AJCSE
ASIAN JOURNAL OF COMPUTER SCIENCE ENGINEERING

www.ajcse.info

Asian Journal of Computer Science Engineering 2016;1(1):17-19

REVIEW ARTICLE

An Overview On Data Warehousing

Kruti Mangal[1], *Saniya Jain[1]

CSE Department Mandsaur Institute of Technology Mandsaur [M.P.]

Received on: 20/07/2016, Revised on: 26/10/2016, Accepted on: 30/10/2016

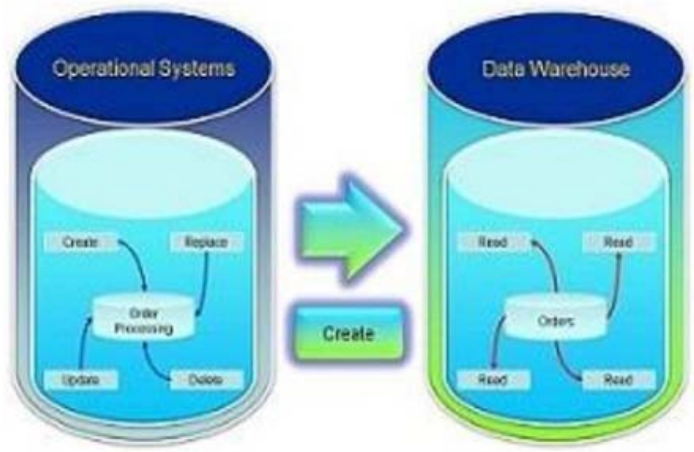
ABSTRACT

Data are crucial raw material in the information age, and data storage and management have become the focus of database design and implementation. A new data storage facility, called a data warehouse was developed to extract or to obtain its data from operational databases as well as from external sources, providing a more comprehensive data pool . In this we explore the main concepts of data warehouse and examine the tools that make using data warehouse information simpler and more effective.

Keywords - Data warehouse, Data analysis, Data schema etc.

INTRODUCTION

The term "Data Warehouse" was first coined by Bill Inmon in 1990. According to Inmon, a data warehouse is a subject oriented, integrated, time-variant, and non-volatile collection of data. This data helps analysts to take informed decisions in an organization. It is designed for query and analysis rather than for transaction processing, and usually contains historical data derived from transaction data, but can include data from other sources. Data warehouses separate analysis workload from transaction workload and enable an organization to consolidate data from several sources. This helps in:



- Maintaining historical records
- Analyzing the data to gain a better understanding of the business and to improve the business.

A data warehouses provides us generalized and consolidated data in multidimensional view. Along with generalized and consolidated view of data, a data warehouses also provides us Online Analytical Processing (OLAP) tools. These tools help us in interactive and effective analysis of data in a multidimensional space.

Comparison of Data Warehouse and Operational Database

	Operational Database	Data Warehouse
Purpose	For data retrieval, updating and management	For data analysis and decision making
Systems/ Applications	OLTP (Online Transaction Processing System)	Analytical Software like Data Mining Tools, Reporting Tools and OLAP tools
Format	▪ Normalised ▪ Relational Database ▪ Lowest level of granularity (e.g. individual transactions)	▪ Denormalised and integrated ▪ Multi-dimensional arrays or relational format ▪ Subject-Oriented ▪ Granularity level depends on subject
Time Frame	Current / Real-Time	Historical

***Corresponding Author:** Kruti Mangal,Saniya Jain**, Email:** gaurimangal8@gmail.com

FEATURES

The key features of a data warehouse are:

Subject Oriented - A data warehouse is subject oriented because it provides information around a subject rather than the organization's on going operations. These subjects can be product, customers, suppliers, sales, revenue, etc. A data warehouse does not focus on the ong oing operations, rather it focuses on modelling and analysis of data for decision making.

Integrated - A data warehouse is constructed by integrating data from heterogeneous sources such as relational databases,flat files, etc. This integration enhances the effective analysis of data.

Time Variant - The data collected in a data warehouse is identified with a particular time period. The data in a data warehouse provides information from the historical point of view.

Non-volatile - Non-volatile means the previous data is not erased when new data is added to it. A data warehouse is kept separate from the operational database and therefore frequent changes in operational database is not reflected in the data warehouse.

DATA MART

A data mart is a simple form of a data warehouse that is focused on a single subject (or functional area), hence they draw data from a limited number of sources such as sales, finance or marketing. Data marts are often built and controlled by a single department within an organization. The sources could be internal operational systems, a central data warehouse, or external data. Denormalization is the norm for data modeling techniques in this system. Given that data marts generally cover only a subset of the data contained in a data warehouse, they are often easier and faster to implement.

- Often holds only one subject area- for example, Finance, or Sales.
- May hold more summarized data (although many hold full detail).
- Concentrates on integrating information from a given subject area or set of source systems.
- Is built focused on a dimensional model using a star schema.

Difference between data warehouse and data mart

Data warehouse	Data mart
Enterprise-wide data	Department-wide data
Multiple subject areas	Single subject area
Difficult to build	Easy to build
Takes more time to build	Less time to build
Larger memory	Limited memory

Types of data marts

- Dependent data mart : In a top-down approach a data mart development "dependants" on enterprise data warehouse hence data mart are know as dependent data mart.
- Independent data mart : In a bottom-up approach a data mart development is "Independent" of enterprise data warehouse. Hence such data mart are know as independent data mart.

OLAP

The need for more intensive decision support prompted the introduction of a new generation of tools. Those new tools , called **online analytical processing (OLAP)** , which create an advanced data analysis environment that supports decision making, business modeling and operations research. To provide better performance, some OLAP systems merge the data warehouse and data mart approaches by storing small extracts of data warehouse at end-user workstations.

OLAP systems share four main characteristics ;

- Use multidimensional data analysis techniques.
- Provide advanced data support.
- Provide easy-to- use end-user interfaces.
- Support client/server architecture.

Data warehouse Design and Implementation

ACKNOWLEDGEMENT

We sincerely thank our CSE lecturer Miss Priyanka Mangal for her guidance and encouragement.

REFERENCES

1. website[online].Availablefrom:https://www.tutorialspoint.com/dwh/dwh_architecture.htm.
2. Book Referred : DATABASE SYSTEMS design,implementation and management [seventh edition].
3. MIT website [online]. Available from: https://en.wikipedia.org/wiki/Data_warehouse.
4. Book Referred : Data warehousing fundamentals for IT professionals
5. MIT website [online]. Available from http://db.csail.mit.edu/madden/high_perf.pdf.