2.4. DATA WAREHOUSING

2.4.1. Introduction

Data warehousing is combining data from multiple and usually varied sources into one comprehensive and easily manipulated database. Common accessing systems of data warehousing include queries, analysis and reporting.

Data warehousing techniques and tools include DW appliances, platforms, architectures, data stores, and spreadmarts; database architectures, structures, scalability, security, and services; and DW as a service.

Data warehousing incorporates data stores and conceptual, logical, and physical models to support business goals and end-user information needs. A data warehouse (DW) is the foundation for a successful BI program.

Creating a DW requires mapping data between sources and targets, then capturing the details of the transformation in a metadata repository. The data warehouse provides a single, comprehensive source of current and historical information.

2.4.2. Meaning and Definition of Data Warehouse

A data warehouse is a place where data is stored for archival, analysis and security purposes.

The term Data Warehouse was coined by Bill Inmon in 1990. He is considered as the father of data warehousing.

According to Bill Inmon, "A warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process".

He defined the terms in the definition as follows:

1) **Subject Oriented:** Data that gives information about a particular subject instead of about a company's ongoing operations.
2) **Integrated:** Data that is gathered into the data warehouse from a variety of sources and merged into a coherent whole.
3) **Time-variant:** All data in the data warehouse is identified with a particular time period.
4) **Non-Volatile:** Data is stable in a data warehouse. More data is added but data is never removed. This enables management to gain a consistent picture of the business.

According to Ralph Kimball, "Data warehouse is a copy of transaction data specifically structured for query and analysis".

Hence, data warehouse is a collection of data prepackaged or summarized according to specific business rules and designed to support management decision making.

2.4.3. Need for a Data Warehouse

The data warehouse (DW) concept sprang from the growing competitive need to quickly analyse information. Existing operational systems cannot meet this need because,

1) They lack on-line historical data.
2) The data required for analysis resides in different operational systems.
3) The query performance is extremely poor, which in turn impacts performance of operational systems.
4) The operational DBMS designs are inadequate for **decision support**.

As a result, information stored in operational systems is inaccessible to planner & decision-makers. A data warehouse eliminates these problems by storing current and historical data

From disparate operational systems that decision-makers need in a single consolidated system. This makes data readily accessible to all in the organisation who needs it without interrupting on-line operational workloads.