

Advanced Database System

Sushmita, Jagjeet Kaur¹, Simranjot Kaur¹, Prof. C.K. Raina²

¹CSE Depaerment, Adesh Institute of Technology, Mohali, Punjab, India

²Associate Professor, CSE, Adesh Institute of Technology, Mohali, Punjab, India

ABSTRACT

Data is collection of raw facts and figure which is basic for decision making. Anything which is unorganized is called data. Database is collection of data, which is organized in a way that allows for easy data retrieval and manipulation. Database system is the way in which databases are arranged. DBMS is the software component or logical tool to handle databases. This software lets users keep, sort, update, retrieve and modify their records in a single database; for example they can keep and update profiles in a client base. DBMS apps are widely used in business to model and manage business objects within corporate databases. The idea of using database management systems in business appeared years ago, and today this idea remains very popular among companies. After being popular, new features are now added to it.. So this research paper is about some of these new features.

Keywords: DBMS, Query Processing And Optimization, Data Wharehouse

I. INTRODUCTION

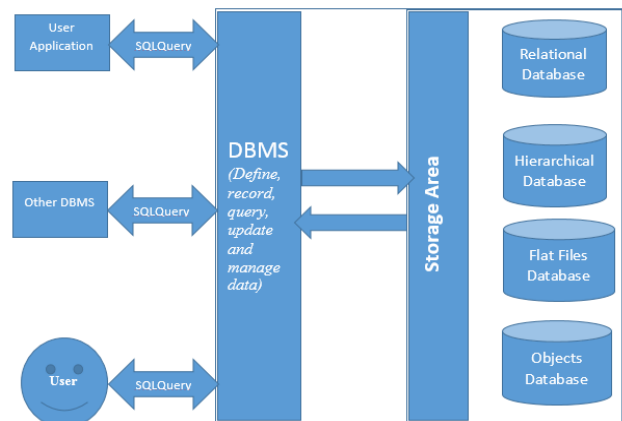
A database management system (DBMS) is a software package designed to define, manipulate, retrieve and manage data in a database. A DBMS generally manipulates the data itself, the data format, field names, record structure and file structure. Database is the back end of an application. A DBMS receives instruction from a database administrator and accordingly instructs the system to make the necessary changes. These commands can be to load, retrieve or modify existing data from the system. There are four main types of database organization:

Relational Database: Data is organized as logically independent tables. Relationships among tables are shown through shared data. The data in one table may reference similar data in other tables, which maintains the integrity of the links among them. This feature is referred to as referential integrity – an important concept in a relational database system. Operations such as "select" and "join" can be performed on these tables. This is the most widely used system of database organization.

Flat Database: Data is organized in a single kind of record with a fixed number of fields. This database type

encounters more errors due to the repetitive nature of data.

Object-Oriented Database: Data is organized with similarity to object-oriented programming concepts. An object consists of data and methods, while classes group objects having similar data and methods.



Hierarchical Database: Data is organized with hierarchical relationships. It becomes a complex network if the one-to-many relationship is violated.

Here are some most popular Database Systems:

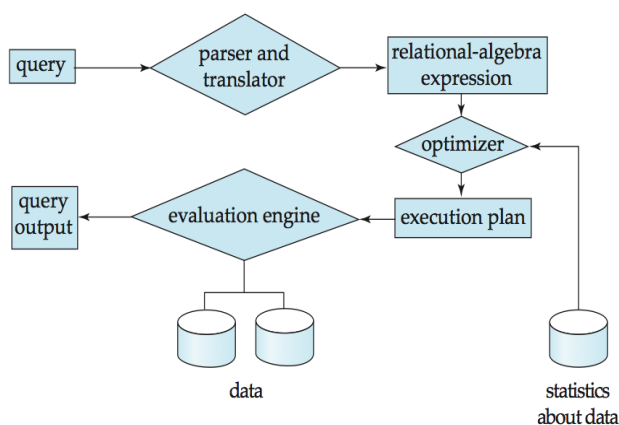
1. Oracle RDBMS
<https://www.oracle.com/database/index.html>
2. Microsoft SQL Server
<https://www.microsoft.com/en-us/sql-server/>
3. IBM DB2
<https://www.ibm.com/analytics/us/en/technology/db2>
4. Teradata
<http://www.teradata.com/>
5. MySQL
<https://www.mysql.com/>

II. QUERY PROCESSING AND OPTIMIZATION

Query processing is a set of activities involving in getting the result of a query expressed in high-level language. The basic steps are:

1. Parsing and translation: Translate the query into its internal form. This is then translated into relational algebra. Parser checks syntax, verifies relation.

2. Optimization: SQL is a very high level language: The users specify what to search for- not how the search is actually done The algorithms are chosen automatically by the DBMS. For a given SQL query there may be many possible execution plans. Amongst all equivalent plans choose the one with lowest cost. Cost is estimated using statistical information from the database catalog.



3. Evaluation: The query evaluation engine takes a query evaluation plan, executes that plan and returns the answer to that query.

2.1 EXAMPLE

SELECT Ename FROM Employee WHERE Salary > 5000;

Translated into Relational Algebra Expression

$\sigma_{Salary > 5000} (\pi_{Ename} (Employee))$

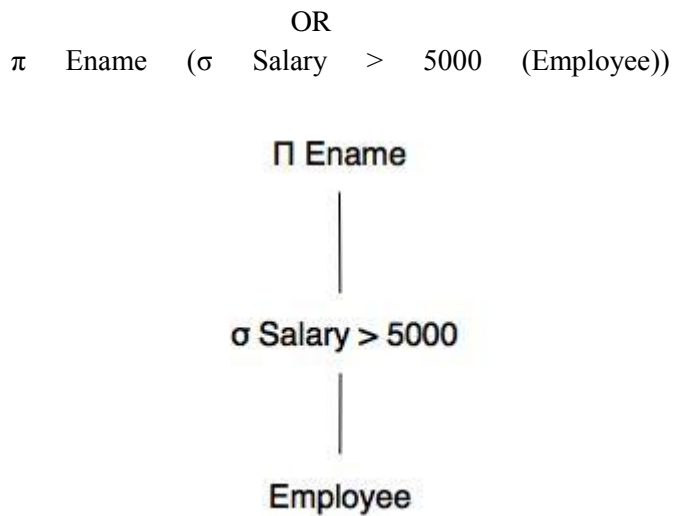


Figure . Query Execution Plan

III. DATA WAREHOUSING

A data warehouse is a relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources.

There are two approaches to data warehousing, top down and bottom up. The top down approach spins off data marts for specific groups of users after the complete data warehouse has been created. The bottom up approach builds the data marts first and then combines them into a single, all-encompassing data warehouse. These technologies help executives to use the warehouse quickly and effectively. They can gather data, analyze it, and take decisions based on the information present in the warehouse. The information gathered in a warehouse can be used in any of the following domains –

- 1. Tuning Production Strategies:** The product strategies can be well tuned by repositioning the products and managing the product portfolios by comparing the sales quarterly or yearly.
- 2. Customer Analysis:** Customer analysis is done by analyzing the customer's buying preferences, buying time, budget cycles, etc.
- 3. Operations Analysis:** Data warehousing also helps in customer relationship management, and making

environmental corrections. The information also allows us to analyze business operations.

3.1 Application of Data Warehouse

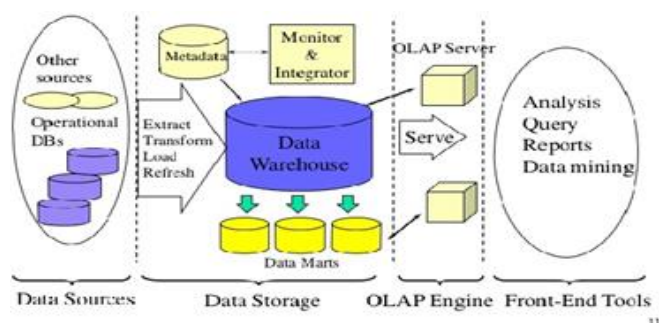
1. Information Processing: Querying, basic statistical analysis and reporting using crosstabs, table, cahrts or graphs. A current trend in data warehouse information processing is to construct low cost web based accessing tools that are then integrated with web browsers.

2. Analytical Processing: A data warehouse is often used as the basis for a decision support system. Data can be analyzed by means of basic OLAP operations, including slice and dice, drill-down. It generally operates on historical data in both summarized and detailed forms.

3. Data Mining: It is also called knowledge discovery. There are two main kinds of models is data mining. One is predictive models and another one is descriptive models.

3.2 Online Analytical Processing

OLAP is computer processing that enables a user to easily and selectively extract and view data from different point of view. OLAP is best known technology that allows a user to slice and dice data or drill-down into data. OLAP is technology that uses a multidimensional view of aggregate data to offer fast access to strategic for the purpose of advanced analysis, deeper understanding.



3.3 Role of the OLAP Server in the Data Warehouse

OLAP servers deliver warehouse applications such as performance reporting, sales forecasting product line and customer profitability, sales analysis, marketing analysis, what-if analysis and manufacturing mix analysis — applications that require historical,

projected and derived data. With OLAP servers' robust calculation engines, historical data is made vastly more useful by transforming it into derived and projected data. Users gain broader insights by combining standard access tools with a powerful analytic engine.

3.4 Data Marts

A data mart is a simple form of data warehouse that is focused on a single subject like sales, Finance, Marketing. Data marts are often built and controlled by a single department with in an organization. Data marts usually takes data from only a few sources because of single subject. The sources could be internal operational systems a central data warehouse or external data. Data marts improve end-user response time ny allowing user to have access to the specific type of data.

3.5 Difference between Data mining & OLAP

Data Mining	OLAP
Used to predict the future.	Used to analyze the past.
Prepare data, launch mining tool and sit back	Multidimensional, drill-down and slice and dice.
It has large number of dimensions.	It has limited number of dimensions.
It deals with summary of data.	It deals with detailed transaction level data.

IV. CONCLUSIONS

The advanced database system is nothing but are some new features in the database system. As we discussed about the query processing and optimization and data ware house, these are the new feature of database system and are in trends. One of the most critical functional requirements of a DBMS is its ability to process queries in a timely manner, query process and optimization are used to process query in timely. And we also discuss the three phases of its through which a query is process and optimized where data ware house is subject oriented, integrated, time –variant and non-volatile collection of data in support of management's decision making process.

V. REFERENCES

- [1]. J. C., Freytag D, and Maier G (eds.) Vossen Query Processing for Advanced Database Systems.Morgan Kaufmann, 1994
- [2]. "Data, data everywhere".The Economist.25 February 2010.Retrieved 9 December 2012.
- [3]. "Community cleverness required".Nature 455 (7209): 1.
4 September 2008.doi:10.1038/455001a.
- [5]. "Sandia sees data management challenges spiral".HPC Projects.4 August 2009.