

## Module 01: Introduction to Database Concepts

### Contents ...

- Basic Concepts of Data,
- Database and DBMS,
- Applications of Databases,
- Advantages of Databases over File Processing System,
- Three Level Architecture of Database System,
- Data Abstraction and Data Independence
- Database Languages,
- Database Users,
- Database Administrator and its functions,
- Overall System Structure.

### 1.1 Basic concepts of Data

Data is collection of facts and figures. It is *bits* of information, but not the information itself. When data is processed, organized, interpreted and converted into meaningful or useful data then it is called information. For example, a list of dates data is meaningless without the information that makes the dates relevant (dates of holiday).

### 1.2 Database and DBMS

A database is a collection of inter related data. For example, student's data such as student name roll number, address, branch etc.

DBMS stands for **Database Management System (DBMS)**. It is a collection interrelated data and set programs to manipulate (store, retrieve, update and delete) this data. DBMS is software used to manage the database. Example of DBMS are SQL server, [MySQL](#), [Oracle](#), etc . These are the very popular commercial database software used in different industries. DBMS allows users to following operations:

1. Creating Database
2. Storing of Data
3. Updating of Database
4. Retrieval Database
5. Managing Database (Monitoring users, maintain data integrity, enforcing data security, dealing with concurrency control, monitoring performance and recovering Data)

### 1.3 Applications of DBMS

1. **Colleges and universities:** To maintains the information of students, teachers, staff, courses, results, grades and other resources.
2. **Banking:** To maintains the information of customers and their payment transactions such as transfer, deposit, load, withdraw etc.
3. **Online shopping:** To maintains the information of the product, customers, credit details and shipping information.
4. **Telecommunication:** DBMS helps to keep call records, bills, balances, network usage, customer details etc.
5. **Airlines:** It helps to for reservations and scheduling information of flights and passenger.
6. **Finance:** Used to store information about the stocks, sales, and purchases of stocks and bonds, etc.
7. **Manufacturing:** It is used to maintain and tracking the records of supply chain, production of items and Inventories.
8. **Sales:** It is used for storing the customer, product & sales information.
9. **Human resources:** Used to maintain the employee records, salaries, tax deductions etc.

### 1.4 Advantages and Disadvantages of DBMS over File Processing System

A File system is collection of files stored on operating system. It allows access to single files or tables at a time. In a File System, data is directly stored in set of files. It contains set of files that have no relation to other files. Example of file is storing data in text file, word file, excel file etc. Following are the advantages of storing data in DBMS over the file system:

1. **Concurrent access of data:** In File system, only one user can access same data at the same time. In DBMS, more than one user can access the same data at the same time in DBMS.
2. **Data redundancy:** Data redundancy is the duplication of data. It often leads to higher storage costs and poor access time. In file system there is more data redundancy compared to DBMS.
3. **Data inconsistency:** Redundancy may leads to inconsistency (Incorrect data) of data in file system in file system. DBMS controls redundancy by maintaining a single repository of data. Therefore there is no or less redundancy, data remains consistent in DBMS.
4. **Data integrity:** It refers to the accuracy and consistency of data. DBMS provides constraints on the data before inserting it in to the database. The file system does not provide any constraints before data insertion.
5. **Data Security:** DBMS provides more security from unauthorized access of users than file system.
6. **Data Isolation:** Data isolation refers to data scattered in various files, and files may be in different formats and therefore writing new application programs to retrieve the appropriate data is difficult. In DBMS there is no data isolation whereas files system suffers from data isolation.
7. **Atomicity:** Atomicity refers to all or nothing. In DBMS either all operations in transaction execute or none operation. File system does not provide atomicity.

Following table difference between DBMS and Files sytem

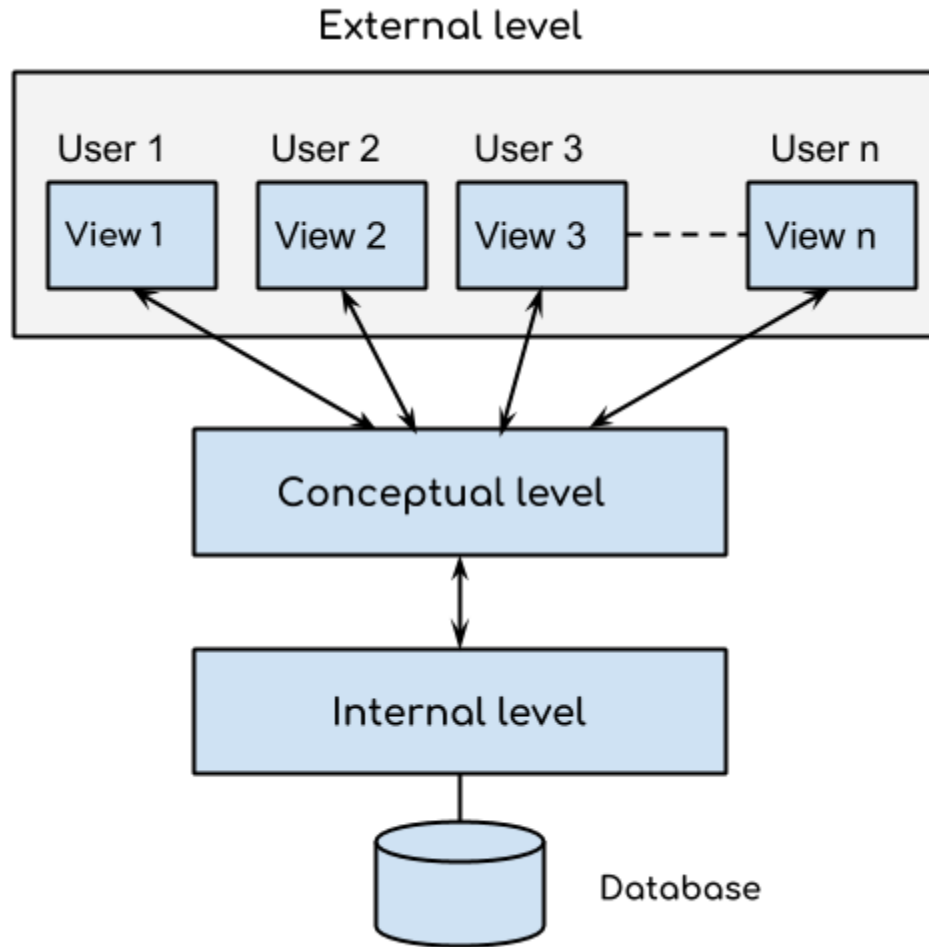
Features	DBMS	File System
Concurrent access	Yes	No
Data redundancy	Less	More
Data inconsistency	Less	More
Data integrity	Yes	No
Data Security	More	Less
Data Isolation	No	Yes
Atomicity	Yes	No
Examples	SQL SERVER, MYSQL, ORACLE	Text File, Word file, Excel file

### 1.5 Three Level Architecture of Database System or 3 levels of data abstraction

DBMS is made up of complex data structures. Therefore to ease the user interaction with database, it is required hide irrelevant details from user . This process of hiding is called data abstraction. This abstraction is achieved in the three level architecture of DBMS. This architecture is also called ANSI/SPARC architecture. It also used to separate the user applications and physical database. In this architecture, database is divided into three levels:

1. External level
2. Conceptual level
3. Internal level

Figure shows the three level architecture of DBMS:



1. **External Level:** External Level or view level is highest level and closest to the user. In this level, several users can view their desired data. The external level only shows the relevant database content to the users in the form of different views and hides the rest of the data. In this level different users can see the database as a different view as per their requirements.
2. **Conceptual Level:** Conceptual level or logical level is higher level than the physical level. It describes how the database appears to the users conceptually and the relationships between various data tables.
3. **Internal Level:** Internal level or physical level is the lowest level in the architecture. It describes how data is actually stored in the database. The physical level also responsible for level is also responsible for allocating space, compression and encryption of data.

## 1.6 Data Independence :

Data independence refers characteristic of DBMS where user is able to modify the schema at one level without affecting the schema at the next level. It helps to keep data separated from all programs which are using it. There are two types of data independence:

1. Logical data independence
2. Physical data independence

1. **Logical Data Independence:** It refers characteristic of being able to change the conceptual schema without having to change the external schema. This data independence is used to separate the external level from the conceptual view
2. **Physical data independence:** It refers to characteristic of being able change the internal schema without affecting the conceptual schema. It used to separate conceptual levels from the internal levels. Compared to Logical Independence, it is easy to achieve physical data independence.

## 1.7 Database Languages

Database language provides interface for writing the different database queries. Database language is used to implement operations such as creating database, tables, inserting data, updating data, deleting data. Following are the database languages:

1. Data definition Language
2. Data Manipulation Language
3. Data Control Language
4. Transaction Control Language

1. **Data Definition language (DDL):** Data Definition Language is used to define database, tables, indexes, constraints, etc. in the database. Following are the some of the operations under DDL.
  - **Create database/Table:** It is used to create database or table.
  - **Alter:** It is used to alter the structure of the database.
  - **Drop:** It is used to delete the database.
  - **Truncate:** It is used to remove all records from a table.
  - **Rename:** It is used to rename an table.
2. **Data manipulation Language (DML) :** DML is used for accessing and manipulating data from the database. It is used to insert, retrieve, update delete data in the database. Following are the some of the operations 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.

3. **Data Control Language (DCL):** It is used granting and revoking user access on a database. Here are some operations under DCL:
  - **Grant:** It is used to give privileges to users on the particular database.
  - **Revoke:** It is used to take back privileges from the user.
4. **Transaction Control Language (TCL):** TCL is used to control transactions. It used to run the changes made by the Data manipulation language. Following are operations under the TCL :
  - **Commit:** It is used to persist the changes made by DML on the database.
  - **Rollback:** It is used to rollback the changes made in the database

### 1.8 Database Users

Database users are the person who really use and take the benefits of database. There will be different types of users depending on their need and way of accessing the database.

Following are the users of DBMS:

1. **Application Programmer** – Application programmer is the developer who interacts with the database by means of DML queries.
2. **Sophisticated User** – Sophisticated User is also database developer, who writes SQL queries to select/insert/delete/update data. It does not use any application or programs to request the database.
3. **Specialized User** – Specialized User is also sophisticated users, who write special database application programs.
4. **Naïve User:** It is unsophisticated user who has knowledge of DBMS. Naïve user just works on database applications and get expected results.
5. **End User** – End user is stand alone user who actual use database for their personal use.

### 1.9 Database Administrator

Important steps in the life cycle of DBMS are designing, implementing to administration. Person who manages administration and maintenance of database is called as Database Administrator or DBA. He is person who controls the design and the use of the database. Following are the roles of DBA:

1. **Schema Definition:** DBA is responsible for writing definition the logical Schema of the database. A Schema refers to the overall logical structure of the database.
2. **Storage Structure and Access Method Definition:** the DBA decides how the data is to be represented and accessed in the database.
3. **Assisting Application Programmers:** The DBA provides assistance to application programmers to develop application programs.
4. **Physical Organization Modification:** The DBA modifies the physical organization of the database to reflect the changing needs of the organization and to improve performance of system.
5. **Approving Data Access:** The DBA determines which user needs access to which part of the database.
6. **Monitoring Performance:** The DBA monitors performance of the system.
7. **Backup and Recovery:** The DBA ensures this periodically backing and recovery of database.

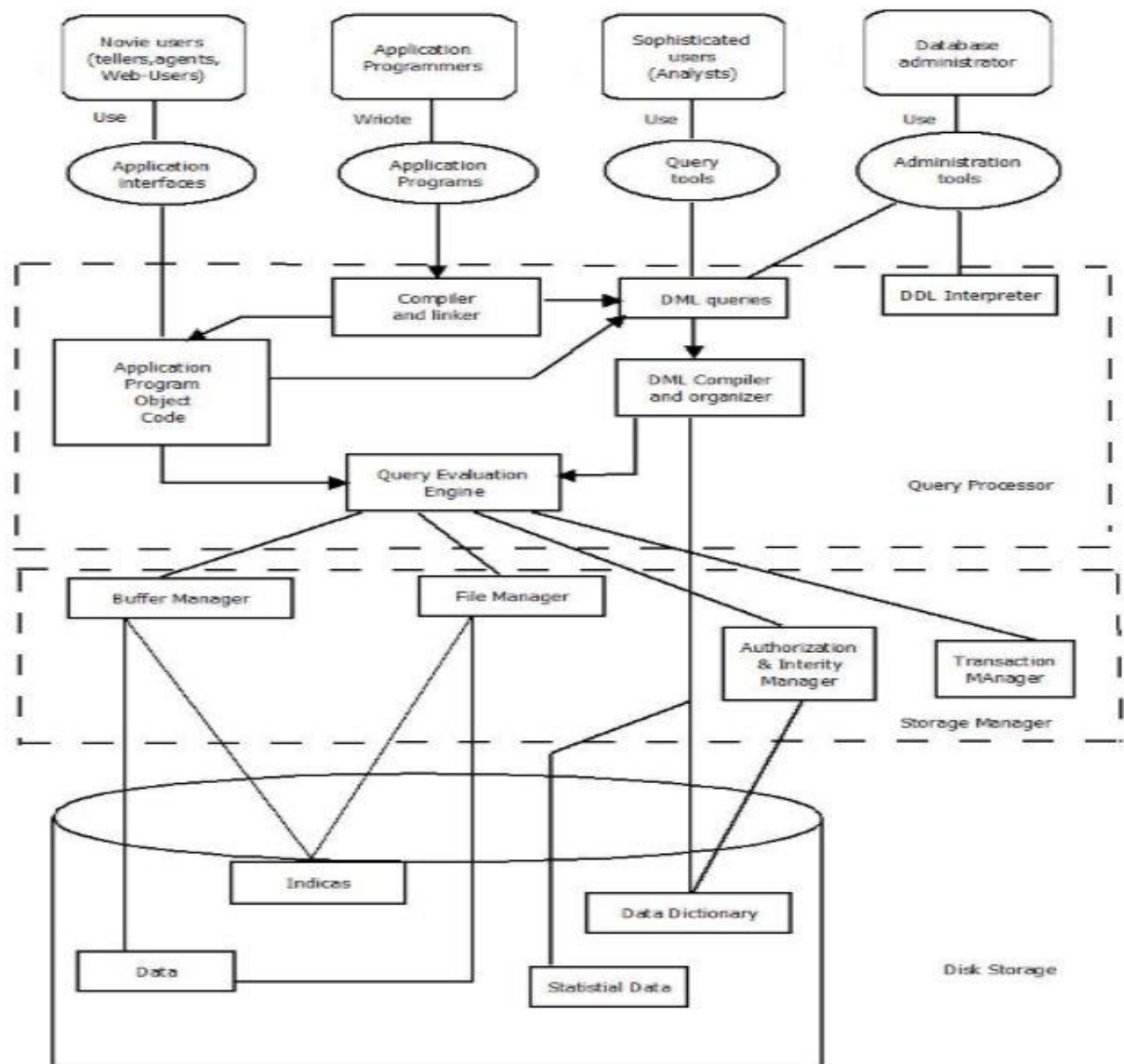
## 1. 10 Instance and schema in DBMS

**1. Schema:** It is the overall design of a database. There are three types of schema: physical schema, logical schema and view schema. Physical level describes how the data stored in blocks of storage .therefore design at this level is called **physical schema**. Logical level certain types of data records get stored in data structures Design of database at logical level is called **logical schema**. View level generally describes end user interaction with database systems. Design of database at view level is called **view schema**.

**2. Instance:** The data stored in database at a particular moment of time is called instance of database. Database schema defines the variable declarations in tables that belong to a particular database; the value of these variables at a moment of time is called the instance of that database.

## 1.11 Overall System Structure

The architecture of DBMS is influenced by the computer system on which the database is running. Following figure shows the overall architecture of DBMS:



**Fig1. Database System Architecture.**

The architecture is divided into following parts:

- 1. Database Users**
- 2. Query Processor**
- 3. Storage Manager**

**1. Database Users :** based on roles and responsibility, users are divided into :



- Application Programmer
- Sophisticated User
- Specialized User
- Naïve User:
- End User
- DBA

2. **Query Processor:** It accepts different queries from user and solves it by accessing the database. It consist of following components :

- **DDL interpreter:** It interprets DDL statements and fetch the definitions in the data dictionary.
- **DML compiler:** It translates DML statements into low level instructions that the query evaluation engine understands.
- **Query evaluation engine:** It execute low-level instructions generated by the DML compiler.

3. **Storage Manager: It is a** program module which acts like interface between the data stored in a database and the application programs and queries. It is responsible for storing, retrieving and updating data. Following are the components of storage manager :

- **Authorization and integrity manager:** It checks integrity constraints and authority of users to access required data.
- **Transaction manager:** It ensures consistency of the database in case of system failures.
- **File manager:** It is responsible for managing the space and data structures used to represent information stored on disk.
- **Buffer manager:** It is responsible for retrieving data from disk storage into main memory. It enables the database to handle data sizes that are much larger than the size of main memory.
- **Data structures implemented by storage manager.**
- **Data files:** Stored in the database itself.
- **Data dictionary:** It stores metadata about the structure of the database.
- **Indices:** It is used to provide fast access to data in DBMS.

## ➤ Question Bank

### Multiple Choice Questions

1. **Database is collection of \_\_\_\_ data.**  
a) Relative                      b) Interrelated                      c) Different.
2. **\_\_\_\_\_ is function of database administrator.**  
a) schema definition      b) Downloadingc) Network Formation

3. **There are \_\_\_\_\_ levels in database architecture.**  
 a) one                                      b) two                                      c) three
4. **\_\_\_\_\_ is Database Language.**  
 a) Multicasting                              b) DML                                      c) Procedural Language
5. **DDL stands for**  
 a) Data Definition Language                                      b) Data Duplication Language  
 c) Data Dilation Language
6. Insertion is Used in  
 a) DDL                                      b) DML
7. **\_\_\_\_\_ is Database application.**  
 a) Banking                                      b) Device switching
8. **\_\_\_\_\_ is Database User**  
 a) Cumulative                                      b) GSM                                      c) Naïve
9. **Redundancy is**  
 a) Mismatching of Data      b) Duplication of data      c) Internet work
10. **There are \_\_\_\_\_ database users**  
 a) one                                      b) Two                                      c) four
11. The view of total database content is  
 a) Conceptual view                                      b) Internal view      c) External view
12. ODBC stands for  
 a) Object Database Connectivity.                                      b) Oracle Database Connectivity.  
 c) Open Database Connectivity.
13. An entity set that does not have sufficient attributes to form a primary key is a  
 a) strong entity set.                                      b) weak entity set.                                      c) simple entity set.
14. In an E-R diagram attributes are represented by  
 a) rectangle                                      b) square                                      c) ellipse
15. The statement in SQL which allows to change the definition of a table is

### Subjective Questions

- Q1. Explain Disadvantages of file processing System.
- Q2. Explain Overall architecture of DBMS with diagram.
- Q3. Explain Different Database Users.
- Q4. Explain Functions Of Database Administrator.
- Q5. Explain Instances And Schema.
- Q6. Explain Different Database Languages.
- Q7. Define Database and database management System.