



Data Warehouse

Presented by
C.Namrata Mahender



What is Data Warehouse?

- “A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management’s decision-making process.” —W. H. Inmon
- Data warehousing:
 - The process of constructing and using data warehouses.
- Defined in many different ways
 - A decision support database that is maintained separately from the organization’s operational database
 - Support information processing by providing a solid platform of consolidated, historical data for analysis.



Understanding Data Warehouse

- The Data Warehouse is that database which is kept separate from the organization's operational database.
- There is no frequent updation done in data warehouse.
- Data warehouse possess consolidated historical data which help the organization to analyze it's business.
- Data warehouse helps the executives to organize, understand and use their data to take strategic decision.
- Data warehouse systems available which helps in integration of diversity of application systems.
- The Data warehouse system allows analysis of consolidated historical data analysis.



Why Data Warehouse Separated from Operational Databases

- The operational database is constructed for well known tasks and workload such as searching particular records, indexing etc but the data warehouse queries are often complex and it presents the general form of data.
- Operational databases support the concurrent processing of multiple transactions. Concurrency control and recovery mechanisms are required for operational databases to ensure robustness and consistency of database.



Why Data Warehouse Separated from Operational Databases

- Operational database query allow reading, modifying operations while the OLAP query need **read only** access of stored data.
- Operational database maintain the current data on the other hand data warehouse maintain the historical data.



Data Warehouse Features

- **Subject Oriented** - The Data Warehouse is Subject Oriented because it provides us the information around a subject rather than the organization's ongoing operations. These subjects can be product, customers, suppliers, sales, revenue etc. The data warehouse does not focus on the ongoing operations. Rather, it focuses on modelling and analysis of data for decision making.
- **Integrated** - Data Warehouse is constructed by integration of data from heterogeneous sources such as relational databases, flat files etc. This integration enhances the effective analysis of data.



Data Warehouse Features

- **Time-Variant** - The Data in Data Warehouse is identified with a particular time period. The data in data warehouse provide information from historical point of view.
- **Non Volatile** - Non volatile means that the previous data is not removed when new data is added to it. The data warehouse is kept separate from the operational database therefore frequent changes in operational database are not reflected in data warehouse.



Data Warehouse Applications

- Financial services
- Banking Services
- Consumer goods
- Retail sectors.
- Controlled manufacturing



Data Warehouse Types

- **Information processing** - Data Warehouse allow us to process the information stored in it. The information can be processed by means of querying, basic statistical analysis, reporting using crosstabs, tables, charts, or graphs.
- **Analytical Processing** - Data Warehouse supports analytical processing of the information stored in it. The data can be analysed by means of basic OLAP operations, including slice-and-dice, drill down, drill up, and pivoting.



Data Warehouse Types

- **Data Mining** - Data Mining supports knowledge discovery by finding the hidden patterns and associations, constructing analytical models, performing classification and prediction. These mining results can be presented using the visualization tools.



Data Warehouse Types

- **Data Mining** - Data Mining supports knowledge discovery by finding the hidden patterns and associations, constructing analytical models, performing classification and prediction. These mining results can be presented using the visualization tools.



Data Warehouse vs OLTP

SN	Data Warehouse (OLAP)	Operational Database(OLTP)
1	This involves historical processing of information.	This involves day to day processing.
2	OLAP systems are used by knowledge workers such as executive, manager and analyst.	OLTP system are used by clerk, DBA, or database professionals.
3	This is used to analysis the business.	This is used to run the business.
4	It focuses on Information out.	It focuses on Data in.
5	This is based on Star Schema, Snowflake Schema and Fact Constellation Schema.	This is based on Entity Relationship Model.
6	It focuses on Information out.	This is application oriented.
7	This contains historical data.	This contains current data.



Data Warehouse vs OLTP

8	This provides summarized and consolidated data.	This provide primitive and highly detailed data.
9	This provides summarized and multidimensional view of data.	This provides detailed and flat relational view of data.
10	The number or users are in Hundreds.	The number of users is in thousands.
11	The number of records accessed is in millions.	The number of records accessed is in tens.
12	The database size is from 100GB to TB	The database size is from 100 MB to GB.
13	This are highly flexible.	This provide high performance.



Thank you