Note: Most of the contents of these notes have been extracted from the book "Database Systems A practical approach to Design, Implementation, and Management" by Thomas Connolly and Carolyn Begg sixth edition.

Introduction to Databases:

Introduction:

1. database is a collection of related data

2. database management system (DBMS) is a software that manages and controls access to the database.

3. database application is a program that interacts with the database at some point in its execution.

You interact with a database when:

1. purchasing goods from supermarket

2. purchasing goods using credit card

3. borrowing book from a library

4. using internet

Traditional File-Based Systems:

Why do we need to know about file-based systems?
1. Understanding their problems may prevent us from repeating these problems in database systems
2. Such knowledge is extremely useful when converting an existing file-based system to database system.

What is file-based system?

It is a collection of programs, where each program define and manages its own data.

When file-based system breaks down? It breaks down **when we have to cross reference large amounts of data** stored in hundreds of files.

Drawbacks/ Limitations of File-Based approach:

1. Separation and Isolation of data.

2. Duplication of data.

3. Program Data dependence: physical structure and storage of data files and record are defined in application code.

4. Incompatible file formats.

5. No facility for asking unplanned queries.

6. No provision for security or integrity.

7. No existence of recovery in case of hardware or software failure.

8. No provision for shared access to files.

To overcome above mentioned limitations a new approach has emerged, and is known as database and data base management system.

Formal definition of Database:

A shared collection of logically related data and its description, designed to meet the information needs of an organization.

Description of the data is known as system catalog, or **data dictionary** or meta data.

DBMS: A software system that enable users to:

- define

- create

- maintain, and

- control

access to the database.

DBMS provides the following languages for the tasks mentioned above:

1. **Data Definition Language(DDL):** allows user to define the database by specifying:

- data types
- structures
- constraints on data.

2. **Data Manipulation Language (DML):** allows user to insert, update, delete, and retrieve data. It also provides an inquiry facility to data, called

a **query language**. The most common query language is the **Structured Query Language(SQL)**.

In addition **DBMS also provides following features:**

- a security system to **prevent unauthorized users from accessing** database.

- an integrity system to maintain the **consistency** of stored data.

- a concurrency control system to **allow shared access** to database.

- a recovery control system to **restore the database to previous consistent state** in case of software or hardware failure.

- **a user accessible catalog** which contains descriptions of data in the database.

Database Application Programs

Users interact with the database through a number of application programs that are used to create and maintain the database and to generate information. These programs can be:

- conventional batch applications, or
- online applications

Application programs are usually written in **higher-level fourth generation** programming language.

Views

A DBMS provides another facility known as a **View**, which allows each user to have his/ her own view of the database. A view is some subset of the database. Views have several benefits:

1. Views provide a level of security: Views can be set up to exclude data that some users should not see.
2. Views provide a mechanism to customize the appearance of the database.
3. A view can present a consistent, unchanging picture of the structure of the database, even if the underlying database is changed.
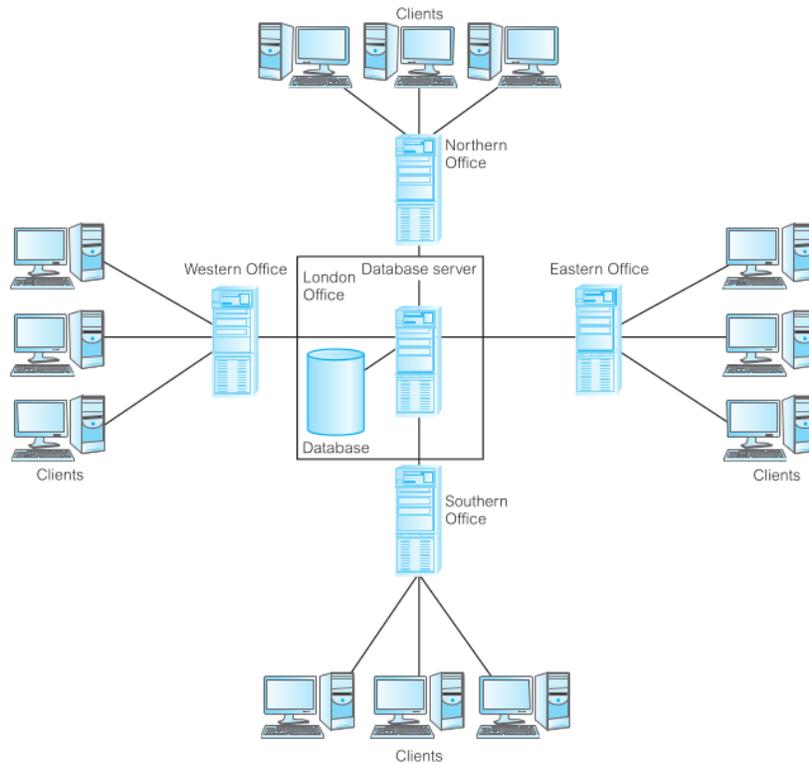
Above mentioned functionalities of DBMS differ from product to product.

Components of the DBMS Environment
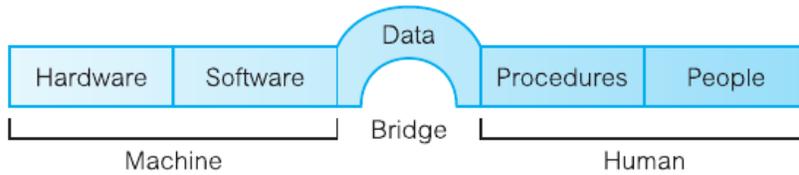
Hardware

The DBMS and the applications require hardware to run. The hardware can range from a **single personal computer** to a **single mainframe** or a **network of computers**.

A simplified hardware configuration for a DBMS is illustrated in figure above. It consists of a network of small servers, with a central server located in London running the **backend** of the DBMS, that is, the part of the DBMS that manages and controls access to the database. It also shows several computers at various locations running the **frontend** of the DBMS, that is, the part of the DBMS

4

that interfaces with the user. This is called a **client–server architecture**: the backend is the server and the frontends are the clients.

Software

The software consists of the following:

1. DBMS software itself.
2. Application programs.
3. Operating system.

Application programs are typically written in 3rd or 4rth generation languages. Some DBMS provide tools that allow rapid development of applications through:

- nonprocedural query languages
- reports generators
- forms generators
- graphics generators
- application generators

Data

It is the most important component of DBMS environment. The database contains both the **operational data**, and **meta data**. The structure of the database is called **schema**. Schema usually consists of more than one **table**. Each table consists of more than one **fields or attributes**.

Procedures

Procedure refer to the instructions and rules that govern the design and use of the database. These instructions describe how to:

- Log on to the DBMS
- Use a particular DBMS facility or application program.
- start and stop the DBMS.
- make backup copies of the database.
- handle hardware or software failures.
- Change the structure of a table, re-organize the database across multiple disks, improve performance, or archive data to secondary storage.

People / Roles in the Database Environment.

Four distinct type of people participate in the DBMS environment.

1. Data/ Database Administrator
2. Database designer
3. Application developer
4. End user

Data Administrator is responsible for the management of the data resource, including database planning; development and maintenance of standards, policies and procedures; and conceptual/logical database design.

The **Database Administrator (DBA)** is responsible for the physical realization of the database, including physical database design and implementation, security and integrity control, maintenance of the operational system, and ensuring satisfactory performance of the applications for users.

The role of the DBA is more technically oriented than the role of the DA, requiring detailed knowledge of the target DBMS and the system environment. **In some organizations there is no distinction between these two roles**.

Database Designer large database design projects have two types of database designers, logical database designer and physical database designer.

1. logical database desginer is concerned with:

    1.1. identifying the data,
    1.2. the relationships between the data, and
    1.3. the constraints on the data that is to be stored in the database.

2. physical database designer decides how the logical database design is to be physically realized.

Application Developer Once the database has been implemented, the application programs that provide the required functionality for the end-users must be implemented. This is the responsibility of the application developers.

End-User The end-users are the "clients" of the database.