DATA WAREHOUSE- A KEY TO MANAGEMENT DECISIONS

Amneet Kaur, Ph. D.

Assist Prof. GGSCW -26 Chandigarh

――――――――――――――――― **Abstract** ―――――――――――――――――

A data warehouse is a subject-oriented, integrated, time-variant, non-volatile *collection of data in support of management's decision making process. Data warehousing is the data management and analysis technology. Data in the data warehouse is preprocessed and presented such that it facilitates the cross functional monitoring and assessment of the overall direction of the organization. Thus, it is the hub for an intelligent management decision support. Successful implementation of a data warehouse requires a high-performance, scale able combination of hardware and software, which can integrate easily with existing systems, so that users can use data warehouse to improve their decision-making. A data warehouse is incomplete until it provides the exploitation tools that enable end users to view analyze and report on data in ways that support their decision-making. Data marts, data mining, data modeling and metadata are some other important concepts attached with data warehousing, the knowledge of which helps to a great extent in data warehouse implementation. Business intelligence data warehouse is the combination of two terms frequently used in the data mining and analysis field. When used together, vendors are trying to emphasize that they offer more services than just data warehousing alone. Because business intelligence can encapsulate a wide variety of services and tools, combining the terms into business intelligence data warehouse reflects that the vendor offers everything an organization needs, including the basics of data warehousing. Many organizations prefer to purchase business intelligence tools and data warehousing as a combined business intelligence data warehouse system in order to increase the efficiency of the integrated system. The key advantages of using a Data warehouse business intelligence system is that it makes it much simpler to analyze and report on the information extracted from the data entered into the system. In the cleansing stage of the data warehousing process, all inconstancies in the data are determined and taken care of before the data enters the ETL stage. The effectiveness of the data warehouse application intensifies especially when the operational data resides in distributed, non-homogenous systems and replace manual data gathering and reconciliation procedures. Therefore, data warehousing can be useful to not only commercial sectors but also to sectors like government, health care, insurance, manufacturing, finance, distribution, education The goal of a data warehouse application in an organization is to increase the effectiveness of the Decision-making and direction setting process. Data in the data warehouse is preprocessed and presented such that it facilitates the cross functional monitoring and assessment of the overall direction of the organization. Thus, it is the hub for an intelligent management decision support. A data warehouse is incomplete until it provides the exploitation tools that enable end users to view analyze and report on data in ways that support their decision-making. The aim of the paper is to generate the knowledge of data warehousing concept and its advantages in the era of globalization. To study the structure and application of data warehousing. Acquiring right information at the right time from the huge ocean of data is becoming difficult day by day. Data warehouse helps to convert the information overload in to meaningful strategic information. It is very useful to overcome the limitation of traditional tools by displaying different dimensions as required. Thus, a well defined and implemented data warehouse can be used to understand trends and make better forecasting decisions,*

bring better products to market in a more timely manner and analyze daily information and make quick decisions that can significantly affect the organization's performance. In this paper I have tried to cover all advantages of data warehousing related to management decision making and hindrances in the usage of business intelligence data warehousing.

__Keywords__: Business Intelligence Data Warehousing, data warehousing application, structure of data warehousing.

Introduction

"A data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management's decision making process." By __Bill Inmon.__

__Subject-oriented__ means the data warehouse focuses on the high-level entities of the business; in higher education's case, subjects such as students, courses, accounts and employees. This is in contrast to operational systems, which deal with processes such as student registration or rolling up financial accounts.

__Integrated__ means the data is stored in a consistent format (i.e., naming conventions, domain constraints, physical attributes, and measurements). For example, production systems may have several unique coding schemes for ethnicity. In the data warehouse, there is only one coding scheme.

__Time-variant__ In contrast to the operational data that focus on current transactions, the Warehouse Data represent the flow of data through time. In short, the Data Warehouse contains data that reflect what happened last week or last year. The Data Warehouse can even contain projected data generated through statistical and other models. It is also time-variant in the sense that once data are periodically uploaded to the Data Warehouse, all time-dependent aggregations are recomputed.

__Non-volatile__ Data in a data warehouse is used only for queries. Once data enter

The Data Warehouse, they never removed. Because the data in the Warehouse

represent the enterprise's entire history, the operational data, representing the

near-term history, are always added to it. Because data are never deleted and new

data are always added, the Data Warehouse is always growing, that's why the DSS DBMS

must be able to support multi gigabytes and even multi terabyte size

Databases and multiprocessor hardware.

Accessible: The primary purpose of a data warehouse is to provide readily accessible information to end-users.

Process-Oriented: It is important to view data warehousing as a process for delivery of information. The maintenance of a data warehouse is ongoing and iterative in nature.

Kimball, in 1997, stated that 'the data warehouse is nothing more than the union of all the data Marts", Kimball indicates a bottom-up data warehousing methodology in which individual data marts providing thin views into the organizational data could be created and later combined into a larger all-encompassing data warehouse. Data warehouse is simply a single, complete and consistent store of data obtained from variety of sources and made available to end users in a way they can understand and use it in business decision making. In other words data warehousing is creating an infrastruction for reusing the data in numerous way.

Data Mart

Data Mart is a subset of the enterprise-wide data warehouse. Data Marts are highly focused sets of information that are designed in the same way as Data Warehouses, but are implemented to address the specific needs of a defined set users who share common characteristics. Data marts fall into two broad categories: 1.Subset data marts created from a parent data warehouse or parent data mart. 2. Incremental data marts used as independent information resources or as data warehouse building blocks.
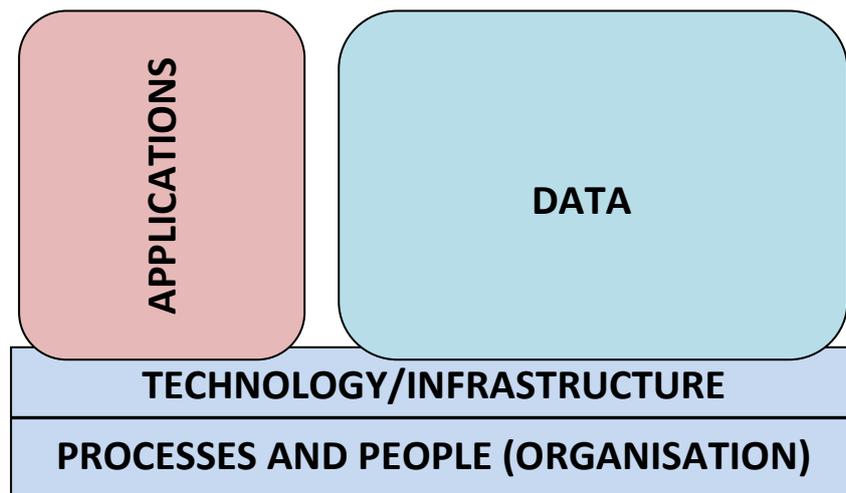
Data Mining

Banks use data warehousing and data mining methodology to build long-term relationships with their customers and because it helps banks to quickly and smoothly adapt to business changes. To make data useful, bank enterprises collect data from almost every platform and data format; clean and transform data into information that users will understand; and stores the information in an open and efficient data warehouse structure. Data analysis is the ability to look at the same information in a variety of ways. This is where data mining comes in. To explore the information stored in data warehouse structure, data mining, OLAP, query and reporting, statistical analysis, data visualization and application-development interfaces are included and are mostly client/server and web enabled. **Data mining** can be formally defined as the process of extracting hidden and interesting implicit information (or knowledge) from large databases and then using the knowledge to make crucial business decision.

Components of Data warehouse

Data is the fundamental component: cleaned, organized data, mostly extracted from the campus' operational systems. Making that data useful to a variety of campus personnel, though, requires some applications to deliver and explain it. These applications range from predefined reports through query tools to complex tools for analysis and modeling. Delivering data and applications and securing the data as specified by campus data stewards requires a set of technology, most of it centralized in secure computer locations. The architecture for the data warehouse is described in terms of four inter-related dimensions:

 1. Applications (or the business intelligence layer).

 2. Data.

 3. Technology and security.

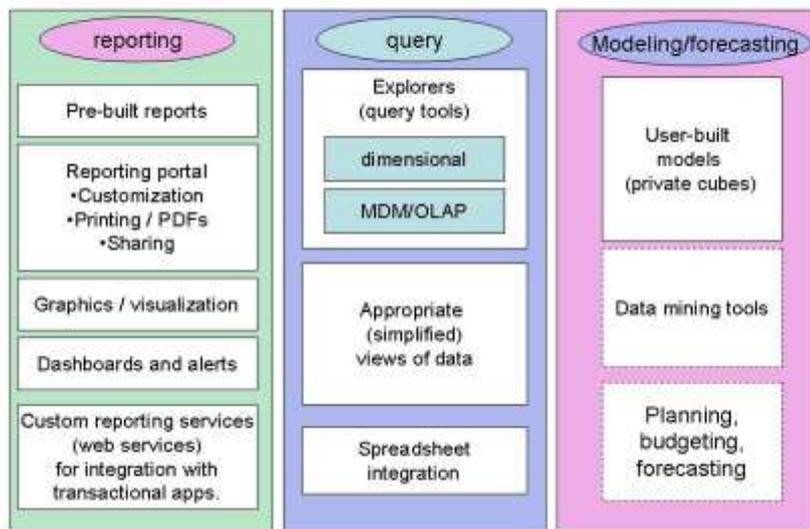 4. Support—processes and organization.

Components of data warehouse

Fig: UC-Berkeley Data Warehouse Roadmap

Business Intelligence: Applications Architecture

Business intelligence usually refers to the information that is available for the enterprise to make decisions on. A data warehousing (or data mart) system is the backend, or the infrastructural, component for achieving business intelligence. Business intelligence also includes the insight gained from doing data mining analysis, as well as unstructured data (thus the need for content management systems). For our purposes here, we will discuss business intelligence in the context of using a data warehouse infrastructure. The architecture provides three kinds of applications:

1. Reporting, or information delivery, including variants such as dashboards and alerts.

2. Query.

3. Modeling, Planning Forecasting

Fig: Business Intelligence Applications

Reporting.

Reporting applications deliver information in a form which is useful to users. In their simplest Form—fixed, printed reports—these tools are as old as computers. They imply a partnership between users who need information and specialists who help design reports, displays and Graphics to deliver the required information. The best reports provide just the information a user needs for a specific purpose, delivered in a way which makes the information usable and actionable.

Dashboards and alerts.

In the interest of delivering just the information most useful to support decision-making or Action, it will be useful to complement conventional reports with information dashboards and alerts. Dashboards communicate information with quickly-comprehended graphics such as dials and meters. Typically dashboards are used to report on established performance indicators, measured at predefined intervals.

Reporting integrated with production applications.

Some reporting will be most useful when it is delivered in the context of some other computer based process. For example, a student who is registering for courses might benefit from

knowing about other students who took similar courses—how many took the same set of courses at once and how well they succeeded; how many completed the courses and their average grades.

Query and analysis

People who have analytical skills and jobs requiring analysis will need the ability to explore the information in the warehouse. Enabling analysis of this kind is one of the great powers of the EDW. Using that power requires understanding the information in the warehouse and knowing how to select data, summarize it or drill down for further detail, and particularly how to combine information across subject areas.

Advanced applications: modeling, forecasting and planning

The solution for this need involves collaboration between the data warehouse and a transactional application for recording plans. Once plans are recorded, they are migrated into the data warehouse, where they are available. Data Warehouse Roadmap Data Warehouse Architecture for reporting and query in combination with actual activity

Data Warehousing Applications

The applications served by data warehousing can be classified in one of three main categories.

• "Personal productivity applications" such as spreadsheets, statistical packages and graphics tools are useful for manipulating and presenting data on individual PCs. Though they are developed for a standalone environment, they address applications requiring only small volume of warehouse data.

• "Data query and reporting applications" deliver warehouse-wide data access through simple, list oriented queries, and the generation of basic reports. These reports provide a view of historical data but do not address the need of organization for in-depth analysis and planning.

• "Planning and analysis applications" address such essential management requirements as budgeting, forecasting and customer profitability, financial consolidations and applications that use historical, projected and derived data.

Design and Implementation of a data warehouse

A. Requirement analysis and specification

The requirements phase addresses the high-level needs of the entire warehouse

environment. Both business and technical (including the infrastructure) requirements are gathered during this phase. Interviews, workshops, and analysis of existing documents and systems may be used to gather and confirm the necessary facts. The resulting requirements

document is reviewed by all affected parties. That document includes identification of business objectives as well as a technical feasibility analysis. Recommendations may include the number of project builds and the priority of each build.

B. Data Warehouse Design

The design phase focuses on one project build at a time. Activities for this phase include:

1. Detailed analysis and requirements for the selected build.

2. Detailed design for the data model.

3. Detailed specification of the process model for extraction, transformation, and

4. loading.

5. Selection of Software and Hardware

6. Creation of the application model or selection of exploitation tools.

7. Design of additional aspects such as the security and metadata models.

Both the project manager and the warehouse architect can expect to be involved full-time in this phase. Additional resources, such as business content experts, the data administrator, warehouse administrator, construction manager, and IT personnel are consulted during this phase.

C. Warehouse Implementation:

During the implementation phase, implementation teams code and populate the

Warehouse with data and develop the applications for end-user analysis and reporting. Business users and the IT organization rigorously test the warehouse and applications to verify that all acceptance criteria are met.

D. Deployment

The deployment phase is the rollout of the data warehouse and end-user applications to end-users and IT staff. Making sure that users are well trained and that applications and data are readily accessible helps promote widespread acceptance of the entire project

E. Review

Three executions of the review phase are conducted with each build:

1. Following the construction phase to access the implementation process and learn from successes and setbacks.

2. 3-6 months later to review the deployment phase and ensure that the transition to support has gone smoothly and those users have access to the warehouse.

3. 18-24 months after initial construction to measure any tangible benefits, calculate ROI, and ensure that the warehouse environment is continuing to meet the business community requirements.

F. Maintenance and Administration Once implemented, the warehouse requires on-going maintenance. It is essential that attention be given during the construction process to this on-going element of the warehouse life cycle.

Data warehousing A key to management decision in banks/business organization/insurance.

1. The Warehouse infrastructure can support a wide range of applications and reports to meet exact business needs. In Banking, the most important of data warehousing is building **Risk Management** Systems. Risk Management System will identify the risks associated with a given set of assets.

2. Most organizations launch many different kinds of promotional campaigns for many different products using many different media. This application enhances the organization's understanding of the entire process from selecting customers to be targeted to analyzing how they responded. **Campaign Analysis** allows you to measure the responsiveness to campaigns by households and by individual customers

3. **Customer Profiling** allows organizations to distinguish, in the mass of customers, the many micro segments that make up the whole. Increasingly, customer segmentation is forming an essential element of marketing strategy as markets become more fragmented especially where customer segments exhibit distinct and different characteristics.

4. **The Loyalty Analysis** application allows you to measure customer loyalty from different viewpoints such as duration of relationship; range of services and products consumed; and the demographic, psycho-graphic and geographic influences on customer attrition. Customers interact with organizations in many ways using different touch points to initiate inquiries, provide feedback or make suggestions.

5. The **Business Performance Analysis** application for banking exploits the industry specific, transaction-level data in a typical retail banking enterprise. Analyzing a bank's business performance requires an understanding of customer behavior, including their usage patterns of the different services the bank offers.

6. The **Sales Analysis** application allows analysis of sales from a variety of viewpoints such as sales by channel, outlet or organizational unit; sales by product, product category or

group; and sales by region and by season. This application offers organizations an integrated perspective on sales results and enables sales managers to understand the underlying trends and patterns in their sales data

7. In any organization, it is essential to understand **profitability** in order to determine pricing, award discounts, allocate resources or develop strategy. But profitability is a many-faceted concept and can be considered in the context of an organization, a channel, a product, a product category, a brand, a customer or a customer segment.

Conclusions

A data warehouse is incomplete until it provides the exploitation tools that enable end users to view, analyze and report on data in ways that support their decision-making. Data marts, data mining, data modeling and metadata are some other important concepts attached with data warehousing, the knowledge of which helps to a great extent in data warehouse implementation. Data in the data warehouse is preprocessed and presented such that it facilitates the cross functional monitoring and assessment of the overall direction of the organization. Thus, it is the hub for an intelligent management decision support. Successful implementation of a data warehouse requires a high-performance, scale able combination of hardware and software, which can integrate easily with existing systems, so that users can use data warehouse to improve their decision-making. Thus, a well defined and implemented data warehouse can be used to understand trends and make better forecasting decisions, bring better products to market in a more timely manner and analyze daily information and make quick decisions that can significantly affect the organization's performance

REFRENCES

Douglas Hackney, "Understanding and implementing successful DataMarts, Addison- Wesley Developers Press".

DKMS Brief No. Six: Data Warehouses, Data Marts, and Data Warehousing: New Definitions and New Conceptions

Enterprise Data Warehouse (EDW) UC Berkeley Peter Cava

Harry S. Singh, "Data Warehousing – concepts, Technologies,Implementations, and Management ", Prentice Hall PTR, New Jersey.

Inmon, W.H. Building the Data Warehouse. USA: John Wiley

Report of the Dr. Vasudevan Committee of Technology Upgradation in the Banking Sector.

http://www.dw-institute.com

http://www.datawarehouse.org

*www.**ibm**.com/software/**data**/infosphere/**warehouse***

www.1keydata.com