

Database Management System

Unit:1 Database Concept

Definitions:

Data: The term data means groups of information that represent the attributes.

Information: It is the groups of words that form the Data.

Database: Database is a collection of files or tables or persistent data that is used by systems of some given enterprise.

DBMS:

It is computerized record keeping system that manages the databases.

Primary goals of DBMS are:

1. To store and retrieve database information
2. Accept DML (insert, update, delete) requests and analyze it.
3. To manage large and small part of information. It define structures for Storage of information and providing mechanism for manipulation of information.
4. It should provide security for stored information; provide system crashes or security from unauthorized access; provide recovery of data if data is lost.
5. Data is shared among several users

Components of DBMS

There are five components of the DBMS given as below.

Data:

- Data is a bridge between Machine components and Human components.
- Data is a raw material using it we can create database.
- Two types of database

☐ Shared: Here database can be shared among the different users, means different users can access the same data. ☐

Integrated: Here database is a unification of several distinct files.

Hardware:

It consists following hardware components. - Storage devices like magnetic disc, hard disk etc. - I/O devices, I/O channels, Processor, RAM etc.

Procedures:

- It is used to design, manage and to use the database. - It is used to be log in, to start and stop the database, to change table structure, to take database backup etc.

Software:

- Software is a collection of programs in computer. - Software are given as below:

1. Operating System: It manages hardware components.
2. DBMS: It manage the database (Ex. MS Access, SQL server, IBM's DB2 etc)
3. Application Programs: It is used to access data, make reports etc.

End-users:

- It includes all users of the database system. - There are five types of users given as below.

1. end users.
2. DBA- DataBase Administrator.
3. Database designer.
4. Application Programmers
5. stand alone end user
6. sophisticated end user

DB users or User's Role in DBMS environment

There are five users are available in DBMS environment.

1. Database Administrator (DA): ☐ In enterprise there is a person who has the central responsibility for the data, this person is called DA. This person is a manager not technician.
2. DA is a person who understands the data, needs of data, what data should be stored, create policies to deal with that data.

3. Database Administrator (DBA):
 4. DBA is the technical person who implements the decisions of the DA.
 5. DBA create actual DB and Check performance of database.
 6. Database Designer: Design the database as per need of users.
4. Application Programmers: This person write database application programs in some programming language
5. End-users: This person access the given database also called client

Functions of DBMS

1. Data Dictionary Management

Data Dictionary Management is the one of the most important function of database management system.

DBMS stores definitions of the data elements and their relationships (metadata) in a data dictionary.

2. Data Storage Management

One of the DBMS functionality is creating and managing the complex structures required for data storage, thus relieving you from the difficult task of defining and programming the physical data characteristics.

A modern DBMS system provides storage not only for the data, but also for related data entry forms or screen definitions, report definitions, data validation rules, procedural code, structures to handle video and picture formats, and so on.

Data storage management is also important for database performance tuning.

3. Data transformation and presentation

The DBMS transforms entered data in to required data structures.

That is, the DBMS formats the physically retrieved data to make it conform to the user's logical expectations.

For example, imagine an enterprise database used by a multinational company.

An end user in England would expect to enter data such as July 11, 2009, as "11/07/2009."

In contrast, the same date would be entered in the United States as "07/11/2009."

Regardless of the data presentation format, the DBMS system must manage the date in the proper format for each country.

4. Security Management

Security Management is another important function of Database Management System(DBMS).

The DBMS creates a security system that enforces user security and data privacy.

Security rules determine which users can access the database, which data items each user can access, and which data operations (read, add, delete, or modify) the user can perform.

This is especially important in multiuser database systems.

5. Multi User Access Control

Multiuser access control is another important DBMS Function.

To provide data integrity and data consistency, the DBMS uses sophisticated algorithms to ensure that multiple users can access the database concurrently without compromising the integrity of the database.

6. Backup and Recovery Management

The DBMS provides backup and data recovery to ensure data safety and integrity.

Current DBMS systems provide special utilities that allow the DBA to perform routine and special backup and restore procedures.

Recovery management deals with the recovery of the database after a failure, such as a bad sector in the disk or a power failure.

7. Data Integrity Management

Data integrity management is another important DBMS function.

The DBMS promotes and enforces integrity rules, thus minimizing data redundancy and maximizing data consistency.

The data relationships stored in the data dictionary are used to enforce data integrity

8. Database Access Languages and Application Programming Interfaces

The DBMS provides data access through a query language.

A query language is a non procedural language—one that lets the user specify what must be done without having to specify how it is to be done.

Structured Query Language (SQL) is the defacto query language and data access standard supported by the majority of DBMS vendors.

9. Database Communication Interfaces

Current-generation DBMS's accept end-user requests via multiple, different network environments.

For example, the DBMS might provide access to the database via the Internet through the use of Web browsers such as Mozilla Firefox or Microsoft Internet Explorer.

In this environment, communications can be accomplished in several ways:

DBMS Advantages and Disadvantages

Advantages of DBMS:

1. Data can be shared.
2. Data isolation can be solved.
3. Data Independence: Hide the details of data representation and storage.
4. Redundancy and Inconsistency can be reduced.
5. Security can be applied.
6. Data Integrity can be maintained.
7. Provided transaction support.
8. Reduced Application Development time.
9. Concurrent access of the data.
10. Recover the data if data is lost.
11. Inconsistency can be avoided.

Disadvantages of a DBMS:

1. Setup of the database system requires more knowledge, money, skills, and time.
2. The complicated database may result in poor performance.

Redundancy: In two files person's name and address repeat frequently. It can be solve by creating new database.

Data isolation: In database data is stored at one location so we don't need to access it from more location.

Inconsistency: In database two or more sets of same file or data known as Data Inconsistency. When we make some changes in one database and changes not done in another database, it creates Data

Inconsistency. Data Integrity (Data consistent): Data must be correct and accurate with valid range and format.

Concurrent access: It means access of the data in such a way that the data is being accessed by only one user a time.

Advantages of DBMS

The database management system has promising potential advantages, which are explained below:

1. **Controlling Redundancy:** In file system, each application has its own private files, which cannot be shared between multiple applications. This can often lead to considerable redundancy in the stored data, which results in wastage of storage space.
2. By having centralized database most of this can be avoided.
3. It is not possible that all redundancy should be eliminated.
4. Sometimes there are sound business and technical reasons for maintaining multiple copies of the same data.
5. In a database system, however this redundancy can be controlled.

For example: In case of college database, there may be the number of applications like General Office, Library, Account Office, Hostel etc. Each of these applications may maintain the following information into own private file applications:

General Office	Library	Hostel	Account Office
Roll No Name Class Father_Name Date_of_Birth Address Phone No Previous Record Attendance Marks etc.	Roll No Name Class Address Date at Birth Phone No No of books issued Fine etc	Roll No Name Class Father_Name Date of Birth Address Phone No Mess bill RoomNo etc.	Roll No Name Class Address Phone No Fee Installments Discount Balance Total etc.

It is clear from the above file systems, that there is some common data of the student which has to be mentioned in each application, like Rollno, Name, Class, Phone_No~ Address etc.

This will cause the problem of redundancy which results in wastage of storage space and difficult to maintain, but in case of centralized database, data can be shared by number of applications and the whole college can maintain its computerized data with the following database:

General Office	Library	Hostel	Account Office
Rollno Name Class Father_Name Address Phone_No Date_of_birth Previous_Record Attendance Marks etc.	Rollno No_of_books_issued Fine etc.	Rollno RoomNo Mess_Bill etc.	Rollno Fee Installments Discount Balance Total etc.

It is clear in the above database that Rollno, Name, Class, Father_Name, Address, Phone_No, Date_of_birth which are stored repeatedly in file system in each application, need not be stored repeatedly in case of database, because every other application can access this information by joining of relations on the basis of common column i.e. Rollno.

Suppose any user of Library system need the Name, Address of any particular student and by joining of Library and General Office relations on the basis of column Rollno he/she can easily retrieve this information.

Thus, we can say that centralized system of DBMS reduces the redundancy of data to great extent but cannot eliminate the redundancy because RollNo is still repeated in all the relations.

2. Integrity can be enforced: Integrity of data means that data in database is always accurate, such that incorrect information cannot be stored in database.

In order to maintain the integrity of data, some integrity constraints are enforced on the database.

A DBMS should provide capabilities for defining and enforcing the constraints.

For Example: Let us consider the case of college database and suppose that college having only BTech, MTech, MSc, BCA, BBA and BCOM classes.

3. Inconsistency can be avoided : When the same data is duplicated and changes are made at one site, which is not propagated to the other site, it gives rise to inconsistency and the two entries regarding the same data will not agree.

At such times the data is said to be inconsistent. So, if the redundancy is removed chances of having inconsistent data is also removed.

Let us again, consider the college system and suppose that in case of General_Office file it is indicated that Roll_Number 5 lives in Amritsar but in library file it is indicated that

Roll_Number 5 lives in Jalandhar. Then, this is a state at which the two entries of the same object do not agree with each other (that is one is updated and other is not). At such time the database is said to be inconsistent.

An inconsistent database is capable of supplying incorrect or conflicting information. So there should be no inconsistency in database. It can be clearly shown that inconsistency can be avoided in centralized system very well as compared to file system ..

Let us consider again, the example of college system and suppose that RollNo 5 is shifted from Amritsar to Jalandhar, then address information of Roll Number 5 must be updated, whenever Roll number and address occurs in the system.

4. Data can be shared: As explained earlier, the data about Name, Class, Father_name etc. of General_Office is shared by multiple applications in centralized DBMS as compared to file system so now applications can be developed to operate against the same stored data.

The applications may be developed without having to create any new stored files.

6. **Standards can be enforced :** Since DBMS is a central system, so standard can be enforced easily may be at Company level, Department level, National level or International level.
7. The standardized data is very helpful during migration or interchanging of data.
8. The file system is an independent system so standard cannot be easily enforced on multiple independent applications.
7. **Restricting unauthorized access:** When multiple users share a database, it is likely that some users will not be authorized to access all information in the database.
8. For example, account office data is often considered confidential, and hence only authorized persons are allowed to access such data.
9. In addition, some users may be permitted only to retrieve data, whereas others are allowed both to retrieve and to update.
10. Hence, the type of access operation retrieval or update must also be controlled. Typically, users or user groups are given account numbers protected by passwords, which they can use to gain access to the database.
11. A DBMS should provide a security and authorization subsystem, which the DBA uses to create accounts and to specify account restrictions. The DBMS should then enforce these restrictions automatically.

7. Solving Enterprise Requirement than Individual Requirement: Since many types of users with varying level of technical knowledge use a database, a DBMS should provide a variety of user interface. The overall requirements of the enterprise are more important than the individual user requirements. So, the DBA can structure the database system to provide an overall service that is "best for the enterprise".

For example: A representation can be chosen for the data in storage that gives fast access for the most important application at the cost of poor performance in some other application. But, the file system favors the individual requirements than the enterprise requirements

9. **Providing Backup and Recovery:** A DBMS must provide facilities for recovering from hardware or software failures.
10. The backup and recovery subsystem of the DBMS is responsible for recovery.

11. For example, if the computer system fails in the middle of a complex update program, the recovery subsystem is responsible for making sure that the .
12. database is restored to the state it was in before the program started executing.

9. Cost of developing and maintaining system is lower: It is much easier to respond to unanticipated requests when data is centralized in a database than when it is stored in a conventional file system.

Although the initial cost of setting up of a database can be large, but the cost of developing and maintaining application programs to be far lower than for similar service using conventional systems.

The productivity of programmers can be higher in using non-procedural languages that have been developed with DBMS than using procedural languages.

11. Concurrency Control : DBMS systems provide mechanisms to provide concurrent access of data to multiple users.

Disadvantages of DBMS

The disadvantages of the database approach are summarized as follows:

1. Complexity : The provision of the functionality that is expected of a good DBMS makes the DBMS an extremely complex piece of software. Database designers, developers, database administrators and end-users must understand this functionality to take full advantage of it.

2. Size : The complexity and breadth of functionality makes the DBMS an extremely large piece of software, occupying many megabytes of disk space and requiring substantial amounts of memory to run efficiently.

3. Performance: Typically, a File Based system is written for a specific application, such as invoicing. As result, performance is generally very good.

4. Higher impact of a failure: The centralization of resources increases the vulnerability of the system. .

Cost of DBMS: The cost of DBMS varies significantly, depending on the environment and functionality provided. There is also the recurrent annual maintenance cost.

6. Additional Hardware costs: The disk storage requirements for the DBMS and the database may necessitate the purchase of additional storage space. Furthermore, to achieve the required performance it may be necessary to purchase a larger machine, perhaps even a machine dedicated to running the DBMS.

7. Cost of Conversion: In some situations, the cost of the DBMS and extra hardware may be insignificant compared with the cost of converting existing applications to run on the new DBMS and hardware..

DBMS USERS

Application Programmers

application programmers are the one who writes application programs that uses the database.

These application programs are written in programming languages like COBOL or PL (Programming Language 1), Java and fourth generation language.

These programs meet the user requirement and made according to user requirements.

Retrieving information, creating new information and changing existing information is done by these application programs.

End user

End users are those who access the database from the terminal end.

They use the developed applications and they don't have any knowledge about the design and working of database.

DBA (Database Administrator)

DBA can be a single person or it can be a group of person.

Database Administrator is responsible for everything that is related to database.

He makes the policies, strategies and provides technical supports.

System Analyst

System analyst is responsible for the design, structure and properties of database.

All the requirements of the end users are handled by system analyst.

Feasibility, economic and technical aspects of DBMS is the main concern of system analyst.

Application of DBMS

1. Banking – For customer information, accounts, and loans, and banking transactions. [All transactions]
2. Airlines – For reservation and schedule information. [Reservations, schedules]
3. Universities – For student information, results, course registrations, and grades. [Registration, Grades]

Advantages Disadvantages

1. Greater flexibility
1. Difficult to learn
2. Greater processing power
2. Slower processing speeds
3. Storage for all relevant data
3. Expensive
4. Fits the needs of many medium to large sized organizations
4. Requires skilled administrators

4. Credit Card Transactions – For purchases on credit card and generation of monthly Statements.
5. Telecommunication – For keeping records of calls made, generating monthly bills, Maintaining balances on prepaid calling cards and storing information about Communication networks.
6. Finance – For storing information about holdings, sales, and purchases of financial Instruments such as stocks and bonds.
7. Sales – For customer, product, and purchase information.
8. Manufacturing – For management production of items in factories, inventories of items in warehouses/stores, and orders for items. [Production, inventory, orders, supply chain]
9. Human Resources – For information about employees, salaries, payroll taxes and Benefits and generation of paychecks. [Employee records, salaries, tax deduction]

Difference between File System and DBMS :

S.NO.	FILE SYSTEM	DBMS
1.	File system is a software that manages and organizes the files in a storage medium within a computer.	DBMS is a software for managing the database.
2.	Redundant data can be present in a file system.	In DBMS there is no redundant data.
3.	It doesn't provide backup and recovery of data if it is lost.	It provides backup and recovery of data even if it is lost.
4.	There is no efficient query processing in file system.	Efficient query processing is there in DBMS.
5.	There is less data consistency in file system.	There is more data consistency because of the process of

S.NO.	FILE SYSTEM	DBMS
		normalization.
6.	It is less complex as compared to DBMS.	It has more complexity in handling as compared to file system.
7.	File systems provide less security in comparison to DBMS.	DBMS has more security mechanisms as compared to file system.
8.	It is less expensive than DBMS.	It has a comparatively higher cost than a file system.

Functions of Data Base Administrator

Data Base Administrator (DBA) is a person or group in charge for implementing **DBMS** in an organization. **Database Administrator's** job requires a high degree of technical expertise and the ability to understand and interpret management requirements at a senior level. In practice the DBA may consist of team of people rather than just one person

- **Makes decisions concerning the content of the database:**

It is the DBA's job to decide exactly what information is to be held in the database-in other words, to identify the entities of interest to the enterprise and to identify information to be recorded about those entities.

- **Plans storage structures and access strategies:**

The DBA must also decide how the data is to be represented in the database, and must specify the representation by writing the storage structure definition (using the internal data definition language).

- **Provides support to users:** It is the responsibility of the DBA to provide support to the users, to ensure that the data they require is available, and to write the necessary external schemas (using the appropriate external data definition language).

- **Defines security and integrity checks:** DBA is responsible for providing the authorization and authentication checks such that no malicious users can access database and it must remain protected.

DBA must also ensure the integrity of the database.

- **Interprets backup and recovery strategies:** In the event of damage to any portion\ of the database-caused by human error, say, or a failure in the hardware or supporting [operating system](#)-it is essential to be able to repair the data concerned with a minimum of delay and with as little effect as possible on the rest of the system.

The DBA must define and implement an appropriate recovery strategy to recover the database from all types of failures.

- **Monitoring performance and responding to changes in requirements:**

The DBA is responsible for so organizing the system as to get the performance that is "best for the enterprise," and for making the appropriate adjustments as requirements change.