

Introduction to DBMS

≡ Proficiency

Dreadful

The File Based System

▼ What is File Based System?

One way to keep information on a computer is to store it in permanent files. Application is designed to manipulate data files.

▼ Disadvantages of File Based System

Data Redudancy - same data is present in multiple places or files.

Example - when we want to store a student's phone number, we will have to store it everywhere we have that student's information object.

Data Inconsistency - multiple copies of same data doesn't match with each other.

Example - when we want to update a student's phone number, we will have to update it everywhere and it may happen that we miss one or two instances.

Diffult Data Access - user needs to know exact location of files where data is stored.

Example - to search a student's data from 10k unsorted files.

Security Problem - may lead to unauthorized access of data.

Example - a student may gain access to his / her grades and can update them.

Difficult Concurrent Access - when one application opens a file, it is locked for all other application which leads to extra wait time.

Database System

▼ What is Data?

Data is small chunks of information. It can be in form such as text, numbers, images etc.

▼ What is Database?

Database is a **collection of inter-related** data which helps in **efficient retrieval and modifications**. Database organizes data in the form of tables, views, schemas or reports.

The main purpose of database is to operate a large amount of information by storing, retrieving, and managing data.

Examples - **MySQL**, MongoDB, SQL Server etc.

▼ What is Database Management System?

Database Management System is a software for **storing and retrieving user's data** and also maintains **appropriate security measures**. It consists of a group of program which manipulates the database.

The DBMS system takes user's request for data from an application and instructs the operating system to provide the specific data.

▼ Why do we need a Database Management System?

A Data Base Management System is a system software for easy, efficient and reliable data processing and management. It is used for -

1. Creation of database
2. Retrieval of information from the database
3. Updating the database
4. Managing a database

Main Features -

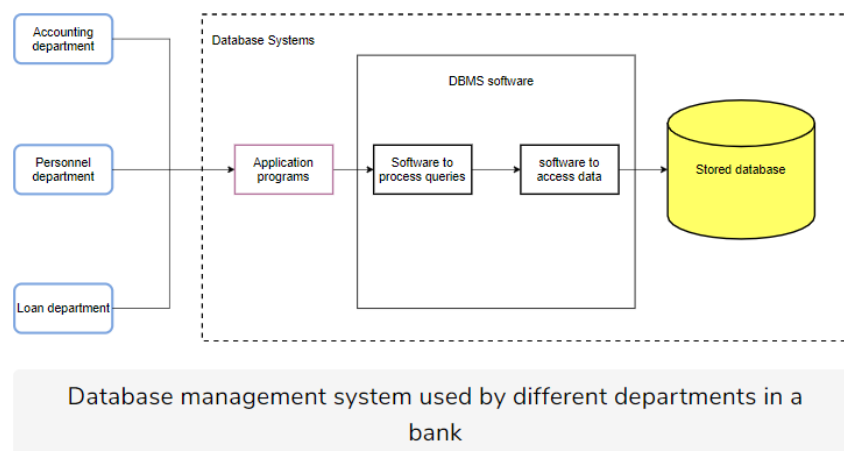
1. **Processing Queries and Object Management** - We can directly store objects in DBMS, because in real life we have objects.
2. **Controlling redundancy and inconsistency** - In DBMS we take care of repeated data by using **normalization** to avoid redundancy and duplicates.
3. **Efficient memory management and indexing** - Object indexing takes place efficiently through database schema based on any attribute of the data.

4. **Concurrency control and transaction management** - Allows multiple user to simultaneously access data. It implements ACID (Atomicity, Durability, Isolation, Consistency) properties to ensure efficient transactions without any corruption.
5. **Access control and ease in accessing data** - Security and can provide access to different parts of data to various users as per their granted permissions.
6. **Integrity Constraints** - Allows us to put different constraints to different fields, so that we don't get anything out of blue.
7. **Multiple User Interface**
8. **Data scalability, expandability, and flexibility**
9. **Security**

▼ Example

▼ 🖱️ Traditional Banking System

With the database approach, we can have the traditional banking system as shown in the diagram below:



▼ What are the functions of DBA (Database Administrator)?

▼ Database Languages

▼ 1. DDL (Data Definition Language)

Used to define the database schema, deals with descriptions of the database schema and is used to create and modify the structure of database objects.

CREATE - is used to create a database or its objects (table, index, function)

DROP - is used to delete objects

ALTER - is used to alter the structure of data

TRUNCATE - is used to remove all the records and the space allocated

COMMENT - is used add comments in data dictionary

RENAME - is used to rename an object

▼ 2. DQL (Data Query Language)

DQL is used to perform queries.

SELECT - is used to retrieve data

▼ 3. DML (Data Manipulation Language)

The SQL commands that deals with manipulation of data.

INSERT - is used to insert data

UPDATE - update existing data

DELETE - delete records

▼ 4. DCL (Data Control Language)

Deals with the rights and permissions of the database system.

GRANT - gives user's access privileges to the database

REVOKE - withdraw those privileges

▼ 5. TCL (Transaction Control Language)

deals with the transaction within the database

COMMIT - commits a transaction

ROLLBACK - rollbacks a transaction in case of any error

SAVEPOINT - sets a savepoint within transaction

SET TRANSACTION - specify characteristics for transaction

Data Modeling

▼ What is a Data Model?

Data model is a collection of concepts describing **data, data relationships, data semantics and data constraints**.

▼ Types of Data Model

▼ High Level Conceptual Data Model

shows data in the form of how people perceive data. Example - Entity Relationship Mode, which uses concepts like entities, attributes, and relationships.

▼ Record-Based Logical Data Model

Users can understand but are similar to the way data is stored in computer.

Examples -

1. Hierarchical Model - data is organized into a **tree-like structure**. Only one parent for a single node.
2. Network Model - expands upon hierarchical structure, each node can have multiple parents and children. **Graph**.
3. Relational Model - represents data as a **relation or tables**.

▼ Physical

represents how **data is stored in computer memory**. It also highlights how tables are built and related to each other in the database.

▼ What is Database Schema?

A schema is a **blueprint of database**. The names of tables, columns, datatype, functions, and other objects are included in the schema.

We use schema diagram to display the schema of a database.

Note - Schema diagram and schema is not the same thing. Schema diagram only displays few necessary details.

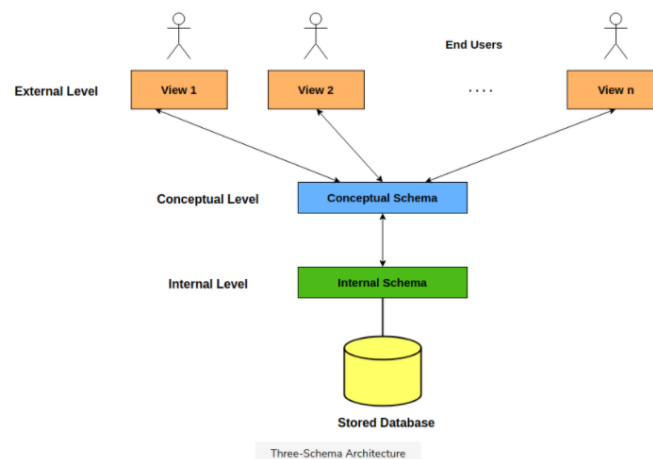
▼ What is Database Instance?

An instance is **information collected in a database at some specific moment in time, also known as database state**. It is like a snapshot of the current state.

Each time there is any modification operation in the database, the state of the database changes.

▼ The three-schema architecture and how is it different from two-schema?

The three-schema architecture is used to separate user application from the physical database.



▼ External Schema (User Level)

It only describes the **part of the database that a specific user is interested in**. It hides any other unrelated details of the database, for example - how data is retrieved or how the data is stored. It is different for everyone as per the needs.

▼ Conceptual Schema

It describes the database structure for the whole community of users. **Record-Based Logical Model**. It hides how the database is stored in physical level but shows information like data types, entities, relationships etc.

▼ Internal Schema

It describes how the database is stored on **physical storage devices**. It uses a **physical data model** and describes the complete details of data storage.

▼ Advantages -

1. User doesn't need to deal with the physical storage directly.
2. DBA can do modifications without disturbing user's view.
3. Every user can access data and see a customized view of data.

▼ Difference with Two-Schema Architecture

Three Schema

1. Web-based Application
2. The application logic resides in the middle-tier, it is separated from user interface or data.
3. Three Layers - Client, Business, Data
4. Easy to build
5. Runs faster
6. Secured
7. Example - Registration Form

Two Schema

1. Client-Server Architecture
2. The application logic is either buried inside the user interface or within database (server) or both.
3. Two Layers - Client and Data
4. Complex to build and maintain
5. Runs slower
6. Less secured
7. Example - Railway Reservation System

▼ What is Data Abstraction?

Data abstraction refers to the process of hiding irrelevant details from the user. We only display the data relevant to the user.

There are mainly three levels of data abstraction to achieve Data Independence.

Data Independence - User and data should not interact with each other directly.

▼ Three levels of Data Abstraction?

▼ View Level

This level tells the application about how the data should be shown to the user. The details required by the user and the data that the user can modify or view.

This increases the security of the database. Different users will have different view as per their authorization.

▼ Conceptual Level

This level shows how the data is actually stored and structures. We have different data models by which we can store data. It shows the entity, relationships and data types.

▼ Physical Level

It tells the location where the data is stored, along with all the details. The DBA decides where to store which fragment of data.