

## Unit-1

### Database and Database Users:-

Data → Data are the raw facts that are found after some experiment, observation or after some experiments. Data itself do not provide any meaning but after processing it becomes information.

Database → The collection of data organized in some specific manner is known as database. For example, university database for maintaining information about students, courses and grades in university.

Database Management System → The database, its processing methods and the set of rules and conditions to be followed is collectively known as database management system (DBMS). It is the way to store data and also provides mechanism for manipulation of data. It is basically just a computerized record-keeping system. The primary goal of DBMS is to store and retrieve data in both convenient and efficient manner. Oracle, MySQL, SQL-Server, MS Access etc. are some examples of DBMS.

#### \* Applications of DBMS:

- Banking → To store information about customers, their account number, balance etc.
- Airlines → For reservations and schedule information.
- Telecommunication → To keep records of customers, call made, balance left, generating monthly bills etc.
- Sales → To keep information of customers, products list, purchase information etc.
- Universities → To keep record of students, courses, marks of students etc.
- Human Resources → To keep record of employee, their ~~salat~~ salary etc.
- Manufacturing → To store orders, tracking production of items etc.



## ⊗ Characteristics of the Database Approach:

In past traditional file processing approach was used in which each user defines and implements the files needed for a specific software application as part of programming the application. But in database approach, a single repository maintains data that is defined once and then accessed by various users repeatedly through queries, transactions and application programs. The main characteristics of the database approach versus the file-processing approach are as follows:-

### 1) Self-describing nature of a database system:

The database system contains not only the database itself but also a complete description. The definition is stored in the DBMS catalog which contains information such as structure of each file, the type and storage format of each data item. The information stored in the catalog is called meta-data and it describes the structure of the primary database.

The newer types of database systems do not require meta-data. Rather the data is stored as self-describing data that includes the data item names and data values together in one structure. These database systems are known as NOSQL systems.

### 2) Insulation between programs and data, and data abstraction:

In traditional file processing, the structure of data files is embedded in the application programs, so any changes to the structure of a file may require changing all programs that access the file. But DBMS access programs do not require such changes in most cases. The structure of data files is stored in the DBMS catalog separately from the access programs. We call this property program-data-independence.



### iii) Support of Multiple Views of the Data:

A database, typically has many types of users, each of whom may require a different perspective or view of the database. A view may be a subset of the database or it may contain virtual data. Some users may not need to be aware of whether the data they refer to is stored or derived. A multuser DBMS whose users have a variety of distinct applications must provide facilities for defining multiple views. For example one user of database may be interested only in accessing and printing the transcript and a second user who is interested only in checking that students have taken all the prerequisites of each course.

### iv) Sharing of data and multuser transaction processing:

A DBMS must allow multiple users to access the database at the same time. This is essential if data for multiple applications is to be integrated and maintained in a single database. The database must include concurrency control software to ensure that several users trying to update the same data.

DBMS should ensure that concurrent transactions operate correctly and efficiently. These types of applications are generally called online transaction processing (OLTP) applications. The concept of ~~a~~ a transaction has become central to many database applications. A transaction is an executing program or process that includes one or more database accesses such as reading or updating of database records.

### ⊗. Actors on the Scene:-

For a small personal database such as the list of addresses, typically one person defines, constructs and manipulates the database and there is no sharing. But in large organizations many people are involved in the design, use and maintenance of a large database with hundreds or thousands of users. The people whose jobs involve the day-to-day use of a large database are called actors on the scene.



i) Database Administrators → In database environment the primary resource is database itself and secondary resource is the DBMS and related software. Administrating these resources is the responsibility of the database administrator (DBA). The DBA is responsible for authorizing access to the database, coordinating and monitoring its use, and acquiring software and hardware resources as needed. The DBA is also accountable for problems such as security breaches and poor system response time.

ii) 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. It is the responsibility of database designers to communicate with all prospective database users in order to understand their requirement and to create a design that meets the requirements. The database design must be capable of supporting the requirements of all user groups.

iii) End Users → End users are the people whose jobs require access to the database for querying, updating and generating reports. There are several categories of end users:

- Causal end users → They occasionally access the database, but they may need different information each time. They use a sophisticated database query interface to specify their requests.
- Naive end users → They constantly query and update the database using standard types of queries and updates called canned transactions that have been carefully programmed and tested.
- Sophisticated end users → It includes engineers, scientists, business analysts and others in order to meet their complex requirements.



• Standalone users → They maintain personal databases by using ready-made, program packages that provide easy-to-use menu-based or graphics-based interfaces.

### iv) System Analysts and Application Programmers (Software Engineers):

System analysts determine the requirement of end users specially naive and parametric end users and develop specifications for standard canned transactions that meet the requirements. Application programmers implement these specifications as programs then they test, debug, document and maintain these canned transactions.

### ⊗ Workers Behind the Scene:

Those who work to maintain the database system environment but who are not actively interested in the database contents as part of their daily job are called workers behind the scene. They include following categories:

i) DBMS system designers and implementors → They design and implement the DBMS modules and interfaces as a software package. A DBMS is very complex software consisting many components including modules for implementing catalog, query language processing, interface processing, handling data recovery and security etc. The DBMS must interface with other system software such as operating system and compilers for various programming languages.

ii) Tool developers → They design and implement tools which are software packages that facilitate database modeling and design and improve performance. Tools are optional packages that are often purchased separately. In many cases, independent software vendors develop and market these tools.

iii) Operators and maintenance personnel → They are also called system administration personnel and are responsible for the actual running and maintenance of the hardware and software environment for the database system.



## Advantages of using the DBMS Approach:-

The DBA must utilize the different capabilities to accomplish a variety of objectives related to the design, administration and use of a large multiuser database. Following are some advantages of using DBMS and the capabilities that a good DBMS should possess.

i) Controlling Redundancy:- The file based management systems contained multiple files that were stored in many different locations in a system or even across multiple systems. Because of this, there were sometimes multiple copies of the same file which lead to data redundancy.

This is prevented in database as there is a single database and any change in it is reflected immediately. Because of this, there is no chance of encountering duplicate data.

ii) Restricting unauthorized access:- Data Security is vital concept in a database. Only authorised users should be allowed to access the database and their identity should be authenticated using a username and password. Unauthorised users should not be allowed to access the database under any circumstances as it violates the integrity constraints.

iii) Providing Persistent Storage:- Programming languages typically have complex data structures such as struct or class definitions in C++ or Java. The values of program variables or objects are discarded once a program terminates unless the programmer explicitly stores them in permanent files, which often involves converting these complex structures into a format suitable for file storage. When need arises to read this data, once more the programmer must convert the file.

The DBMS software is compatible with these programming languages such as C++ and Java and automatically performs necessary conversions. Hence a complex object can be stored in an object-oriented DBMS. Such an object is said to be persistent, since it survives the termination of program execution and can be later directly retrieved by another program.



4.

v) Providing Backup and Recovery:—The backup and recovery subsystem of DBMS is responsible for recovery. For example, if the computer fails in the middle of a complex update transaction, the recovery system is responsible for making sure that the database is restored to the state it was in before the transaction started executing. Disk backup is also necessary in case of a catastrophic disk failure.

v) Providing Multiple User Interfaces:—Many types of users with varying levels of technical knowledge use a database, so a DBMS should provide a variety of user interfaces. These include apps for mobile users, query language for casual users, programming language interfaces for application programmers, forms and commands for parametric users, and menu-driven interfaces and natural language interface for standalone users. Both forms-style interfaces and menu-driven interfaces are commonly known as graphical user interfaces (GUIs).

vi) Flexibility:—It may be necessary to change the structure of a database as requirements change. For example, a new user group may emerge that needs information necessary to add a file to the database or to extend the data elements in an existing file. Modern DBMSs allow certain ~~database without~~ types of evolutionary changes to the structure of the database without affecting the stored data and the existing application programs.

vii) Availability of Up-to-Date Information:—As soon as users update is applied to the database, all other users can immediately see that update. This availability of up-to-date information is essential for many transaction-processing applications such as reservation systems or banking databases.



viii) Economics of Scale: - The DBMS approach reduces the amount of wasteful overlap between activities of data-processing personnel in different projects or departments. This enables the whole organization to invest in more powerful processors, storage devices, rather than having each department purchase its own low performance equipment. This reduces overall costs of operation and management.

### ⊗. Purpose of DBMS/Why DBMS is required?

Traditionally, file processing system was used to manage information. It used to store data in various files of different application programs to extract or insert data to appropriate file.

File processing system has several drawbacks due to which DBMS is required. DBMS removes the problems found in file processing system. Some major problems of file processing system are:

- i) Data redundancy and inconsistency.
- ii) Difficult in accessing data.
- iii) Data isolation.
- iv) Integrity problem.
- v) Security problem
- vi) Atomicity problem etc.