Unit 1
DBMS Concept

1. Data: Data is a collection of raw facts and figures which are collected from various sources. But data do not give meaning because they are incomplete in themselves.

2. Information: Information are processed data. When data are processed by a computer that give meaning.

3. Database: A database is the collection of data and information that is organized in a such way that its content can be easily accessed, updated and managed. It is a space where any kind of information is stored. A database can be computerized or not computerized.

4. DBMS: A set of program that manages database. It is a special software that allows creating, retrieving and updating data. Some examples of DBMS are Microsoft Access, Oracle, MySQL, SQL server, Fox Pro etc

5. Advantages of DBMS

- Data stored in database can be shared.
- It reduce data redundancy.
- Integrity (Reliability) can be maintained.
- It can provide multiple user interface.
- Data can be easily accessed.
- Security can be provided to database.

6. Disadvantages of DBMS

- Initial cost is high.
- Proper training is required for staff to run database properly.
- Maintenance, backup and recovery is difficult.
- Database is complex in nature so well trained manpower is needed.

7. Database Model

Database Model describes the structure of database. It defines how data are stored and accessed, it describes data relationship etc.

a. Hierarchical Database Model

In Hierarchical Database Model record are organized in hierarchy (tree pattern) relationship order. The top level record is called root record or parent record and each parent record have more child records.

b. Network Database Model

The modified version of Hierarchical Database Model is Network Database Model. In this model each child nodes may have several parents records.

c. Relational Database Model

In relational database model records are stored in different tables and all the tables may have relationship between all the tables stored in the database. One table have relationship between another one table or many tables.

8. RDBMS (Relational Database Management System)

RDBMS (Relational Database Management System) is a software where database is stored in different tables and contain relationship between the tables. To remove drawback of DBMS RDBMS was developed.

9. Normalization

The process of avoiding repetition of data (Data redundancy), loss of information, inability to represent information and presenting data to normal form is called Normalization. It improve database by reducing data redundancy. A poor table structure can ruin database so, normalization decompose big table into smaller and simplest table and improve database design.

10. Three forms of Normalization

a. First Normal Form (1NF)
b. Second Normal Form (2NF)
c. Third Normal Form (3NF)

a. First Normal Form (1NF): A table is said to be in 1Nf if all attributes are atomic.The purpose of 1NF is to eliminate repeating group of attributes in an entity. Such attributes are decomposed into different rows of same table.

b. Second Normal Form (2NF):The table is said to be in second normal form when it is in first normal form and each attribute is functionally dependent on primary key and there must not be partial dependency.

c. Third Normal Form (3NF):The table is said to be in third normal form if it is in second normal form and doesn't have transitive dependency. All the attributes that are not dependent on primary key must be eliminated

11. SQL (Structured Query Language)

SQL is a database computer language stands for Structured Query Language. It is used in relational database system for storing, manipulating and retrieving data stored in database. The commonly used two types of SQL are as follows:

a. DDL (Data Definition Language)
b. DML (Data Manipulation Language)

a. DDL (Data Definition Language)

DDL is used by the database designers specify the content and structure of the database. It contains commands that can add, remove or modify tables in the database. It has a predefined syntax.

b. DML (Data Manipulation Language)

DML is used to retrieve, sort, display and delete records or data in the database. It provides query commands that can easily update and retrieve data in the database.

12. Centralized and Distributed Database

a. Centralized database: A centralized database is a type of database that is stored, located as well as maintained at a single location. This type of database is modified and managed from that location. The centralized database is accessed via LAN or internet. This centralized database is mainly used by small institutions or organizations.

Advantages Centralized database:

- The centralized database is cheaper than other types of databases as it requires less power and maintenance.
- Since all the data is in one place, there can be stronger security measures around it. So, the centralised database is much more secure.
- All the information in the centralized database can be easily accessed from the same location and at the same time.
- Data is easily portable because it is stored at the same place

Disadvantages Centralized database:

- Failure of server will fail whole database
- Since all the data is at one location, it takes more time to search and access it. If the network is slow, this process takes even more time.

- There is a lot of data access traffic for the centralized database. This may create a bottleneck situation.

b. Distributed database: A distributed database is a type of database which consists of multiple databases that are connected with each other and are spread across different physical locations. The communication between databases at different physical locations is done by a internet.

Advantages Distributed database :

- This database can be easily expanded as data is already spread across different physical locations.
- Since database is located in different locations, there is no data traffic. This will not create a bottleneck situation.
- The distributed database can easily be accessed from different networks.
- Performance is better than centralized database.

Disadvantages Distributed database :

- This database is very costly and it is difficult to maintain because of its complexity.
- Since database is located in different locations, It expensive

13. Database Security

Database security refers to the measures used to protect and secure a database or database management software from against compromising of their confidentially, integrity, other illegal use and malicious cyber threats and attacks.

Some of the ways database security is analyzed and implemented include:

- Restricting unauthorized access / Access controls.
- ensure it does not crash of user overload.
- Physical security of the database server and backup equipment from theft and natural disasters.
- Regular data backups can be planned and multiple copies can be stored off-site to provide redundancy and emergency recovery.
- Data encryption can provide an additional layer of security to protect the integrity and confidentiality of data.
- Data integrity

-0-