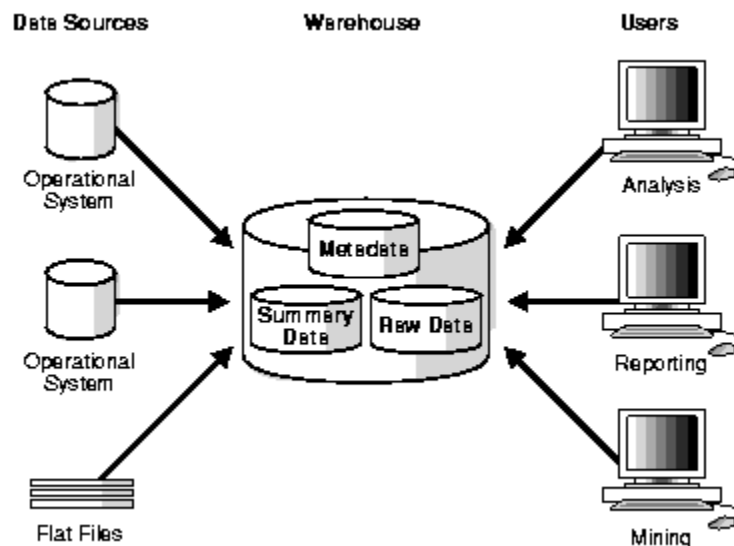
Figure 1: Architecture of Data warehouse

In basic data warehouse architecture, operational data needs to be cleaned and processed before putting it into the warehouse. Cleaning and processing the data can be done programmatically, although most data warehouses use a staging area to facilitate the process. A staging area simplifies building summaries and general warehouse management as shown in figure 2. Although the architecture in Figure 2 is quite common, data warehouse's architecture can be customized for different groups within an organization. This can be done by adding data marts, which are systems designs designed for a particular line of business.
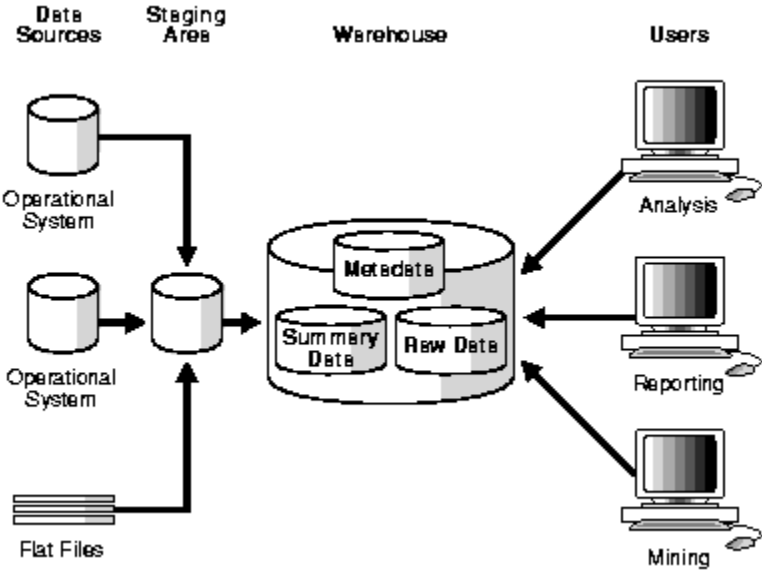
Figure 2: Data Warehouse Architecture with a Staging Area

Figure 3 illustrates an example where purchasing, sales, and inventories are separated. In this example, a financial analyst might want to analyze historical data for purchase and sales.
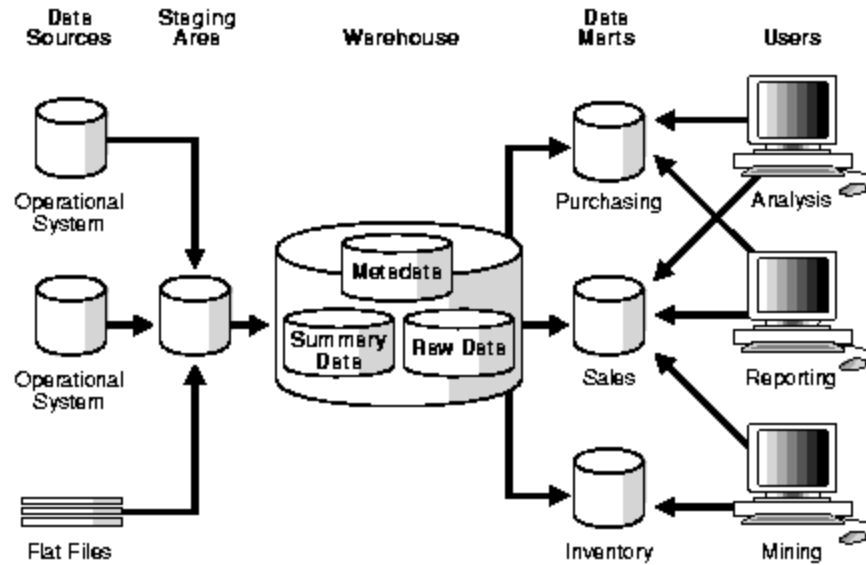
Figure 3: Data warehouse with Staging Area and Data Marts.

There are several advantages of this particular data warehouse architecture which are as follows (Handley, 2002):

- Provides an Organizing framework: The architecture draws the lines on the map in terms of what the individual components are, how they fit together, who owns what parts.

- Improved flexibility and maintenance: Allows us to quickly add new data sources, interface standards allow plug and play, and the model and metadata allow impact analysis and single-point changes.

- Faster development and reuse: Warehouse developers are better able to understand the data warehouse process, database contents, and business rules more quickly.

- Management and communications tool: Define and communicate direction and scope to set expectations, identify roles and responsibilities, and communicate requirements to vendors.

- Coordinate parallel efforts: Multiple, relatively independent efforts have a chance to converge successfully. Data marts without architecture become the stovepipes of tomorrow.

Furthermore, Handley also described other benefits that can include: Standardizing data across the organization, improving turnaround time for analysis and reporting, sharing data and allowing others to easily access data, supporting ad hoc reporting and inquiry, reducing the development burden of IS/IT, and removing information processing load from transaction-oriented databases.

Data warehousing improves the productivity of corporate decision-making by creating an integrated database of consistent, subject-oriented, historical data. It is integrated data from multiple incompatible systems into a form that provides one consistent view of the organization. By transforming data into meaningful information, a data warehouse allows corporate decision-makers to perform more substantive, accurate, and consistent analysis (Connolly & Begg, 2005).

Although there are many advantages of a data warehouse, there are some disadvantages to building a data warehouse. The major disadvantage is that a data warehouse can be costly to maintain and it can become a problem if the warehouse is underutilized (Power, 2010). Sometimes managers have unrealistic expectations about benefits of having a data warehouse.

Moreover, a data warehouse is a very complex system. Data warehousing can complicate business process significantly. For example, small changes in a transaction processing system may have major impacts on data warehouse. Sometimes, adding, deleting, or changing the data and the transaction can cause time consuming. The administrator needs to control and check the correctness of changing transactional data and the impact on other transactions. Therefore, complexity of data warehouse prevents the organizations from changing the data or transaction that are necessary to make.

Increasingly companies are recognizing the value of an enterprise data warehouse. A data warehouse provides a single 360-degree view of the business and a powerful platform for a wide spectrum of business intelligence tasks ranging from predictive analysis to near real-time strategic and tactical decision support throughout the organization.

ETL process in data warehousing

ETL stands for extract, transform, and load, the processes that enable companies to move data from multiple sources, reformat and clean it, and load it into another database, a data mart or a data warehouse for analysis, or on another operational system to support a business process.

Companies know they have valuable data lying around throughout their networks that needs to be moved from one place to another – such as from one from one business application to another or to a data warehouse for analysis. The only problem is that the data lies in all sorts of heterogeneous system, and therefore in all sorts of formats. For instance, a CRM system may define a customer in one way, while a back-end accounting system may define the same customer differently (Songini, 2004). ETL extracts data from outside sources and transforms the data to fit the business needs by placing it in a data warehouse.

Due to the wide range of possible data inconsistencies and the sheer volume of organization's data, cleansing is considering to be one of the biggest problem for the data warehousing (Rahm, & Do, 2000). During the ETL process, illustrate in Fig. 4, further data transformations deal with schema/data translation and integration, and with filtering and aggregating data to be stored in the warehouse. As indicated in fig. 4, all data cleaning is typically performed in a separate data staging area before loading and transformed data into warehouse. A large number of tools of varying functionality are available to support these tasks, but often a significant portion of the cleansing and transformation work has to done manually or by low-level programs that are difficult to write and maintain.
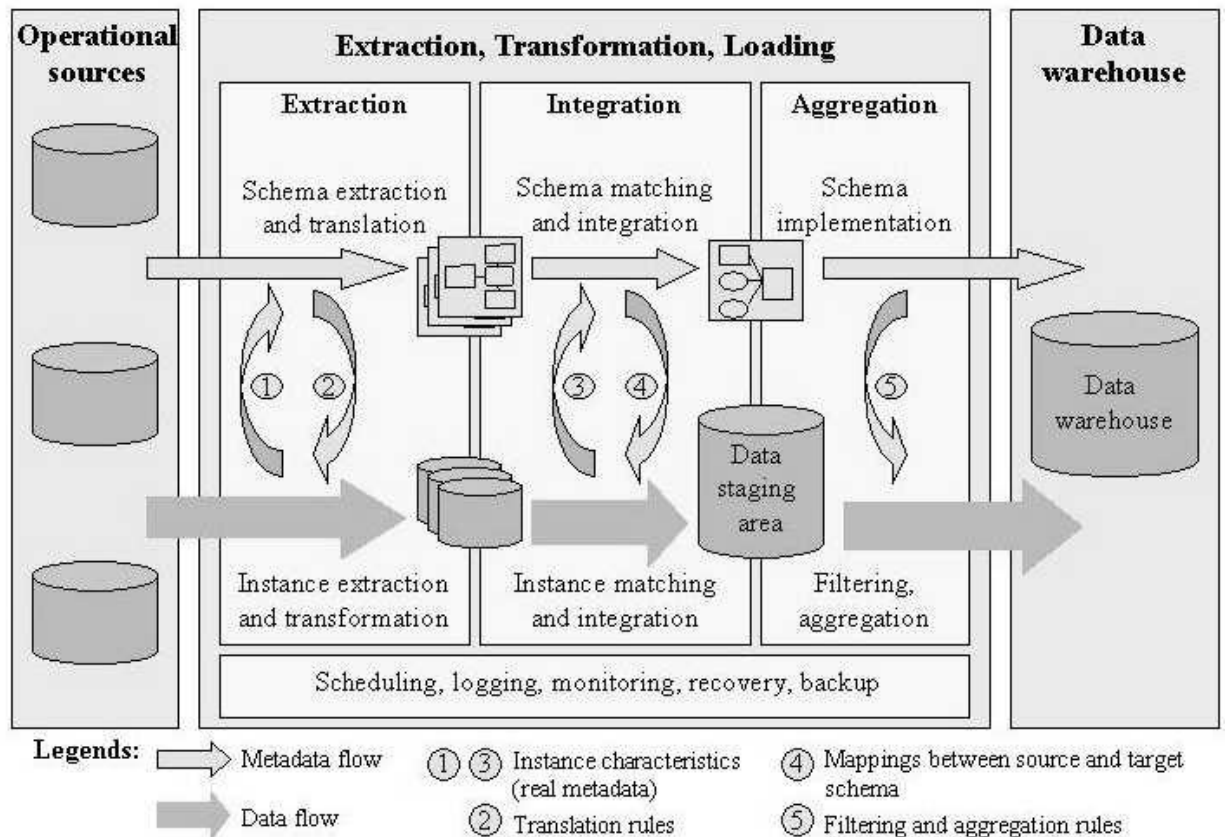
Figure 4: Steps of building a data warehouse: the ETL process

In ETL process, the first step is to extract data from different source systems. Each source system may use a different format and different types of data organization. The commonly used data sources format are relational database and flat files (Codd 1990). In this step data is extracted in a format for transformation processing. In the next step of transform phase, a serous of rules are applied data to drive the data to be loaded. Manipulation of data is done in different transformation type depending on the data sources. In the last phase, data is loaded into the data warehouse depending on the requirements of the company. This process may vary from organization to organization. In some instances, the ETL may overwrite the old information with new data and in some cases may maintain a history and audit trail of all changes of the data.

Data Warehouse Schema Architecture

For the Data Warehouse environment, data is usually transformed from relational data model into some special architectures. The following are the most commonly used schema models designed for data warehousing:

- Star schema

- Snowflake schema

The determination of which schema model should be used for a data warehouse should be based upon the analysis of project requirements, accessible tools and project team preferences.
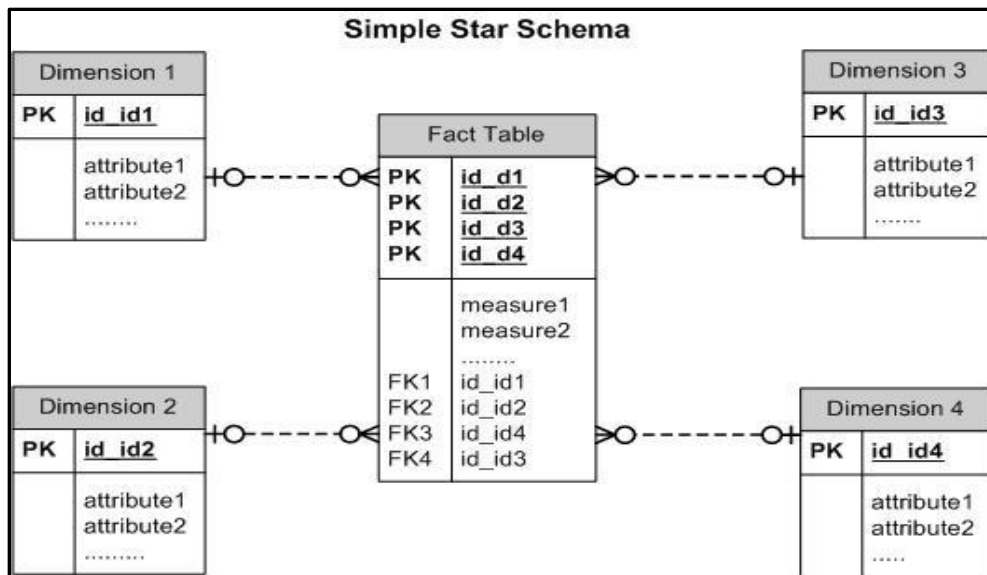
Figure 5: Simple Star Schema

The star schema architecture is the simplest data warehouse schema. Star schema is called a star schema because the entity relationship diagram of this schema resembles a star in the center which consists of a large fact table and end points that are dimensional tables (Kimball & Ross, 2005).

The snowflake schema is a more complex variation of the star schema in a data warehouse, because the tables which describe the dimensions are normalized.
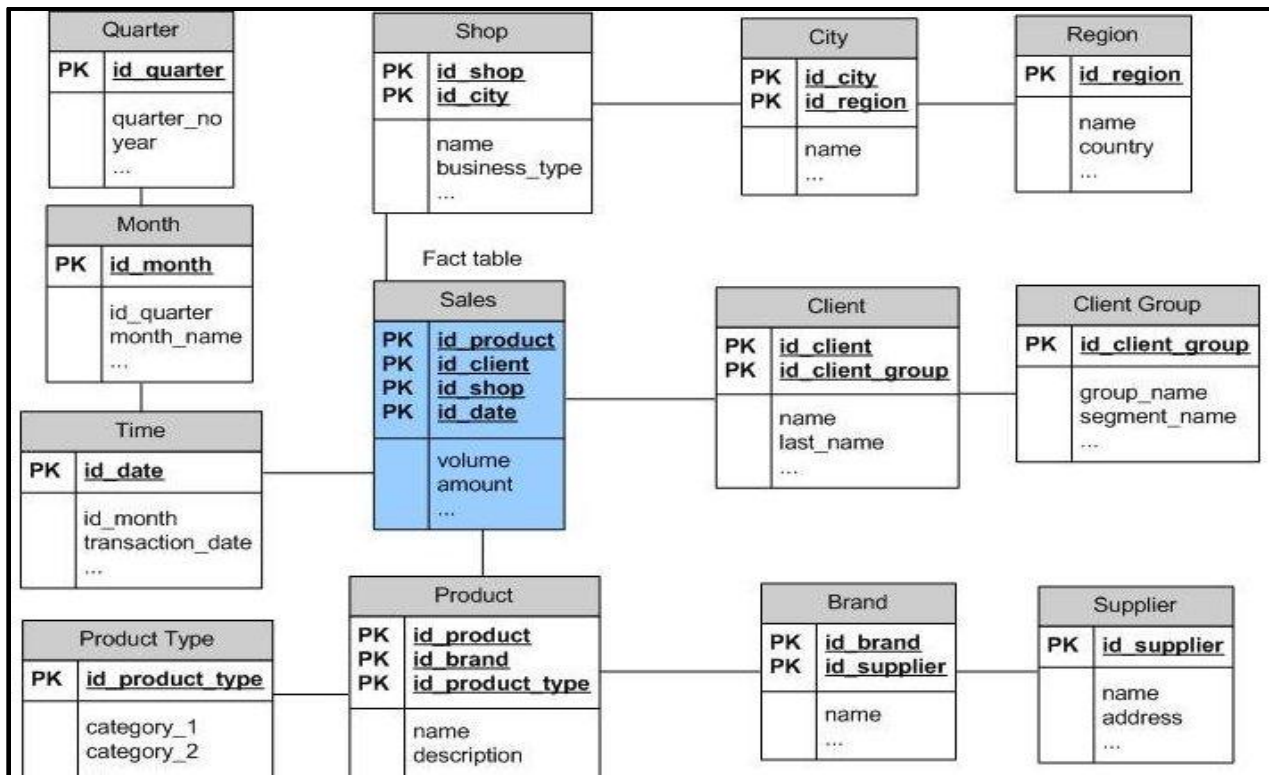
Figure 6: Snowflake schema.

Snowflake schemas provide a refinement of star schemas where the dimensional hierarchy is explicitly represented by normalizing the dimension tables (Chaudhuri & Dayal 1997), as shown in figure 6. A snowflake schema is a star schema with all hierarchies explicitly shown. A Snowflake schema can be formed from a star schema by expanding out (normalizing) the hierarchies in each dimension (Moody & Kortink, 2000).

Business Intelligence

Business intelligence (BI) concept is popular and is a powerful technology used to turn data into meaningful information. BI is closely related to data warehousing. Business intelligence has applied the functionality, scalability, and reliability of modern database management systems to build ever-large data warehouses, and to utilize data mining techniques to extract business advantage from vast amount of available enterprise data (Cody, Kreulen, Krishna, & Spangler, 2002). BI usually refers to the information that is available for the enterprise that facilitates

decision making. BI is now a widely used term, to describe analytic applications. BI is currently

the top-most priority on many chief information officers because it provides present business

information that supports the business decision making process (Watson, H. J. & Wixom, B. H.,

2007). For our purpose, we will discuss business intelligence in the context of using a data

warehouse infrastructure. To achieve business intelligence, the proper tools must be selected.

The most commonly used Business Intelligence technologies are Online Analytical Processing

(OLAP) and reporting tools for analyzing the data. These technologies allow as user to make

tactical decision for the better performance of the organization, and provide quick and fast access

to an end user request.

Figure 7 shows that BI is a process that includes two primary activities: getting data in

and getting data out.
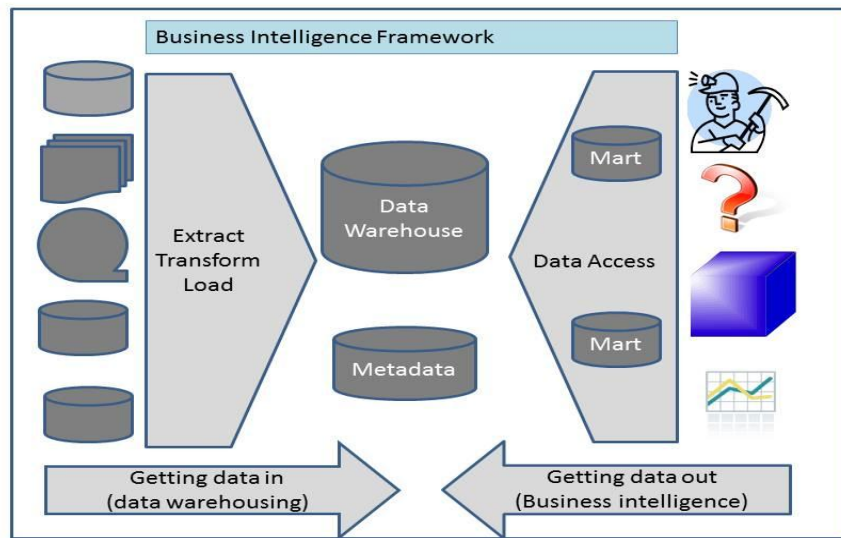
Figure 7: Business intelligence framework.

Getting data in delivers limited value to an enterprise; only when users and applications

access the data and use it to make decisions, does the organization realize the full value from its

data warehouse. Thus, getting data out receives the most attention from organizations. This

second activity, which is commonly referred to as BI, consists of business users and applications according to data from the data warehouse to perform enterprise reporting, OLAP, querying, and predictive analytics. BI lets companies analyze large amounts of information collected while doing business to capture trends, gain insights, and draw conclusions about the organization.

The term BI is used to characterize a range of technologies, software platforms, specific application and processes. Since BI is an all-round concept, it can be viewed from three different perspectives:

- Provides a reasonable approach to management

- Data can be converted into information

- Allows for making better and faster decisions

Most of the venders have front end presentation layer which allows users to create reports. Users should check these functionalities while evaluating reporting tools:

- Data source connecting capabilities

- Scheduling and distribution capabilities

- Security features

- Customization

- Export capabilities

Business intelligence applications have become among highest-priority technology investments for most Chief Information Officers (CIOs) because organizations see BI as more than just a way of measuring and understanding the past performance of an isolated business activity (Hostman, Rayner, & Friedman, 2006). The role of BI is becoming more pervasive and is affecting the way information is used, analyzed and applied. As a result, organizations can lead, decide, measure, manage and optimize performance to achieve greater efficiency and

financial benefits. BI is now becoming key to the better management of performance associated with multiple dimensions of an organization and its business processes.

Online Analytical Processing (OLAP)

Online analytical processing (OLAP) is a business intelligence tool that allows a business person to analyze and understand particular business drivers in factual terms. OLAP allows us to do analysis very quickly. The speed is achieved by the transformation of data to a highly indexed, compact format that is purpose-built to optimize the performance of multidimensional queries. Bedell (1998) stated that the core challenge associated with enterprise OLAP are how to structure the database and how to build a decision support architecture that allows for complex analysis and qualification. He further mentioned that OLAP tools must remove data size limitations and data structures requirements while providing a means of delivering advanced analytical functionality even for large data volumes. The key to overcoming these challenges lies in achieving tighter integration between analysis tools and databases.

OLAP tools are geared towards slicing and dicing the data. OLAP tools are usually used by advanced users, and the tools are frequently used in business, science and health care to extract useful knowledge from massive databases. Chaudhuri & Dayal (1997) found that data warehousing and online analytical processing are essential elements of decision support, which has increasingly become a focus of the database industry. Many commercial products and services are now available, and the principle database management system vendors now have offerings in these areas. Watson and Wixom (2007) further stated that business intelligence is a process that includes the activities of getting data in and getting data out. The activity of getting data out consists of business users and applications accessing data from the data warehouse to perform enterprise reporting, OLAP, querying, and predictive analytics. Lawrence (2006) described an important and difficult optimization problem in OLAP data warehouses is the view

selection problem, consisting of selecting a set of aggregate views of the data for speeding up future query processing. A common variant of the view selection problem addressed in the literature minimizes the sum of maintenance cost and query time on the view set. To meet the needs of business users, information can be sliced and diced through warehouse data with any means and ability. A dynamic report allows users to view warehouse data from different angles, at different levels of detail. The ready availability of different data views also improves business analysis by reducing the time and effort to collect format and filter information from data.

More recently, relational database vendors have also sold their databases as tools for building data warehouses. A data warehouse stores tactical information that answers "who?" and "what?" questions about past events. For example, a typical query submitted to a Data Warehouse is: "What was the total revenue for the eastern region in the third quarter?"

While OLAP systems have the ability to answer "Who?" and "What?" questions, it is their ability to answer "What if?" and "Why?" that sets them apart from Data Warehouses. OLAP enables decision-making about future actions. A typical OLAP calculation is more complex than simply summing data. For example: "What would be the effect on soft drink costs to distributors if syrup price went up by $.10/gallon and transportation costs went down by $.05/mile?" OLAP reporting integrates the complex issues of an enterprise data structures, procedures and algorithms into dimensional data structures. Using OLAP tools, business users can easily navigate the predefined reports and read the business data to build new reports with minimum support from IT professionals.

OLAP Applications

There are many examples of OLAP applications in various functional areas as listed in Table 1 (OLAP Council, 2001).

Functional area	Example of OLAP applications
Finance	Budgeting, activity-based costing, financial performance analysis, and financial modeling
Sales	Sales analysis and sales forecasting
Marketing	Market research analysis, sales forecasting, promotions analysis, customer analysis, and market /customer segmentation
Manufacturing	Production planning and defect analysis

Table 1: Example of OLAP applications in various functional areas.

OLAP applications span a variety of organizational functions. Finance departments use OLAP for applications such as budgeting, activity-based costing (allocations), financial performance analysis, and financial modeling. Sales analysis and forecasting are two of the OLAP applications found in sales departments. Among other applications, marketing departments use OLAP for market research analysis, sales forecasting, promotions analysis, customer analysis, and market/customer segmentation. Manufacturing uses OLAP applications for production planning, defect analysis, demand planning, and forecasting. OLAP tool can be useful to all departmenst of an organization to reach the organizational goals.

Important to all of the above applications is the ability to provide managers with the information they need to make effective decisions about an organization's strategic direction. The key indicator of a successful OLAP application is its ability to provide information as needed, i.e., its ability to provide "just-in-time" information for effective decision-making.

In OLAP systems, the data model must be flexible. Any truly flexible data model ensures that OLAP systems can respond to changing business requirements as needed for effective decision making. Although OLAP applications are found in widely divergent functional areas, they all require the following key features (Forsman 1997):

- Multidimensional views of data

- Calculation-intensive capabilities

- Time intelligence

Multidimensional views represent the actual business model and they are limited to less than three dimensions. Managers usually look at the financial data by scenario that is actual verses budget and at sales by product and time. This view of data provides more than the ability to slice and dice, and gives flexible access to information for analytical processing. For example, database design should not influence which operations are to be performed on a dimension and managers must be able to analyze data at any level with ease.

The OLAP database can be verified by its ability to perform complex calculations; these databases must be able to do more than simple aggregation. For example, calculation of percentage of total from a top down prospective. While doing aggregation it should follow hierarchy analyses rather than simple data roll-ups. The key performance of a database involves algebraic equations, supporting algorithms such a moving averages and percentage growth of a given company.

Finally, time is an essential element of almost any analytical applications. A true OLAP system understands the sequential nature of time. Business performance is almost always judged over time, for example, this month vs. last month, this month vs. the same month last year. The time hierarchy is not always used in the same manner as other hierarchies. For instance, a manager might ask to see the sales for May or the sales for first five months of 1995. The same manager might also ask to see the sales for blue shirts but would never ask to see the sales for the first five shirts. Concepts such as year-to-date and period over period comparisons must be easily defined in an OLAP system.

OLAP Benefits

Successful OLAP applications increase the productivity of business managers, developers, and entire organizations. The inherent flexibility of OLAP systems means business users of OLAP applications can become more self-sufficient. Managers are no longer dependent on IT to make schema changes, to create joins, or worse. Perhaps more importantly, OLAP enables managers to model problems that would be impossible using less flexible systems with lengthy and inconsistent response times.

IT developers can be benefit by using the right OLAP software. Although it is possible to build an OLAP system using software designed for transaction processing or data collection, it is certainly not a very efficient use of developer resources. Faster delivery of applications using OLAP tools to reduce the applications backlog.

OLAP reduces the applications backlog still further by making business users self-sufficient enough to build their own models. However, unlike standalone departmental applications running on PC networks, OLAP applications are dependent on Data Warehouses and transaction processing systems to refresh their source level data. As a result, IT gains more self-sufficient users without relinquishing control over the integrity of the data. OLAP provides more efficient operations. By using software designed for OLAP, IT reduces the query drag and network traffic on transaction systems or the data Warehouse (Forsman 1997).

OLAP enables the organization as a whole to respond more quickly to market demands by providing the ability to model real life business problems and the more efficient use of human resources. OLAP applications have been most commonly used in the financial and marketing areas and their uses do extend to other functions. Data rich industries have been the most typical users (consumer goods, retail, financial services and transport) for the obvious reason that they had large quantities of good quality internal and external data available, to which they needed to add value. However, there is also potential to use OLAP technology in other industries. The

applications will often be smaller, because of the lower volumes of data available. Therefore,

OLAP applications can offer high-performance access to large amounts of pre-summarized data.

OLAP applications give users the power to retrieve answers to multidimensional business

questions quickly and easily.

OLAP System Components

An OLAP system is comprised of multiple components. A top-level view of the system

includes a data source, an OLAP server and a client. The data source is the source of data to be

analyzed. Data from the source are transferred or copied into the OLAP server where it is

organized and prepared to provide short query times. The Client is the user interface to the

OLAP server.

The data source in an OLAP system is the server that supplies the data to be analyzed.

Depending on the use of OLAP product, the source can be a data warehouse, a legacy database

housing corporate data, a collection of spreadsheets that hold financial data, or a combination of

any of the above.

The main component of OLAP is the OLAP server, which sits between a client and

database management systems (DBMS). The OLAP server understands how data is organized in

the database and has special functions for analyzing the data.

The client is what is used to view and manipulate data in the database. A client can be as

simple as a spreadsheet that incorporates OLAP features such as pivoting and drilling. It can be a

specialized yet still simple report viewer, or it can be as complicated as a custom-built

application designed for more complicated data manipulation. The World Wide Web is one form

of client.

An OLAP query tool allows companies to get to the bottom of complicated queries

through quick consolidating of information in data-rich matrices. The query tools have an easy-

to-use graphical interface, and allow users to create reports by dragging and dropping objects into the reports.

An OLAP system consists of multiple components, but client/server architecture is important for an OLAP product in the sense that it is the best way to deploy information to many users. In order for all analysts to produce coherent results, they should have access to the same data. It is also more efficient from the view of the database to have all of the data consolidated in one place.

OLAP Tools

There are many varieties of OLAP tools available in the marketplace. This choice has resulted in some confusion, with much debate regarding what OLAP actually means to a potential buyer and in particular what are the available architectures for OLAP tools. In 1993, E. F. Codd formulated twelve rules as the basis for selecting OLAP tools. The publication of these rules was the outcome of research carried out on behalf of Arbor Software and has resulted in a formalized redefinition of the requirements for OLAP tools (Connolly & Begg, 2005). Codd worked up these 12 rules for defining OLAP, and industry of software and data processing which allows consolidation and analysis of data in a multi-dimensional space. The following are the detailed description of Codd's twelve rules (Codd, Codd, & Sally, 1993):

1. Multi-dimensional conceptual view: OLAP tools should provide users with a multi-dimensional model that corresponds to users' views of the enterprise and is intuitively analytical and easy to use. User-analysts would view an enterprise as being multidimensional in nature – for example, profits could be viewed by region, product, time period, or other scenario (such as actual, budget, or forecast).

2. Transparency: The OLAP technology, the underlying database and architecture, and the possible heterogeneity of input data sources should be transparent to users. The

user should not be exposed to the source of the data supplied to the OLAP tool, which may be homogeneous or heterogeneous.

3. Accessibility: The OLAP tool should be able to access data required for the analysis from all heterogeneous enterprise data sources such as relational, non-relational, and legacy system.

4. Consistent reporting performance: Performance of the OLAP tools should not suffer significantly as the number of dimensions is increased.

5. Client-server architecture: The server component of OLAP tools should be sufficiently intelligence that the various clients can be attached with minimum effort. The server should be capable of mapping and consolidating data between disparate databases.

6. Generic dimensionality: Every data dimension should be equivalent in its structure and operational capabilities.

7. Dynamic sparse matrix handling: The OLAP server's physical structure should have optimal sparse matrix handling.

8. Multi-user support: OLAP tools must provide concurrent retrieval and update access, integrity and security.

9. Unrestricted cross-dimensional operations: The OLAP system must be able to support a group of users working concurrently on the same or different models of the enterprise's data.

10. Intuitive data manipulation: Slicing and dicing, drill down, and consolidation, and other manipulation should be accomplished via direct actions (point and click, drag and drop) on the cells of the cube.

11. Flexible reporting: Reporting facilities should present information in any way the user wants to view it.

12. Unlimited dimensions and aggregation levels: Depending on the business
requirements, an analytical model may be numerous dimensions, each having
multiple hierarchies. The OLAP system should not impose any artificial restrictions
on the number of dimensions or aggregation levels.

The above guidelines proposed by Dr. Codd form the yardstick for measuring any sets of
OLAP tools and product. A true OLAP system must conform to these guidelines. When an
organization is looking for OLAP tools, they can use these guidelines and select tools that meet
this set of criteria.

OLAP Architecture

OLAP tools are categorized according to the architecture used to store and process multi-
dimensional data. OLAP products are classified as either multidimensional OLAP (MOLAP), or
relational OLAP (ROLAP) based on the underlying architecture of the system. When combining
the strengths of ROLAP and advantages of MOLAP, a third option for implementing OLAP can
be created OLAP (HOLAP). Desktop OLAP (DOLAP) is based on the idea that a user can
download a section of the data from the database or source, and work with that dataset locally on
their desktop.

MOLAP is the more traditional way of OLAP analysis. In MOLAP, data is stored in a
multidimensional cube. The storage is not in the relational database, but in proprietary formats.
The advantages of this mode is that it provides excellent query performance and the cubes are
built for fast data retrieval. All calculations are pre-generated when the cube is created and can
be easily applied while querying data. The disadvantages of this model are that it can handle only
a limited amount of data. Since all calculations have been pre-built when the cube was created,
the cube cannot be derived from a large volume of data. This deficiency can be bypassed by
including only summary level calculations while constructing the cube. This model also requires

huge additional investment as cube technology is proprietary and knowledge base may not exist

in the organization. The typical architecture for MOLAP tools is shown in figure 8.
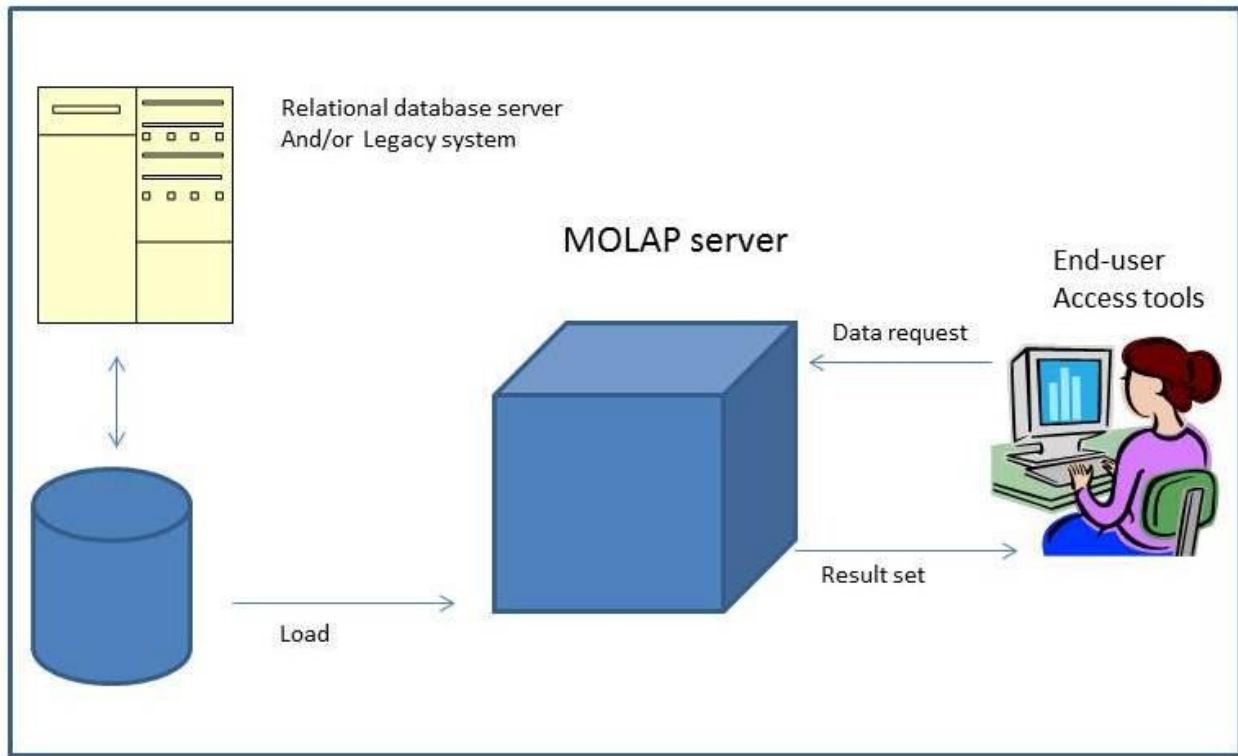
Figure 8: Architecture for MOLAP tools.

ROLAP is the fastest growing type of OLAP tool. The underlying data in this model is

stored in relational databases. Since the data is stored in relational database this model gives the

appearance of traditional OLAP's slicing and dicing functionality. The advantage of this model

is it can handle a large amount of data and can leverage all the functionalities of the relational

database. The disadvantages are that the performance is slow and each ROLAP report is an SQL

query with all the limitations of the genre. It is also limited by SQL functionalities. ROLAP

vendors have tried to mitigate this problem by building into the tool out-of-box complex

functions as well as providing the users with an ability to define their own functions. The typical

architecture for ROLAP tools is shown in Figure 9.
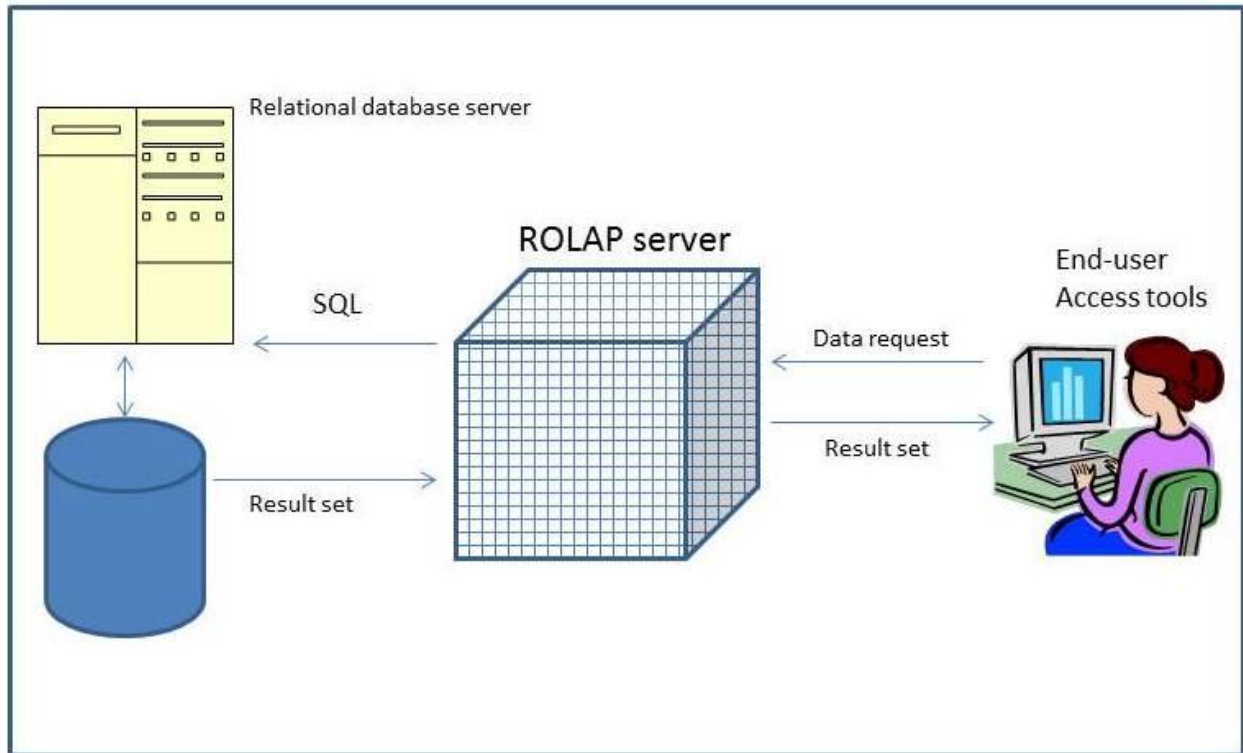
Figure 9: Architecture for ROLAP tools.

HOLAP technology tries to combine the strengths of the above two models. For summary type information HOLAP leverages cube technology and for drilling down into details it used the ROLAP model. It provides limited analysis capability, either directly against RDMS products, or by using an intermediate MOLAP server. HOLAP tools deliver selected data directly from the DBMS or via a MOLAP server to the desktop (or local server) in the form of data cube, where it is stored, analyzed, and maintained locally. The typical architecture of HOLAP tool is shown in Figure 10.
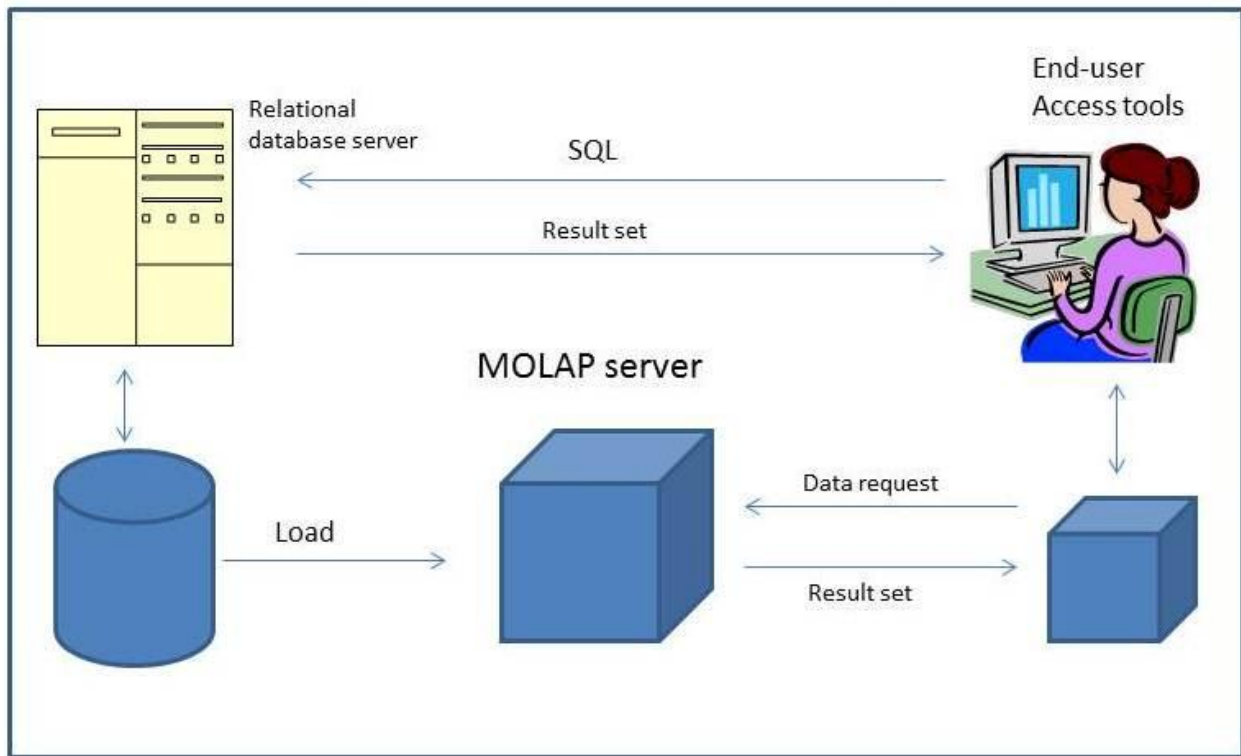
Figure 10: Architecture for HOLAP tools.

DOLAP tools store the OLAP data in client-based files and support multi-dimensional processing using a client multi-dimensional engine. Most DOLAP systems were developed as extensions of production system report writers, while others were developed in the early days of client/server computing to take advantage of the power of the emerging PC desktop. It is popular and typically requires relatively little IT investment to implement. They also provide highly mobile OLAP operations for users who may work remotely or travel extensively. However, most are limited to a single user and lack the ability to manage large data sets. The typical architecture of DOLAP tools is shown in Figure 11.
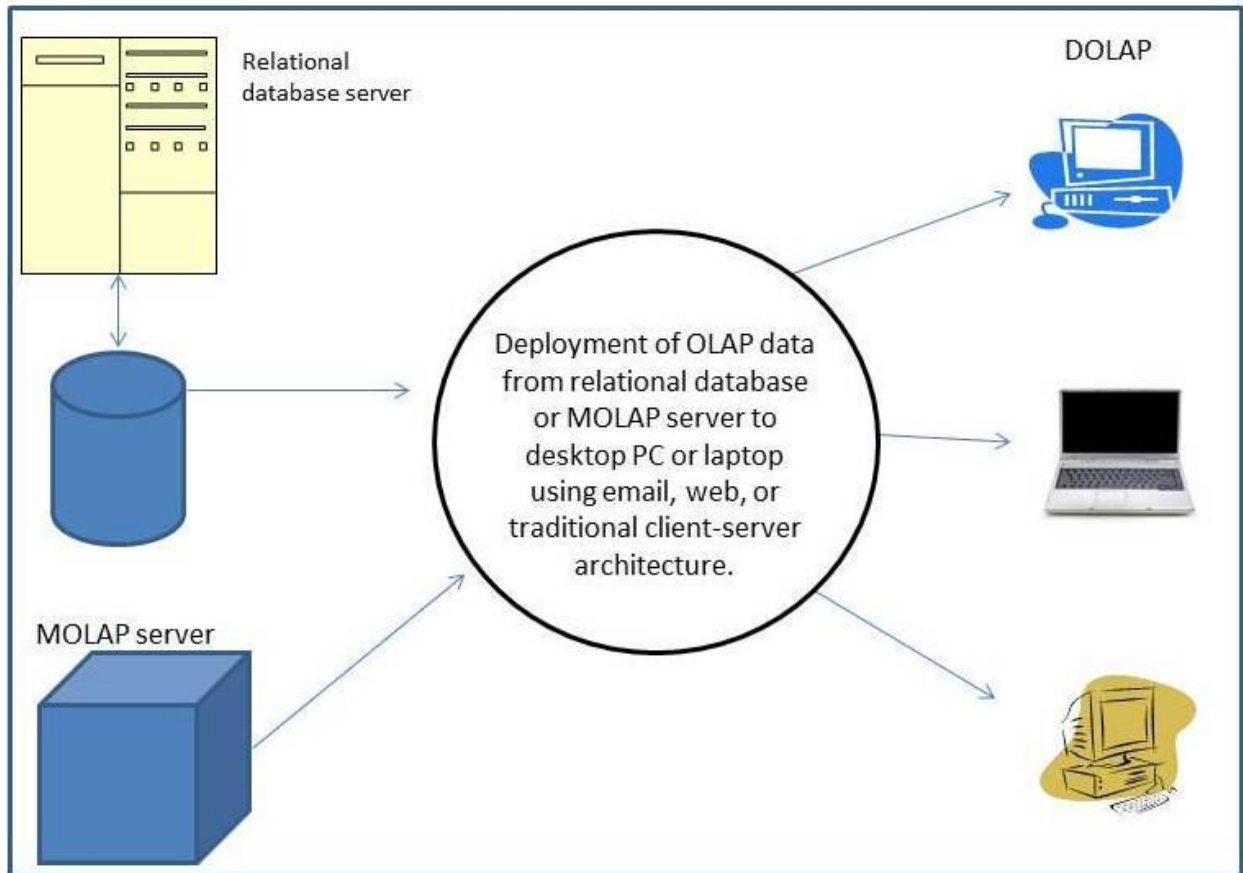
Figure 11: Architecture for DOLAP tools.

Chapter – 3 Research Methodologies

In chapter 2, data warehouse concepts, existing BI tools, OLAP, and OLAP tools were

discussed. To be successful in today's competitive, fast-paced business environment, it is

imperative that the right content is aggregated and delivered to the right people at the critical

movement when a decision must be made. This study will focus on business intelligence and

theories involved with it. The study also conducts comparative analysis of commonly used

OLAP tools. This study provides new information to the management and users to select OLAP

tools from a diverse and complex market place. In this chapter, the methods will be introduced

that were used to test and compare the OLAP tools.

Research Method

The thesis is divided into two parts. The first part is to conduct a literature review about data warehousing, business intelligence, and OLAP and ETL processing. The high-level review will help the researcher to understand background knowledge associated with various business problems. The second phase is to explore analytical and reporting functionality of some of the popular business intelligence tools such as Business Objects, Cognos, Hyperion, and MicroStrategy. For the first phase, a thorough literature study will be conducted. Literature sources like Google scholar, ACM Digital Library, IEEE explorer, and relevant articles and white papers from vendors will be used. For the second phase, the researcher will then explore the analytical and reporting functionality of BI tools. To accomplish this goal, the type of research will be explorative. The researcher will explore the tools and test the functionality by generating report. At the end of the research study, tools will be analyzed against selected criteria defined previously. The following approach will be used in the overall effort:

```
┌──────────────┐      ┌──────────────┐      ┌──────────────────┐
│ Problem      │ ───► │ Literature   │ ───► │ Design the strategy│
│ Formulation  │      │ Review       │      │                  │
└──────────────┘      └──────────────┘      └──────────────────┘
                                                     │
                                                     ▼
                                      ┌────────────┐      ┌────────────┐
                                      │ Execution  │ ───► │ Analysis   │
                                      └────────────┘      └────────────┘
```
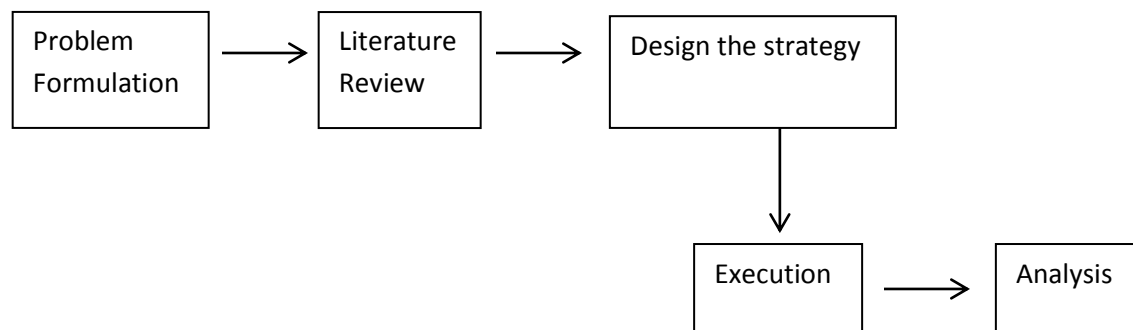
The strategy design phase will serve the purpose of creating factors to be analyzed at the end of the study and then planning how and which tools will be explored. In the execution phase, tools will be installed and configured for the evaluation process. The analyses were performed by correlated results of execution phase with selected factors.

Reporting Tools

Online Analytical Processing (OLAP) data analysis tools have drill-down, slice and dice and drill through capabilities. OLAP allows for dimensional analysis of our data and gives the analyst a sandbox for analyzing a series of measures across multiple dimensions. The most common applications of reporting tools are business operations reporting and dashboards. In this study, the researcher will implement different available OLAP tools. These are commonly used by advanced business users to look at the data from multiple dimensions. The main purpose of OLAP tools are slicing and dicing of data and they require a strong metadata layer and front end flexibility. For OLAP tools, the researcher will choose the most commonly used tools such as Cognos, Hyperion, Micro Strategy and Business Object (Pendse, 2008).

Oracle's Hyperion Essbase

In 2007, Oracle acquired Hyperion, a leading provider of performance management software. Oracle Essbase is the industry-leading multi-dimensional online analytical processing (OLAP) server, providing a rich environment for effectively developing custom analytic and enterprise performance management applications. Hyperion Essbase supports extraction from a wide range of datasources, including Oracle Database 11g and other relational database, and spreadsheets, and provides query access through several APIs and languages, including XML/A and MDX (Rittman, 2008). It is designed to support custom analytic and business intelligence applications, with a particular focus on enterprise performance management (EPM) applications. By leveraging its self-managed, rapid application development capabilities, business users can quickly model complex business scenarios. It supports forecasting, variance analysis, root cause identification, scenario planning and what-if modeling for both custom and package applications (Rayner, Chandler, & Decker, 2007). It can be tightly integrated with multiple data sources and the information generated can be delivered through a wide variety of reporting options. Oracle

Essbase supports extremely fast query response times for vast number of users, large data sets, and complex business models.
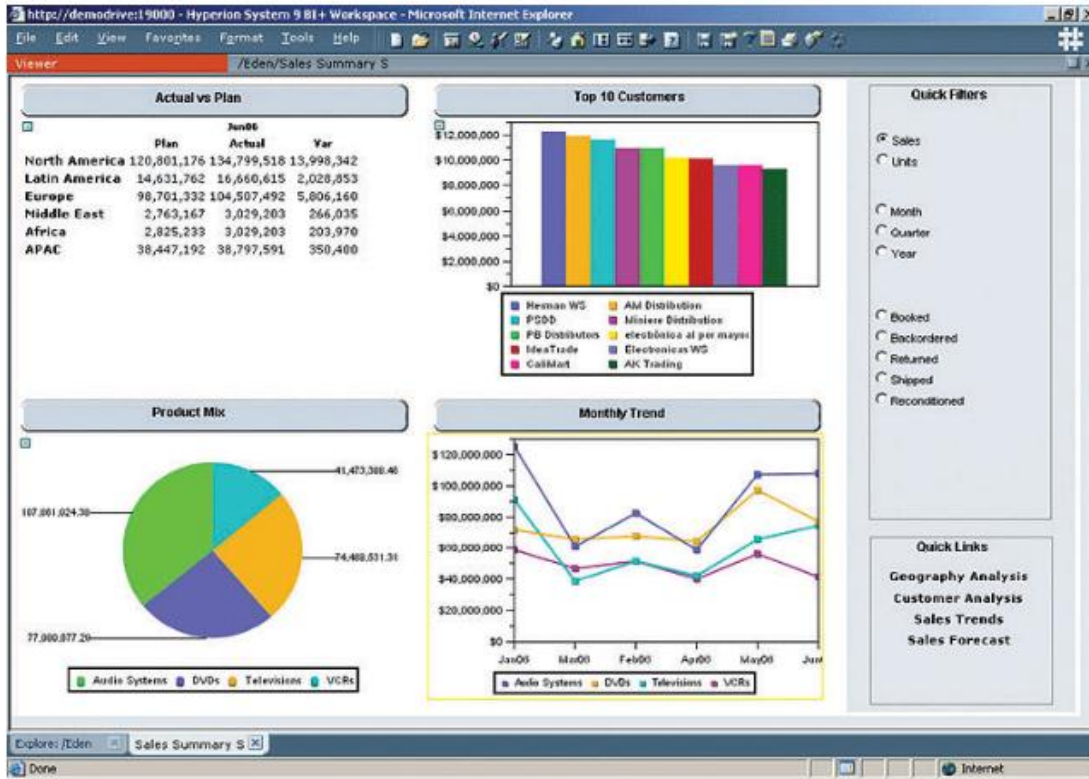
Figure 12: Oracle Essbase's Summary report.

Essbase not only stores data, but also incorporates one of the leading calculation engines for dong financial calculations such as allocations and eliminations. By putting the calculations in Essbase instead of in reports, everyone in an organization gets same answer because of there is a single version of the truth. Additionally, results automatically are displayed in reports, dashboards and ad-hoc queries.

Oracle's Hyperion Essbase tools are oriented for a business user, rather than a highly trained technician. It also provides the line of business with capabilities to build custom analytic applications. The typical buyer of Oracle's Hyperion Essbase, the line of business is the direct beneficiary of the EPM application and has the business expertise to design and implement it. The line of business benefits from a solution that is easily managed by business users and can be

deployed without extensive support from IT. The LoB that deploys an EPM application has special interest in:

1. Performance management applications.

2. What-if capabilities such as write-back and allocations.

3. Pre-built applications and Microsoft Excel query and data entry.

4. Tools appropriate for line of business users.

5. The ability to consolidate data from disparate systems, with limited support from IT.

The typical end user of Oracle's Hyperion Essbase comes from the business side, such as business analyst, financial analyst, quantitative end user, or professional in business or executive management. The end users of Oracle's Hyperion Essbase applications and business intelligence tools can be include anyone in the organization ranging from operations to executive management.

The typical front end of Oracle's Hyperion Essbase is Microsoft Office, including Excel, OBIEE+ (Oracle's Hyperion Reporting tools) and /or search engines such as Google OneBox or Oracle Secure Enterprise Search.

The data management of Oracle's Hyperion Essbase allows for the consolidation of data from multiple data sources including many different databases, flat file, excel and other financial sources.

The following are a few areas where savvy users of Oracle Essbase can extend it horizontally across multiple business functions:

- Operations: inventory control, forecast accuracy analysis

- Sales & Marketing: product profitability, spending planning

- HR: compensation trend analysis, timesheet tracking

- Manufacturing: build plans, bill of material analysis

- Finance: days sales outstanding analysis.

The following are the key benefits of Hyperion Essbase:

- Capable of storing huge amount of data with extremely fast response time

- Supports large number of end-users doing read-write queries

- Comes with over 350 out-of-the-box functions that accelerate development

- Built-in financial intelligence.

Oracle Interactive Reporting Studio

Essbase Studio is a single environment to update data source administration and access and supports the modeling of Essbase cubes from relational databases, OBI Suite EE, flat files, and the Dimension Server Library in EPM Architect. In addition, Essbase Studio supports several drill-through options: relational databases, Oracle Business Intelligence Enterprise Edition, URLs, custom SQL, and Java methods. Drill-through functionality is supported from data cells and member cells and is dynamically linked to cubes with matching metadata context with no interruption to application up time.

Essbase Studio helps organizations manage change through impact analysis and determine the data lineage of the metadata artifacts. With the support for lineage tracking through a rich graphical view of the metadata relationships, Essbase Studio allows users to follow application lineages to their metadata components and through to the data sources from which they were sourced. Essbase Studio supports multiple end-user roles and each role has its own access privileges on various types of studio catalog objects.

There are varieties of ways of report presentation. We need to think ahead about how we want our data to be displayed. The audience for our report may determine the best way to present

data. The following are the procedure for opening a new report in Hyperion Interactive Report

Studio:

Open Hyperion Interactive Reporting Studio by selecting Start/ All Programs/ Hyperion/

Reporting and Analysis/ Interactive Reporting/ Studio. The Welcome dialog box displays which

is shown in Figure 13. In the dialog box, click browse button to the lower right of the Recent

Documents area. Navigate to the Sample folder and open the file Sample1mod.bqy that shows in

figure 14. After clicking the open tab, the Sample1mod document opens in Hyperion. Select the

RegionReport section. If section catalog is not available, select the menupath
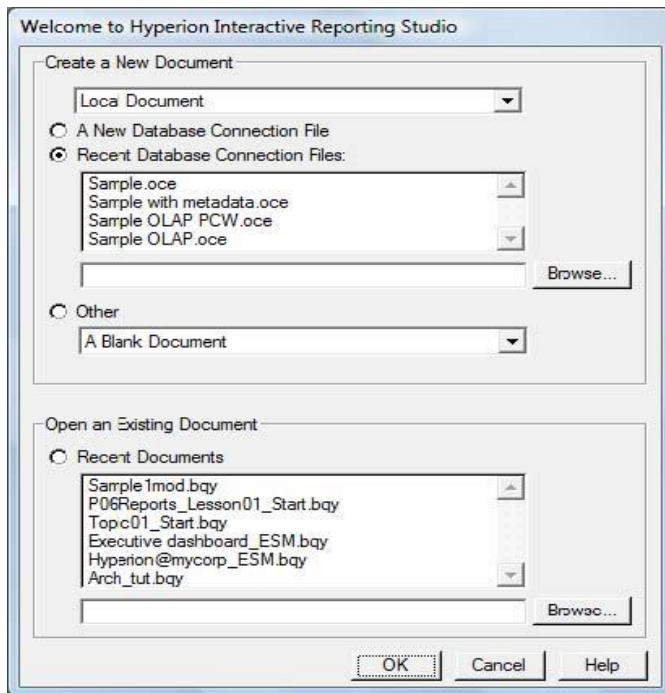
View/Section/Catalog.
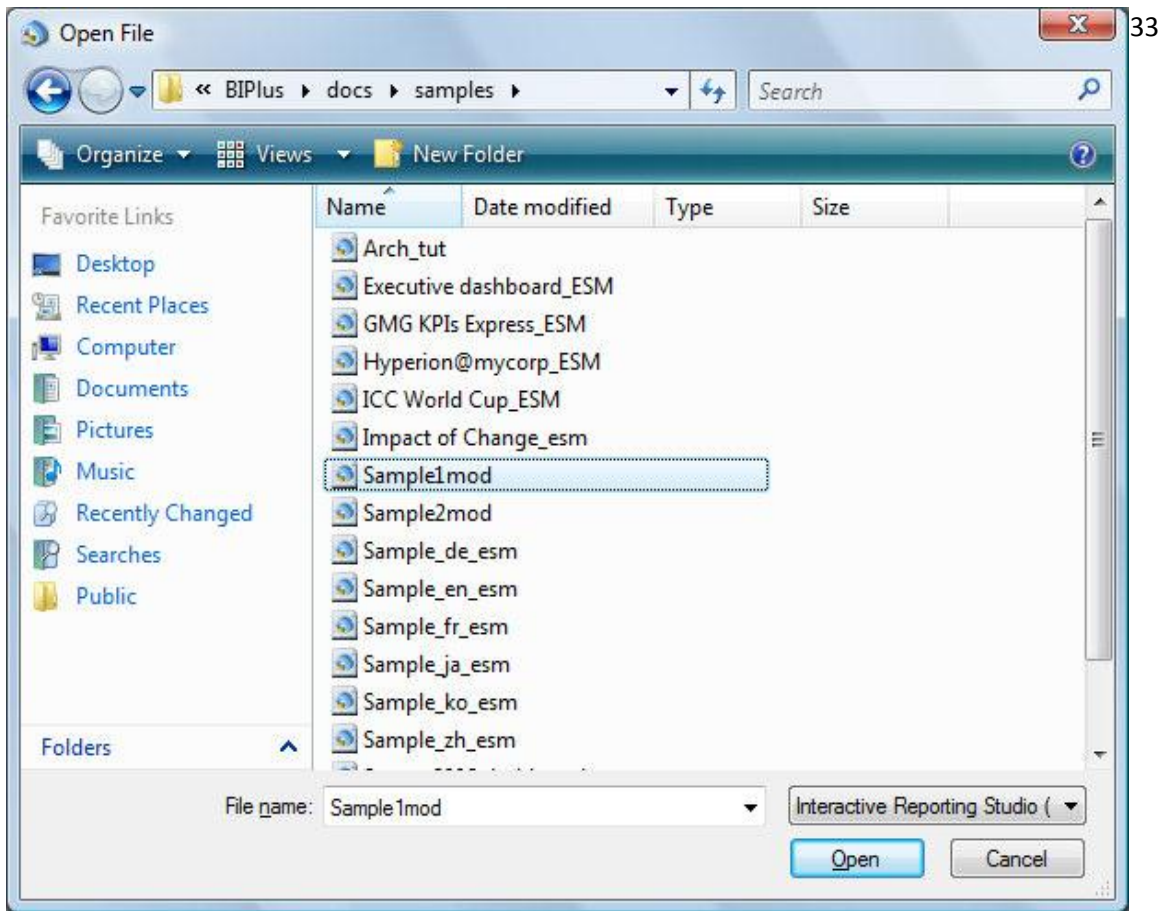

Figure 13: Welcome dialog box.
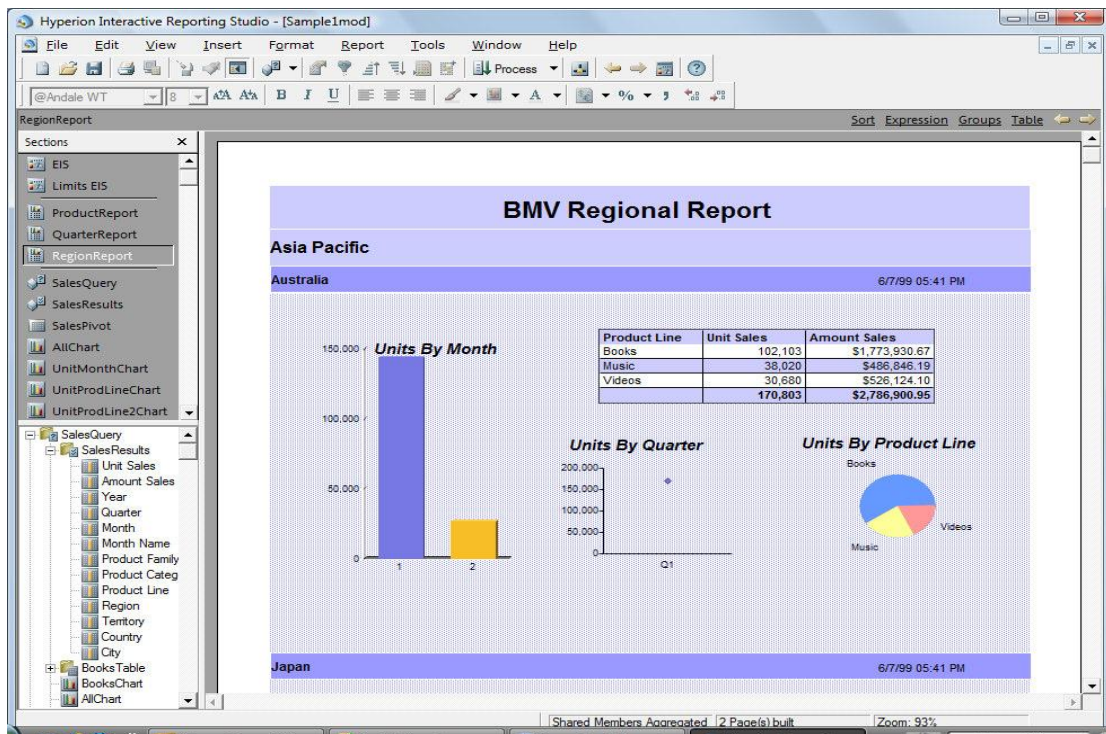
Figure 14: Open file dialog box.

Figure 15: Regional Report

Oracle's Hyperion Interactive Reporting helps individual and departments transform data from different sources into meaningful business intelligence. With powerful user-directed query and analysis capabilities via an easy-to-use interface with custom designed dashboards, we can get the information we need to drive good business decisions.

IBM Cognos Express

Cognos is IBM's business intelligence (BI) and performance management software suite. The software is designed to enable business users without technical knowledge to extract corporate data, analyze it and assemble reports. It delivers the essential reporting, analysis, dashboard, scoreboard, planning, budgeting and forecasting capabilities that midsize companies need at a price they can afford (IBM, 2010). Everything is included in a pre-configured solution that is easy to use and easy to buy. It allows us to unlock the silos of data across our company to provide consistent, meaningful information, so we can drive deeper and uncover insights and new opportunities then take action by aligning the right resources to capitalize on those opportunities.

IBM Cognos Express is a complete and modular solution. Start anywhere with the capabilities we need today and add other modules over time when we are ready. We can choose from IBM Cognos Express (IBM, 2009):

- Reporter for self-service reporting and ad hoc query
- Advisor for analysis and visualization
- Xcelerator for Excel-based business analysis and optimization, or
- Planner for planning, budgeting and forecasting.

Dashboard: IBM Cognos Express dashboards provide personalized, highly visual summaries of the information we need to monitor and manage our business (Wilson, 2011). It let us see how our company is doing and then drill down to analyze the reasons "why". In the figure 16, we see a personalized dashboard that contains a variety of reports, charts and gauges. The individual elements of the dashboard are also interactive. For example, click on a section of the pie chart and that section expands into its own pie chart with additional details. We can drill up or down to get the level of information we need. We can move sections around with simple drag-and-drop actions and resize items easily to give emphasis to different sections.

Figure 16: IBM Cognos Dashboard

Reporting: IBM Cognos Express has self-service reporting capability that empowers business users to access, modify and author their own reports quickly and easily. It offers broad report coverage with financial production, operational, transactional, managed or ad hoc reports to meet the reporting needs of any user. We can create reports quickly with easy-to-use prompts. Prompted reports provide a single reporting screen that lets users select the information they

want. As a business user, we have flexibility to view a report in a variety of formats, we can open

it as a PDF file, send it via email or open it in a Microsoft Excel. Using the ad hoc query

capability, novice users can also create reports of their own design from scratch.
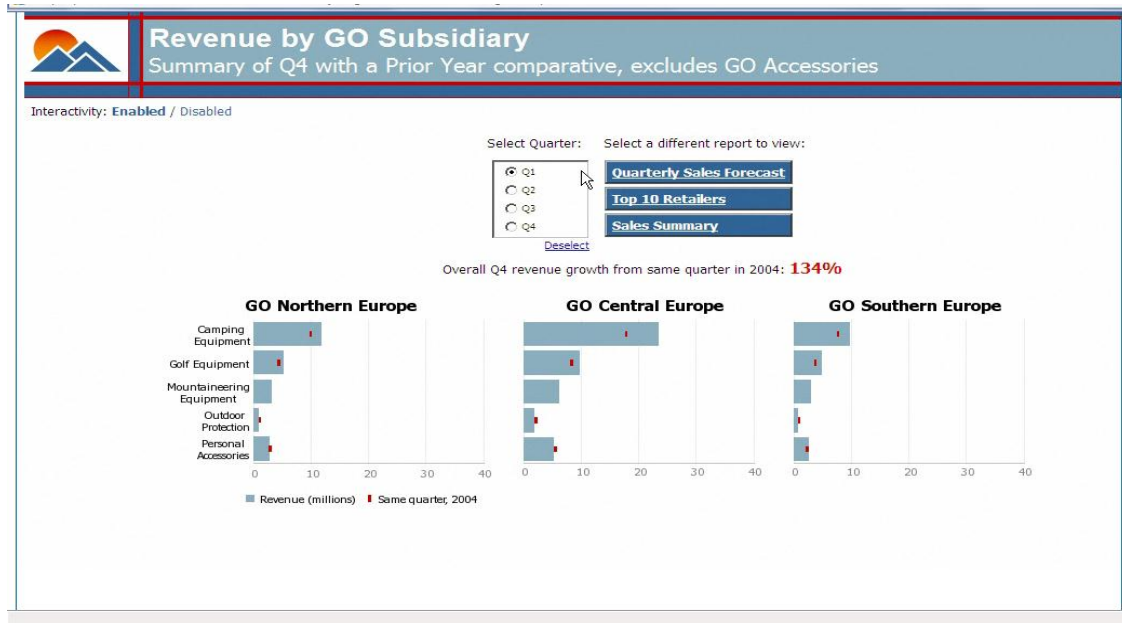
Figure 17: IBM Cognos Express, Summary Report.

Analysis: Analysis gives us the ability to answer the questions such as "Which are our

most profitable products, and why?" by analyzing complex data and then communicating

business insights with high-impact visualization. It also offers a multi-dimensional view of our

business. For example, we can group information about products, regions, customers or

employees- and use these dimensions in any combination to get the insight we need. We can

slice and dice data, display dimensions in any combination – and drill down for more detail.

Information is displayed the way we do business. Figure 17 shows a simple presentation of sales

in a chart and bar graph format. We can ask for a trend line to show gross margin percentages by

month- and then swap in another dimension (in this case, products), and drill down to see more

detail on individual product sales. We can publish results to the included Web portal by emailing

to a colleague or even by embedding live content into Microsoft Office applications such as

PowerPoint.

What-If Modeling: What-If Modeling transforms the familiar Excel front-end with powerful, in memory analytics engine that gives us fast access to centrally managed data that is trusted and secure. The modeling provides the options that can evaluate business options with multi-dimensional "what-if" analysis and optimize business outcomes by linking the insights gained through analysis to precise, action-based planning.

Planning: It automates the planning, budgeting and forecasting processes throughout company. It also provides a managed contribution model and workflow to streamline information gathering, aggregation and analysis. It transforms a company's planning processes, enabling business users to react quickly to changing business conditions and to grow their business effectively.
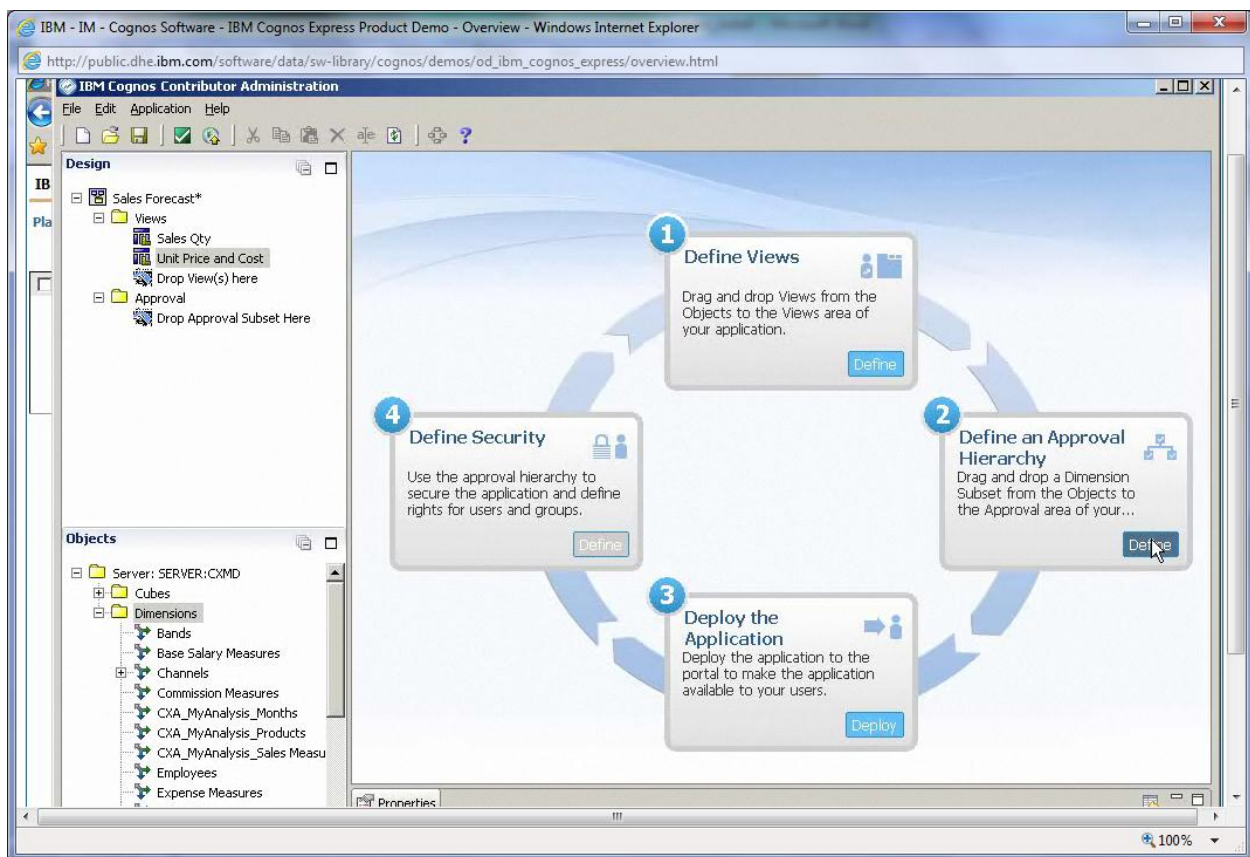

Figure 18: IBM Cognos Contributor Administrator.

Simple Management: It includes everything we need for immediate use in a preconfigured solution with no assembly required. It plugs into our existing infrastructure and

can be deployed within an hour. A simple Web-based console manages all administrative tasks.

There is no extra burden on IT staff and no additional investment in software infrastructure. It is

based on a set of integrated a core capabilities that are shared across all product modules. A

centralized, Web-based console manages all administrative aspects of installation, deployment

and ongoing management with push-button ease. The following figure shows a screen shot of

IBM Cognos Express Manager.
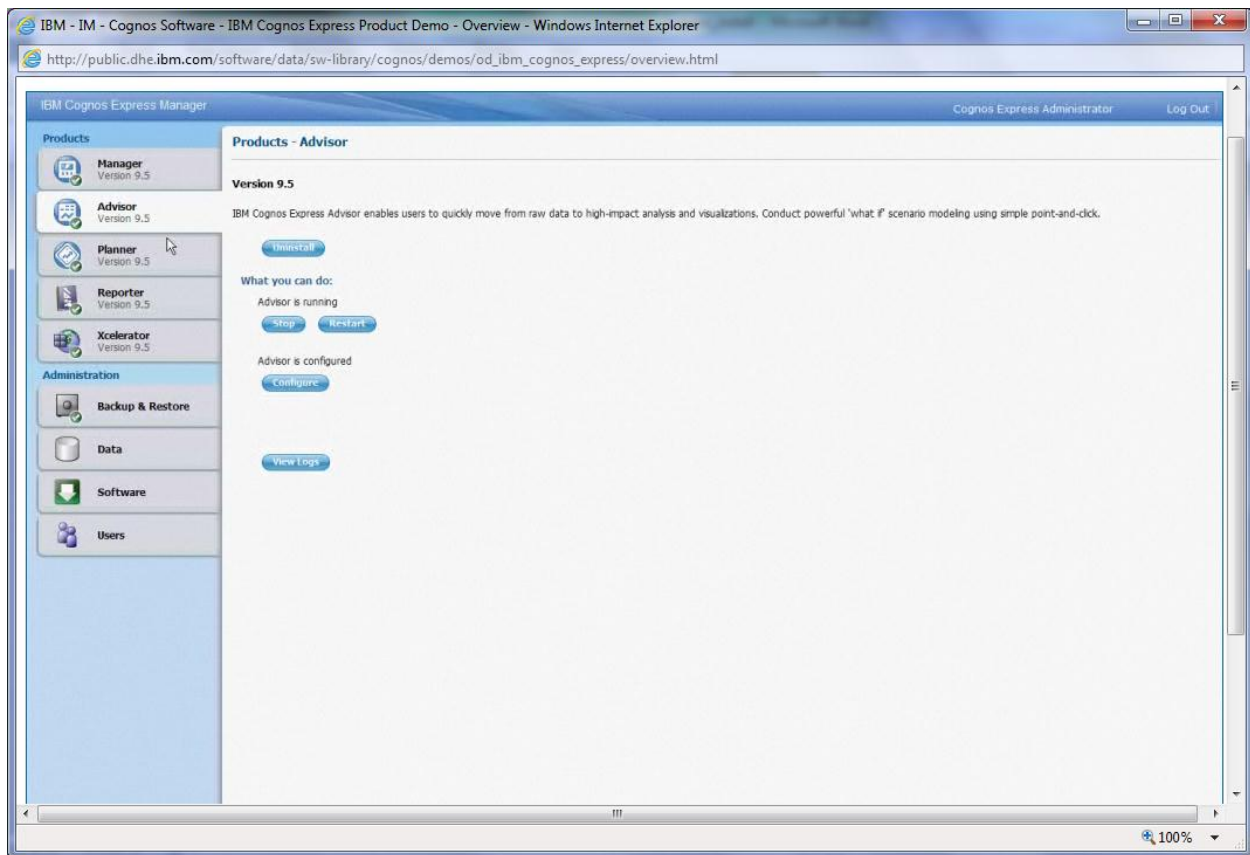
Figure 19: IBM Cognos Express Management.

MicroStrategy

MicroStrategy is a global leader in Business Intelligence. With MicroStrategy's BI

platform companies create reporting, analysis, and monitoring solutions to make more informed

decisions every day (Cover, 2011). MicroStrategy 8 inherits the industrial – strength attributes

that have become the hallmark of MicroStrategy technology- including the highest user

scalability, highest data scalability, supporting all 5 styles of Business Intelligence as plug-n-play

components, and an enterprise-caliber of non-stop operations and powerful administration (MicroStrategy, 2011). These five styles represent the complete spectrum of BI functionality required to support the monitoring, reporting, and analytical needs of each and every business user. The MicroStrategy platform delivers five styles of BI functionality to every user across the enterprise. The MicroStrategy provides the following five styles of BI functionality to an organization:

1. Scoreboards & Dashboards: It provides pixel-perfect report with the broadest appeal that conveys information at a glance for managers and executive.

2. Enterprise reporting: Print perfect report formats that delivers more detailed operational information than a scoreboard or dashboard in a concise format that can be consumed by all the business users.

3. OLAP Analysis: Slice and dice analysis with drilling, pivoting, page by and sorting capabilities for the manager or business user who need to perform analysis beyond standard operational reports.

4. Advance & Predictive Analysis: Full investigative query against the data warehouse down to the transaction level, allowing power users and professional analysts to perform extensive predictive and statistical analyses.

5. Alerts & Proactive Notification: Alerts and Reports are delivered on schedules or event triggers of database.

The infrastructure to support the five styles of Business Intelligence is provided by the sixteen different products. MicroStrategy Web is one of the products of MicroStrategy that allows us to report, analyze and monitor real world data from application modules provided in MicroStrategy BI Developer Kit.

The Following demonstration allows us to quickly see how MicroStrategy Web can be used in a business setting. The demonstration works with the States Force Analysis Modules that

is one of the analytic modules provided in this demonstration. There is a Sales Force Analysis

Module that will provide the sales-related data we need. Understanding our organization's sales

process is a vital aspect of an effective and successful business environment. This model is

designed to provide deep insight into the entire sales process. This insight in turn allows us to

increase lead conversions, optimize product lines, take advantage of our organization's most

successful sales practices and improve our sales organization's effectiveness.
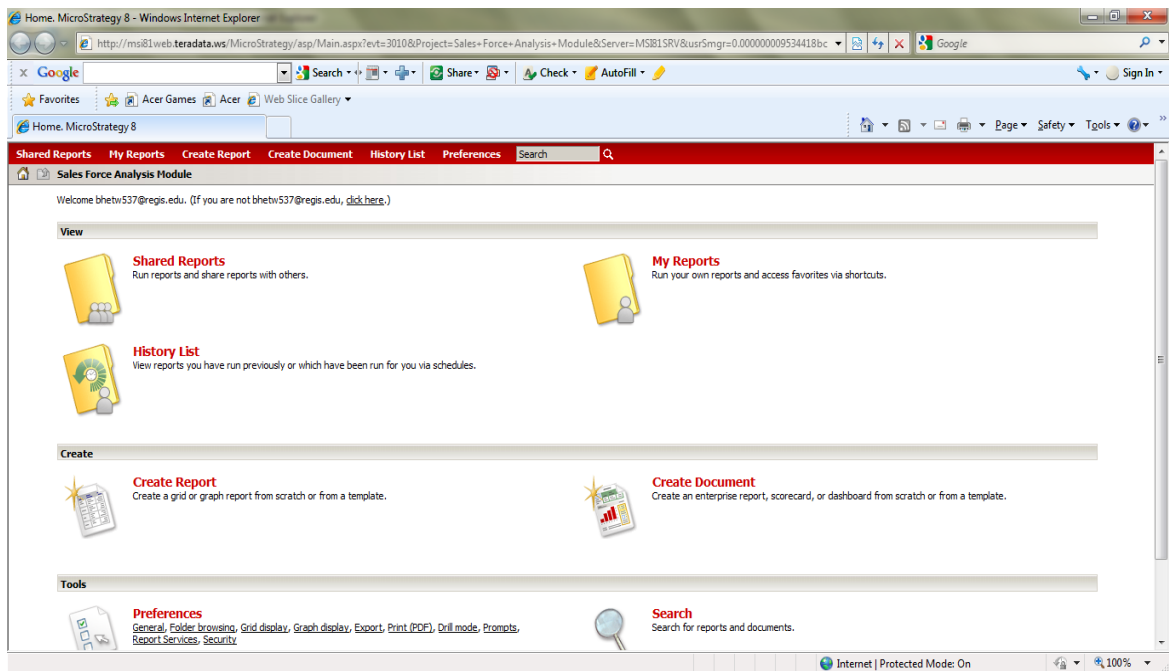
Figure 20: Sales Force Analysis Model.

The type of information we desire is available in Shared Reports, and the reports shared

by all users in our company. Now, we can click in the link Shared Reports to get the desire

reports. In the above figure 20, we see the icons Shared Reports and My reports. The Shared

Reports is the location where reports and documents are saved and shared with everybody, and

the My Reports is the location we can use for saving personal reports and documents. Personal

reports cannot be seen by any other user.

The model also provides the report for different time frame. We can choose the data

criteria for our report. We can choose date and quarter as the report parameters to see the report.

After clicking the Run Report, the report is displayed in our browser. The following figure 21 shows the report layout and display:

Figure 21: Report Layout and Display.

The module also provides drill-down facility for further into the data by clicking on the column-heading. Drilling allows us to see data at levels other than that of the Original report, enabling us to further investigate data in our report quickly and easily. We can save report and that saved report can be open later for review. By clicking in My Reports folder, we can see our saved reports in a personal folder. The following figure shows my report folder that includes two report files:
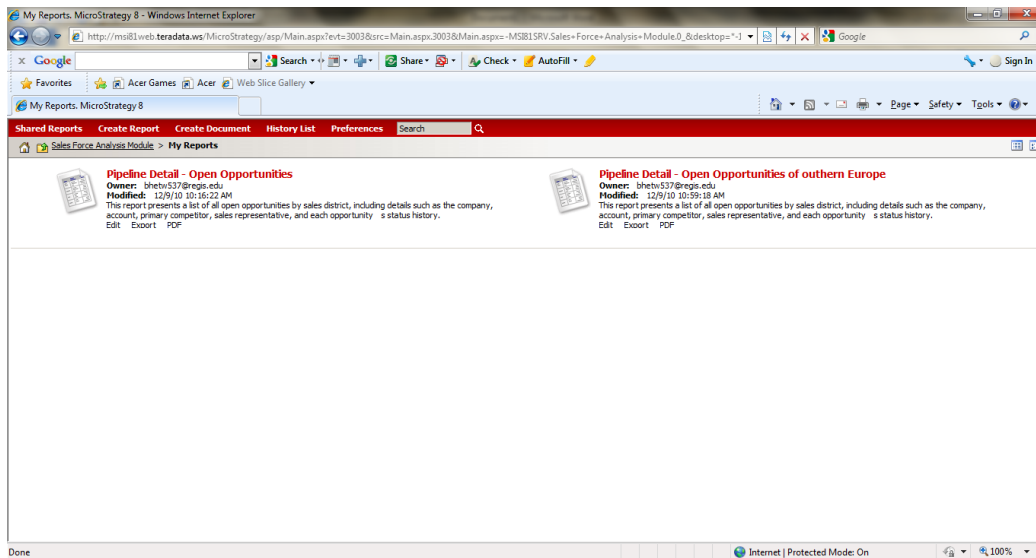
Figure 22: My Report Folder.

Business Objects (BO)

Business Objects is a software suite of report creation, viewing, and distribution tools. Business Objects is an independent worldwide BI vender and is recognized as a market leader by analyst organizations such as Gartner, Meta Group, and IDC. Business Objects has over 2,000 development and customer support employee worldwide (Kannan, 2008). Business Objects is composed of various components that are accessible through the Web or reside on our desktop computer. The major components are InfoView, WebIntelligence, and Business Objects. Business Objects delivers extreme insight through specialized end user tools on a single, trusted business intelligence platform. Business Objects include the industry's best reporting, query and analysis, and performance management dashboards, scoreboards, and applications. These integrated end user interface are supported by the most reliable, scalable, flexible, and manageable BI platform, and complemented by unified data integration (Cunningham, 2004). Business Objects XI is a complete suite of performance management, information management, reporting, and query and analysis tools built on Business Objects Enterprise XI.

Business Objects XI is the latest major release of the BI product suit from Business Objects, an SAP company. Building on the Business Objects XI platform, it delivers the

industry's most connected, interactive, and open BI solution for any organization. Designed to shed light on business problems, it provides substantial functional improvements and innovations across the Business Objects XI offers an easy to use and adopt solution that simplifies the way users share insights. It is open and works within any organization's environment, whether that is an SAP or Oracle application or a Teradata database.

SAP Business Objects BI solutions break down our existing data and process silos to provide business users with trustworthy and timely information no matter where our data resides. Even if the deployed software environment is highly heterogeneous, the solutions bring all our data together so it is both highly accessible and manageable.

SAP Business Objects help us achieve remarkable results by giving everyone in our organization the means to be highly effective. The solutions provide self-service access to critical information, so our people can achieve new insight, make better decisions, and act boldly like never before with minimal help from IT.

Each Business Objects user is assigned certain rights by Decision Support. The rights assigned to each user define the user's profile which defines access to:

- document lists

- document editors

- universes for creating and editing queries

- saving documents on the server

- sending documents to other users

- database connections

The Universe

A universe is a database interface, which maps everyday business terms to the data in database. The universe simplifies the process of certain reports by providing an easy way to see

and understand the data. It also provides automatic joins between database tables based on key

values. Universes consist of classes and objects. Objects represent fields in a database table.

Using the Report Wizard

The first time we use Business Objects we will be greeted with the Welcome screen of

the New Report Wizard. This wizard walks us through creating a new document.

1. Uncheck the Run this Wizard at Startup check box in the lower left-hand corner.

2. Select Generate a standard report for the report layout.

3. Click Begin to go to the Specific Data Access screen.
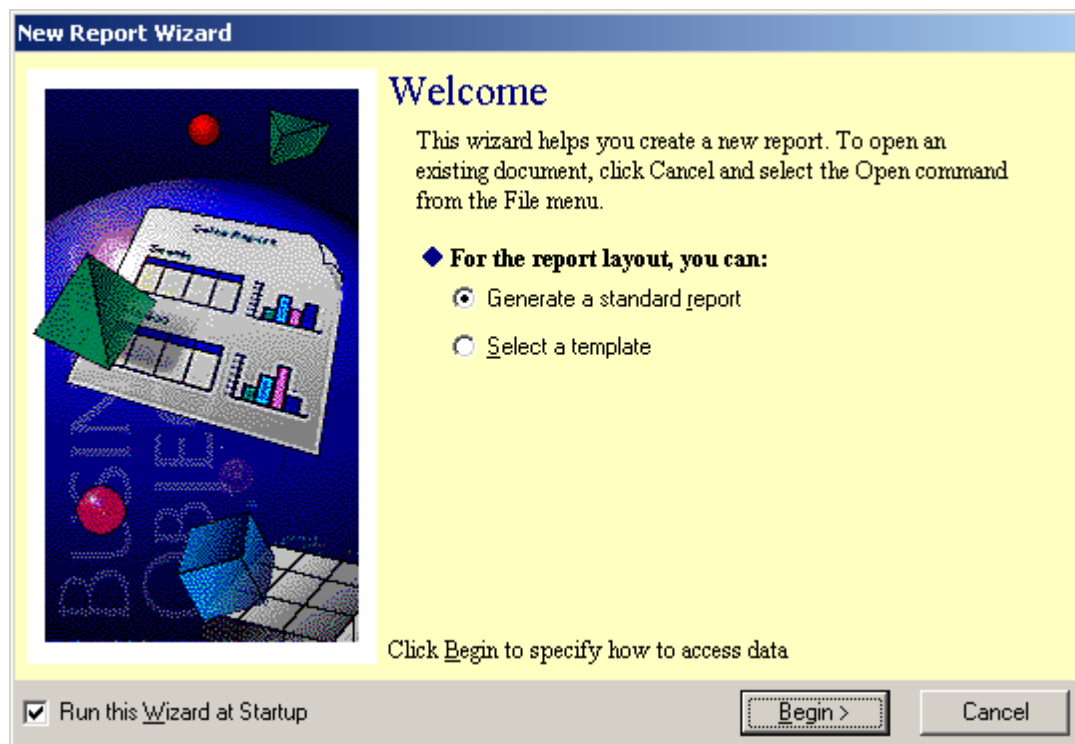
Figure 23: Welcome screen.

4. Select Universe as the way we want to access data. The select Universe screen will

    appear. The universes listed are ones we can access based on our security profile.
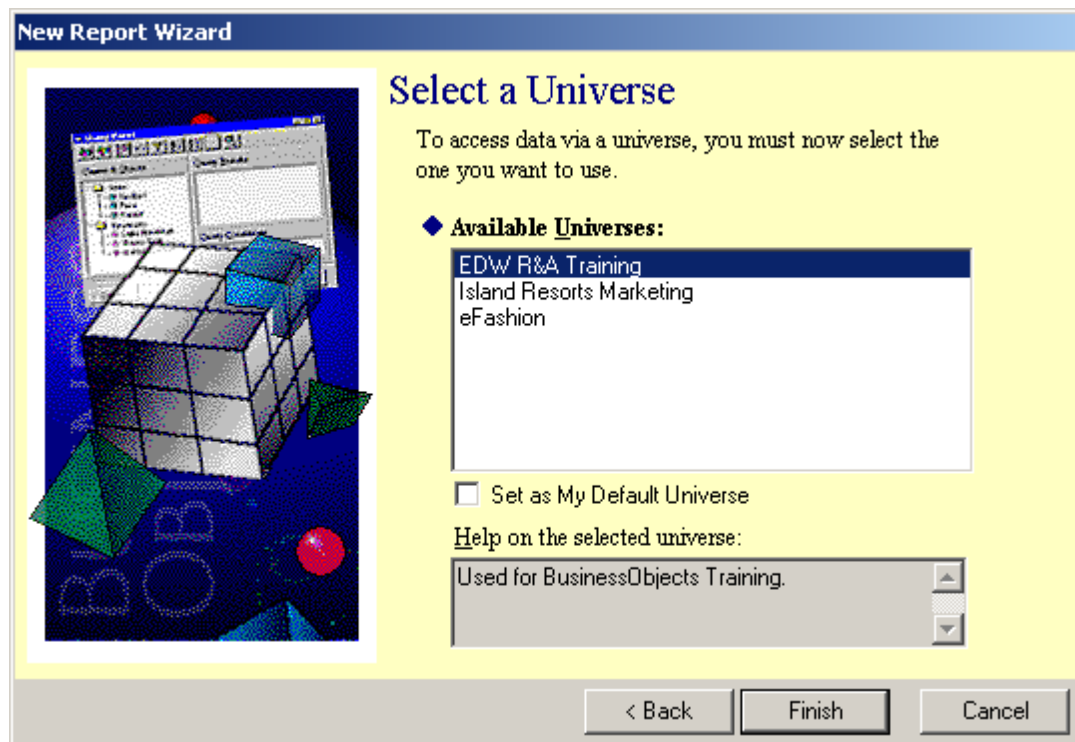
Figure 24: Select Universe Screen.

5. Select the eFashion universe. The Query Panel will appear. The query Panel is used to

    create the query that will retrieve data from the Enterprise Data Warehouse (EDW).

    The results of the query will be called a Data Provider. The data provider is stored in

    the document and becomes the source of the data displayed in the report.

The query Panel also known as the Data Provider window. The aim of the building a

query is to retrieve the data that we want to view and analyze in a report. The query panel also

lets build combined queries and sequences, as well as view, edit and print our query's SQL.

The query panel has three panels:

- Class and Objects: The universe is displayed in this panel. The classes are represented

    as folders. The class can be expanded by clicking the + or collapse by clicking the –

    located to the left of the folder.

- Result Objects: This panel holds the objects that are to be included in the report. Objects are moved from the Class and Objects panel to this panel.

- Conditions: A Selection criterion for the query is built in this panel. Clicking the funnel radio button at the button left-hand corner of the Query Panel will display pre-defined conditions. These are conditions that have been created by the universe designers and are stored in the universe.

Figure 25 shows the screen shut of the Query Panel.
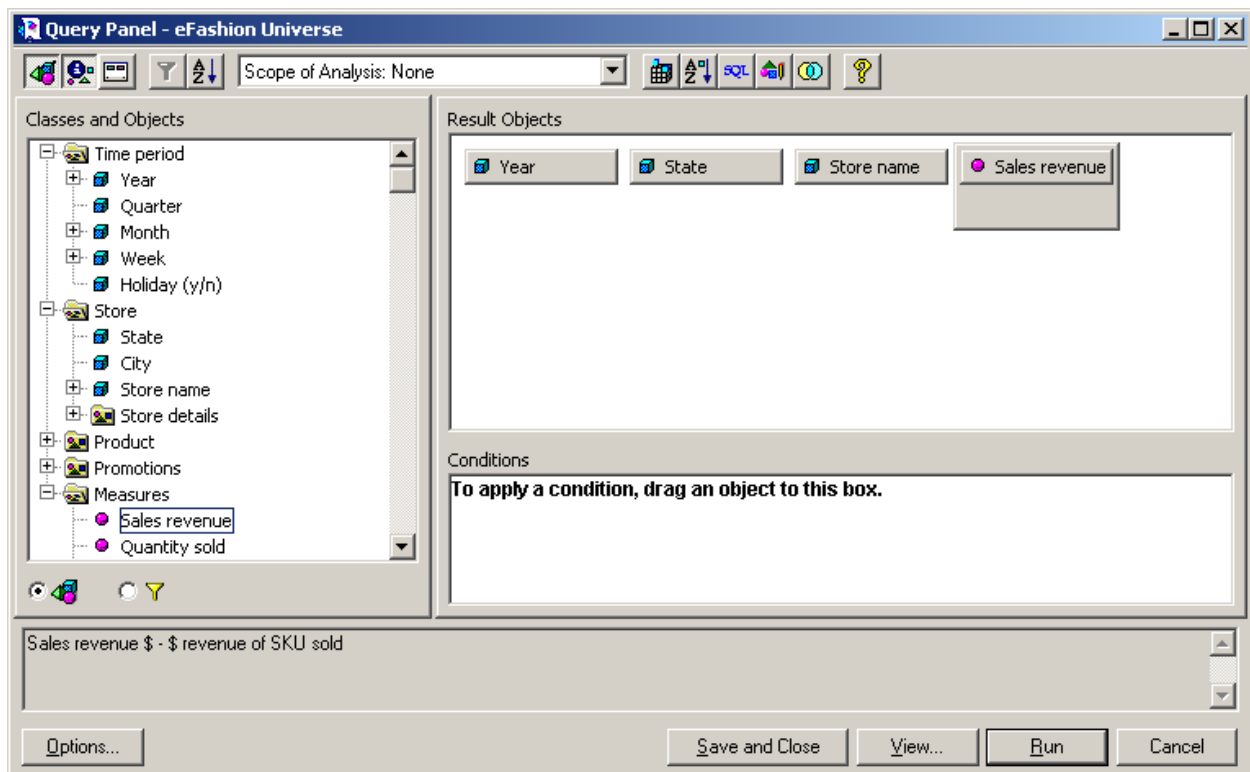
Figure 25: Query Panel.

In query panel, click the Run button to run the query. The query should run and the results will be displayed in the Report panel.
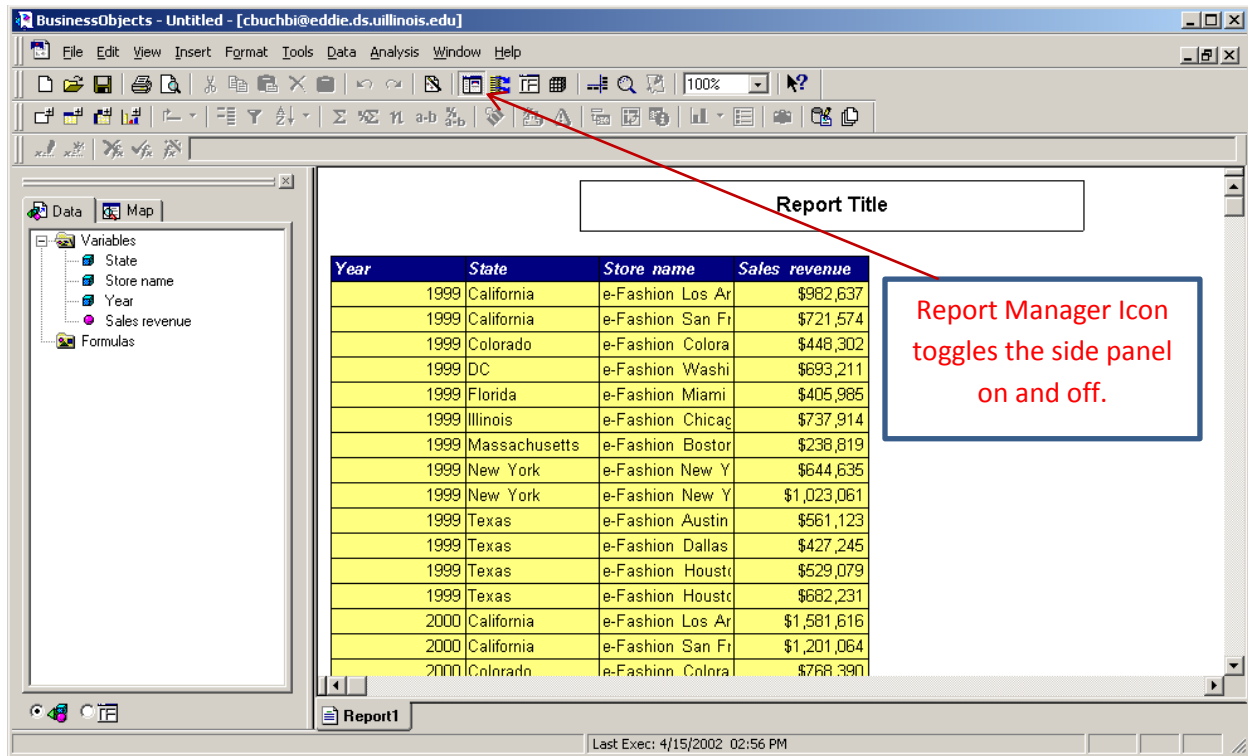
Figure 26: Report Panel.

Universal Integration

Business Object's Crystal Reports XI delivers powerful authoring, extensive distribution, and flexible integration. It allows us to access and format data, and embedded reporting into Java, .NET, and COM applications using comprehensive set of software developer kits (SDKs). In fact, Crystal Reports is already the de facto reporting standard with over 500 independent software vendors, and is embedded in leading software from Microsoft, SAP, Borland, BEA, PeopleSoft, and IBM. When delivered within the BusinessObjects Enterprise XI business intelligence platform, we can publish Crystal reports to the web, view them in portals and applications, and manage and deploy them using a standardized information infrastructure. With BusinessObjects Enterprise XI, organization can deploy Crystal Reports on Window, UNIX, and Linux platforms.

Query and Analysis

BusinessObjects XI query and analysis tools enable users to identify trends and find root causes in data from relational and OLAP data sources. A range of users- from experienced analysts to maintain business users- can easily create queries and perform calculations without understanding complex database schemas or SQL. Because these clients share the single, trusted BusinessObjects Enterprise BI platform, users can share and even have threaded discussions about this information with others across the organization, and beyond.

Analysis Method

After installing and generating reports in each tool, we can analyze which tool is more profitable for the company. The following factors will be correlated with results in order to conclude about the best available OLAP tool.

1. To check which tool can generate variety of graphs and reports and get detailed transactions from the database.

2. Analyzing target customers.

3. Multiple OS and language support.

4. Analyzing present market requirements.

5. Percentage of usage.

6. Installing and troubleshooting problems.

7. Scalability of the tool.

8. To check the ease of use and learn ability of each tool.

Database

The database taken for the purpose of this study is of the food chain supermarket which sells food and beverage. The store has lot of varieties of brands and many products under each brand available in the store. The total numbers of different products sold are grouped into one

category and were divided into sub-categories based on some condition. For example, there is a product category called "Seafood" which has been subdivided into Shellfish, Fresh Fish etc.

There are gigabits of records and information available in the database. The company uses Microsoft Access for database. The company's sales figures for two years are available in the database. Each year nearly one hundred thousand new records added into the database. This is what was required for the successful implementation of project with the help of different OLAP tools. This gives a scope to design problem identification from the database and to solve the problem successfully, and that in turn was helpful in making better and realistic decisions, faster and easier, which is the objective of the study.

Chapter – 4 Results

As the final step of the project, we have implemented our tools according to the methodologyw, in which these are some sample results in each tool.

Oracle Hyperion

Oracle Hyperion Essbase is easy to install. After installing the Java Runtime Environment and initializing wizard, we can choose the language setting. It provides the option of typical or custom installation. It has broad capabilities of ad hoc query and analysis, reporting, dashboards and scorecards, OLAP, Predictive analysis, mobile and cloud deployment. Oracle Hyperion can be plugged seamlessly into existing infrastructure, and it scales from simple departmental needs to Enterprise wide implementation.

Cognos

Cognos is world leader in business intelligence and performance planning software for company. It delivers managed reporting for consistent and fast based decision making. It is a sophisticated report authoring tool to create wide variety of reports. It enables authors to create

reports from any data sources that can deliver to report customers. It is used to create business context reports from simplest lists to most complex data grouping and can quickly design and delivered to the customers. It has framed based reporting interface which ease the report authors can authenticate virtually any report on the web.

The installation of Cognos was not easy and simple. It was difficult to give appropriate parameters to installation manager during installation process without the help of installation guides. Database connection process was also a little bit complex. We generated different reports in cognos, it was difficult to learn cognos environment and to generate reports and graphs. Particularly while generating graphs we encountered a lot of technical problems. These problems were caused by lack of technical experience and difficult working environment.

Business Objects

It was fairly simple to install BO, and in the same way it went with learning the environment and generating reports graphs. The configuration steps were easy and self-driven. Connection with database was also an easy process. More importantly, it is a business representation of data warehouse or transactional database.

MicroStrategy

Installation of MicroStrategy was not complicated. Before installation, we needed to create the database schema where log information is stored. If we are using Oracle database, we need to use SQL Plus or SQL worksheet to create a user and schema, and grant "connect" and "resource" to this user. Set up the ODBC data source names (DSNs) in order to import logs into the database, configure the MicroStrategy Intelligence Server enable MicroStrategy Intelligence Server to read the logs. We can install the MicroStrategy Server using the installation CD.

Chapter – 5 Discussions

Business intelligence is a powerful tool in a company's competitive arsenal, so perhaps it is not surprising that such software has been regarded as expensive to purchase and complicated to use. Business intelligence software generally relies on a platform of medium-to-large scale accounting or Enterprise Resource Planning (ERP) software to accurately capture detailed transactions. But recent innovations have broken through those barriers, making Business Intelligence practical for small to medium enterprises and their accountants. To address the analysis needs of the sales department, we examined the functionality of different OLAP technologies like Congos, Business Objects, Oracle, and Microstrategy. We found that the Business Objects OLAP technology would be best suited for the recurring analysis performed by the sales department answering known questions from summary to detailed data. While OLAP technologies seemed appropriated for the analysis needs, we conclude that data mining technology would help the sales department find previously unknown patterns in detailed data, patterns that could prove to be significant business value.

Cognos business intelligence has capabilities of reporting, analysis, score carding, dashboards, business event management, and data integration. In the reporting module, the user interface supports more than 10 languages and reports are delivered in more than 25 languages and there is multiple export formats of reports (IBM, 2009). It provides full UNICODE support. In this module, reports can be generated against any data source. Support is given for Windows, Linus, Unix, and mixed platform deployments and Web based deployment.

Analysis model is a robust analysis tool, easy to use and easy to deploy. Score carding helps us to align team and tactics with strategy, communicate goals consistently, and monitor performance against targets. It ensures ownership and accountability for performance. It makes it possible to align decision and tactics with own strategy and to execute against it.

Business dashboards communicate complex information quickly. Dashboards translate information from various corporate systems and data into visually rich presentations using gauges, maps, charts, and other graphical elements to show multiple results together. With dashboards we can have simple version of business issues at glance, real world view with timely data refreshers and unite data from different core areas of the company.

Business event management tracks the significant events that need attention, The time to action and resolution is reduce by continuous monitoring of events and uses decision process and business process automation. For high performance business intelligence, it is necessary that data integration is done throughout enterprise wide to deliver data warehouses ready for business reporting and analysis.

Business Objects XI allows nontraditional users of BI tools to easily retrieve structured questions of business information. The new release includes Desktop intelligence which is the new name of Business Objects full client. By using this customers can protect their existing investment and benefit from the latest platform. The new features include support for Crystal Report, Web Intelligence, OLAP Intelligence, Dashboard Manager, Desktop Intelligence, Performance Manager, Auditing, Publishing, Live Office, and Cristal Reports Explorer. It also include, integrated scheduling, full cross-platform support, auditing of all system interaction and support for Encyclopedia, Process Tracker, and Discussions.

Installation of Oracle Database on Windows is very easy. By running setup CD of Oracle for windows, we can invoke a setup and by clicking NEXT button we can install Oracle Software and Database without any problem. But in Linux it's different. If we want to install Oracle Database on Linux OS, we should follow some prerequisite actions. Oracle Database XE support connections between SQL command Line (SQL*Plus) and the database either locally, or remotely over a TCP/IP network. Connecting locally means running SQL Command Line on the same host computer where Oracle Database XE is installed and then initiating a database

connection from SQL command Line. To connect locally, we must supply only a database user name and password. To making a local connection on Linux, we must set environment variables. Connecting remotely means running SQL Command Line on a computer other than the Oracle Database XE host computer, and then initiating a database connection from SQL Command Line over the network. The remote computer must have Oracle client software installed.

Oracle Reports, a component of Oracle Fusion Middleware is Oracle's award-winning, high-fidelity enterprise reporting tool. It enables businesses to provide instant access to information to all levels within and outside of the organization in a scalable and secure environment. Oracle Reports consists of Oracle Reports Developer - a powerful, declarative WYSIWYG reports design tool and a J2EE 5.0 based Oracle Reports Server with multi-tier architecture to access any data sources, generate reports in any popular format for web and paper, and to burst and distribute reports to any destination. Oracle remains committed to the development of this technology, and to the ongoing release as a component of the Oracle Fusion Middleware platform.

Oracle products can be found in organizations of all sizes, both public and private sector, across most industries and geographies. Oracle has a direct sales force which engages primarily with its larger customers, as well as a network of partners who engage with customers of all sizes - a sizeable chunk of Oracle's revenue comes through its partner channel, Oracle Partner Network (Walker, 2011). Oracle supports most of the operating system available in the market such as Linux, UNIX, OS/390, Windows, Mac OSX, and OpenVMS. We can use the dbms_utility.port_string function to determine on what hardware platform, Operating system and TWO TASK PROTOCAL version our Oracle database in running on. The database enables us to store, process, and retrieve data in native languages. The languages that can be stored in a database are all languages written in scripts that are encoded by Oracle-supported character sets.

Through the use of Unicode databases and datatypes, the Oracle database supports most

contemporary languages.

In the business intelligence marketplace, MicroStrategy competes vigorously with

vendors such as SAP BusinessObjects. Some MicroStrategy products are available in two

versions. The one is Window only version and the other one is the universal version. The

window only version of intelligence server can be used with MicroStrategy Web Universal.

MicroStrategy Intelligence Server Universal, Web Universal are compatiable with Windows as

well as UNIX/Linux platforms. The MicroStrategy Installation Wizard provides steps to guide us

through installing one or more MicroStrategy products in Window and other platform. The DSN

parameters required for ODBC drivers to connect MicroStrategy to various databases. DSNs can

be created using the MicroStrategy Connectivity Wizard, and in UNIX and Linux, we can

configure ODBC parameters with the odbc.ini file. In MicroStrategy, installing and connecting

to the database is not simple and easy.

MicroStrategy Report Services is the dashboard and enterprise reporting engine of the

MicroStrategy BI platform. With fully integrated reporting, analysis and monitoring as well as

intuitive What-You-See-Is-What-You-Get (WYSIWYG) report design features, business users

can easily create and analyze a pixel-perfect dashboard, scorecard, or production report from

multiple sources of data. Report Services also employs advanced visualizations to quickly

identify trends or anomalies within the data. MicroStrategy provides integrated reporting,

analysis, and monitoring software that helps leading organizations worldwide make better

business decisions every day. Companies choose MicroStrategy for its advanced technical

capabilities, sophisticated analytics, and superior data and user scalability (Cover, 2011).

The factors discussed in method section are applied in the following table. These factors

are chosen from common factors used in analyzing usability, facilities (in term of option offered

to user), and form factors used identify coverage and wideness of software applications. All the

mentioned factors are selected from brochures used to market software systems, and a careful

attention has been given to select factors to cover the whole aspect of a software system.

Factor	Oracle	Business Objects	Cognos	Microstrategy
Installation problem	Fairly easy	Fairly easy	Rather difficult	Rather difficult
Choice of report & graph format and style	Offer variety of choice of report format and styles.	Offer variety of choice of report format and styles.	Offer variety of choice of report format and styles.	Offer variety of choice of report format and styles.
Target Customers	All kind of organization small to large	Small and Medium	Small and Medium	Medium and large
Multiple OS support	Linux, UNIX, OS/390, Windows, Mac OSX, and OpenVMS	Windows, HP-UX, Sun Solaris, AIX and Linux.	Window, Linux, and Unix	Unix, Linux, & Windows.
Language support	Support most contemporary languages.	10 languages.	25 Languages.	9 languages.
System requirements	P4 or equivalent, 2 GB RAM, Disk Space 200MB	P3 700 MHz or equivalent, 512 MB RAM, 1 GB	Minimum 1 GB ram, Disk space minimum	300MHZ Pentium or equivalent, 512

		Recommend, Disk Space 3 GB, CD-ROM	1.5 GB	MB RAM, 1GB disk space
Percentage of use (*Pendse, 2008)	21.7%	17.8%	16.6%	7.3%
Learn ability	Good	Very Good	Good	Good
Scalability	Web access, Microsoft Office Integration.	Web support and access.	Web support, managing more concurrent request threads.	Web Access, Support 100,000 active users.

Table 2: Comparative factors of OLAP tools.

As it is clear from the above table that Oracle and Business Objects are better than other two because these are easy to install. Moreover, these are easy to setup and connect to the database, and also the process of setting up environment was guided and self-explanatory. These enable us to setup and learn the system environment very quickly. Furthermore, these are simple to change different setting in reports and changing graphs at run time.

It has been observed that installation problems greatly impacts on user experience and retains the impression throughout the life of the software. On the other hand, software which is easy to install gets user attention and confidence from the beginning. However from the above discussion, it should not be assumed that easy installation process of Business Objects have created a bias in our comparison.

Chapter – 6 Conclusions

I have limited resources in terms of technical support, tool and time. Analyzing business intelligence tools performed in this thesis is an initial effort made to facilitate business and technical managers in deciding about to purchase OLAP tool.

Recommending the best available Online Analytical Processing Tool is a difficult task for us initially. When we compare Cognos, Oracle, BO, and MicroStrategy, it is very easy to say that Oracle and BO supersedes other two in most of the factors. Oracle will be better for a large organization and BO will fit to small and medium sized organizations. However, this analysis cannot be regarded as a comprehensive comparison, since all of the available tools in the market were not selected for evaluation.

References

Bedell, J. A. (1998). Outstanding Challenges in OLAP. Retrieved on April 2, 2011 from ACM

Digital Library.

Bonifati, A., Cattaneo, F., Ceri, S., Fuggetta, A., & Paraboschi, S. (2001). Designing data marts

for data warehouses. Retrieved on July 13 2011 from ACM Digital Library.

Breslin, Mary (2004). Data Warehousing Battle of the Giants: Comparing the basics of the

Kimball and Inmon Models. Retrieved on March 30 2011 from http://www.bi-

bestpractices.com/view-articles/4768

Cahudhury, S., & Dayal U. (1997). An overview of data warehousing and OLAP technology.

Retrieved on March 8, 2011 from ACM Digital library.

Codd, E. F. (1990). Relational model for database management, Version 2. New York: Addison

Wesley, Pearson Education, Inc.

Codd E. F., Codd S. B., & Sally C. T. "Providing OLAP (On-line Analytical Processing) to

user-analysis: An IT Mandate". Codd & Date, Inc 1993. Retrieved on August 2, 2011

from http://www.fpm.com/refer/codd.html

Cody, W. F., Kreulen, J. T., Krishna, V., & Spangler, W. S. (2002). The integration of business

intelligence and knowledge management. IBM System Journal, 41, 697 – 713.

Connolly, T. & Begg, C. (2005). Database Systems; a practical approach to design,

implementation, and management (5th ed.). New York: Addison Wesley, Pearson

Education, Inc.

Cover, W. (2011). MicroStrategy recognizes four companies for excellence in enterprise-class

business intelligence. Retrived on September 11, 2011 from

http://www.microstrategy.com/news/pr_system/press_release.asp?ctry=167&id=1746

Cunningham, D. (2004). Business Objects XI: Redefining how information empowers users. San

Jose, California: Business Objects Americas, Business Objects.

Forsman, S. (1997). OLAP council white paper. Retrieved on August 8, 2011 from

http://www.olapcouncil.org/research/whtpaply.htm

Hadley, L. (2002). Developing a data warehouse architecture. Retrieved on July 28, 2011 from

http://www.users.qwest.net/~lauramh/resume/thorn.htm

Hostman, B., Rayner, N., & Friedman, T. (2006). Gartner's Business Intelligence and

Performance Management Framework. Retrieved on August 5, 2011 from ACM Digital

Library.

Inmon, W. H. (2000). Building the data warehouse: getting started. Retrieved on August 15,

2011 from http://inmoncif.com/inmoncif-old/www/library/whiteprs/ttbuild.pdf

Inmon, W. H. (1995). What is a Data Warehouse? Prism, Volume 1, Number 1, 1995.

IBM (2009). Multilingual reporting with IBM Cognos 8 BI. Retrieved on August 15, 2011 from

ftp://ftp.software.ibm.com/software/data/sw-

library/cognos/issuepapers/ip_c8v3_multilingual_reporting.pdf

IBM Software Group (2010). IBM white paper on IBM Cognos 10: Intelligence unleashed.

Retrieved on August 16, 2011 from http://www-

05.ibm.com/il/events/cognos/pdf/IBM_Cognos_10_White_Paper.pdf

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). The data warehouse
lifecycle toolkit (2nd ed.). Indianapolis, Indiana: Wiley Publishing, Inc.

Kimball, R. (1996). The Data Warehouse Toolkit: Practical techniques for building dimensional
data warehouses. New York: John Wiley & Sons, Inc.

Kannan, S. (2008). Business Objects XI. Retrieved on August 20, 2011 from
http://it.toolbox.com/wiki/index.php/Business_Objects_XI

Lawrence, M. (2006). Multiobjective genetic algorithms for materialized view selection in
OLAP data warehouse. Retrieved on March 11, 2011 from ACM Digital Library.

MicroStrategy (2011). Introduction to MicroStrategy: Evaluation guide. Retrieved on July 15,
2011 from http://www.microstrategy.com/producthelp/manuals/9.2.1/en/mstrintro.pdf

Mimno, P. (1997). Comparing and evaluating data warehousing products & tools. Retrieved on
March 9, 2011 from http://www.dciexpo.com/dwprod/

Moody, D. L. & Kortink, M.A.R. (2000). From Enterprise Models to Dimensional Models: A
Methodology for Data Warehouse and Data mart Design. Retrieved on April 2, 2011
from
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.33.511&rep=rep1&type=pdf

OLAP Council (2001). OLAP Council White Paper. Retrieved on August 1, 2011 from
http://www.olapcouncil.org/research/whtpapco.html

Oracle (2002). Data Warehousing Concepts. Retrieved on July 29, 2011 from
http://download.oracle.com/docs/cd/B10500_01/server.920/a96520/concept.htm

Pendse, N. (2008). OLAP market share analysis. Retrieved on April 28, 2011 from

http://www.bi-

verdict.com/fileadmin/dl_temp/bf0b9dfbc0f995b35e1a583f13fdedfb/market.htm?user_id

Power, D. (2010). What are advantages and disadvantages of data warehouses? Retrieved on July

29, 2011 from http://dssresources.com/faq/index.php?action=artikel&id=180

Rahm, E. & Do, H., H. (2000). Data Cleaning: Problems and Current Approaches. Retrieved on

August 3, 2011 from Data Engineering (Vol. 23 No.4, pp. 3 - 12). Washington, D. C.:

IEEE Computer Society.

Rayner, N., Chandler, N., & Decker, J. E. (2007). Magic Quadrant for CPM Suites, 2007.

Retrieved on June 25, 2011 from

http://www.alarconyharris.com/_prensa/tagetik/tagetik_pdf/080128_tagetik_cuadrantema

gico.pdf

Slaughter, A. H. (1999). OLAP System Component. Retrieved on August, 4 2011 from

http://web.mit.edu/profit/PDFS/SlaughterA.pdf

Songini, M.L. (2004). QuickStudy: Extract, Transform and Load (ETL). Retrieved on August 1,

2011 from

http://www.computerworld.com/s/article/89534/QuickStudy_ETL?taxonomyId=9&page

Number=1

Walker, S. (2011). What does Oracle sell, and who are its target customers? Retrieved on

September 5, 2011 from http://www.quora.com/Oracle/What-does-Oracle-sell-and-who-

are-its-target-customers

Watson, H.J., & Wixon, B.H. (2007). The current state of business Intelligence. Retrieved on

March 6, 2011 from IEEE.

Wilson, J. (2011). IBM Cognos 8 Business Intelligence and IBM Cognos 8 Planning. Retrieved

on August 15, 2011 from http://www.information-

management.com/issues/20_2/ibm_cognos_8_bi_ibm_cognos_8_planning-10017429-

1.html