DATABASE MANAGEMENT SYSTEM

UNIT- 1

Database: The database is a collection of inter-related data which is used to retrieve, insert and delete the data efficiently. It is also used to organize the data in the form of a table, schema, views, and reports, etc.

For example: The college Database organizes the data about the admin, staff, students and faculty etc.

Using the database, you can easily retrieve, insert, and delete the information.

Database Management System:

Database management system is a software which is used to manage the database.

For example: MySQL, Oracle, etc. are a very popular commercial database which is used in different applications.

DBMS provides an interface to perform various operations like database creation, storing data in it, updating data, creating a table in the database and a lot more.

It provides protection and security to the database. In the case of multiple users, it also maintains data consistency.

DBMS allows users the following tasks:

1. **Data Definition:** It is used for creation, modification, and removal of definition that defines the organization of data in the database.
2. **Data Updation:** It is used for the insertion, modification, and deletion of the actual data in the database.
3. **Data Retrieval:** It is used to retrieve the data from the database which can be used by applications for various purposes.
4. **User Administration:** It is used for registering and monitoring users, maintain data integrity, enforcing data security, dealing with concurrency control,

monitoring performance and recovering information corrupted by unexpected failure.

Characteristics of DBMS:

1. **Real-world entity:** Should be able to store all kinds of data that exists in this real world.
2. **Relation-based tables:** DBMS allows entities and relations among them to form tables.
3. **Isolation of data and application:** Data and application should be isolated.
4. **Less redundancy:** There should not be any duplication of data in the database.
5. **Consistency:** Consistency is a state where every relation in a database remains consistent.
6. **Query Language:** DBMS has a strong query language.
7. **ACID Properties:** DBMS follows the concepts of Atomicity, Consistency, Isolation, and Durability (normally shortened as ACID).
8. **Multiuser and Concurrent Access:** Multiple users should be able to access the same database, without affecting the other user.
9. **Multiple views:** It supports multiple views to the user, depending on his role.
10. **Security:** Database should also provide security.

Advantages of DBMS:

1. **Controls database redundancy:** It can control data redundancy because it stores all the data in one single database file and that recorded data is placed in the database.
2. **Data sharing:** In DBMS, the authorized users of an organization can share the data among multiple users.
3. **Easily Maintenance:** It can be easily maintainable due to the centralized nature of the database system.
4. **Reduce time:** It reduces development time and maintenance need.
5. **Backup:** It provides backup and recovery subsystems which create automatic backup of data from hardware and software failures and restores the data if required.

6. **multiple user interface:** It provides different types of user interfaces like graphical user interfaces, application program interfaces.

Disadvantages of DBMS:

1. **Cost of Hardware and Software:** It requires a high speed of data processor and large memory size to run DBMS software.
2. **Size:** It occupies a large space of disks and large memory to run them efficiently.
3. **Complexity:** Database system creates additional complexity and requirements.
4. **Higher impact of failure:** Failure is highly impacted the database because in most of the organization, all the data stored in a single database and if the database is damaged due to electric failure or database corruption then the data may be lost forever.

Basic terminology:

1. **Database:** A database is a collection of information that is organized so that it can be easily accessed, managed and updated.
2. **Command:** A command is a string that you send to the server.
3. **Query:** A query is a type of command that retrieves data from the server.
4. **Table:** A table is a collection of data elements organised in terms of rows and columns.
5. **Column (field, attribute):** A single unit of named data that has a particular data type.
6. **Row (record, tuple):** A row is a collection of column values.
7. **View:** A view is an alternative way to present a table (or tables).
8. **Client:** A client is an application that makes requests of the PostgreSQL server.
9. **Server:** The PostgreSQL server is a program that services commands coming from client applications.
10. **Postmaster:** Postmaster creates a new server process in the host operating system.
11. **Transaction:** A transaction is a collection of database operations that are treated as a unit.

12. **Commit:** A commit marks the successful end of a transaction.
13. **Rollback:** A rollback marks the unsuccessful end of a transaction.
14. **Index:** An index is a data structure that a database uses to reduce the amount of time it takes to perform certain operations.
15. **Result set:** When you issue a query to a database, you get back a result set.

Difference between File System and DBMS: -

File System	DBMS
A file system is a software that manages and organizes the files in a storage medium. It controls how data is stored and retrieved.	DBMS or Database Management System is a software application. It is used for accessing, creating, and managing databases.
The file system provides the details of data representation and storage of data.	DBMS gives an abstract view of data that hides the details.
Storing and retrieving of data can't be done efficiently in a file system.	DBMS is efficient to use as there are a wide variety of methods to store and retrieve data.
It does not offer data recovery processes.	There is a backup recovery for data in DBMS.
Protecting a file system is very difficult.	DBMS offers good protection mechanism.
Data inconsistency is higher in the file system.	Data inconsistency is low in a database management system.
Not provide support for complicated transactions.	Easy to implement complicated transactions.
These systems doesn't offer concurrency.	DBMS system provides a concurrency facility.
It doesn't offer backup and recovery of data if it is lost.	DBMS system provides backup and recovery of data even if it is lost.
The centralization process is hard in File Management System.	Centralization is easy to achieve in the DBMS system.

Data Independence:

Data independence refers characteristic of being able to modify the schema at one level of the database system without altering the schema at the next higher level.

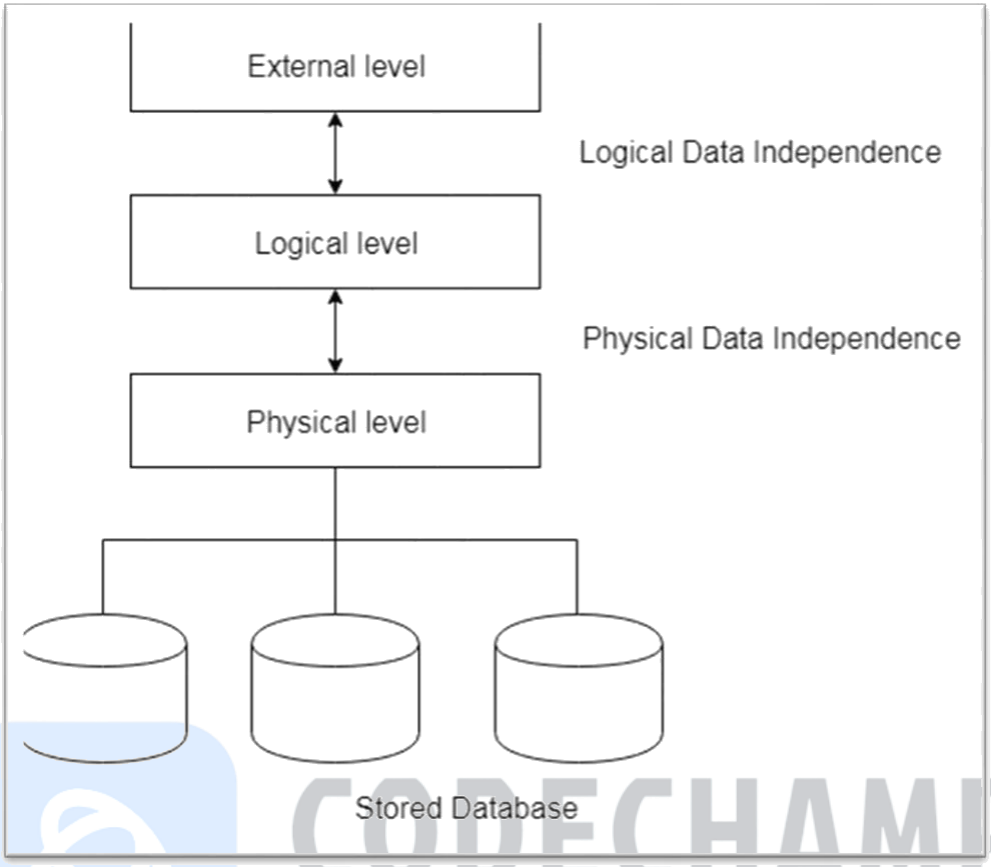
There are two types of data independence:

1. Logical Data Independence

- Logical data independence refers characteristic of being able to change the conceptual schema without having to change the external schema.
- Logical data independence is used to separate the external level from the conceptual view.
- If we do any changes in the conceptual view of the data, then the user view of the data would not be affected.
- Logical data independence occurs at the user interface level.

2. Physical Data Independence

- Physical data independence can be defined as the capacity to change the internal schema without having to change the conceptual schema.
- If we do any changes in the storage size of the database system server, then the Conceptual structure of the database will not be affected.
- Physical data independence is used to separate conceptual levels from the internal levels.
- Physical data independence occurs at the logical interface level.
- Data independence can be explained using the three-schema architecture.



DBMS Architecture:

The DBMS design depends upon its architecture. The basic client/server architecture is used to deal with a large number of PCs, web servers, database servers and other components that are connected with networks.

The client/server architecture consists of many PCs and a workstation which are connected via the network.

DBMS architecture depends upon how users are connected to the database to get their request done.

Types of DBMS Architecture:

Database architecture can be seen as a single tier or multi-tier. But logically, database architecture is of two types like: 2-tier architecture and 3-tier architecture.

1. 1-Tier Architecture

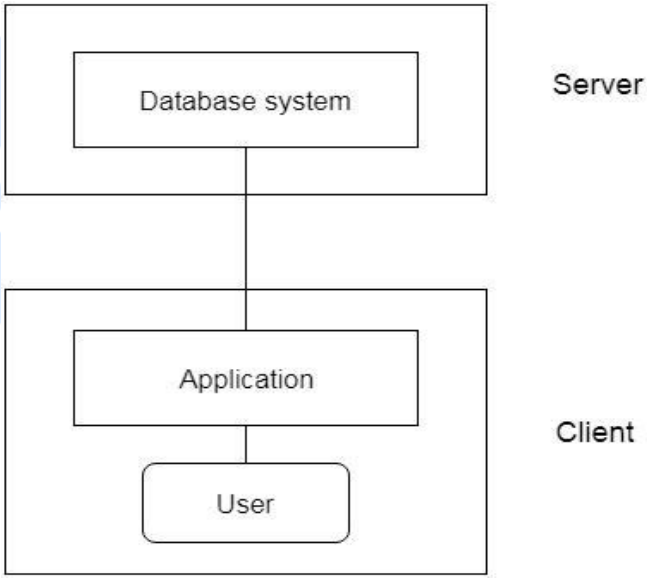
In this architecture, the database is directly available to the user. It means the user can directly sit on the DBMS and uses it.

Any changes done here will directly be done on the database itself. It doesn't provide a handy tool for end users.

The 1-Tier architecture is used for development of the local application, where programmers can directly communicate with the database for the quick response.

2. 2-Tier Architecture

The 2-Tier architecture is same as basic client-server. In the two-tier architecture, applications on the client end can directly communicate with the database at the server side. For this interaction, API's like: ODBC, JDBC are used.



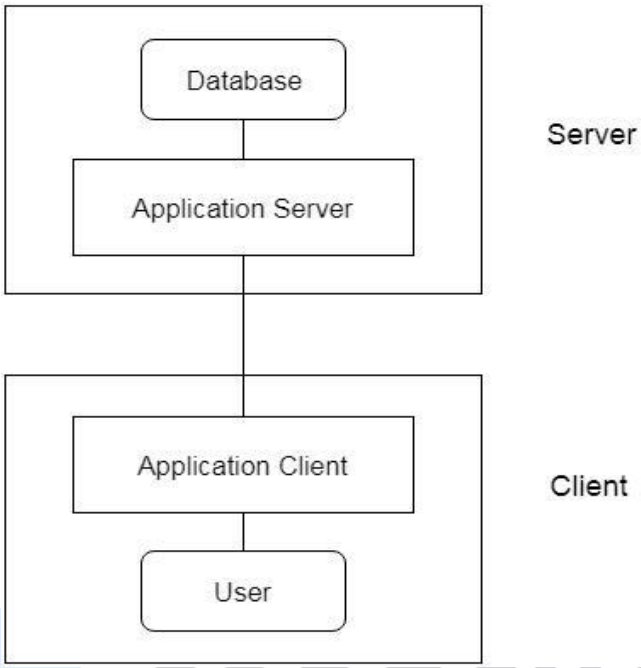
The user interfaces and application programs are run on the client-side.

The server side is responsible to provide the functionalities like: query processing and transaction management.

To communicate with the DBMS, client-side application establishes a connection with the server side.

3. 3-Tier Architecture

The 3-Tier architecture contains another layer between the client and server. In this architecture, client can't directly communicate with the server.



The application on the client-end interacts with an application server which further communicates with the database system.

End user has no idea about the existence of the database beyond the application server. The database also has no idea about any other user beyond the application.

The 3-Tier architecture is used in case of large web application.

Database Schema:

A database schema is the skeleton structure that represents the logical view of the entire database.

It can be categorized into three parts.

1. Physical Schema: Physical schema can be defined as the design of a database at its physical level. In this level, it is expressed how data is stored in blocks of storage.

2. Logical Schema: Logical schema can be defined as the design of database at logical level. In this level, the programmers as well as the database administrator (DBA) work.

3. View Schema: View schema can be defined as the design of database at view level which generally describes end-user interaction with database systems.

Instances:

A database is generally used by many users where insertion and deletion of data occurs frequently. Overall information stored in a database at a particular moment is called the instance of the database.

DBMS languages:

Database languages are used for read, update and store data in a database.

Types of DBMS languages

Data Definition Language (DDL):

DDL is used to define the database structure or schema.

1. CREATE - to create objects in the database
2. ALTER - alters the structure of the database
3. RENAME - rename an object

Data Manipulation Language (DML):

DML is used for accessing and manipulating data in a database.

1. SELECT - Retrieve data from the a database
2. INSERT - Insert data into a table
3. UPDATE - Updates existing data within a table
4. DELETE - deletes all records from a table, the space for the records remains

Data Control Language (DCL):

DCL is used to control the user access to the database.

1. GRANT – To grant access to user
2. REVOKE – To revoke access from user

Transaction Control Language (TCL):

TCL is used to manage the changes made by DML statements.

1. COMMIT - save work done
2. ROLLBACK - restore database to original since the last COMMIT

Database Administrator (DBA):

A Database Administrator, Database Analyst or Database Developer is the person responsible for managing the information within an organization.

Functions and responsibilities of DBAs:

1. Installing and Upgrading an SQL Server
2. Monitoring performance
3. Using Storage Properly
4. Working with Developers
5. Transferring Data, etc.

Data Models:

Data Model is the modeling of the data description, data semantics, and consistency constraints of the data. It provides the conceptual tools for describing the design of a database at each level of data abstraction.

Types of Data Models in DBMS:

Data Models in DBMS include the following:

1. Hierarchical Model:

This concept uses a hierarchical tree structure to organise the data. The hierarchy begins at the root, which contains root data, and then grows into a tree as child nodes are added to the parent node.

2. Network Model:

The main difference between this model and the hierarchical model is that any record can have several parents in the network model. It uses a graph instead of a hierarchical tree.

3. Entity-Relationship Model:

The real-world problem is depicted in visual form in this model to make it easier for stakeholders to comprehend. The ER diagram also makes it very simple for developers to comprehend the system.

4. Relational Model:

The data in this model is kept in the form of a table that is two-dimensional. All of the data is kept in the form of rows and columns. Tables are the foundation of a relational paradigm.

5. Object-Oriented Data Model:

Both the data and the relationship are contained in a single structure that is known as an object in this model. We can now store audio, video, pictures, and other types of data in databases, which was previously impossible with the relational approach.

6. Object-Relational Data Model:

It is a hybrid of relational and object-oriented models. This model was developed to bridge the gap between the object-oriented and relational models.