STYUDY MATRIAL

OF

DATABASE MANAGEMENT SYSYTEM

Prepared by
Smt.Archana Tripathy
Sr.Lecturer in CSE
Govt.Polytechnic , Bhubaneswar

1.0 BASIC CONCPETS OF DBMS

What is Database?

A database is a collection of information that is organized so that it can be easily accessed, managed and updated. Computer databases typically contain aggregations of data records or files, containing information about sales transactions or interactions with specific customers.

What is DATABASE ?

Database

• A collection of related data with

✷ logically coherent structure

✷ Inherent meaning

✷ Purpose, for intended users and applications

✷ Varying size

✷ Scope, content of varying breadth

✷ Physical organization of varying complexity

✷ Various applications with possibly-conflicting objectives

✷ Persistence, existence over a long period of time

Database Management System (DBMS):

It is a collection of programs that enables user to create and maintain a database. In other words it is general-purpose software that provides the users with the processes of defining, constructing and manipulating the database for various applications. Disadvantages in File Processing Data redundancy and inconsistency. Difficult in accessing data. Data isolation. Data integrity. Concurrent access is not possible. Security Problems.

DBMS:-

- A database system is nothing more than a computer based record keeping system i.e. whose purpose is to record and maintains the data or information.
- A DBMS is a software system that allows accessing data contained in a database.
- The objective of DBMS is to provide a convenient and effective method of defining, storing and retrieving information contained in the database.

- The DBMS interfaces with application programs so that the data contained in the database can be used by multiple applications and users.
- A database system involves four major components namely data, hardware, software and user.
- **Data:** - The fundamental unit of database is data. A database is therefore nothing but a repository for stored data. Every data must have two properties namely integrated and shared. Integrated means that the data can be uniquely identified in the database and shared means the data can be shared among several different users.
- Hardware: - The hardware consists of secondary storage volumes on which the database resides.
- Software: - Between the physical database and the users of the system there is a layer of software usually called as database management system (DBMS). All requests from users for access to the database are handled by DBMS.
- Users: - The users are the application programmers responsible for writing application programs that use the database. The application programmers operate on the data in all the usual ways, i.e. retrieving information, creating new information, deleting or changing existing information. The second classes of users are the end users whose task is to access the data of a database from a terminal. The end users use a query language to invoke user written application programs as per the commands from the terminal. The third classes of users are the database administrators or DBA who has control over the whole system.

1.1 PURPOSE OF DATABASE SYSTEMS:-

- Reduction of redundancies:-

- Centralized control of data by the Database administrator avoids unnecessary duplication of data and effectively reduces the total amount of data storage required. It also eliminates inconsistency and extra processing necessary to trace the data in large mass of data.
- Shared data:-

  A database allows the sharing of data under its control by any number of application programs or users.
- Integrity:-

  Centralized control can also ensure that adequate checks are incorporated in the DBMS to provide data integrity. Data integrity means that the data contained in the database is both accurate and consistent.
- Security:-

  Data is of vital importance. These should not be accessed by unauthorized persons. DBMS can ensure that proper access procedures are followed including proper authentication schemes for access to DBMS and additional checks before permitting access to sensitive data. Different levels of security could be implemented for various types of data and operations.

- Conflict resolution:-
  The conflicting requirements of various users and applications are solved by DBMS and best file structure and access methods are chosen to get optimal performance.
- Data independence:-

- DBMS supports both physical and logical data independence. Data independence is advantageous in the database environment since it allows for changes at one level of the database without affection other levels.

Advantages of DBMS:

- 1Data Independence.

- Efficient Data Access.

- Data Integrity and security.

- Data administration.

- Concurrent access and Crash recovery.

- Reduced Application Development Time

APPLICATIONS OF DATABASE SYSTEMS:-

- Databases are widely used. Some of them are as follows.
- Banking-> For customer information, accounts, loans and banking transactions.
- Airlines-> For reservation and scheduled information.
- Universities-> For student, course and grade information.
- Credit card transactions-> For purposes on credit cards and generation of monthly statements.
- Telecommunications-> For keeping records of calls made, generating monthly bills, maintaining balances on prepaid calling cards and storing information about the communication network.
- Finance-> For storing information about holdings, sales and purchases of financial instruments such as stocks and bonds.
- Manufacturing-> For management of supply chain and for tracking production of items in factories, inventory of items in warehouses and orders for items.
- Human resources-> For information about employees, salaries, payroll taxes& benefits and for generation of payments

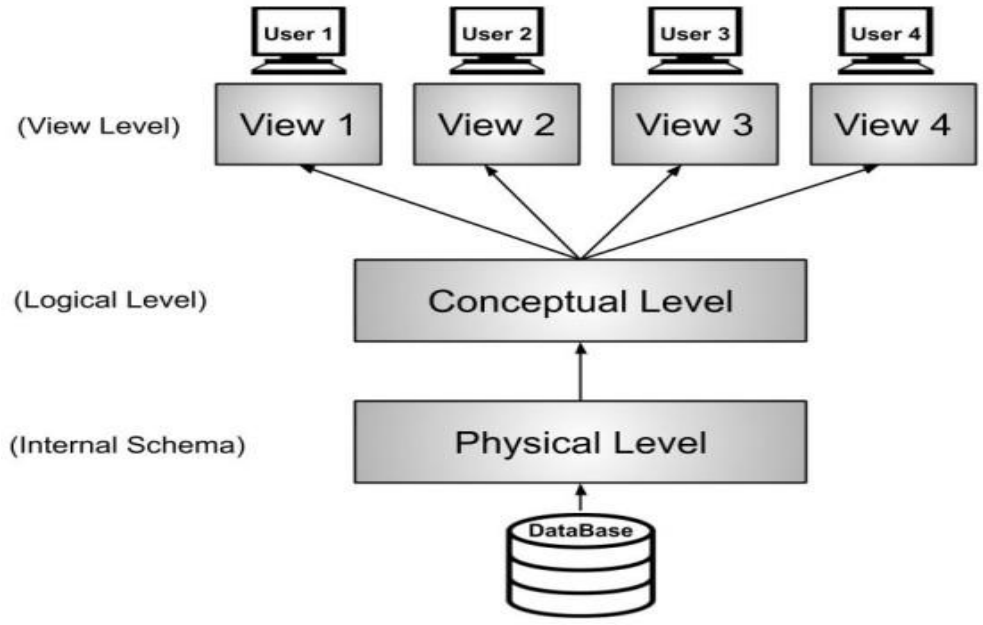
1.2 DATA ABSTRACTION (THREE LEVEL ARCHITECTURE OF DATABASE SYSTEM):-

This architecture is used to provide framework which is extremely useful in describing general database concepts and for explaining the structure of individual systems.

The purpose of designing a generalized database is so that it must have the capability to transform the query asked by the user into programming form so that the system can understand it and will be able to retrieve back the answer of the query.

It is divided into three levels.
I. External level or view level or user view
II. Conceptual level or logical level or global view
III. Physical level or internal level or internal view

The view at each of these levels is described by a schema. A schema is an outline or a plan that describes the records and relationships existing in the view. The schema also describes the way in which entities at one level of abstraction can be mapped to next level.



Levels of Data Abstraction

1. **Physical Level:-**

The lowest level of abstraction describes how the data are actually stored. The physical level describes complex low level data structure in detail. This level is expressed by physical schema which contains the definition of stored record, the method of representing the data fields and access aids used.

The internal level is the one closest to the physical storage. Whenever an external user query a database and the response of the query is available at the conceptual level then it is

provided to the user. If the response is not available at the conceptual level then it is retrieved from the internal level.

2. Logical Level

1. The next higher level of abstraction describes what data are stored in the database and what relationship exists among those data. The logical level thus describes the entire database in terms of a small number of relatively simple structures.
2. It is defined by the logical schema. It describes all the records and relationships. There is one logical schema per database.
3. It also includes features that specify the checks to retain data consistency and integrity.
4. The conceptual schema is basically derived from the internal schema and it can be updated as per the demand of the users.

3. View Level:-

1. The highest level of data abstraction describes only part of the entire database. The system may provide many views for the same database.
2. Each view is described by a schema called external schema. The external schema consists of the definition of logical records and relationships in the view level.
3. The external schema also contains the method of deriving the objects in the external view from the objects in the conceptual view.
4. It is the level closest to the users i.e. it deals with the way in which data is viewed by the users.
5. All the types of database users use a particular language to query the database.
6. The three level architecture is designed in such a way that each and every level maintains their abstraction by their own. The DBMS controls all the levels and the DBMS is basically controlled by the DBA.

1.3 DATABASE USERS:-

There are five different types of database system users differentiated by the way they expect to interact with the system.

- **Naive Users:-**

- They are unsophisticated users who interact with the system by invoking one of the application programs that have been written previously.

Example:- ATM user v The typical user interface for a naive user is a forms interface where the user can fill in appropriate fields of the form. Naive users may also simply read reports generated from the database.

- **Application Programmers:-**

- ❖ Application programmers are computer professionals who write application programs.

- Application programmers can choose from many tools like RAD tools, programming languages, fourth generation languages etc. to develop user interfaces.

- **Sophisticated Users:-**

- Sophisticated users interact with the system without writing programs. Instead they form their requests in database query language. They submit each such query to query processor whose function is to break down DML statements into instructions that the storage manager understands.

- Analysts who submit queries to explore data in the database fall in this category .

Database Administrator (DBA):-

- A person who has central control over the entire database system is called database administrator (DBA).
- The functions of DBA include
- Schema definition
- Storage structure and access method definition
- Schema and physical organization modification
- Granting of authorization for data access
- Routine maintenance

OR

Database Administrators (DBA):

The DBA is responsible for authorizing access to the database, for Coordinating and monitoring its use and for acquiring software and hardware resources as needed. These are the people, who maintain and design the database daily.

DBA is responsible for the following issues.

a. Design of the conceptual and physical schemas: The DBA is responsible for interacting with the users of the system to understand what data is to be stored in the DBMS and how it is likely to be used. The DBA creates the original schema by writing a set of definitions and is Permanently stored in the 'Data Dictionary'.

b. Security and Authorization: The DBA is responsible for ensuring the unauthorized data access is not permitted. The granting of different types of authorization allows the DBA to regulate which parts of the database various users can access.

c. Storage structure and Access method definition: The DBA creates appropriate storage structures and access methods by writing a set of definitions, which are translated by the DDL compiler.

d. Data Availability and Recovery from Failures: The DBA must take steps to ensure that if the system fails, users can continue to access as much of the uncorrupted data as possible. The DBA also work to restore the data to consistent state

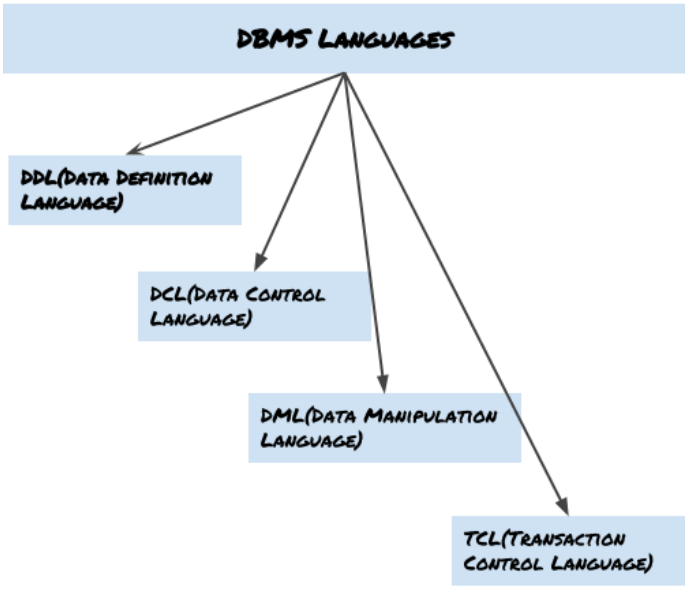
Database Designers:

Database designers are responsible for identifying the data to be stored in the database and for choosing appropriate structures to represent and store this data.

DATA BASE LANGUAGES:-

- A DBMS has appropriate languages and interfaces to express database queries and updates.
- Database languages can be used to read, store and update the data in the database.

Types of Database Language



There are four types of database languages:-

- Data Definition Language (DDL)
- Data Manipulation Language (DML)
- Data Control Language (DCL)
- Transaction Control Language (TCL)

1. Data Definition Language

- **DDL** stands for **D**ata **D**efinition **L**anguage. It is used to define database structure or pattern.
- It is used to create schema, tables, indexes, constraints, etc. in the database.

- Using the DDL statements, you can create the skeleton of the database.
- Data definition language is used to store the information of metadata like the number of tables and schemas, their names, indexes, columns in each table, constraints, etc.

Here are some tasks that come under DDL:

- **Create:** It is used to create objects in the database.
- **Alter:** It is used to alter the structure of the database.
- **Drop:** It is used to delete objects from the database.
- **Truncate:** It is used to remove all records from a table.
- **Rename:** It is used to rename an object.
- **Comment:** It is used to comment on the data dictionary.

These commands are used to update the database schema that's why they come under Data definition language.

2. Data Manipulation Language

DML stands for **D**ata **M**anipulation **L**anguage. It is used for accessing and manipulating data in a database. It handles user requests.

Here are some tasks that come under DML:

- **Select:** It is used to retrieve data from a database.
- **Insert:** It is used to insert data into a table.
- **Update:** It is used to update existing data within a table.
- **Delete:** It is used to delete all records from a table.
- **Merge:** It performs UPSERT operation, i.e., insert or update operations.
- **Call:** It is used to call a structured query language or a Java subprogram.
- **Explain Plan:** It has the parameter of explaining data.
- **Lock Table:** It controls concurrency.

3. Data Control Language

- **DCL** stands for **D**ata **C**ontrol **L**anguage. It is used to retrieve the stored or saved data.
- The DCL execution is transactional. It also has rollback parameters.

  (But in Oracle database, the execution of data control language does not have the feature of rolling back.)

Here are some tasks that come under DCL:

- **Grant:** It is used to give user access privileges to a database.
- **Revoke:** It is used to take back permissions from the user.

There are the following operations which have the authorization of Revoke:

CONNECT, INSERT, USAGE, EXECUTE, DELETE, UPDATE and SELECT.

4. Transaction Control Language

TCL is used to run the changes made by the DML statement. TCL can be grouped into a logical transaction.

Here are some tasks that come under TCL:

- **Commit:** It is used to save the transaction on the database.
- **Rollback:** It is used to restore the database to original since the last Commit.

1.5 DATA DICTIONARY:-

- Information regarding the structure and usage of data contained in the database, the metadata maintained in a data dictionary. The term system catalogue also describes this metadata. The data dictionary which is a database itself documents the data.
- Each database users can consult the data dictionary to learn what each piece of data and various synonyms of data fields mean.
- In an integrated system (i.e. in system where the data dictionary is a part of the DBMS) the data dictionary stores information concerning the external, conceptual and internal levels of the database. It contains the source of each data field value, the frequency of its use and an audit trail concerning updates, including the who and when of each update.