Sub code:- 313302

#### Unit 1- database management system

What is Database



The database is a collection of inter-related data which is used to retrieve, insert and delete the data efficiently. It is also used to organize the data in the form of a table, schema, views, and reports, etc.

For example: The college Database organizes the data about the admin, staff, students and faculty etc.

Using the database, you can easily retrieve, insert, and delete the information.

### Database Management System

- o Database management system is a software which is used to manage the database. For example: <u>MySQL</u>, <u>Oracle</u>, etc are a very popular commercial database which is used in different applications.
- o DBMS provides an interface to perform various operations like database creation, storing data in it, updating data, creating a table in the database and a lot more.
- o It provides protection and security to the database. In the case of multiple users, it also maintains data consistency.

#### DBMS allows users the following tasks:

- o **Data Definition:** It is used for creation, modification, and removal of definition that defines the organization of data in the database.
- o **Data Updation:** It is used for the insertion, modification, and deletion of the actual data in the database.
- o **Data Retrieval:** It is used to retrieve the data from the database which can be used by applications for various purposes.
- o **User Administration:** It is used for registering and monitoring users, maintain data integrity, enforcing data security, dealing with concurrency control,

monitoring performance and recovering information corrupted by unexpected failure

### Characteristics of DBMS

- o It uses a digital repository established on a server to store and manage the information.
- o It can provide a clear and logical view of the process that manipulates data.
- o DBMS contains automatic backup and recovery procedures.
- o It contains ACID properties which maintain data in a healthy state in case of failure.
- o It can reduce the complex relationship between data.
- o It is used to support manipulation and processing of data.
- o It is used to provide security of data.
- o It can view the database from different viewpoints according to the requirements of the user.

# Advantages of DBMS

- o **Controls database redundancy:** It can control data redundancy because it stores all the data in one single database file and that recorded data is placed in the database.
- o **Data sharing:** In DBMS, the authorized users of an organization can share the data among multiple users.
- o **Easily Maintenance:** It can be easily maintainable due to the centralized nature of the database system.
- o **Reduce time:** It reduces development time and maintenance need.
- Backup: It provides backup and recovery subsystems which create automatic backup of data from <u>hardware</u> and <u>software</u> failures and restores the data if required.
- o **multiple user interface:** It provides different types of user interfaces like graphical user interfaces, application program interfaces

# Disadvantages of DBMS

- o **Cost of Hardware and Software:** It requires a high speed of data processor and large memory size to run DBMS software.
- o Size: It occupies a large space of disks and large memory to run them efficiently.
- o **Complexity:** Database system creates additional complexity and requirements.

o **Higher impact of failure:** Failure is highly impacted the database because in most of the organization, all the data stored in a single database and if the database is damaged due to electric failure or database corruption then the data may be lost forever.

#### Advantages of DBMS file processing system:-

- 1. Redundancy can be reduced
- 2. Inconsistency can be avoided
- 3. Data can be shared
- 4. Centralized control data
- 5. Standards can be enforced
- 6. Security restrictions can be applied
- 7. Integrity can be maintained
- 8. Data independence can be provided
- 9. New applications

#### **Applications of DBMS:-**

- 1. Cost of operation
- 2. System complexity
- 3. Requirement of technical staff
- 4. Database failure
- 5. Extra hardware
- 6. Size of data
- 7. Data conversion
- 8. Maintenance cost

### What is Data?

Data is a collection of a distinct small unit of information. It can be used in a variety of forms like text, numbers, media, bytes, etc. it can be stored in pieces of paper or electronic memory, etc.

Word 'Data' is originated from the word 'datum' that means 'single piece of information.' It is plural of the word datum.

In computing, Data is information that can be translated into a form for efficient movement and processing. Data is interchangeable

### What is Database?

A database is an organized collection of data, so that it can be easily accessed and managed.

You can organize data into tables, rows, columns, and index it to make it easier to find relevant information.

**Database handlers** create a database in such a way that only one set of software program provides access of data to all the users.

The **main purpose** of the database is to operate a large amount of information by storing, retrieving, and managing data.

There are many dynamic websites on the World Wide Web nowadays which are handled through databases. For example, a model that checks the availability of rooms in a hotel. It is an example of a dynamic website that uses a database.

There are many databases available like MySQL, Sybase, Oracle, MongoDB, Informix, PostgreSQL, SQL Server, etc.

Modern databases are managed by the database management system (DBMS).

**SQL** or Structured Query Language is used to operate on the data stored in a database. SQL depends on relational algebra and tuple relational calculus.

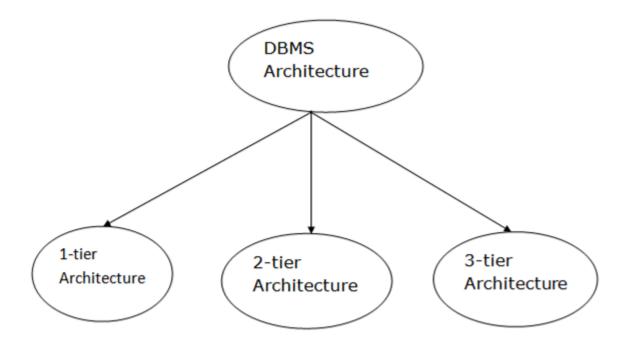
A cylindrical structure is used to display the image of a database.



### **DBMS Architecture**

- o The DBMS design depends upon its architecture. The basic client/server architecture is used to deal with a large number of PCs, web servers, database servers and other components that are connected with networks.
- o The client/server architecture consists of many PCs and a workstation which are connected via the network.
- o DBMS architecture depends upon how users are connected to the database to get their request done.

# Types of DBMS Architecture



Database architecture can be seen as a single tier or multi-tier. But logically, database architecture is of two types like: **2-tier architecture** and **3-tier architecture**.

### 1-Tier Architecture

- o In this architecture, the database is directly available to the user. It means the user can directly sit on the DBMS and uses it.
- o Any changes done here will directly be done on the database itself. It doesn't provide a handy tool for end users.
- o The 1-Tier architecture is used for development of the local application, where programmers can directly communicate with the database for the quick response.

### 2-Tier Architecture

- o The 2-Tier architecture is same as basic client-server. In the two-tier architecture, applications on the client end can directly communicate with the database at the server side. For this interaction, API's like: **ODBC**, **JDBC** are used.
- o The user interfaces and application programs are run on the client-side.
- o The server side is responsible to provide the functionalities like: query processing and transaction management.
- o To communicate with the DBMS, client-side application establishes a connection with the server side.

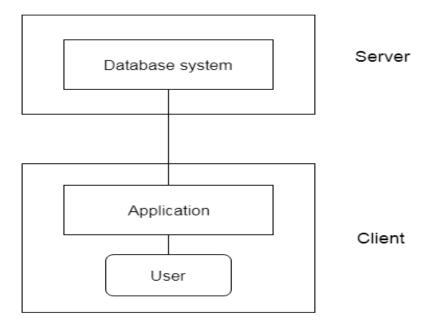


Fig: 2-tier Architecture

### **3-Tier Architecture**

- o The 3-Tier architecture contains another layer between the client and server. In this architecture, client can't directly communicate with the server.
- o The application on the client-end interacts with an application server which further communicates with the database system.
- o End user has no idea about the existence of the database beyond the application server. The database also has no idea about any other user beyond the application.
- o The 3-Tier architecture is used in case of large web application.

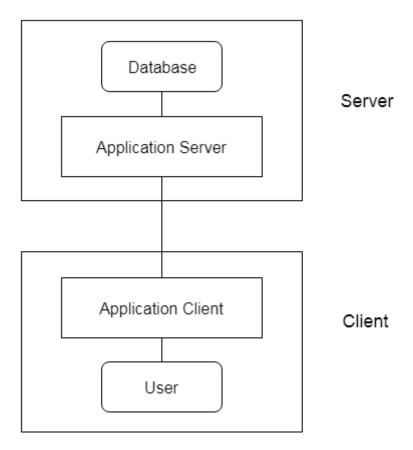
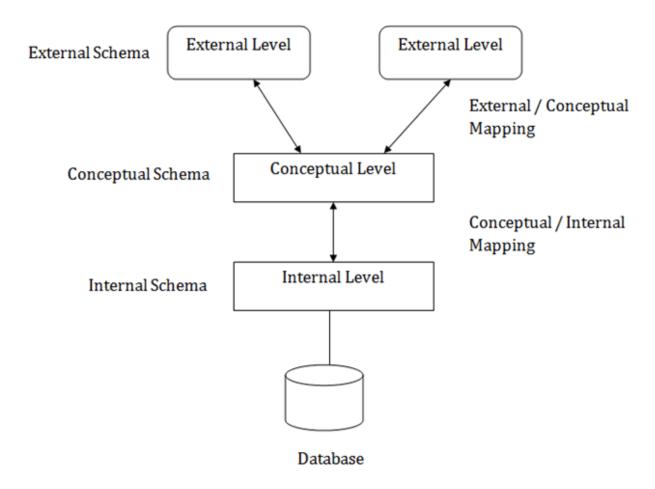


Fig: 3-tier Architecture

# **Three schema Architecture**

- o The three schema architecture is also called ANSI/SPARC architecture or three-level architecture.
- o This framework is used to describe the structure of a specific database system.
- o The three schema architecture is also used to separate the user applications and physical database.
- o The three schema architecture contains three-levels. It breaks the database down into three different categories.

The three-schema architecture is as follows:



#### In the above diagram:

- o It shows the DBMS architecture.
- o Mapping is used to transform the request and response between various database levels of architecture.
- o Mapping is not good for small DBMS because it takes more time.
- o In External / Conceptual mapping, it is necessary to transform the request from external level to conceptual schema.
- o In Conceptual / Internal mapping, DBMS transform the request from the conceptual to internal level.

# Objectives of Three schema Architecture

The main objective of three level architecture is to enable multiple users to access the same data with a personalized view while storing the underlying data only once. Thus it separates the user's view from the physical structure of the database. This separation is desirable for the following reasons:

- o Different users need different views of the same data.
- o The approach in which a particular user needs to see the data may change over time.
- o The users of the database should not worry about the physical implementation and internal workings of the database such as data compression and encryption techniques, hashing, optimization of the internal structures etc.
- o All users should be able to access the same data according to their requirements.
- o DBA should be able to change the conceptual structure of the database without affecting the user's
- o Internal structure of the database should be unaffected by changes to physical aspects of the storage.

### 1. Internal Level

STORED\_EMPLOYEE record length 60

Empno : 4 decimal offset 0 unique

Ename: String length 15 offset 4
Salary: 8,2 decimal offset 19

Deptno : 4 decimal offset 27

Post : string length 15 offset 31

Internal view

- o The internal level has an internal schema which describes the physical storage structure of the database.
- o The internal schema is also known as a physical schema.
- o It uses the physical data model. It is used to define that how the data will be stored in a block.
- o The physical level is used to describe complex low-level data structures in detail.

The internal level is generally is concerned with the following activities:

o Storage space allocations.

For Example: B-Trees, Hashing etc.

o Access paths.

**For Example:** Specification of primary and secondary keys, indexes, pointers and sequencing.

o Data compression and encryption techniques.

- o Optimization of internal structures.
- o Representation of stored fields.

### 2. Conceptual Level

Global view

Empno : Integer(4) Key
Ename : String(15)
Salary : String (8)
Deptno : Integer(4)
Post : String (15)

- o The conceptual schema describes the design of a database at the conceptual level. Conceptual level is also known as logical level.
- o The conceptual schema describes the structure of the whole database.
- o The conceptual level describes what data are to be stored in the database and also describes what relationship exists among those data.
- o In the conceptual level, internal details such as an implementation of the data structure are hidden.
- o Programmers and database administrators work at this level.

### 3. External Level



- o At the external level, a database contains several schemas that sometimes called as subschema. The subschema is used to describe the different view of the database.
- o An external schema is also known as view schema.
- o Each view schema describes the database part that a particular user group is interested and hides the remaining database from that user group.
- o The view schema describes the end user interaction with database systems.

# **Data Abstraction in DBMS**

Let us understand the concept of data abstraction.

Data abstraction is the procedure of concealing irrelevant or unwanted data from the end user.

Let us understand this with an example, if you go to a shop to buy a pair of shoes, you ask the shopkeeper to show you the shoes of a certain company, and you also tell the shopkeeper about the size, color, and material you want. Then, you will only see the specified things in the shoes, or will you be asking the shopkeeper questions such as, where are these shoes made? From where does the material come? What is the cost of the material?

The answer to these questions is **NO**. You will not ask these questions because these questions are of no use. You do not care about these questions. You are only concerned about a few things, such as the company, size, color, material, and how the shoes look. That is why these unimportant details are kept hidden from the end user. This is the process we call data abstraction

# What is Data abstraction in Database Management System?

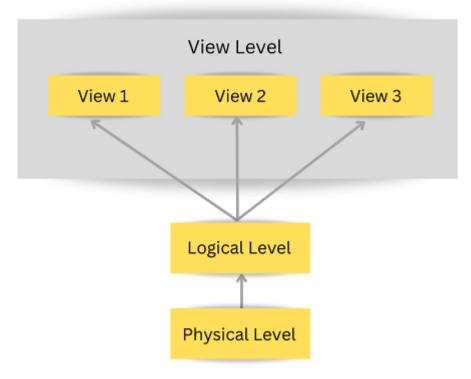
The database system contains intricate data structures and relations. The developers keep away the complex data from the user and remove the complications so that the user can comfortably access data in the database and can only access the data they want, which is done with the help of data abstraction.

The main purpose of data abstraction is to hide irrelevant data and provide an abstract view of the data. With the help of data abstraction, developers hide irrelevant data from the user and provide them the relevant data. By doing this, users can access the data without any hassle, and the system will also work efficiently.

In DBMS, data abstraction is performed in layers which means there are levels of data abstraction in DBMS that we will further study in this article. Based on these levels, the database management system is designed.

### Levels of Data Abstractions in DBMS

In DBMS, there are three levels of data abstraction, which are as follows:



Levels of Data Abstraction in DBMS

### 1. Physical or Internal Level:

The physical or internal layer is the lowest level of data abstraction in the database management system. It is the layer that defines how data is actually stored in the database. It defines methods to access the data in the database. It defines complex data structures in detail, so it is very complex to understand, which is why it is kept hidden from the end user.

Data Administrators (DBA) decide how to arrange data and where to store data. The Data Administrator (DBA) is the person whose role is to manage the data in the database at the physical or internal level. There is a data center that securely stores the raw data in detail on hard drives at this level.

### 2. Logical or Conceptual Level:

The logical or conceptual level is the intermediate or next level of data abstraction. It explains what data is going to be stored in the database and what the relationship is between them.

It describes the structure of the entire data in the form of tables. The logical level or conceptual level is less complex than the physical level. With the help of the logical level, Data Administrators (DBA) abstract data from raw data present at the physical level.

### 3. View or External Level:

View or External Level is the highest level of data abstraction. There are different views at this level that define the parts of the overall data of the database. This level is for the end-user interaction; at this level, end users can access the data based on their queries.

### What is DBMS Schema?

Here the DBMS schema means designing the database. For example, if we take the example of the employee table. The employee table contains the following attributes. These attributes are EMP\_ID, EMP\_ADDRESS, EMP\_NAME, EMP\_CONTACT. These are the schema of the employee table.

Schema is further divided into three types. These three are as follows.

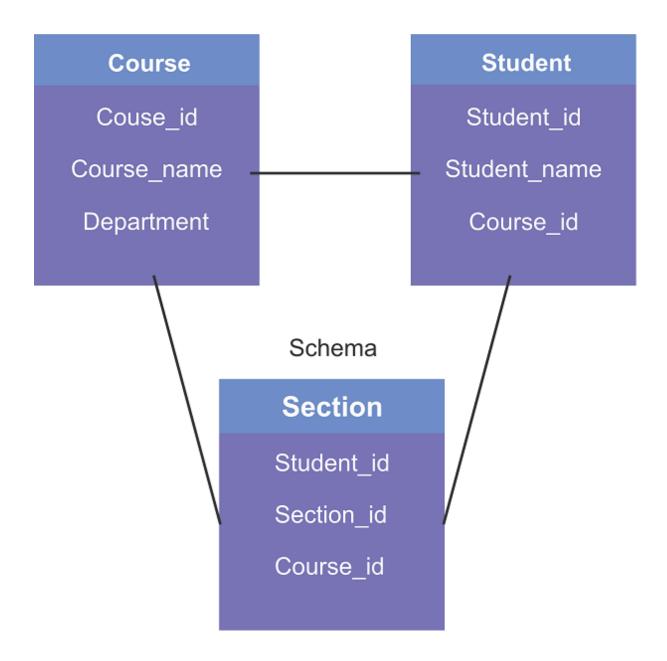
- 1. Logical schema.
- 2. View schema.
- 3. Physical schema.

The schema defines the logical view of the database. It provides some knowledge about the database and what data needs to go where.

In DBMS, the schema is shown in diagram format.

We can understand the relationship between the data present in the database. With the help of this schema, we can implement the DBMS function such as delete, insert, search, update, etc.

Let us understand this by the below diagram. There are three diagrams, i.e., section, course, and student. This diagram shows the relationship between the section and the course diagram. Schema is the only type of structural view of the database that is shown below.



### 1. Physical schema:

In the physical schema, the database is designed at the physical level. At this level, the schema describes how the data block is stored and how the storage is managed.

### 2. Logical schema:

In the logical schema, the database is designed at a logical level. At this level, the programmer and data administrator perform their work. Also, at this level, a certain amount of data is stored in a structured way. But the internal implementation data are hidden in the physical layer for the security proposed.

#### 3. View schema:

In view schema, the database is designed at the view level. This schema describes the user interaction with the database system.

Moreover, Data Definition Language (DDL) statements help to denote the schema of a database. The schema represents the name of the table, the name of attributes, and their types; constraints of the tables are related to the schema. Therefore, if users want to modify the schema, they can write DDL statements.

### What is DBMS Instance?

In DBMS, the data is stored for a particular amount of time and is called an instance of the database. The database schema defines the attributes of the database in the particular DBMS. The value of the particular attribute at a particular moment in time is known as an instance of the DBMS.

For example, in the above example, we have taken the example of the attribute of the schema. In this example, each table contains two rows or two records. In the above schema of the table, the employee table has some instances because all the data stored by the table have some instances.

Let's take another example: Let's say we have a single table student in the database; today, the table has 100 records, so today, the instance of the database has 100 records. We are going to add another 100 records to this table by tomorrow, so the instance of the database tomorrow will have 200 records in the table. In short, at a particular moment, the data stored in the database is called the instance; this change over time as and when we add, delete or update data in the database.

### Differences between Database Schema and Instance

Database Schema	Database Instance
It is the definition of the database, or it is defined as the description of the database.	It is a snapshot of a database at a specific moment.
It rarely changes.	It changes frequently.
This corresponds to the variable declaration of a programming language.	The value of the variable in a program at a point in time corresponds to an instance of the database schema.
Defines the basic structure of the database, i.e., how the data will be stored in the database.	It is the set of Information stored at a particular time.

Schema is same for whole database.	Data in instances can be changed using addition, deletion, updation.
It does not change very frequently.	It changes very frequently

# **Data Independence**

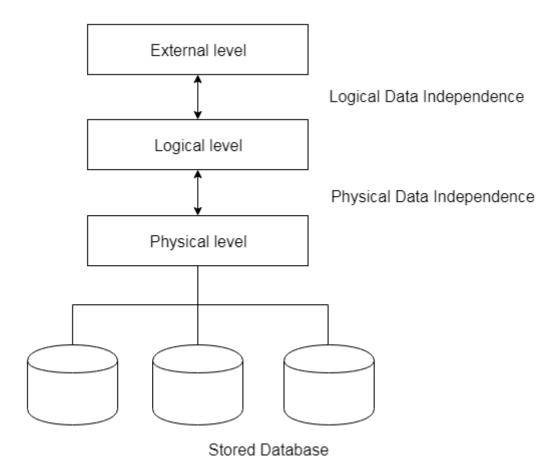
There are two types of data independence:

# 1. Logical Data Independence

- o Logical data independence refers characteristic of being able to change the conceptual schema without having to change the external schema.
- o Logical data independence is used to separate the external level from the conceptual view.
- o If we do any changes in the conceptual view of the data, then the user view of the data would not be affected.
- o Logical data independence occurs at the user interface level.

# 2. Physical Data Independence

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



#### Overall structure of DATABASE:-

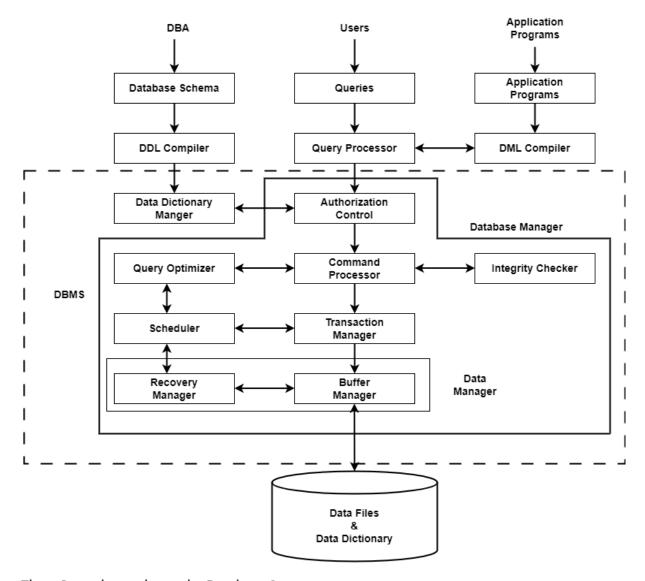
DBMS means Database Management System, which is a tool or software used to create the database or delete or manipulate the database. A software programme created to store, retrieve, query, and manage data is known as a Database Management System (DBMS). Data can be generated, read, updated, and destroyed by authorized entities thanks to user interfaces (UIs).

Because they give programmers, Database Managers, and end users a consolidated view of the data, Database Management Systems are crucial because they relieve applications and end users of the need to comprehend the physical location of the data. Application Programme Interfaces (APIs) manage internet requests and responses for particular sorts of data.

In marketing materials, the phrase "database as a service" (DBaaS) may be used to refer to both relational and non-relational DBMS components that are given via the internet.

Users of DBMSs include application programmers, Database Administrators (DBAs), and end users.

Database Administrators are typically the only people who work directly with a DBMS. Today, end users read and write to databases using front-end interfaces made by programmers, while programmers use cloud APIs to connect with DBMSs.



Three Parts that make up the Database System are:

- o Query Processor
- o Storage Manager
- o Disk Storage

The explanations for these are provided below:

# 1. Query Processor

The query processing is handled by the query processor, as the name implies. It executes the user's query, to put it simply. In this way, the query processor aids the database system in making data access simple and easy. The query processor's primary duty is to successfully execute the query. The Query Processor transforms (or interprets) the user's application program–provided requests into instructions that a computer can understand.

### **Components of the Query Processor**

#### o DDL Interpreter:

Data Definition Language is what DDL stands for. As implied by the name, the DDL Interpreter interprets DDL statements like those used in schema definitions (such as create, remove, etc.). This interpretation yields a set of tables that include the meta-data (data of data) that is kept in the data dictionary. Metadata may be stored in a data dictionary. In essence, it is a part of the disc storage that will be covered in a later section of this article.

#### o DML Compiler:

Compiler for DML Data Manipulation Language is what DML stands for. In keeping with its name, the DML Compiler converts DML statements like select, update, and delete into low-level instructions or simply machine-readable object code, to enable execution. The optimization of queries is another function of the DML compiler. Since a single question can typically be translated into a number of evaluation plans. As a result, some optimization is needed to select the evaluation plan with the lowest cost out of all the options. This process, known as query optimization, is exclusively carried out by the DML compiler. Simply put, query optimization determines the most effective technique to carry out a query.

#### o Embedded DML Pre-compiler:

Before the query evaluation, the embedded DML commands in the application program (such as SELECT, FROM, etc., in SQL) must be pre-compiled into standard procedural calls (program instructions that the host language can understand). Therefore, the DML statements which are embedded in an application program must be converted into routine calls by the Embedded DML Pre-compiler.

#### o Query Optimizer:

It starts by taking the evaluation plan for the question, runs it, and then returns the result. Simply said, the query evaluation engine evaluates the SQL commands used to access the database's contents before returning the result of the query. In a nutshell, it is in charge of analyzing the queries and running the object code that the DML Compiler produces. Apache Drill, Presto, and other Query Evaluation Engines are a few examples.

# 2. Storage Manager:

An application called Storage Manager acts as a conduit between the queries made and the data kept in the database. Another name for it is Database Control System. By applying the restrictions and running the DCL instructions, it keeps the database's consistency and integrity. It is in charge of retrieving, storing, updating, and removing data from the database.

### **Components of Storage Manager**

Following are the components of Storage Manager:

#### o Integrity Manager:

Whenever there is any change in the database, the Integrity manager will manage the integrity constraints.

### o Authorization Manager:

Authorization manager verifies the user that he is valid and authenticated for the specific query or request.

#### o File Manager:

All the files and data structure of the database are managed by this component.

#### o Transaction Manager:

It is responsible for making the database consistent before and after the transactions. Concurrent processes are generally controlled by this component.

### o **Buffer Manager:**

The transfer of data between primary and main memory and managing the cache memory is done by the buffer manager.

# 3. Disk Storage

A DBMS can use various kinds of Data Structures as a part of physical system implementation in the form of disk storage.

## **Components of Disk Storage**

Following are the components of Disk Manager:

#### o Data Dictionary:

It contains the metadata (data of data), which means each object of the database has some information about its structure. So, it creates a repository which contains the details about the structure of the database object.

#### o Data Files:

This component stores the data in the files.

#### o Indices:

These indices are used to access and retrieve the data in a very fast and efficient way.

# **ACID Properties in DBMS**

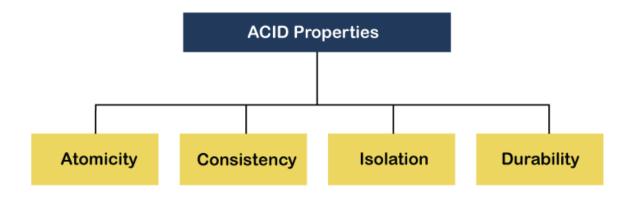
DBMS is the management of data that should remain integrated when any changes are done in it. It is because if the integrity of the data is affected, whole data will get

disturbed and corrupted. Therefore, to maintain the integrity of the data, there are four properties described in the database management system, which are known as the **ACID** properties. The ACID properties are meant for the transaction that goes through a different group of tasks, and there we come to see the role of the ACID properties.

In this section, we will learn and understand about the ACID properties. We will learn what these properties stand for and what does each property is used for. We will also understand the ACID properties with the help of some examples.

# **ACID Properties**

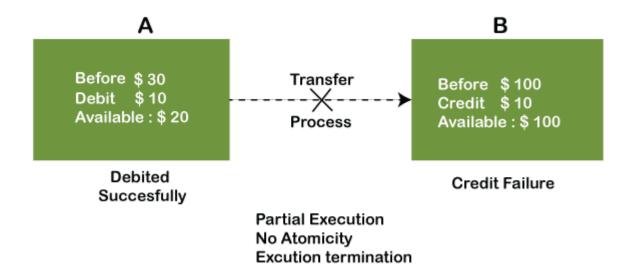
The expansion of the term ACID defines for:



# 1) Atomicity

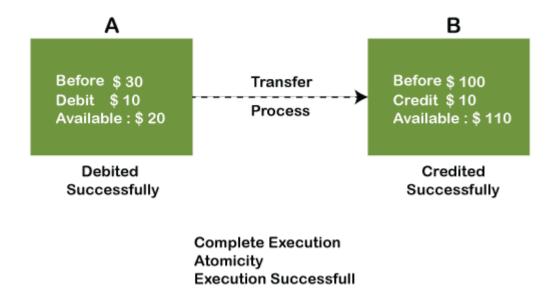
The term atomicity defines that the data remains atomic. It means if any operation is performed on the data, either it should be performed or executed completely or should not be executed at all. It further means that the operation should not break in between or execute partially. In the case of executing operations on the transaction, the operation should be completely executed and not partially.

**Example:** If Remo has account A having \$30 in his account from which he wishes to send \$10 to Sheero's account, which is B. In account B, a sum of \$100 is already present. When \$10 will be transferred to account B, the sum will become \$110. Now, there will be two operations that will take place. One is the amount of \$10 that Remo wants to transfer will be debited from his account A, and the same amount will get credited to account B, i.e., into Sheero's account. Now, what happens – the first operation of debit executes successfully, but the credit operation, however, fails. Thus, in Remo's account A, the value becomes \$20, and to that of Sheero's account, it remains \$100 as it was previously present.



In the above diagram, it can be seen that after crediting \$10, the amount is still \$100 in account B. So, it is not an atomic transaction.

The below image shows that both debit and credit operations are done successfully. Thus the transaction is atomic.



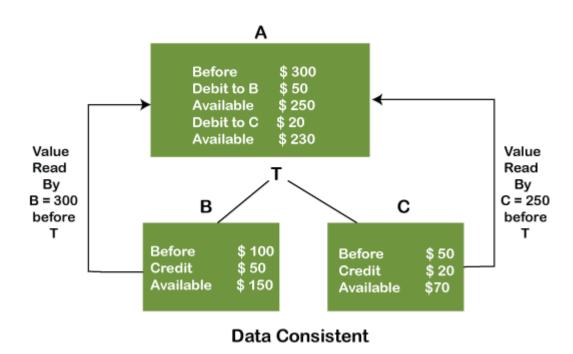
Thus, when the amount loses atomicity, then in the bank systems, this becomes a huge issue, and so the atomicity is the main focus in the bank systems.

# 2) Consistency

The word **consistency** means that the value should remain preserved always. In <u>DBMS</u>, the integrity of the data should be maintained, which means if a change in the database

is made, it should remain preserved always. In the case of transactions, the integrity of the data is very essential so that the database remains consistent before and after the transaction. The data should always be correct.

#### **Example:**

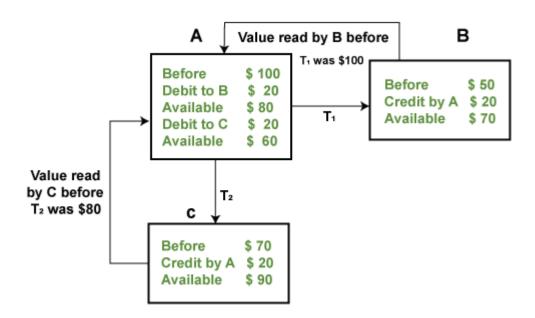


In the above figure, there are three accounts, A, B, and C, where A is making a transaction T one by one to both B & C. There are two operations that take place, i.e., Debit and Credit. Account A firstly debits \$50 to account B, and the amount in account A is read \$300 by B before the transaction. After the successful transaction T, the available amount in B becomes \$150. Now, A debits \$20 to account C, and that time, the value read by C is \$250 (that is correct as a debit of \$50 has been successfully done to B). The debit and credit operation from account A to C has been done successfully. We can see that the transaction is done successfully, and the value is also read correctly. Thus, the data is consistent. In case the value read by B and C is \$300, which means that data is inconsistent because when the debit operation executes, it will not be consistent.

# 3) Isolation

The term 'isolation' means separation. In DBMS, Isolation is the property of a database where no data should affect the other one and may occur concurrently. In short, the operation on one database should begin when the operation on the first database gets complete. It means if two operations are being performed on two different databases, they may not affect the value of one another. In the case of transactions, when two or more transactions occur simultaneously, the consistency should remain maintained. Any changes that occur in any particular transaction will not be seen by other transactions until the change is not committed in the memory.

**Example:** If two operations are concurrently running on two different accounts, then the value of both accounts should not get affected. The value should remain persistent. As you can see in the below diagram, account A is making T1 and T2 transactions to account B and C, but both are executing independently without affecting each other. It is known as Isolation.



Isolation - Independent execution of T<sub>1</sub> & T<sub>2</sub> by A

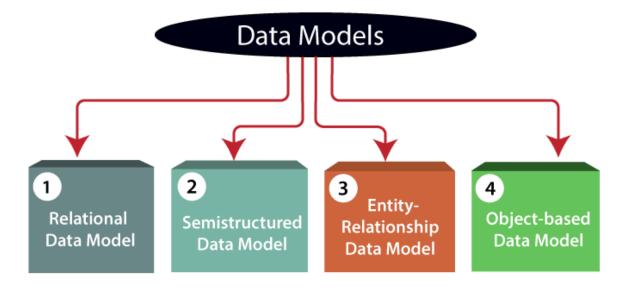
# 4) Durability

Durability ensures the permanency of something. In DBMS, the term durability ensures that the data after the successful execution of the operation becomes permanent in the database. The durability of the data should be so perfect that even if the system fails or leads to a crash, the database still survives. However, if gets lost, it becomes the responsibility of the recovery manager for ensuring the durability of the database. For committing the values, the COMMIT command must be used every time we make changes.

Therefore, the ACID property of DBMS plays a vital role in maintaining the consistency and availability of data in the database.

### **Data Models**

Data Model is the modeling of the data description, data semantics, and consistency constraints of the data. It provides the conceptual tools for describing the design of a database at each level of data abstraction. Therefore, there are following four data models used for understanding the structure of the database:



- 1) Relational Data Model: This type of model designs the data in the form of rows and columns within a table. Thus, a relational model uses tables for representing data and in-between relationships. Tables are also called relations. This model was initially described by Edgar F. Codd, in 1969. The relational data model is the widely used model which is primarily used by commercial data processing applications.
- 2) Entity-Relationship Data Model: An ER model is the logical representation of data as objects and relationships among them. These objects are known as entities, and relationship is an association among these entities. This model was designed by Peter Chen and published in 1976 papers. It was widely used in database designing. A set of attributes describe the entities. For example, student\_name, student\_id describes the 'student' entity. A set of the same type of entities is known as an 'Entity set', and the set of the same type of relationships is known as 'relationship set'.
- 3) Object-based Data Model: An extension of the ER model with notions of functions, encapsulation, and object identity, as well. This model supports a rich type system that includes structured and collection types. Thus, in 1980s, various database systems following the object-oriented approach were developed. Here, the objects are nothing but the data carrying its properties.
- 4) Semistructured Data Model: This type of data model is different from the other three data models (explained above). The semistructured data model allows the data specifications at places where the individual data items of the same type may have different attributes sets. The Extensible Markup Language, also known as XML, is widely used for representing the semistructured data. Although XML was initially designed for including the markup information to the text document, it gains importance because of its application in the exchange of data.

# Relational Model in DBMS

Relational model can represent as a table with columns and rows. Each row is known as a tuple. Each table of the column has a name or attribute.

**Domain:** It contains a set of atomic values that an attribute can take.

**Attribute:** It contains the name of a column in a particular table. Each attribute Ai must have a domain, dom(Ai)

**Relational instance:** In the relational database system, the relational instance is represented by a finite set of tuples. Relation instances do not have duplicate tuples.

**Relational schema:** A relational schema contains the name of the relation and name of all columns or attributes.

**Relational key:** In the relational key, each row has one or more attributes. It can identify the row in the relation uniquely.

**Example: STUDENT Relation** 

NAME	ROLL_NO	PHONE_NO	ADDRESS	AGE
Ram	14795	7305758992	Noida	24
Shyam	12839	9026288936	Delhi	35
Laxman	33289	8583287182	Gurugram	20
Mahesh	27857	7086819134	Ghaziabad	27
Ganesh	17282	9028 9i3988	Delhi	40

- o In the given table, NAME, ROLL\_NO, PHONE\_NO, ADDRESS, and AGE are the attributes.
- o The instance of schema STUDENT has 5 tuples.
- o t3 = <Laxman, 33289, 8583287182, Gurugram, 20>

### **Properties of Relations**

- o Name of the relation is distinct from all other relations.
- o Each relation cell contains exactly one atomic (single) value
- o Each attribute contains a distinct name
- o Attribute domain has no significance
- o tuple has no duplicate value
- o Order of tuple can have a different sequence