INTRODUCTION OF DBMS

As the name suggests, the database management system consists of two parts. They are:

1. Database
2. Management System

Database:

To find out what database is, we have to start from data, which is the basic building block of any DBMS.

Data: Facts, figures, statistics etc. having no particular meaning (e.g. 1, ABC, 19 etc).

Record: Collection of related data items, e.g. in the above example the three data items had no meaning. But
if we organize them in the following way, then they collectively represent meaningful information.
Roll Name Age
1 ABC 19

Table or Relation: Collection of related records.

Roll Name Age
1 ABC 19
2 DEF 22
3 XYZ 28

The columns of this relation are called Fields, Attributes or Domains. The rows are called Tuples or Records.

Database: Collection of related relations. Consider the following collection of tables:

T1 T2
Roll Name Age
1 ABC 19
2 DEF 22
3 XYZ 28
T3 T4

We now have a collection of 4 tables. They can be called a “related collection” because we can clearly find out that there are some common attributes existing in a selected pair of tables. Because of these common attributes we may combine the data of two or more tables together to find out the complete details of a student.

Age and Hostel attributes are in different tables.

A database in a DBMS could be viewed by lots of different people with different responsibilities.

In a database, data is organized strictly in row and column format. The rows are called Tuple or Record. The data items within one row may belong to different data types. On the other hand, the columns are often called Domain or Attribute. All the data items within a single attribute are of the same data type.

Management System:

A database-management system (DBMS) is a collection of interrelated data and a set of programs to access those data. This is a collection of related data with an implicit meaning and hence is a database. The collection of data, usually referred to as the database, contains information relevant to an enterprise. The primary goal of a DBMS is to provide a way to store and retrieve database information that is both convenient and efficient. By data, we mean known facts that can be recorded and that have implicit meaning.

The management system is important because without the existence of some kind of rules and regulations it is not possible to maintain the database. We have to select the particular attributes which should be included in a particular table; the common attributes to create relationship between two tables; if a new record has to be inserted or deleted then which tables should have to be handled etc. These issues must be resolved by having some kind of rules to follow in order to maintain the integrity of the database.

Database systems are designed to manage large bodies of information. Management of data involves both defining structures for storage of information and providing mechanisms for the manipulation of information. In addition, the database system must ensure the safety of the information stored, despite system crashes or attempts at unauthorized access. If data are to be shared among several users, the system must avoid possible anomalous results.

Because information is so important in most organizations, computer scientists have developed a large body of concepts and techniques for managing data.

Advantages

- Reduced data redundancy.
- Also, there is reduced updating errors and increased consistency.
- Easier data integrity from application programs.
- Improved data access to users through the use of host and query languages.
- Data security is also improved.
- Reduced data entry, storage, and retrieval costs.

Disadvantages

- Complexity: Databases are complex hardware and software systems.
- Cost: It requires significant upfront and ongoing financial resources.
- Security: Most leading companies need to know that their Database systems can securely store data, including sensitive employee and customer information.
- Compatibility: There is a risk that a DBMS might not be compatible with a company's operational requirements.

Types of Database:

1. Centralised database: A centralised database is one that operates entirely within a single location. Bigger organisations like corporations or universities typically use centralised databases. The database

itself is on a central computer or database system. Users can access the database through a computer network, but it is the central computer that runs and maintains the database.

2. Cloud database: Cloud databases run on cloud computing networks. Rather than storing data on local devices or purchasing dedicated storage hardware, companies can install software to access large amounts of data in the cloud. Companies often choose cloud databases because they are easy to use. They can easily increase storage capacity and they often back up the data on remote servers.

3. Commercial database: Commercial databases are databases created by companies and sold to other businesses. Businesses typically pay a licensing fee for these. The vendor often provides technical support to the businesses rather than internal IT teams or community support. These are high-performing databases that are reliable and scalable for growing businesses.

4. Distributed database: Distributed databases store data at various physical locations. This can mean companies store data on different physical computers at a single location or hardware across different locations on a network. Distributed databases use processes called replication and fragmentation when storing data. Replication is when the database stores two or more copies of the same file at different locations while fragmentation is when the database divides files into smaller pieces that they then store in different locations.

5. End-user database: End-user is a term used in product development that refers to the person who uses the product. An end-user database is, therefore, a database that is primarily used by a single person. A good example of this type of database is a spreadsheet stored on your local computer if you own a business.

6. Graph database: Graph databases focus on the relationships between data along with the data itself. The database uses structures of connected graphs of nodes and edges for users to run queries. Companies may choose graph databases if they have datasets with complex relationships.

7. Hierarchical database: Hierarchical databases store data in categories that get more specific. The databases store data as tree diagrams, showing the relationship between data elements. The top-level data is the root or parent, below that is the child or level one and below that is level two. For example, if the data is information about a particular technology market, the parent may be a category like computers, the child might be laptops, and the level two could be a specific brand.

8. Network database: Network databases allow the data to be linked to over one parent file. Visually, companies represent this as an upside-down tree where each piece of data links to the primary data. With network databases, users link many data points to other bits of information, creating a web or network. Different from the hierarchical model, relationships span multiple categories or trees.

9. NoSQL database: There are essentially two major types of databases, NoSQL and Relational, with all the others being different versions of these. A NoSQL database has a hierarchy similar to a file folder system and the data within it is unstructured. This lack of structure allows them to process larger amounts of data at speed and makes it easier to expand in the future. Cloud computing regularly makes use of NoSQL databases.

10. Objected-oriented database: Object-oriented databases are ones in which the data is represented as objects and classes. An object is a real-world item, such as a name or phone number, while a class is a group of objects. Objected-oriented databases are a type of relational database. Consider using an object-oriented database when you have a large amount of complex data you want to process quickly.

11. Open-source database: An open-source database is a public database that all people can use for free. Unlike commercial databases, users can download or sign up for open source databases without paying a fee. The term "open source" refers to a program in which users can see how companies make it and they can edit the program to fit their needs. Open-source databases are typically much cheaper than commercial databases, but they can also lack some of the more advanced features found in commercial databases.

12. Operational database: The purpose of an operational database is to allow users to modify data in real-time. Operational databases are critical in business analytics and data warehousing. Companies set these up either as relational databases or NoSQL, depending on needs. Conventional databases rely on batch processing, where they process commands in groups. Operational databases allow you to add, edit and remove data at any moment, in real-time.

13. Personal database: A personal database is one that is designed for a single person. You typically find these on a personal computer and has a very simple design, comprising only a few tables. Personal databases are not typically suitable for complex operations, large amounts of data or business operations.

14. Relational database: Relational databases are the other major type of database, opposite of NoSQL. Relational databases store data in a structured way and about other data. A good representation of a relational database would be a sales associate and their client's purchase history. Companies prefer relational databases when their primary concern is the integrity of their data, or when they are not particularly focused on scalability.

Instances and Schemas

Databases change over time as information is inserted and deleted. The collection of information stored in the database at a particular moment is called an instance of the database. The overall design of the database is called the **database schema**. Schemas are changed infrequently, if at all. The concept of database schemas and instances can be understood by analogy to a program written in a

programming language. A database schema corresponds to the variable declarations (along with associated type definitions) in a program.

Each variable has a particular value at a given instant. The values of the variables in a program at a point in time correspond to an instance of a database schema. Database systems have several schemas, partitioned according to the levels of abstraction. The physical schema describes the database design at the physical level, while the logical schema describes the database design at the logical level. A database may also have several schemas at the view level, sometimes called sub-schemas, which describe different views of the database.

Of these, the logical schema is by far the most important, in terms of its effect on application programs, since programmers construct applications by using the logical schema. The physical schema is hidden beneath the logical schema, and can usually be changed easily without affecting application programs.

Application programs are said to exhibit physical data independence if they do not depend on the physical schema, and thus need not be rewritten if the physical schema changes.

Database Administrators and Database Users

Database Administrators

Database administrators also back up, restore, and troubleshoot database sets and system access, updating and integrating old programs to implement the latest technology.

Database administrators need at least a bachelor's degree in information science or computer science for most entry-level positions. Depending on the size and complexity of their company or governing body, these professionals may need a master's degree in database administration or information technology. All database administrators need fundamental knowledge of structured query language (SQL) and software vendor certifications.

A primary goal of a database system is to retrieve information from and store new information in the database.

People who work with a database can be categorized as database users or database administrators.

Database Users and User Interfaces

There are four different types of database-system users, differentiated by the way they expect to interact with the system. Different types of user interfaces have been designed for the different types of users.

Naive users are unsophisticated users who interact with the system by invoking one of the application programs that have been written previously. For example, a bank teller who needs to transfer $50 from account

A to account B invokes a program called transfer. This program asks the teller for the amount of money to be transferred, the account from which the money is to be transferred, and the account to which the money is to be transferred.