**Name : Hemalatha Karumuri**

**Data Warehouse**

An ordinary database can store MB to GB of data and that too for a specific purpose.

\*For storing TB of data we have shifted to Data warehouse.

\*collection of raw data will be stored

"Data Warehouse is a subject-oriented, integrated, and time-variant store of information in support of management's decisions."

**Ex**: social media data,banking

**Characteristics of Data Warehouse**

## Subject-Oriented

A data warehouse target on the modeling and analysis of data for decision-makers.. This is done by excluding data that are not useful concerning the subject and including all data needed by the users to understand the subject.

## Integrated

A data warehouse integrates various heterogeneous data sources like RDBMS, flat files, and online transaction records.

## Time-Variant

Historical information is kept in a data warehouse.

Ex: one can retrieve files from 3 months, 6 months, 12 months, or even previous data from a data warehouse.

**Non volatile:**

No updations are allowed i.e,data is immutable. Once the data entered into the data warehouse, they are never removed.

**Purpose of Data Warehouse**

1. to perform on enormous amounts of data and make better forecasting decisions.
2. User can effectively perform the operations as the data finds at a common place
3. quick response time while accessing the data.
4. To store non-volatile data
5. for data reporting and analysis.

**Operational Data Store**

1. It is accessed by an Operational System to carry out regular operations of an organization.
2. It uses an OLTP architecture which is optimized for faster transaction processing.
3. OLTP Databases access the data in the form of operations like- Inserting, Deleting, and Updating data.
4. The DSS(Data support system) data which is extracted from multiple sources differ from operational data in three main areas:time span,granularity and dimensionality
5. The data that arrived at data warehouse are first passed to **Operational Data Store (ODS).**

**OLTP Vs Warehouse Applications**

|  |  |
| --- | --- |
| **OLTP** | **OLAP[DATA warehouse]** |
| 1)transaction processing | Decision making by analysing the data |
| 2)stores current data only | Stores historical data |
| 3)Application oriented | Subject oriented |
| 4)Applications:ATM | Applications:markrtingstrategies,weather forecasting etc |

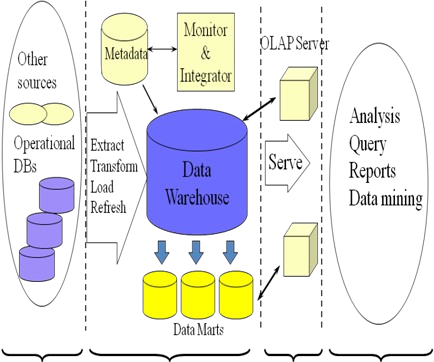
**Data Marts**

* The data in the data warehouse is stored in the form of Data marts.
* It allows the user to access the data in terms of a specific business line or team.
* It is a *subset* of the data warehouse

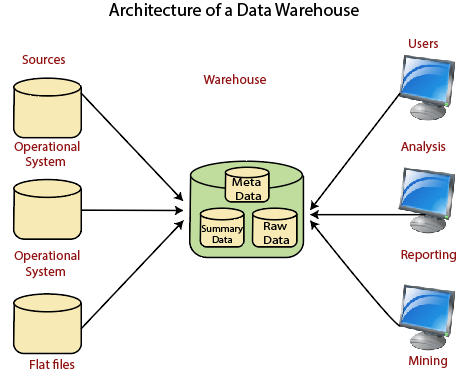
**Data marts Vs Data Warehouses**

|  |  |
| --- | --- |
| **Data marts** | **Data warehouse** |
| 1)subset of data warehouse | Collection of data marts |
| 2)data contained in a summarized form | data contained in a detailed form |
| 3)stores less amount of data | Store huge amount of data4 |
| 4)takes less processing time | Take high time for processing |

**Data Warehouse Life cycle**



The life cycle shows the data is gathered from multiple resources such as structured data ,unstructured data,flat files etc and keeps the data at one place i.e, staging of data. Then using ETL tool or OLAP engine analyse the data and present to the user.



The data is collected from various sources like operational system[for performing day to day transactions],flat files etc, and the data may be meta data,summary data or raw data.This data undergoes into processing,analysis ,reporting,mining and decision making.