

Learning Resource On Database Management Systems

Chapter-1 Introduction to DBMS

**Prepared By:
Kunal Anand, Asst. Professor
SCE, KIIT, DU, Bhubaneswar-24**

Chapter Outcomes

- After the completion of this chapter, the students will be able to:
 - Explain the limitations of traditional file processing system.
 - List out the characteristics of the database approach.
 - Explain different types of database users.
 - Classify different types of databases.
 - Identify the advantages and disadvantages of DBMS.
 - List the areas where DBMS should not be used.

Organization of this Chapter

- Introduction
- Data Processing
- Traditional data storage and its limitations
- Characteristics of the database approach
- Types of Database
- Database Users
- Advantages and Disadvantages of Database approach
- When not to use a DBMS

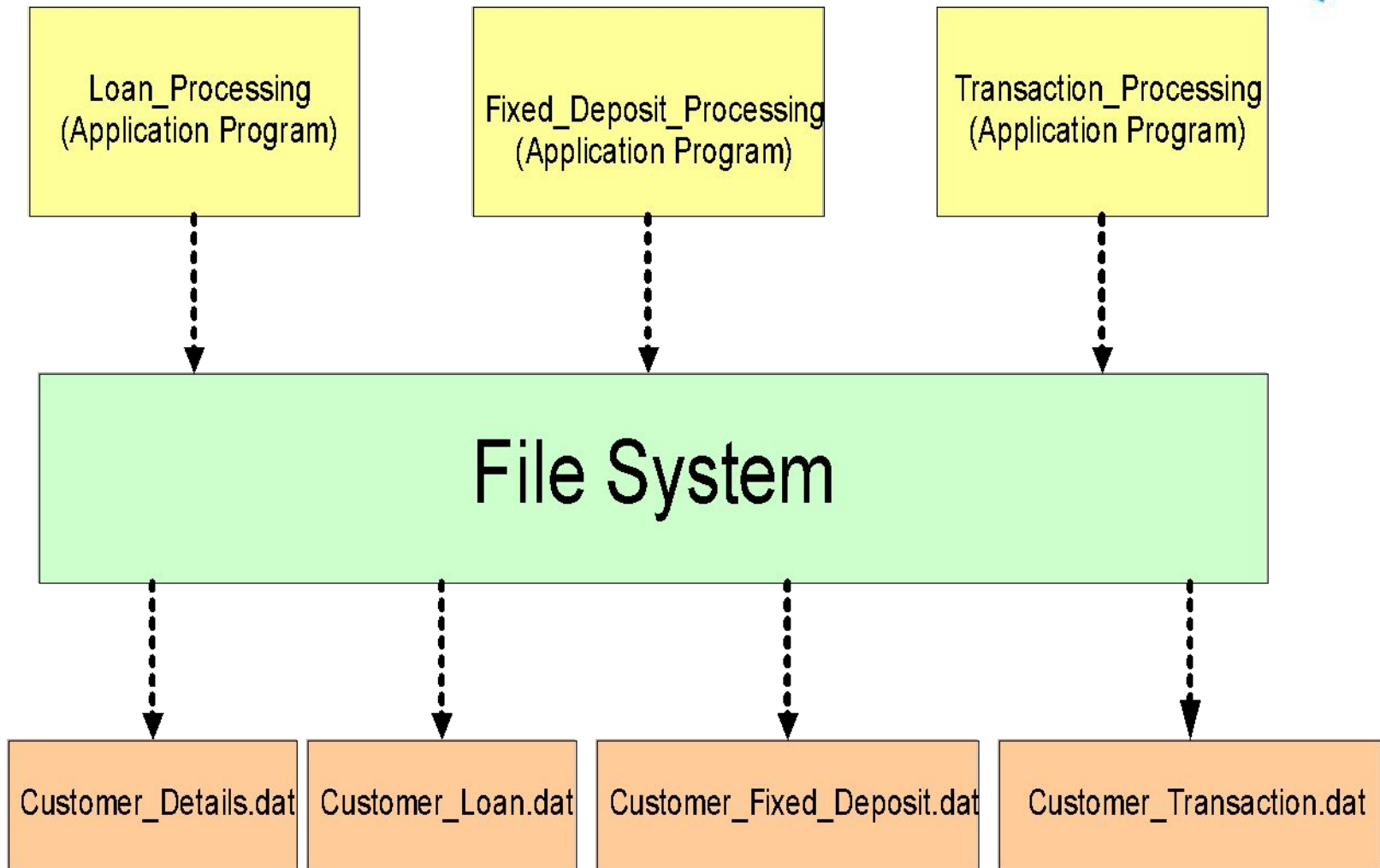
- Database and database systems are an essential part of life in a modern society.
- Most of us encounter several activities on day-to-day basis which involves database and database systems in one or another way.
 - **Example:** Online shopping, Online ticket booking, banking transactions, use of social media etc.
- The above examples are basically the traditional database applications.
- However, with the rapid technological advancements a range of exciting new database applications have come into existence like Geographical Information System, multimedia databases like youtube, hotstar, netflix, amazon prime, etc.

- Any raw or unprocessed fact can be referred to as “**Data**”. Data in its raw form does not hold any importance.
- When the data is processed in order to extract some meaning of any significance, it is referred to as “**Information**”.
- Data processing may involve several operations like data collection, recording, sorting, classifying, retrieving, calculating, summarizing, and communicating.
- Data Management focuses on generation, storage, and retrieval of data.

Traditional Data Storage

- File processing system was an early attempt to computerize the manual filing system that we are all familiar with.
- A file system is a method for storing and organizing computer files and the data they contain, to make it easy to find and access them.
- File systems may use a storage device such as a hard disk or CD-ROM and involve maintaining the physical location of the files.
- The file processing system works well when the number of items to be stored is small. It even works quite adequately when there are large numbers of items and we have only to store and retrieve them. However, the manual filing system breaks down when we must cross-reference or process the information in the files.

Traditional File System



Limitations of Traditional File System

- Data Security
- Data Redundancy
- Data Isolation
- Data Inconsistency
- Program / Data Dependence
- Lack of Flexibility
- Incompatible file formats
- Concurrent Access Anomalies

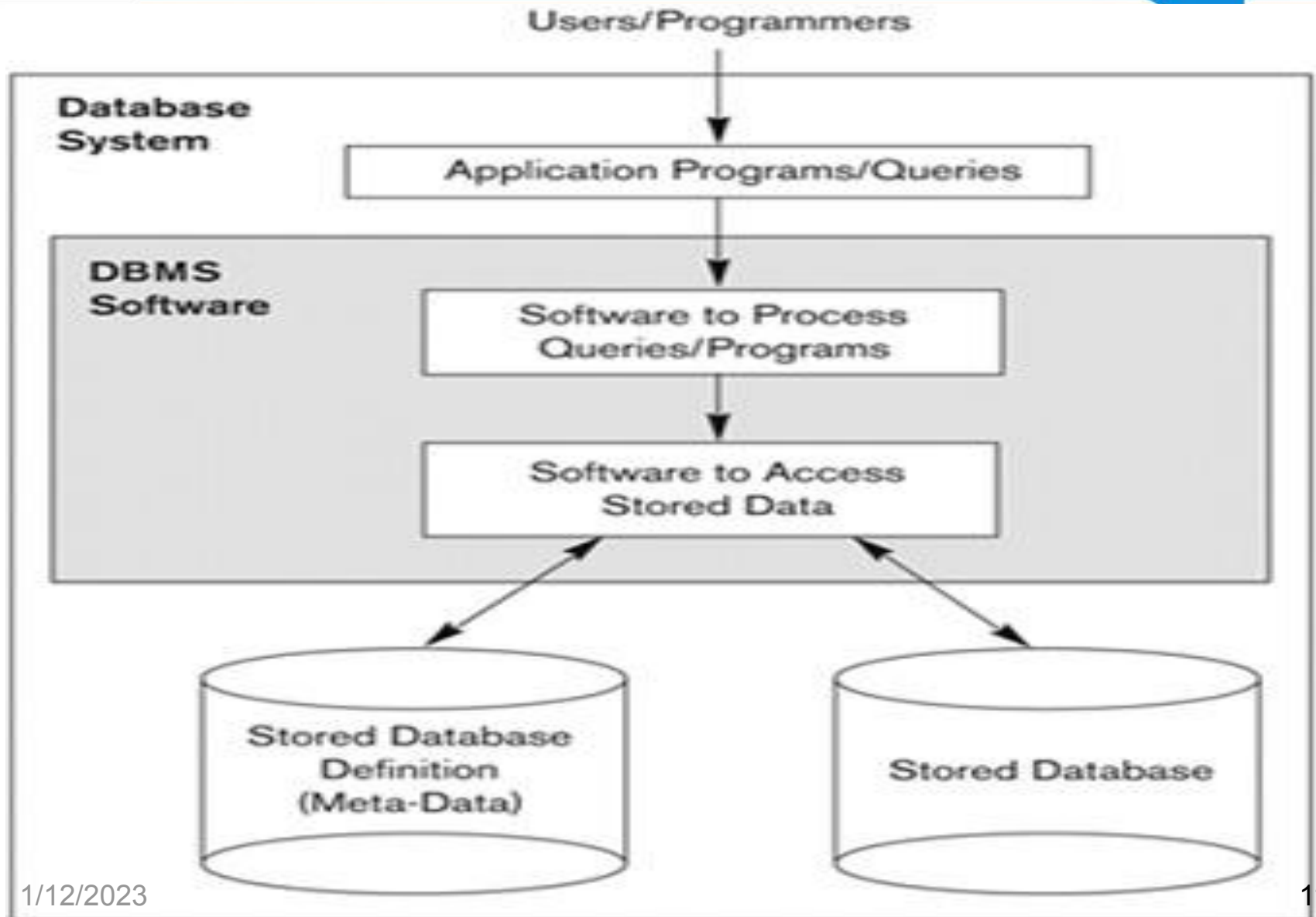
Database (DB) :

- Computer based organized collection of **interrelated data**.
- Records & maintains end users' data and meta data i.e., data through which the end user data are integrated and managed.
- A database is designed, built and populated with data for a specific purpose as it has an intended group of users.
- A database can be of any size and complexity based on the requirement. For example, a database for a university, database of e-commerce companies etc.
- A database may be generated and maintained manually, if the size is small, or the same may be computerized.

Database Management System (DBMS)

- A Database Management System, DBMS, is a collection of programs that enables user to create and maintain a database.
- It is an interface between end users and the DB.
- DBMS is a general-purpose software system that facilitates the process of **defining, constructing, manipulating, and sharing** DB among various users and applications.
- How it works
 - An application program accesses the DB by sending queries to the DBMS.
 - The query causes some data to be retrieved over which a transaction will be performed.
 - A transaction may cause some data to be read and some data to be written into the DB.

DB System Environment



An Example

- **UNIVERSITY** is a DB that maintains information concerned to students, courses, sections, and grades in a university environment.
- Now, this DB can be organized in following files that stores data records of same type.
 - **STUDENT** file stores data on each student,
 - **COURSE** file stores data on each course,
 - **SECTION** file stores data on each section,
 - **GRADE** file stores the grade that student receives.
- To define this DB, we must specify the structure of the records of each file by specifying the different types of data element to be stored in each record.
- Now, to create the UNIVERSITY database, we store data to represent each student, course, section, and grade as a record in the appropriate file.

Characteristics of DB Approach

- The main characteristics of the DB approach is as below:
 - **Self describing:**
 - A DB also contains a complete definition along with the database.
 - This definition is stored in the DBMS catalogue which contains information like structure of each file, its type, storage format of each data item, and various constraints.
 - This catalogue is known as “**Meta-data**”.

- **Insulation between program and data, and data abstraction:**
 - In traditional file system, since the structure of data file is embedded in the application program, so any changes to the structure of a file may require changing all programs that access that file.
 - By contrast, DBMS access programs do not require such changes because the structure of the data file is stored in the meta-data which is separate from the access programs. This is known as **program-data independence**.

- **Support of multiple views of data:**
 - A DB typically has many users, each of whom may require a different view of DB. A multi-user DBMS facilitates for defining multiple views for variety of users.
- **Sharing of data and multi-user transaction processing:**
 - A multi-user DBMS must allow multiple users to access the DB at the same time. This is essential if data for multiple applications is to be integrated and maintained in a single DB.
 - DBMS must include **concurrency control** software to ensure that several users trying to update the same data do so in a controlled manner so that the result of the updates is correct. Ex: Ticket booking system.

Types of Database

- Depending on the number of users accessing the DB, it may be classified as below:
 - **Single-user:**
 - Supports only one user at a time.
 - When used on a personal computer it is known as “Desktop” DB system.
 - **Multi-user:**
 - Supports multiple user at the same time.
 - When used by a relatively small group of users, it is known as “Workgroup” DB system.
 - When used by many users across globe, it is known as “Enterprise” DB system

- Depending on the location of the DB, it may be classified as below:
 - **Centralized:**
 - In this DB system, the data is located at one single location.
 - **Distributed:**
 - It supports data located at several different sites.
 - Here, the same DB is located at different servers at different locations so that even if the original server goes down; the data can be available to users from another server.

– **DB Administrators (DBA):**

- In a DB environment, the primary resource is the DB itself and the secondary resource is the DBMS and the related software.
- Administering these resources is the responsibility of the database administrator.
- The DBA is responsible for authorizing access to the DB, coordinating and monitoring its use and acquiring software and hardware as needed.
- DBA is also responsible for breach of security or poor system response time.
- In large organizations, DBA is assisted by a team of individuals in managing these responsibilities.

- **DB Designers:**

- The DB designers are responsible for identifying the data to be stored in the database and for choosing appropriate structures to represent and store the data.
- They interact with all perspective users of DB in order to understand their requirements so that they can create a design that meets these requirements.
- DB designers may also be assigned some administration related jobs after the design work is over.

- **End Users:** People who access the database for querying, updating and generating reports are known as end users. End users may be categorized as below:
 - **Casual end user:** Occasional user with different requirements every time. These include middle or high-level managers or occasional browsers.
 - **Naive end user:** Sizeable portion of DB end users. Their job includes constantly querying and updating the DB using standard types of queries and updates, known as canned transactions that have been carefully programmed and tested.

- **Sophisticated end user:** It include engineers, scientists, business analysts who thoroughly familiarize themselves with the facilities of DBMS in order to implement their application to meet their complex requirement.
- **Standalone users:** They maintain personal DB by using ready made packages that provide easy to use menu based or graphics-based interfaces.
- **Software Engineers:**
 - They determine the requirements of the end users, especially naive end users and develop specifications for canned transactions to meet these requirements.
 - Application programmers implement these specifications as program, then they test, debug, document and maintain these canned transactions.

- In addition to the actors, others are associated with the design, development, and operation of of DBMS software and system environment. They are known as workers behind the scene.
 - DBMS system developers and implementers
 - Tool developers
 - Operators and maintenance personnel.

Advantages of DBMS

- A DBMS have some significant advantages as mentioned below:
 - Controls redundancy
 - Restricts unauthorized access
 - Allows data sharing
 - provides storage structure and search techniques for efficient query processing
 - provides backup and recovery
 - supports multiple user interface
 - Availability of up-to-date information
 - Represents complex relationship among data.

Disadvantages of DBMS

- Increased cost and complexity
- Technical staff requirement
- Database failure
- Extra cost of hardware
- Large operational size
- Concurrency maintenance

When not to use DBMS

- Simple well defined database applications that are not expected to change very frequently.
- Real time requirements for some program may not meet because of DBMS overhead due to its high investment, for providing security, concurrency control, recovery functions etc.
- No need of multiple user interface.
- Embedded system with limited storage capacity and its not enough to get DBMS fit in.
- Small organizations