→ Database System

Once upon a time, data was stored in journals, libraries & in hundreds of filing cabinets. Everything was recorded via paper. That means it took up space & also hard to find as well as difficult to back up. After that computers became available for better data management.

In the year 1960's **Charles Bachman** designed the first computerised database. It is called Integrated Data Store (IDS). It is devided into 2 models,

1. The Hierarchical model
2. The Network model.

- Hierarchical model was developed by IBM. In it, data is organised like a tree. Each data entry has a parent record, starting with a root record.
- The Network model was different from hierarchical model, in that it allowed a record to have more than one parent and child record.

In the year 1970's **E-F-Codd** (Edgar frank codd) introduced a Relational Database is one that shows the relationship between different data records. It is more space efficient; reduced data storage cost.

Def:-

A database is an organized collection of data, stored & accessed electronically from a computer system. Databases are used for storing, managing & retrieving information

→ Characteristics (Database Vs file System)

File system:

The file system is basically a way of arranging the files in a storage medium like a hard disk. file system organizes the files & helps in the retrieval of files when they are required. file system consists of different files which are grouped into directories. It stores small amount of data.

Database:

It is organized collection of data. intended for easily organizing, storing & retrieving large amount of data. It holds a bundle of organized data.

- There are mainly 15 characteristics of database

1. Real world Entity
2. Self-Describing Nature
3. Support ACID properties
4. concurrent use of Database
5. Insulation between Data & program
6. Transactions
7. Data persistence
8. Backup & Recovery
9. Data Integrity
10. multiple views
11. stores any kind of data
12. Security
13. Represents complex relationship between data
14. Query Language
15. cost

1. Real world Entity:

Real world Entities are used the DBMS architecture & also behaviour. that means if we can take an example of an organization, Banks (or) any field, where Employee is an Entity & his Employee id is an attribute. By DBMS we can store any data of Real world Existed entity.

2. Self-Describing Nature:

DBMS should be of self-describing nature as it not only contains the database itself but also the metadata.

metadata: Data About Data is called meta data.

A meta data defines & describes not only the type, structure,

and format of all data but also relationship between data. This data represents itself that what actions should be taken on it.

3. supports ACID properties:

A - Atomicity
C - consistency
I - Isolation
D - Durability

It means any data should not be lost while performing transactions like delete, insert & update. we can retrieve data fast

Ex:- if an Employee name is updated then it should be make sure that there is no duplicate data & no mismatch of employee information.

4. concurrent use of database / parallel access

There are many chances that many users will be accessing the data at the same time. They may require update the data concurrently. At that time, DBMS supports them to concurrent use database without any problem.

Ex: Railway Reservation System (or) cinema tickets

5. Insulation between data & program

In file management System, structure of data files was defined in the application programs, so user had to change all the programs that are using that particular data file.

But in DBMS, structure of data files is not stored in the program but it is stored in system catalogue. With the help of this feature, internal improvement of data efficiency & any changes in the data do not have any effect on application software.

6. Transactions:

Transactions are bunch of actions that are done to bring data from one state to another state. File system did not have this feature, it can only be completed & uncompleted.

For Example, A person wants to credit money from his account to another person's account. Then transaction will be complete if he sends money, other person receives his money. Anything other than this can lead to an inconsistent transaction.

7. Data Persistence:

Persistence means of the data is not removed Explicitly then all the data will be maintained in DBMS. If any system failure happens then life span of data stored in the DBMS will be decided by the users directly & indirectly. Any data stored in the DBMS can never be lost. If system failure happens in between any transaction then it will be rolled back & fully completed, but data will never be at risk.

8. Backup and Recovery:

There are many chances of failure of the whole database. At that time no one will be able to get the database back and company will be in loss. The only solution is to take backup of database & whenever it is needed, it can be stored back. A database have this characteristic.

9. Data Integrity :

This is one of the most important characteristics of data base management system. It protects unauthorized access to the database & make it more secure.

10. multiple views :

Users can have multiple views of database depending on their department & interest. DBMS supports multiple view of database to the users. For Example, a user of the teaching department will have different view & user of hostel department will have different. this feature helps users to have somewhat Security because users of other departments can not access their files.

11. Stores any kind of Data :

DBMS should be able to store any kind of data. It should not be restricted to employee name, salary & address. any kind of data that exists in the real world can be stored in DBMS. because we need to work with all kinds of data that is present around us.

12. Security :

DBMS provides Security to the data stored in it because all users have different rights to access database. Some of the users can access the whole database while other can access a small part of database.

For Example, A computer network lecture can only access files that are related to computer subject but HOD of the department can access files of all subjects that are related to their department.

13. Represents complex Relationship between Data

Data stored in a database is connected with Each other and a relationship is made in between data.

14. Query Language:

Queries are used to retrieve & manipulate data. Users have the power to retrieve any kind of data they want from the database by applying different sets of queries.

15. Cost:

The cost of the DBMS is high as compared to the other software. But if we consider the long run then DBMS is way far better because its maintenance cost will be less.

→ Database Users (Actors on scene, workers behind the scene)

The people whose jobs involve the day-to-day use of a large database. We call them the Actors on the screen. Scene.

1. Database Administrators (DBA)

In any organization where many people use the same resources, there is a need for a chief administrator to oversee & manage these resources. Administering the resources in the DBMS called Database Administrator (DBA). The DBA is responsible for authorizing access to the database, coordinating & monitoring its use, and acquiring software & hardware resources as needed.

2. Database Designers:

Database designers are responsible for identifying the data to be stored in the database & for choosing appropriate structure to represent and store this data. The responsibility of database designers is to communicate with all database users in order to understand their requirements and to create a design that meets these requirements. In many cases, the designers may be assigned other staff responsibilities after the database design is completed.

3. End users:

End users are the people whose jobs require access to the database for updating, querying & generating reports. There are several categories of end users.

Casual end users:

occasionally access the database, but they may need different information each time. Ex: Browsers

Naive (or parametric) end users:

They make up a sizable portion of database. Their main job function revolves around constantly querying i.e standard types of queries.

Ex: Reservation agents :- check availability for a given request & make reservations.

Sophisticated end users:

These includes engineers, scientists, business analysts & other who thoroughly familiarize themselves with the facilities of DBMS.

Standalone users:

They maintain personal databases by using ready-made programs that provide easy-to-use, menu based (or) graphic based interfaces. Ex: Tax package that stores variety of personal financial data for tax purpose.

4. System Analysts & Application programmers

System analyst determine the requirements of end users, especially naive end users. & develop specifications for standard transactions that meets these requirements

Application programmers implements these requirements as programs then they test, debug, document & maintain.

Such analysts & programmers referred to as Software developers (or) software Engineers.

Workers Behind the Scene:

The persons those who work to maintain the database system but who are not actively interested in the database contents of their daily job, called as workers behind the scene.

1. DBMS system designers & implementers:

design & implement the DBMS modules & interfaces as a software package. A DBMS is a very complex package/software that consists of many components (or) modules.

Ex: modules for implementing catalog, query processing, Accessing, buffering data, controlling, data recovery & Security.

2. Tool Developers:

Tool developers design & implement tools. It facilitates the data base modeling & design, database system design & improved performance. Some tools are purchased separately.

3. Operators & maintenance personnel:

Those are responsible for the actual running & maintenance of the hardware & software environment for the database system.

- These categories of workers behind the scene are making the database system available to end users, they do not use database content for their own purpose.

→ **Advantages of Database Systems**

There are 15 Advantages of database systems. Those are the characteristics of database System. they are

1. Real world Entity
2. Self describing Nature
3. Supports ACID properties
4. concurrent use of database
5. Insulation between data & program
6. Transactions
7. Data persistence
8. Backup & Recovery

9. Data Integrity
10. multiple views
11. stores any kind of data
12. Security
13. Represents complex relationship between data
14. Query Language
15. cost

The main advantage of database is controlling Redundancy i.e controlling duplicates.

1. Real world Entity: We can use database to any existed entity of world either a person, thing (or) any else.
2. Self describing nature: In database we can store data as well as metadata. i.e data about data.
3. Supports ACID properties: Any data access from database that accessing must be accuracy, completeness & durable long lasting.
4. Concurrent use of database: Database provide a feature of more than one user will access same data at a time.
5. Insulation between data & program: If any modification done in data they will not effect the program because they both are independent
6. Transactions: Transaction means a bunch of actions. In database Every transaction information is recorded. i.e Each State what will happened to data is informed to the user.
7. Data persistance: The data will deleted by the User Explicitly then only data deleted. Otherwise data will not delete at any cost, once the data stored in DBMS. If any problem occures then it moves one step back(or) forward.

but data will not delete.

8. Data Backup & Recovery: In Database if any fire Accident or System corrupted occurs then by using Backup we can recovery our valuable information back.

9. Data Integrity: Any unauthorized person can't be changed the data even a single letter.

10. multiple views: Different users have different needs. They access their data depending on their need. No two persons views are same.

11. stores any kind of data: Database is not restricted to stores the same data or particular data. we will store any data in database.

12. Security: Database provide security to our data, so unauthorized people can't access data. Only authorized persons have that right.

13. Represents Complex relationship between data: DBMS provides a relationship between all data by using a small links.

14. Query language :- Any thing access from data or any operation perform on data we use Queries only.

15. Cost :- DBMS Software cost is high but its maintainance is very less. & it runs long time

→ **Database Applications:**

There are different fields where a database management system is utilized.

1. Railway Reservation System:

In this system DBMS used in the update of status of trains appearances, ticket appointments. Additionally, if trains get late information will also update.

2. Library Management System:

There are lot of books in the library so, it is difficult to store the record of the relative books in a register. But by using DBMS we can keep up all the data identified with the name of the book, issue date, its writers etc.

3. Banking:

In Banking customer details, exchange date of client, loan issue dates all are stored by using database only

4. Education Sector:

In any colleges or schools, the student information, marks, results, courses, grades, expenses all are stored in database.

5. Credit card Exchanges:

In this DBMS used for buying on credit cards & age of month to month details are updated by DBMS.

6. Social media sites:

In present days we use different types of social medias like facebook, instagram, google, twitters like & so on. In those all information about many people & connect with many companies all information stored in database.

7. Broadcast communications:

In this field all postpaid bills, call recordings, month to month bills all are stored in database.

8. Account:

In accounts also we use database for storing stocks, prices, bills & so on.

9. Online shopping:

Now a days all are familiar with online shopping. In that all items are sold & stored in the database, Receipt charges, installments, buy data these all are stored using data base only. Ex: Amazon, flipcard, Meesho etc

10. Manufacturing:

Various kinds of items like bills, inventory, amount, furniture details stored in db.

11. Airline Reservation System:

In this the time of flight takeoff, appearances, vacancies, ticket booking details stored in database.

12. Human Resource Management:

Workers details, compensation of workers, their works in the big firms or organizations store information in database.

→ Brief introduction of different data models

Data model:

A datamodel defines the logical design & structure of database & defines how data will stored, accessed and updated in a database management system.

There are mainly 4 types of Datamodels are used

1. Hierarchical model
2. Network model
3. Entity-Relationship model
4. Relational model.

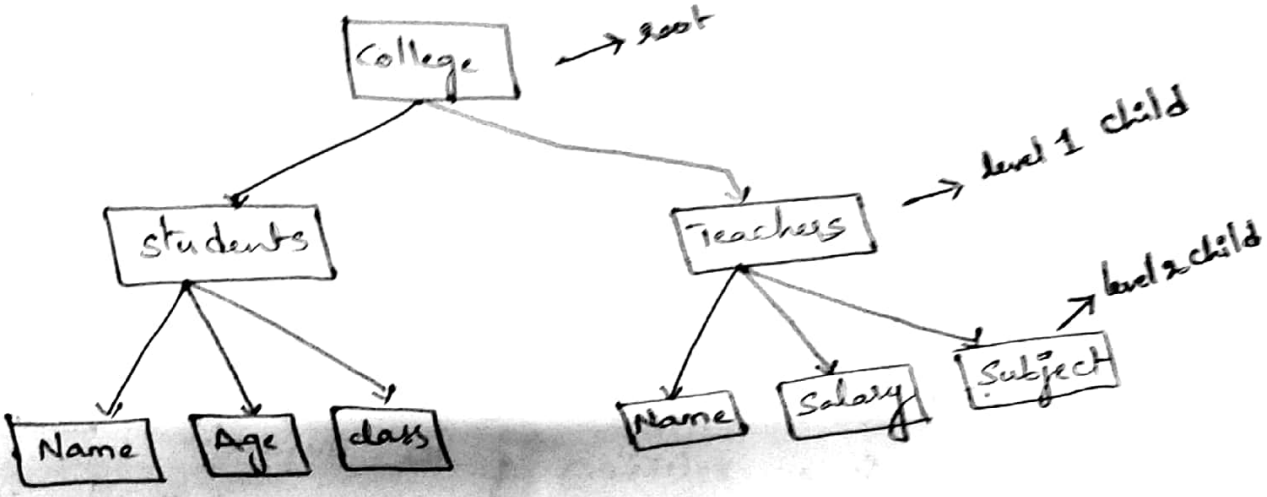
present days Relational model is most widely used data model.

1. Hierarchical Model:

It was the 1st DBMS model. This model organizes the data in the hierarchical tree structure. The hierarchy start from root also called parent node & then it expands in the form of a tree adding child node to the parent node. This model easily represents some of the real-world relationships

Ex: food recipes

Ex: Represent the relationship between college, student & teacher



Hierarchical model

Features:

1. One-to-many relationship: In this model we can represent a one-to-many relationship between data. There can be only one path from parent to any node.
2. parent-child Relationship: Each child node has a parent node but a parent node can have more than one child node multiple parents are not allowed.
3. Deletion problem: If a parent node is deleted then the child node is automatically deleted.
4. pointers: pointers are used to link the parent node with the child node & are used to navigate between the stored data.

Advantages:

- It is very simple & fast to traverse through a tree like structure.
- Any change in the parent node is automatically reflected in the child node also

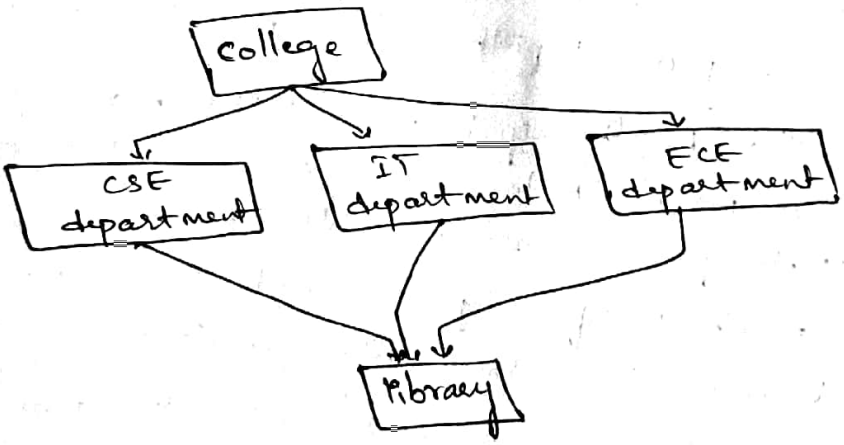
Disadvantages:

- complex relationships are not allowed / supported.
- If a parent node is deleted then the child node is automatically deleted.
- It does not support more than one parent node so if we have some complex relationships where a child node needs to have two parent node then that can't be represented using this model.

2. Network Model:

This model is an extension of the Hierarchical model. It was the most popular model before the relational model. This model is the same as the Hierarchical model, the only difference is that a child can have more than one parent. It replaces the Hierarchical tree with a graph.

Ex:



here library has more than one parent

Features:

1. Merge more relationships: this model has the ability to manage one-to-one relationships as well as many-to-many relationships.
2. Many paths: Here more than one path to the same node it makes data access fast & simple.
3. Circular linked list: the operations on the network model are done with the help of circular linked list.

Advantages :

→ The data can be accessed faster than hierarchical model. because data have more than one path, so data can be accessed in many ways.

→ Any change in parent node is reflected in the child node.

Disadvantages :

→ Here more & more relationships need to be handled so the system might get complex. So, a user must be having detailed knowledge of the model to work with the model.

→ Any change like updation, deletion, insertion is very complex.

3. Entity Relationship model:

It is also called ER model. In this model, we represent the real-world problem in the pictorial form to make it easy for the users to understand. It is also very easy for the developers to understand the system by just looking the ER diagram. ER diagram has the following three components.

1. **Entities:** Entity is a real world thing. It can be a person, place (or) thing.

   Ex:- Teachers, course, Department, Building, Table

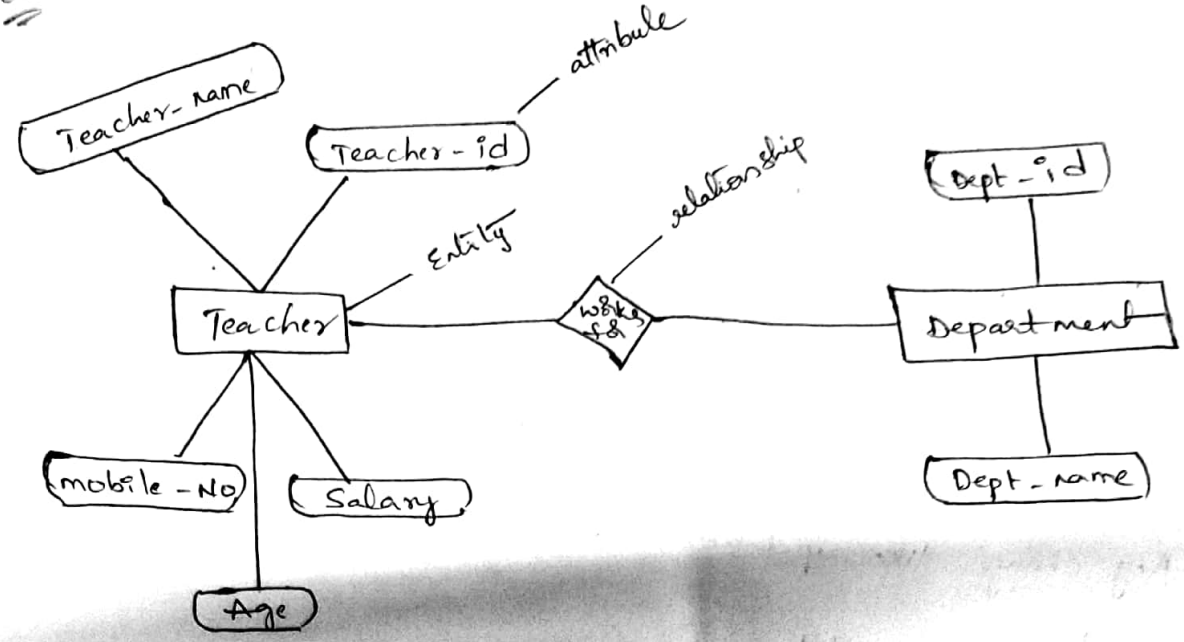
2. **Attributes:** An Entity contains a real-world property called attribute. This is the characteristics of that Entity

   Ex:- Teacher — Entity
   id, Salary, age — Attributes.

3. **Relationship:** Relationship tells how two attributes are related.

   Ex:- Teacher works for a department.

Ex:-



Features :

Graphical Representation for better understanding : It is very Easy & Simple to understand.

ER Diagram : ER diagram is used as a visual tool for representing the model.

Database Design : This model helps the database designers to build the database & is widely used in database design

Advantages :

Simple : ER model is very Easy to build. If we know the relationship between the attributes & the Entities we can Easily build the ER diagram for the model.

Effective Communication Tool : This model is used widely by the database designers for communicating their ideas.

Easy Conversion to any model : This model can converted to any other model Easily like network model hierarchical model etc.

Disadvantages

No Industry standard for Notation: one developes use notation which are not understood by other developes.

Hidden information: Some information might be lost (or) hidden in the ER model.

4. Relational Model:

It is the mostly widely used model. In this model, the data is maintained in the form of a Two-Dimensional table. All the information is stored in the form of row & columns. The basic structure of a Relational model is Tables. so the Tables are called Relations in the Relational model.

Ex:-

Emp_id	Emp_name	Salary	mobile_NO
1	Ramesh	10,000	9111037890
2	Suresh	90,000	9587569124
3	Mahesh	10,00000	789512355

EMPLOYEE TABLE

Features:

Tuples: Each row in the table called "Tuple". A row contains all the information about any instance of the object. In the above table 1st row represents Ramesh details.

Attribute (or) Field: Attributes are the properties which defines the table & relation. The values of the attribute should be from the same domain. In the above Example we have diffrent attributes like salary, mobile_NO, Emp_name etc.

Advantages:

1. Simple: This model is more simple as compared to network & hierarchical model.
2. Scalable: This model can be Easily Scaled as we can add as many rows & columns we want
3. Structural independent: We can make changes in database structure without changing the way to access the data.

→ Concepts of Schema

- A database schema is the logical representation of a database which shows how the data is stored logically in the entire database. It contains list of attributes and instructions that informs the database Engine that how the data is organized & how the elements are related to Each other

- A database schema contains schema objects that may include tables, fields, views, relationships, packages, primary key, foreign key.

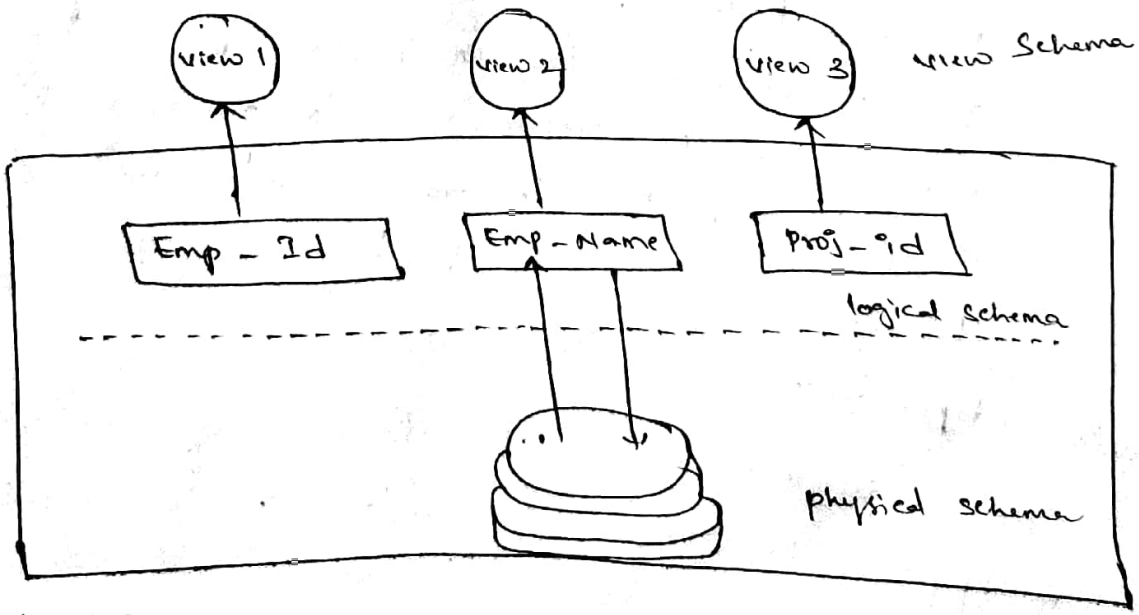
- The data is physically stored in files, but to retrieve it & use it, we need to put it in a structured form. To do this schema is used.

- the schema does not physically contain the data itself it gives information about the shape of data & how it can be related to other tables & modules.

Types of Database Schema

Schema is divided into Three types:

1. Logical Schema
2. physical Schema
3. View Schema.



physical Schema:

It specifies how the data is stored physically on a disk in the form of files. designing a database at the physical level is called a "physical schema"

Logical Schema:

It specifies all the logical constraints that need to be applied to the stored data. It represents how the data is stored in the form of tables & how the attributes of a table are linked together.

Various tools are used to create this schema, these tools defines the relationship between components of our data this process is called **ER modeling.**

View Schema:

The view level design of a database is called as view Schema It describes the end-user interaction with the database systems.

→ **Instance and data Independence**

Instance:

The data stored in database at a particular moment of time is called instance of database.

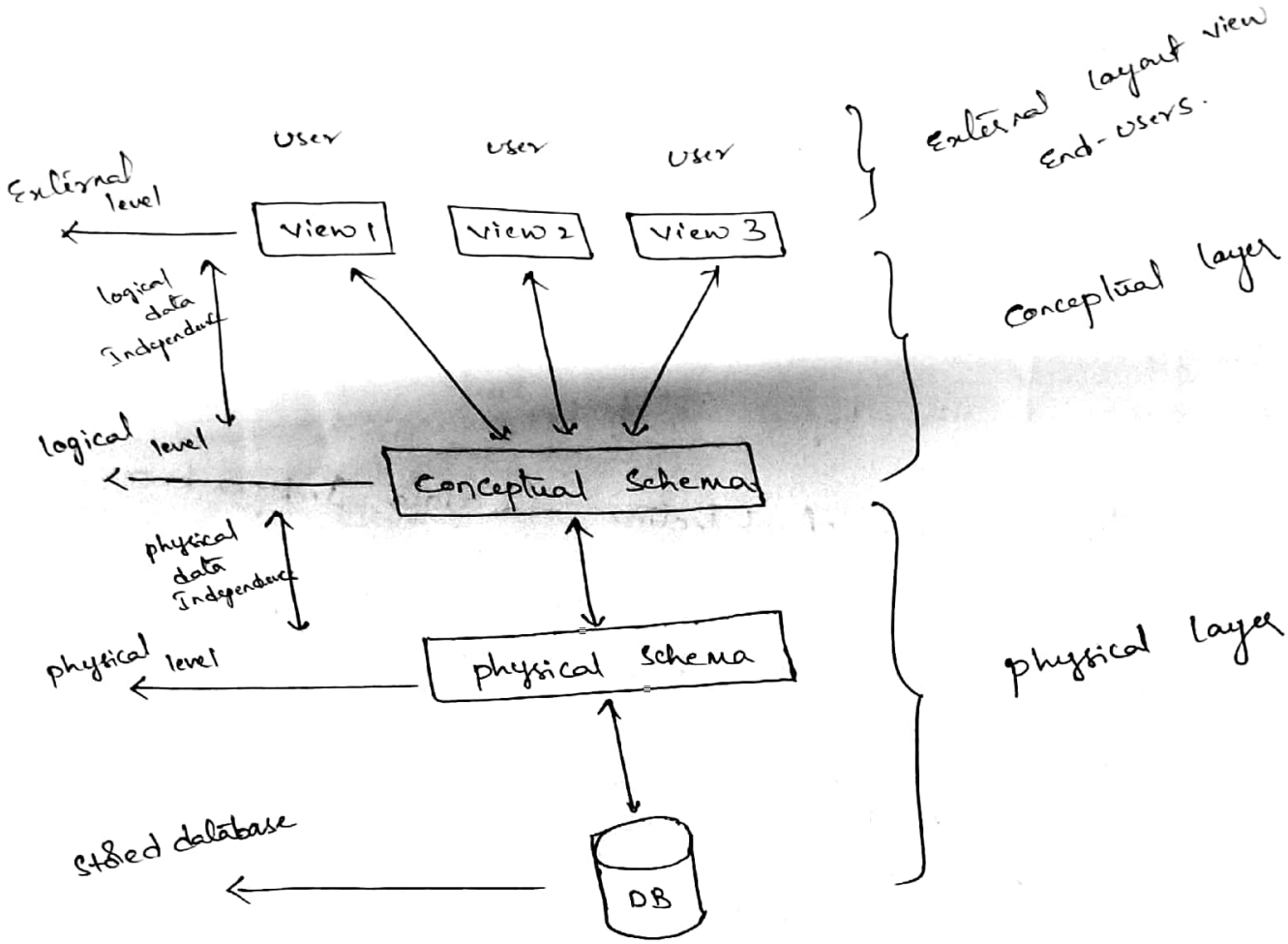
Ex:-

lets say we have table called student in the database present table has 100 records, so today the instance of the database has 100 records. we are going to add another 100 records by tomorrow will have 200 records in table tomorrow. so the instance of database tomorrow will have 200 records.

Data Independence:

Data independence refers characteristic of being able to modify the schema at one level of the database system without altering the schema at the next higher level.

- it can be Explained using the three-schema Architecture



Three Schema Architecture levels

There are two types of data Independence.

1. Logical data Independence
2. physical data Independence

Logical Data Independence:

- It refers characteristic of being able to change the conceptual schema without having to change the External schema.
- It is used to separate the External level from the Conceptual view
- If we do any changes in the conceptual view of the data then the user view of the data would not be affected
- It occurs at the User Interface level.

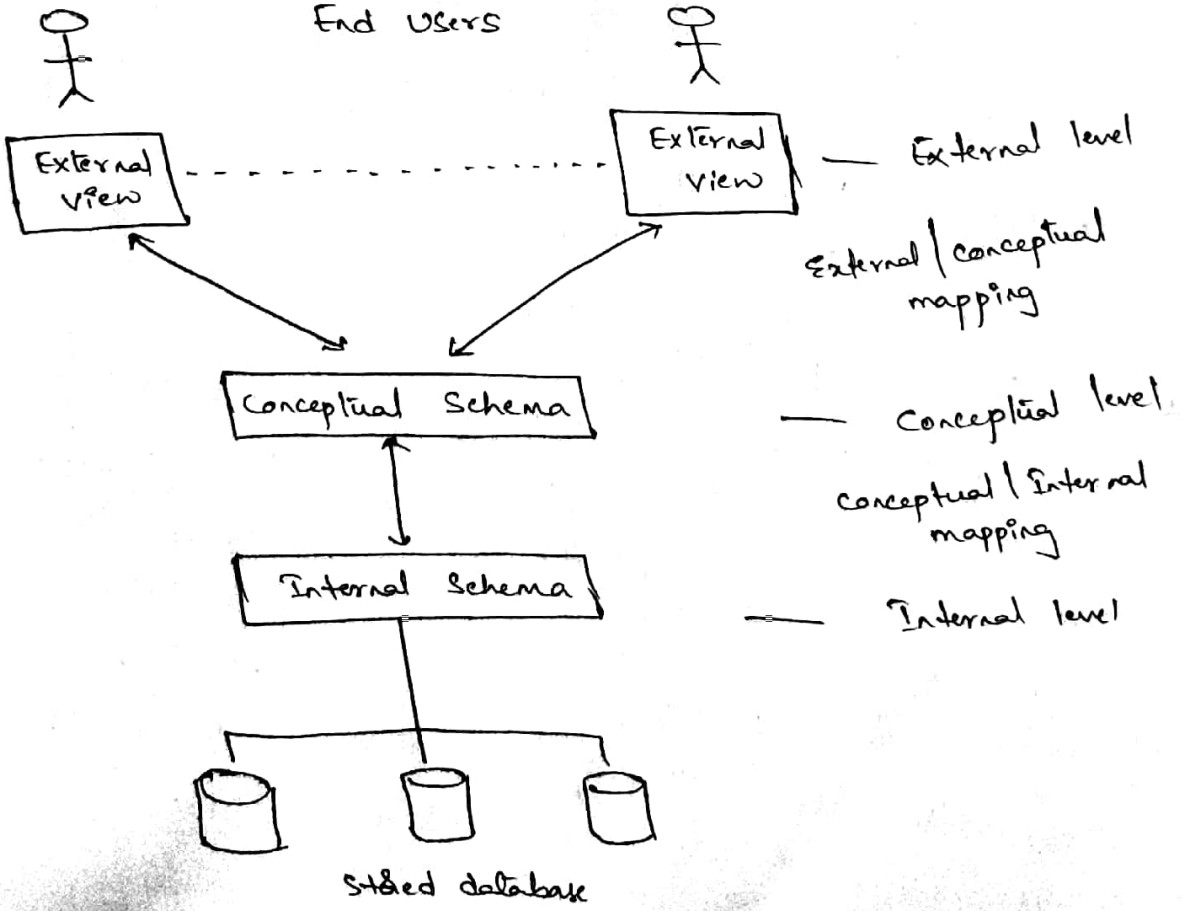
Physical Data Independence:

- It is defined as the capacity to change the internal schema without having to change the conceptual schema.
- If we do any changes in the storage size of database system server, then the conceptual structure of the database will not be affected.
- It is used to separate conceptual level from the internal level.
- It occurs at the logical Interface level.

→ Three Tier Schema Architecture for data Independence

Three tier Schema architecture goal is to separate the user applications from the physical database. In this schema three levels are defined.

1. Internal level / physical
2. Conceptual level / logical
3. External (or view level



• Internal level:

It has an internal schema, which describes the physical storage structure of the database. It describes complete details of data storage & access paths for the database.

Conceptual level:

It has an Conceptual Schema, which describes the structure of the whole database for a community of users. It hides the details of physical storage structures & concentrates on describing Entities, data types, relationships, user operations.

External or view level

It includes a number of External schemas (or) user views. Each External Schema describes the part of the database that a particular user group is interested in & hides the rest of the database from that user group.

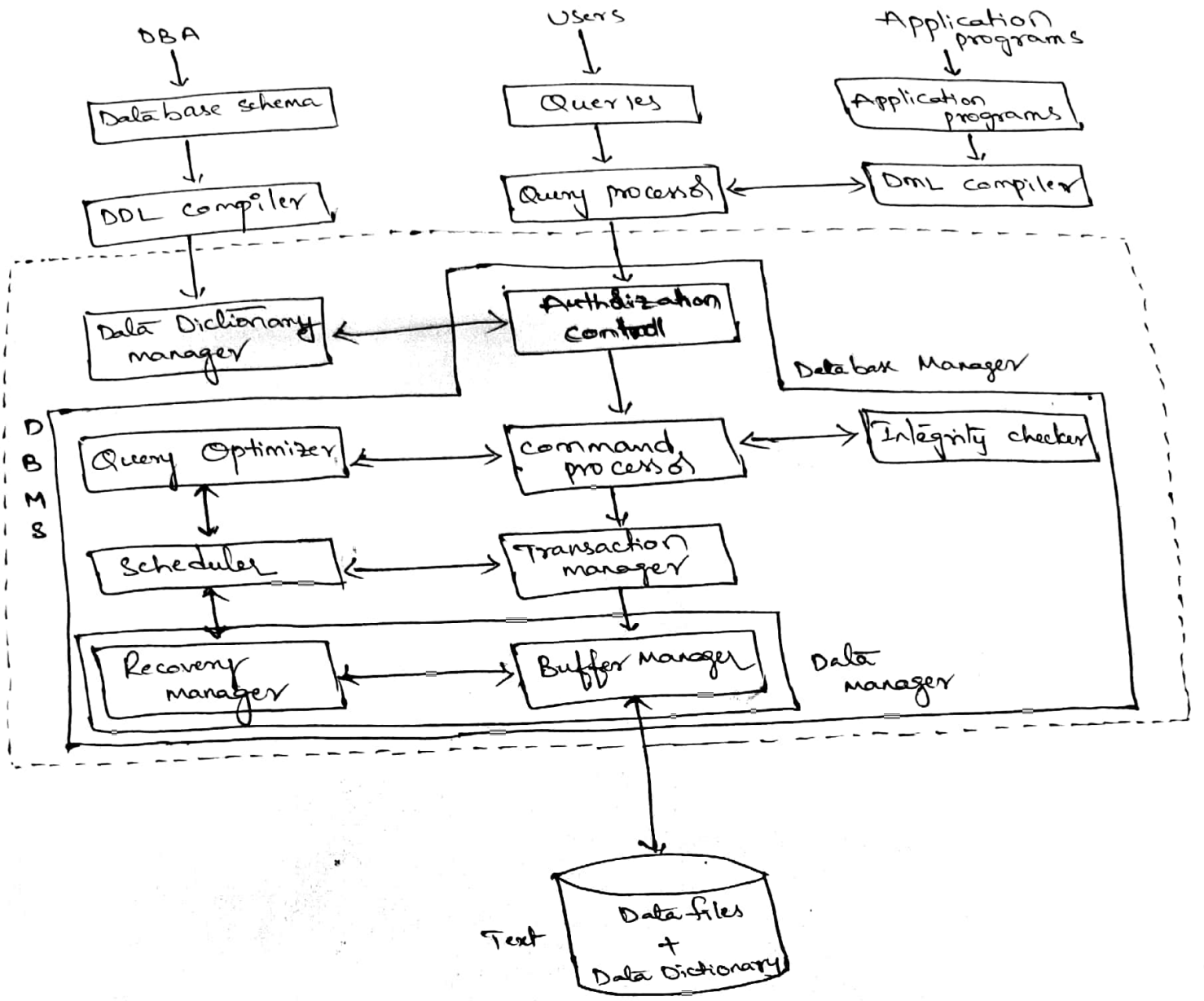
— The Three-schema architecture is a convenient tool with which the user can visualize the schema levels in a database system. Most DBMS's do not separate the three levels completely and explicitly, but support the three-schema architecture to some extent. Some older DBMS's may include physical-level details in the conceptual schema.

Three schemas are only descriptions of data; the stored data that actually exists is at the physical level only. Each user group refers to its own External schema. DBMS must transfer a request against the Conceptual schema & then into a internal schema for processing over the stored database. The process of transforming requests and results between levels are called mappings.

→ Database System Structure

The database system is divided into three components

1. Query processor
2. Storage manager
3. Disk storage



1. Query processor:

It interprets the requests (queries) received from the end user via an application program into instructions. It is also executes the user requests which is received from the DML compiler.

Query processor contains the following components

1. DML Compiler:- It processes the DML statements into machine level instructions, so that they can be executed
2. DDL Interpreter:- It processes DDL statements into a set of table containing meta data.
3. Embedded DML pre-compiler:- It processes DML statements embedded in an application program into procedural calls
4. Query Optimizer:- It executes the instruction generated by DML compiler.

2. Storage Manager:
It is a program that provides an interface between the database & the queries received. It is also called database control system. It executes DCL statements. It is responsible for updating, storing, deleting & retrieving data in the database.
- It contains the following components
1. Authorization manager:- It checks whether the particular person is privileged to perform the requested operation or not
2. Integrity manager:- It checks the integrity constraints when the database is modified
3. Transaction manager:- It ensures that the database remains in the consistent state before & after execution of a transaction.
4. File manager:- It manages the file space & the data structure used to represent information in the database.
5. Buffer manager:- It is responsible for cache memory & the transfer of data between the secondary storage & main memory

3. Disk storage:-
It contains the following components
1. Data Files:- It stores data
2. Data Dictionary:- It contains the information about the structure of any database object.
3. Indices:- It provides faster retrieval of data item.