CHAPTER: 1 DATABASE SYSTEM CONCEPT

Specific Objective

- State the importance of DBMS effectiveness and database tools.
- State the advantages of using database system to store operational data.
- Explain the concept of RDBMS.
- Describe the overall structure of DBMS & Architecture of Client/Server system.
- Explain the concept of data mining and data warehousing.

1.1 An Introduction to Database

1.1.1 What is Data?

- Data can be facts related to any object in consideration.
- User input is known as data.
- Data is a collection of facts, such as numbers, words, measurements, observations or even just descriptions of things.
- Example: - Name, age, height, weight etc.

1.1.2 What is Database?

- A database is a systematic collection of data. It support storage & manipulation of data.
- A database is an organized collection of data. The data is typically organized to model aspects of reality in a way that supports processes requiring information.
- Example:-
  1. An online telephone numbers, other contact details etc.
  2. Facebook it needs to store, manipulate & present data related to members, their friends, member activities, messages etc.

1.1.3 What is DBMS?

Q. What is DBMS? Explain its functions.

- DBMS is a collection of interrelated data & set of programs to access the data.
- Data Base Management Systems (DBMS) are computer software applications that interact with the user, other applications, and the database itself to capture and analyze data.
- DBMS is a collection of programs which enables its users to access database, manipulate data, reporting or representation of data. It also helps to control access to the database.
- **The primary goal of a DBMS is to provide a way to store & retrieve database information that both convenient & efficient.**
- Another important feature of DBMS is security. The data stored can be confidential & important for the further developments.
- Example: - MySQL, FoxPro, Ms-access etc.

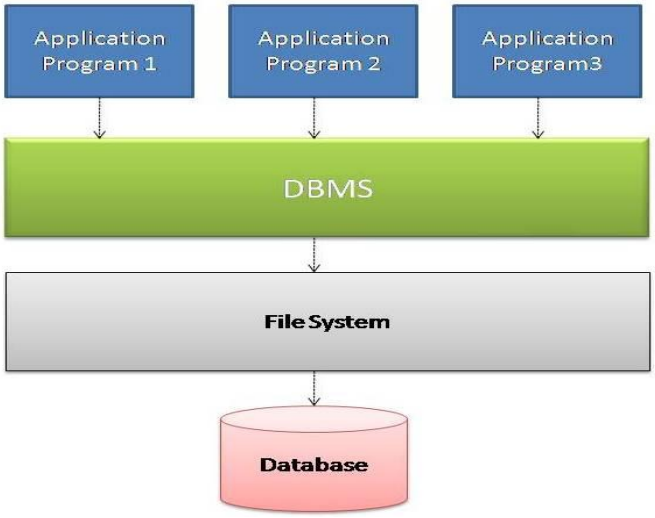


Fig.1. Data Base Management Systems

Functions of DBMS. (IMP)

1. **Data Definition:** The DBMS must be able to accept data definitions in source form & convert them to the appropriate object from.
2. **Data Manipulation:** The DBMS must be able to handle requests from the users to retrieve, update or delete existing data the database, or to add new data to the database.
3. **Database Communication Interfaces:** The end-user's requests for database access are transmitted to DBMS in the form of communication messages.
4. **Authorization / Security Management**: The DBMS protects the database against unauthorized access, either international or accidental. It furnishes mechanism to ensure that only authorized users can access the database.
5. **Backup and Recovery Management:** The DBMS provides mechanisms for backing up data periodically and recovering from different types of failures. This prevents the loss of data.
6. **Concurrency Control Service:** Since DBMSs support sharing of data among multiple users, they must provide a mechanism for managing concurrent access to the database. DBMSs ensure that the database kept in consistent state and that integrity of the data is preserved.
7. **Transaction Management:** A transaction is a series of database operations, carried out by a single user or application program, which accesses or changes the contents of the database. Therefore, a DBMS must provide a mechanism to ensure either that all the updates corresponding to a given transaction are made or that none of them is made.
8. **Database Access and Application Programming Interfaces:** All DBMS provide interface to enable applications to use DBMS services. They provide data access via Structured Query Language (SQL). The DBMS query language contains two components: (a) a Data Definition Language (DDL) and (b) a Data Manipulation Language (DML).

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

1.1.4 Disadvantages of File Processing System:-

1. **Data Redundancy:-**Redundancy is unnecessary duplication of data. For example, if accounts department & registration department both keep student name, number & address.
2. **Data Inconsistency: -** It means different copies of the same data are not matching. That means different versions of same basic data are existing.
3. **Difficulty in accessing data:-**The file processing system do not allow to access data in convenient & efficient way.
4. **Data Isolation:**-Data are scattered in various files, & the files may be in different format, writing new application program to retrieve data is difficult.
5. **Integrity Problems: -** The data values may need to satisfy some integrity constraints. For example, the balance field value must be greater than 5000. We have to handle this through program code in file processing systems. But in database we can declare the integrity constraints along with definition itself.
6. **Atomicity Problem: -** If any failure occurs in the system the transactions which are executing should fully get executed or should not, so that database remains in consistent (correct) state. For example, if a transaction of transferring money from account X to account Y is in process. If a system failure occurs the money from account X to Y should get transferred or should not get transferred. File processing system do not ensures such atomicity.
7. **Security Problems:-**Enforcing security constraints in file processing system is very difficult as the application programs are added to the system in an ad-hoc manner.

1.1.5 Advantages of DBMS over file processing system

Q. State and explain advantages of DBMS over file processing system.

1. Reduction in Redundancy: Duplication of records is reduced.
2. Avoiding Inconsistency: As the redundancy is reduced inconsistency is avoided.
3. Maintaining Integrity: Accuracy is maintained.
4. Sharing of data: Sharing of data is possible.
5. Enforcement of Security: Security can be enforced.
6. Transaction support.

Q. Comparison between File Processing system & DBMS.

S. N.	File Processing System	DBMS
1	PC- Based small systems.	Mini-mainframe based large systems.
2	Single user system.	Multiple user system.
3	Relatively cheaper.	Relatively expensive.

4	Less number of files used.	More number of files used.
5	Data redundancy & inconsistency occur.	Data independent & non-redundant.
6	Data access is difficult.	Data access is efficient.
7	Security problems.	Better security.
8	Data is isolated.	Data is integrated.
9	Transaction concept is not used.	The concept of transaction is important aspect.

1.1.6 Applications of DBMS

Q. Explain different applications of DBMS.

1. **Airlines and Railways:** Online database for reservation, displaying the schedule information.
2. **Banking:** Customer inquiry, accounts, loans, and other transactions.
3. **Education:** Course registration, result, and other information.
4. **Telecommunications:** Communication network, telephone numbers, record of calls, for generating monthly bills, etc.
5. **E-commerce:** Business activity such as online shopping, booking of holiday package, consulting a doctor, etc.
6. **Human resources:** Organizations use database for storing information about their employees, salaries, benefits, taxes, and for generating salary checks.
7. **Web based services:-**For web user's feedback, responses, resource sharing etc.
8. **Sales: -** For customer, product & purchase information.

1.2 Introduction to RDBMS

1.2.1 What is RDBMS?

- RDBMS stands for Relational Database Management System. It organizes data into related rows and columns.
- The principles of the relational model were first outlined by Dr. E. F. Codd in a June 1970 paper called "A Relational Model of Data for Large Shared Data Banks:'
- The data in RDBMS is stored in database object called tables. The table is a collection of related data entries & it consists of columns & rows.
- A record also called a row of data, is each individual entry that exists in a table.
- A column is a vertical entity in a table that contains all information associated with a specific field in a table.
- A field is a column in a table that is designed to maintain specific information about every record in the table.
- It is a program that lets you create, update & administer a relational database.
- The most popular RDBMS are MS-SQL Server, DB2, Oracle, MySQL & so on.

Features of RDBMS:-

1. Provides data to be stored in tables.
2. Provides facility primary key to uniquely identify the rows.
3. Creates indexes for quicker data retrieval.
4. Persists data in the form of rows & columns.

1.2.2 Compare DBMS and RDBMS (any four points-for 4 Marks).

S. N.	DBMS	RDBMS
1	DBMS applications store **data as file**.	RDBMS applications store **data in a tabular form**.
2	Old version of software to handle the databases.	Latest version of software for handling databases.
3	Can relate one table to another table.	RDBMS can relate one database to another database.
4	Data security is low as compare to RDBMS.	Level of data security is very high as compare to DBMS.
5	Data storage capacity is less as compare to RDBMS.	Data storage capacity is very high.
6	Not easy to maintain data integrity.	It can be maintained easily in RDBMS.
7	Works better in single user or few user systems.	Works very efficiently and give good performance over the network.
8	All Codd's 12 rules are not followed.	All Codd's 12 rules are followed.
9	DBMS normalization process will not be present.	RDBMS fully support normalization
10	e.g. FoxPro, Ms-Access	e.g. SQL-server, Oracle,IBM-DB2

1.3 Names of various DBMS & RDBMS Software's

FoxPro FoxProW Dbase MS Access etc.	Oracle RDBMS IBM DB2 Microsoft SQL Server SAP Sybase ASE Teradata ADABAS MySQL FileMaker Microsoft Access Informix etc.

1.4 DBMS Concepts

1.4.1 Data Abstraction (IMP)

Q. Explain three levels of data abstraction with suitable diagram.

OR

Q. Describe data abstraction with neat diagram. ***(Diagram- 1 Mark, Description of level-1 Mark)***

OR

Q. Explain three levels of data abstraction with suitable diagram.

Ans: - Data Abstraction:-

- It can be defined as the process of hiding the complexity of data & representing the data which need to be shown to the user.
- The major purpose of a database system is to provide users with an abstract view of the system.
- The system hides certain details of how data is stored and created and maintained complexity should be hidden from database users.

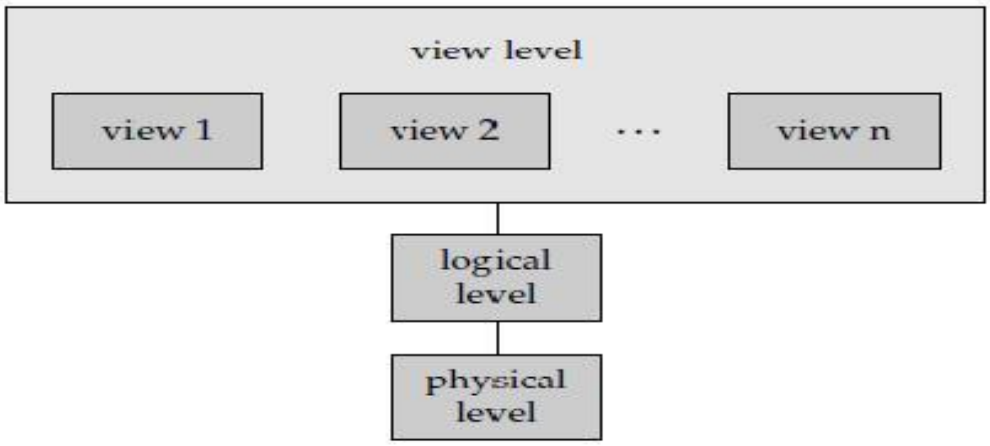


Fig.:- Three levels of data abstraction

1. Physical Level:

- **How the data are stored.**
- Lowest level of abstraction.
- Complex low-level structures described in detail.
- E.g. index, B-tree, hashing.

2. Logical /Conceptual Level:

- Next highest level of abstraction.
- **Describes *what* data are stored.**
- Describes what relationships among data.
- Database administrator level.

3. View Level:

- Highest level.
- **Describes *part* of the database for a particular group of users.**
- Complexity of physical as well as logical level is hidden from the user.
- E.g. tellers in a bank get a view of customer accounts, but not of payroll data.

1.4.2 Database Languages

- **DDL – Data Definition Language**
- **DML- Data Manipulation Language**
- **DCL – Data Control Language**
- **TCL-Transaction Control Language**

1. <u>DDL – Data Definition Language:-</u>

- DDL statements are used to define the database structure or schema.(Schema means overall design of the database)
- For fast retrieval the data itself should be stored in a well-structured manner, so that it easy to find it out from large storage.
- Data Definition Language (DDL) is a standard for commands that define the different structures in a database.
- For describing data and data structures a suitable description tool, a data definition language (DDL), is needed.
- **DDL Commands**
  1. **CREATE**: - To create objects in the database.
  2. **ALTER: -** Alter the structure of the database.
  3. **DROP: -** Delete objects from the database.
  4. **TRUNCATE: -** Remove all records from a table, including all spaces allocated for the records are removed.
  5. **RENAME: -** Rename an object.

2. <u>DML- Data Manipulation Language:-</u>

- Data Manipulation is_
  1. To retrieve the information from the database.
  2. To insert information to the database.
  3. To Delete information from the database.
  4. To modify information from the database.
- Additionally a language for the descriptions of the operations with data like store, search, read, change, etc. the so-called data manipulation, is needed.
- Such operations can be done with a data manipulation language (DML). Within such languages keywords like insert, modify, update, delete, select, etc. are used.

- **There are two types of DML**
  1. Procedural DML: - The users specifies both, what data is needed & how to get the data.
  2. Non Procedural DML: - In this type only a user to specify what data are needed without specifying how to get those data.

- **DML Commands**
  1. **SELECT** - Retrieve data from the a database
  2. **INSERT** - Insert data into a table
  3. **UPDATE** - Updates existing data within a table
  4. **DELETE** - Deletes all records from a table, the space for the records remain

3. <u>DCL – Data Control Language:-</u>

- Data security is an important aspect of any DBMS.
- SQL provides means for protecting data from unauthorized access.
- Used to control access to data stored in a database. In particular, it is a component of Structured Query Language (SQL).
- **DCL commands include**
  1. **GRANT** - gives user's access privileges to database
  2. **REVOKE** - withdraw access privileges given with the GRANT command

4. <u>TCL:-Transaction Control Language:-</u>

- Transaction Control (TCL) statements are used to manage the changes made by DML statements.
- It allows statements to be grouped together into logical transactions.
- **TCL Commands:-**
  1. **COMMIT** - save work done
  2. **SAVEPOINT** - identify a point in a transaction to which you can later roll back
  3. **ROLLBACK** - restore database to original since the last COMMIT.

1.4.3 Instance & Schema (IMP)

<u>Instance :-</u> (Definition)

- The data in the database at a particular moment of time is called an "instance" or a database state.
- Every time we update (i.e. insert, delete or modify) the value of a data item in record, that time one state of the database changes into another state.

<u>Schema:-</u>

- **The overall design of the database is called the database schema.**
- A database schema is the skeleton structure that represents the logical view of the entire database.
- It defines how the data is organized and how the relations among them are associated. It formulates all the constraints that are to be applied on the data.
- A database schema defines its entities and the relationship among them.
- **Types of Schema:-**

1. **Physical Database Schema** − It describes the database design at the physical level.
   - This schema pertains to the actual storage of data and its form of storage like files, indices, etc.
   - It defines how the data will be stored in a secondary storage.
2. **Logical Database Schema** −It describes the database design at the logical level.
   - What data are stored in the database & what relationship exists among those data?
   - It defines tables, views, and integrity constraints.
3. **View or Subschema –**
   - A database may also have several schema at the view level sometime called **"Subschema".**
   - It defines the portion of the database as seen by application programs & application programs can have different view of data stored in the database.

1.4.4 Data Independence (IMP)

Q. What are data independence? What are its types? (Definition of data independence - 1 Mark, list of types - 1 Mark, definition of each type - 1Mark each)

Ans: - Data Independence:

- It is the ability to modify a schema in one level without affecting schema in another level.
- When we change data at one layer, it does not affect the data at another level. This data is independent but mapped to each other.

There are two types of data independence:

1) Logical data independence
2) Physical data independence.

1. Logical data independence:

- It refers to the ability to modify the Logical or Conceptual schema without causing any changes in schema followed at view levels.
- For example, a table (relation) stored in the database and all its constraints, applied on that relation.

2. Physical data independence:

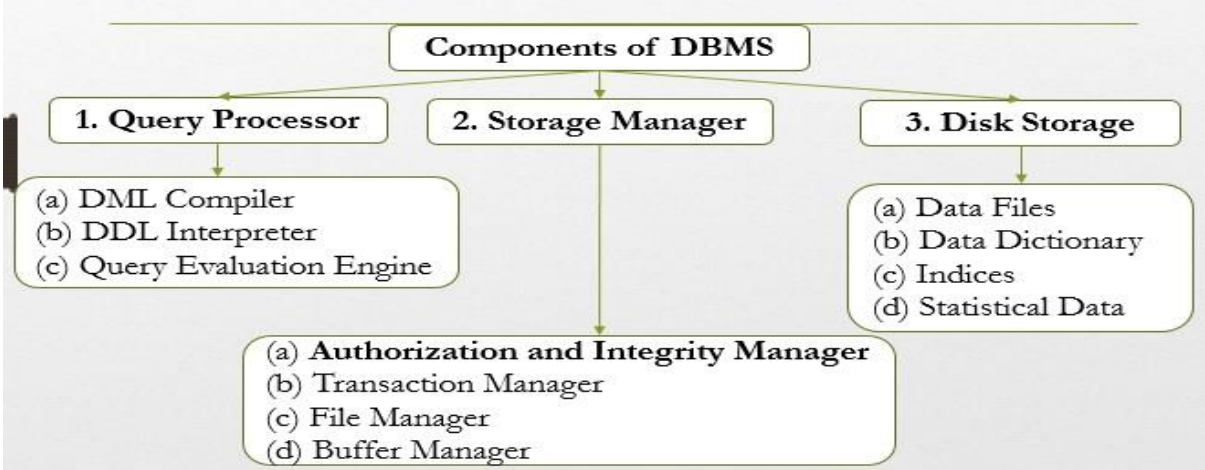
- It refers to the ability to modify the schema followed at the physical level without affecting the schema followed at the logical level/conceptual level.
- For example, in case we want to change or upgrade the storage system itself − suppose we want to replace hard-disks − it should not have any impact on the logical data or schemas.

1.5 Components of DBMS & Overall Structure of DBMS (IMP)

Q. Draw a neat labeled diagram of overall DBMS structure. OR

Q. Explain the overall structure of DBMS with suitable diagram.

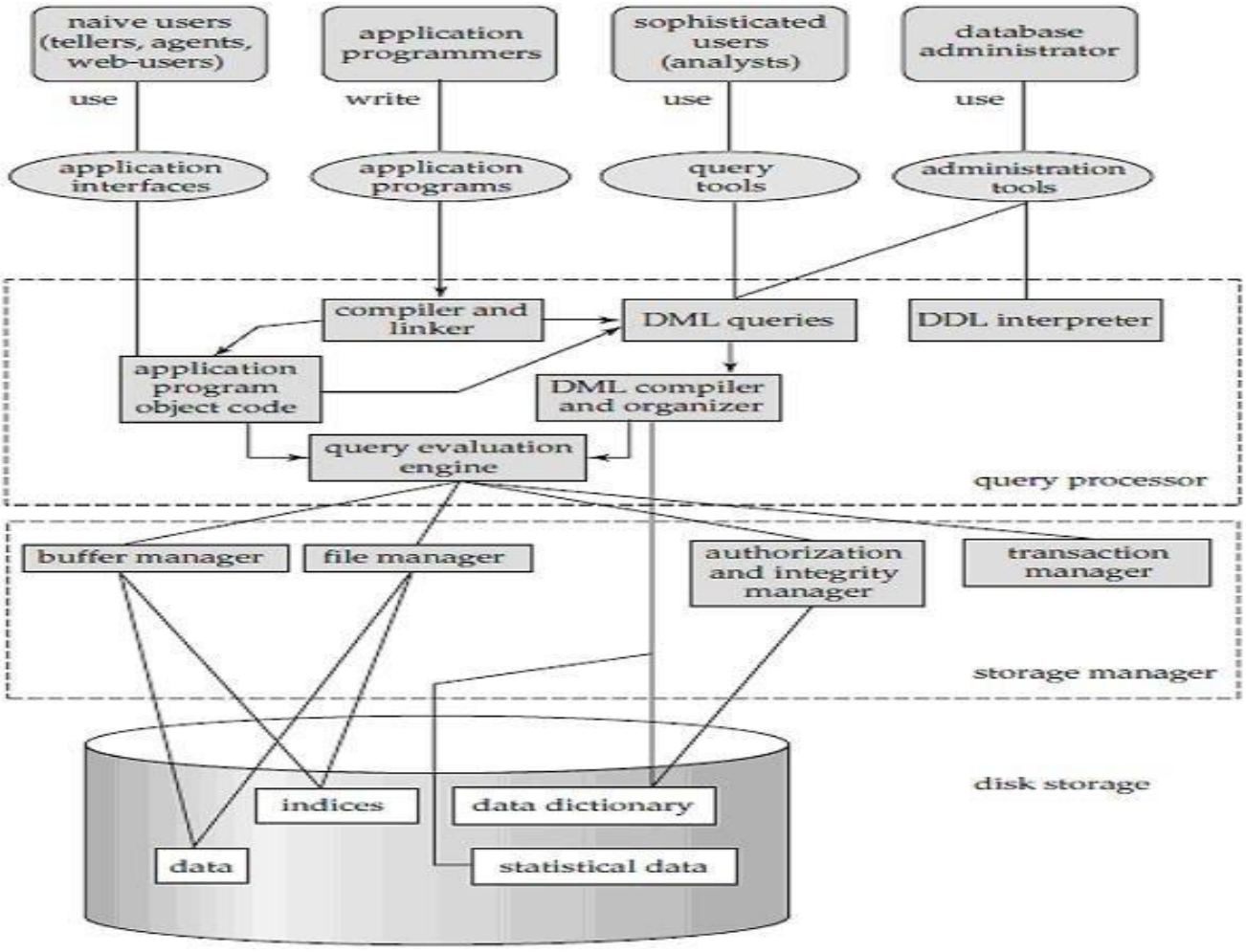
- A database system is partitioned into modules that deal with each of the responsibilities of the overall system.
- The functional components of a database system can be broadly divided into the **Storage manager, Query processor & Disk Storage.**



1. **Query Processor:-**

   a). DDL Interpreter: DDL interpreter converts DDL statements into a set of tables containing metadata or data dictionary.
   b). DML Compiler: It translates DML statements in a query language into low level instructions that query evaluation engine understands.
   c). Query Evaluation Engine: It executes low-level instructions generated by the DML compiler.

2. **Storage Manager:-**
   They provide the interface between the low-level data stored in the database and application programs and queries submitted to the system.
   a). Authorization and Integrity Manager: It tests for the satisfaction of integrity constraints checks the authority of users to access data.
   b). Transaction Manager: It ensures that the database remains in a consistent (correct) state.
   c). File Manager: it manages the allocation of space on disk storage & the data structure used to represent information stored on disk.
   d). Buffer Manager: It is responsible for fetching data from disk storage into main memory and deciding what data to cache in memory.



3. **Data Structures:-**

   a). Data Files: It stores the database.

   b). Data Dictionary: It stores Meta data (data about data) about the structure of the database.

   c). Indices: Provide fast access to data items that hold particular values.

   d). Statistical Data: It stores statistical information about the data in the database. This information is used by query processor to select efficient ways to execute query.

1.5.1 Database Users (IMP)

- There are different types of database users:-

1. **Naive Users: -**

   - Naive means Lacking Experience (untrain) these are the users who need not be aware of the presence of the Data Base System.
   - Example of these type of users is the user of an ATM machine. Because these users only responds to the instructions displayed on the screen (enter your pin number, click here, enter the required money etc.). Obviously operations performed by these users are very limited.

2. **Application Programmers: -**
   - Professional / Application programmers are those who are responsible for developing application programs or user interface. The application programs could be written in a general-purpose programming language or the commands available to manipulate a database.
3. **Sophisticated Users:-**
   - Simply we can say that these are the EXPERIENCED users. These people interact with the system without writing programs. They form requests by writing queries in database query language.
   - Sophisticated users are the users who are familiar with the structure of database & facilities of DBMS.
4. **Specialized Users:-**
   - These are the sophisticated users who write specialized database applications.
   - Among these applications are computer aided design (CAD) systems, knowledge-based and expert systems.
5. **Database Administrator:-**
   - It is responsible for managing the whole database system.
   - He/She designs, creates & maintains the database.
   - He/She manages the users who can access this db & controls integrity issues.
6. **Database Designers:-**
   - Database designers are the computer professionals which responsible for developing db oriented application /software's.

1.5.2 Functions of Database Administrator (DBA) (IMP)

Q. List four function of database administrator. ***(List any four function - 1 Mark each)***

- The DBA (Database Administrator) is a person or group of persons who is responsible for the management of the database.
- The person having central control over the system is called DBA. The functions of DBA are listed below:-

  1. Schema Definition: - The Database Administrator creates the database schema by executing DDL statements. Schema includes the logical structure of database table (Relation) like data types of attributes, length of attributes, integrity constraints etc.

  2. Storage structure and access method definition: - Database tables or indexes are stored in the following ways: Flat files, Heaps, B+ Tree etc.

  3. Schema and physical organization modification: - The DBA carries out changes to the existing schema and physical organization.

  4. Granting authorization for data access:-
  - The DBA provides different access rights to the users according to their level. Ordinary users might have highly restricted access to data.

- DBA is responsible for granting the access to the database.

5. Integrity- Constraint specification: Integrity constraints are written by DBA and they are stored in a special file, which is accessed by database manager, while updating the data.

6. Routine Maintenance: - Some of the routine maintenance activities of a DBA is given below.

a) Taking backup of database periodically
b) Ensuring enough disk space is available all the time.
c) Monitoring jobs running on the database.
d) Ensure that performance is not degraded by some expensive task submitted by some users.
e) Performance Tuning

1.6 Introduction to Client Server Architecture (IMP)

Q. Explain client-server architecture.

- In client-server architectures the application program functions are divided up between clients and servers.
- **Client: - PCs or workstations on which users run applications.**
- **Server: - it is powerful computers dedicated to managing printers or network traffic.**
- Application logic may reside on the client, server or be split up between the two.
- Most networks today use a client-server architecture.
- Client-Server architecture consists of front end & back end. Front end is the GUI & back end is where the db is stored.
- In client-server architectures, remote procedure calls (RPCs) or standard query language (SQL) statements are typically used to communicate between the client & server.
- Example of front End is Visual Basic, VC++ etc. & back end is SQL, Ms-Access, SQL server etc.
- Computer networking allows some task to be executed on a server system and some tasks on client system.
- There are different types of client/server architecture such as two tier, three tier architecture.

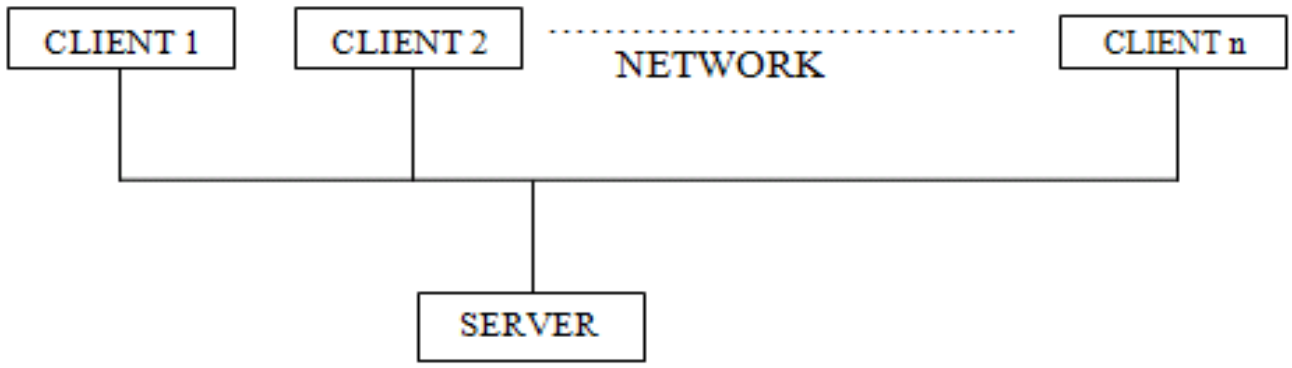
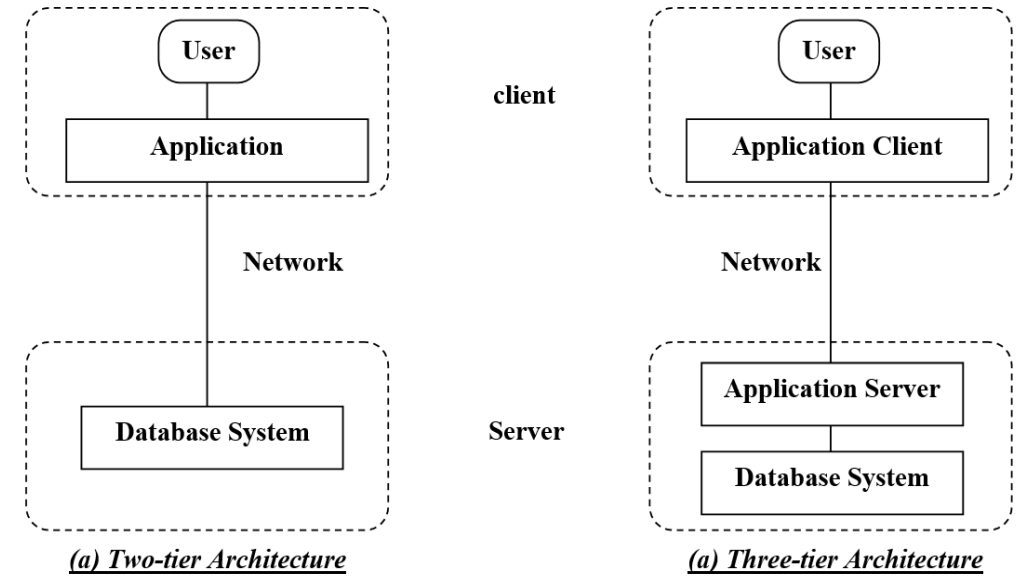


Fig: - Client/Server Architecture

Two & Three Tier Architecture.



(a) Two-tier Architecture *(a) Three-tier Architecture*

1.6.1 Two-Tier Architecture:-

- The two-tier is based on Client Server architecture.
- The two-tier architecture is like client server application.
- The direct communication takes place between client and server.
- There is no intermediate between client and server. Because of tight coupling a 2 tiered application will run faster.
- The above figure shows the architecture of two-tier. Here the communication is one to one.
- Let us see the concept of two tier with real time application.
- For example now we have a need to save the employee details in database. The two tiers of two-tier architecture is.
- **Advantages: -** Understanding & maintenance is easier.
- **Disadvantages: -** Performance will be reduced when there are more users.

1.6.2 Three-Tier Architecture:-

- In three tier architecture the communication taken place from client to application server and then application server to database system to access the data.
- The application server or web server is sometimes called middle layer or intermediate layer.
- The middle layer which processes applications and database server processes the queries.
- This type of communication system is used in the large applications or the world web applications. On WWW all clients requests for data and server serves it.
- There are multiple servers used like fax server, proxy server, mail server etc.

Advantages

- Easy to modify without affecting other modules
- Fast communication
- Performance will be good in three tier architecture.

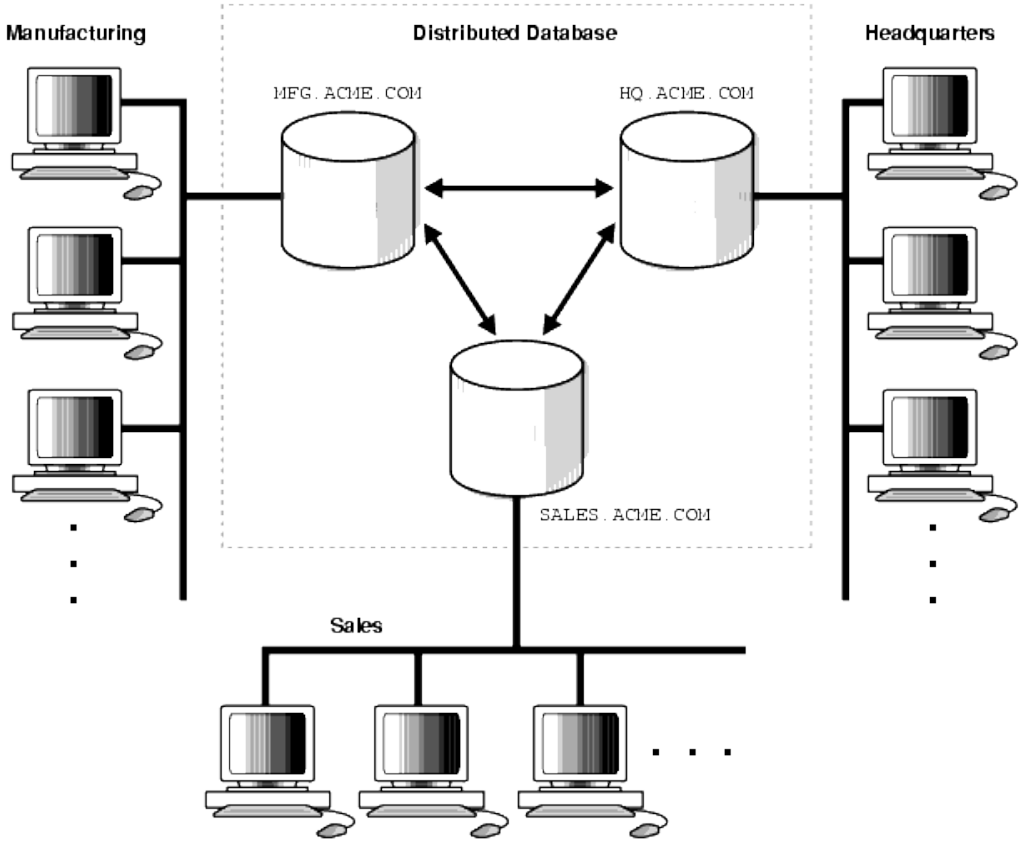
1.7 The 12 Rules (Codd's laws) for fully functional RDBMS. (IMP)

- **Rule 1: Information Rule: -** The data stored in a database, may it be user data or metadata, must be a value of some table cell. Everything in a database must be stored in a table format.
- **Rule 2: Guaranteed Access Rule: -** Every single data element (value) is guaranteed to be accessible logically with a combination of table-name, primary-key (row value), and attribute-name (column value). No other means, such as pointers, can be used to access data.
- **Rule 3: Systematic Treatment of NULL Values:-**The NULL values in a database must be given a systematic and uniform treatment. This is a very important rule because a NULL can be interpreted as one the following − data is missing, data is not known, or data is not applicable.
- **Rule 4: Active Online catalog based on the relational Model: -** The structure description of the entire database must be stored in an online catalog, known as data dictionary, which can be accessed by authorized users.
- **Rule 5: Comprehensive Data Sub-Language Rule: -** A database can only be accessed using a language having linear syntax that supports data definition, data manipulation, and transaction management operations. This language can be used directly or by means of some application. If the database allows access to data without any help of this language, then it is considered as a violation.
- **Rule 6: View Updating Rule: -** All the views of a database, which can theoretically be updated, must also be updatable by the system.
- **Rule 7: High-Level Insert, Update, and Delete Rule: -** A database must support high-level insertion, updation, and deletion.
- **Rule 8: Physical Data Independence: -** The data stored in a database must be independent of the applications that access the database. Any change in the physical structure of a database must not have any impact on how the data is being accessed by external applications.
- **Rule 9: Logical Data Independence:-**The logical data in a database must be independent of its user's view (application). Any change in logical data must not affect the applications using it.
- **Rule 10: Integrity Independence: -** A database must be independent of the application that uses it. All its integrity constraints can be independently modified without the need of any change in the application.
- **Rule 11: Distribution Independence: -** The end-user must not be able to see that the data is distributed over various locations. Users should always get the impression that the data is located at one site only.

- **Rule 12: Non-Subversion Rule: -** If the RDBMS has a language that accesses the information of a record at a time, this language should not be used to bypass the security and integrity constraints.

1.8 Introduction to Distributed database (IMP)

- A distributed database system allows applications to access data from local and remote databases.
- It is database in which portions of the database are stored on multiple computer within a network.
- Distributed database is a collection of multiple interconnected databases, which are spread physically across various locations that communicate via a computer network.
- In a **homogenous distributed system,** each database is an Oracle database. In **a heterogeneous distributed system,** at least one of the databases is a non-Oracle database. Distributed database uses a client-server architecture to process information requests.



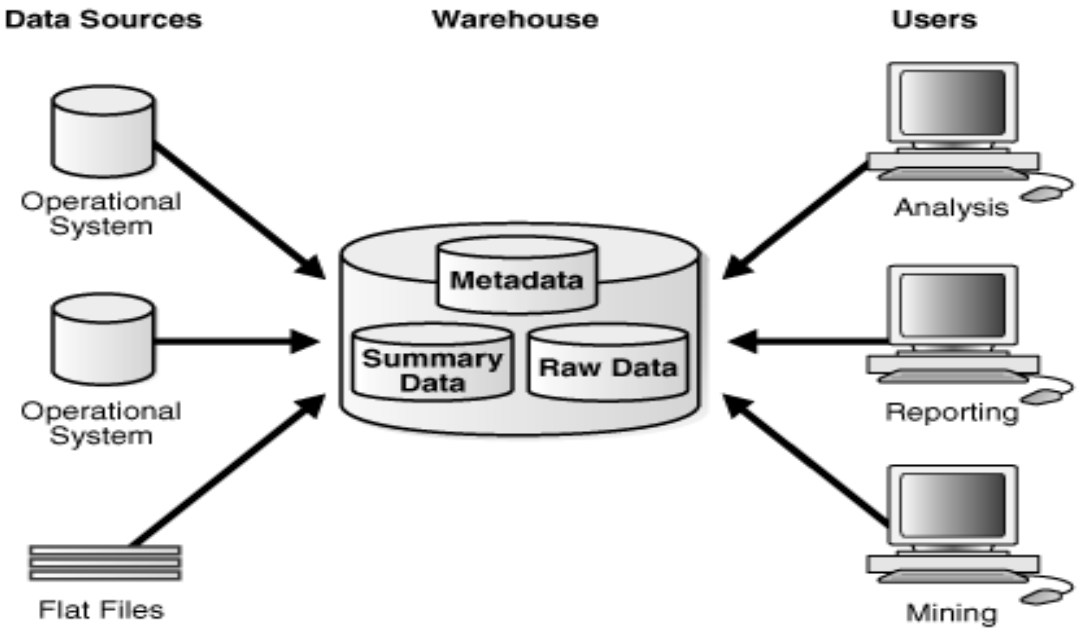
Distributed Database with example.

- Using distributed database technology, a bank may implement their database system on a number of separate computer systems rather than a single, centralized mainframe.
- The computer systems may be located at each local branch office: for example, Mumbai, Pune, and Nagpur. A network linking the computer will enable the branches to communicate with each other, and DDBMS will enable them to access data stored at another branch office.
- Thus a client living in Pune can also check his/her account during the stay in Mumbai or Nagpur

1.9 Introduction to Data Warehousing (IMP)

Q. Explain data warehousing and data mining.

- Data warehouse is storage architecture designed to hold data extracted from transaction systems, operational data stores & external sources.
- A data warehouse is a repository of information gathered from multiple sources, stored under a unified schema, at a single site. Once gathered, data are stored for long time, permitting access to historical data.
- Data warehouses provide the user a single consolidated interface to data, making decision-support queries easier to write.
- Moreover, by accessing information for decision support from a data warehouse, the decision makers ensures that online transaction-processing systems are not affected by decision support workload



- The process of extracting data from source systems & bringing it into the data warehouse is commonly called **ETL, which stands for Extraction, Transformation, & Loading.**

1. **Extraction of data: -** During extraction, the desired data is identified & extracted from many different sources, including database systems & applications.

2. **Transformation of data:-**After data is extracted, it has to be physically transported to the target system or to an intermediate system for further processing.

3. **Loading: -** When data is transported to target system it is loaded to the final destination so that users can query on the database.

1.10 Introduction to Data Mining (IMP)

- **Data Mining**: Data mining is the exploration and analysis of large quantities of data in order to discover valid, novel, potentially useful and ultimately understandable patterns in data.
- It is known as **"Knowledge Discovery in Databases".** When the data is stored in large quantities in data warehouse, it is necessary to dig the data from the warehouse that is useful and required for further use.
- For data mining, different software tools are used to analyze, filter and transfer the data from the data warehouses.
- **Feature of data mining:-**

1. Prediction
2. Identification
3. Classification
4. Optimization

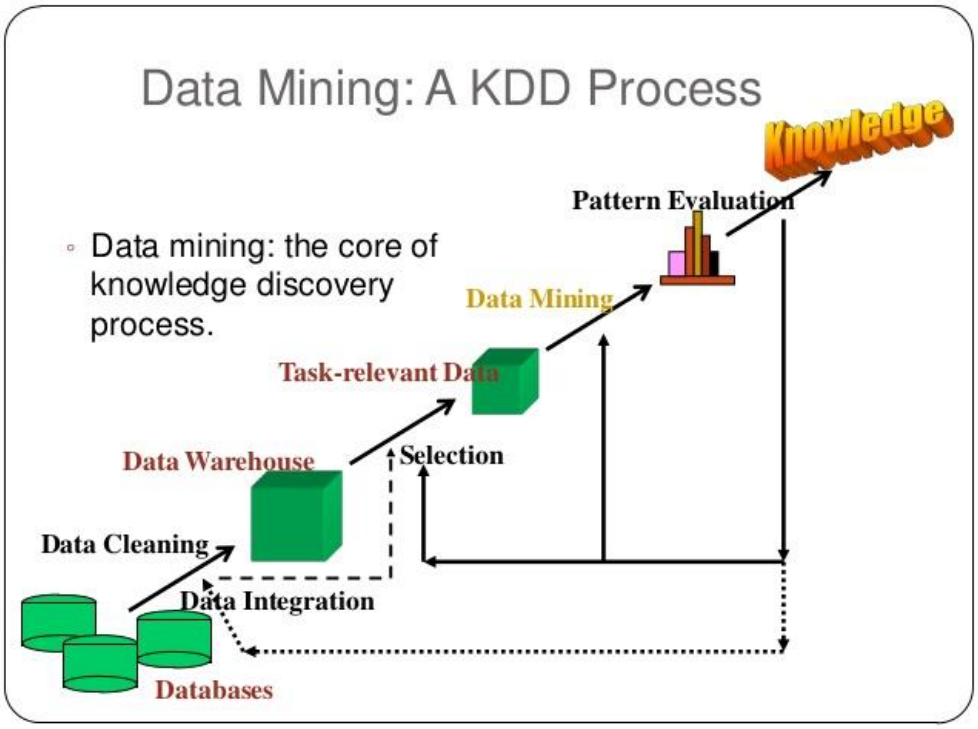


Fig.:- Steps of a KDD Process

Steps of a KDD Process

1. Learning the application domain: Relevant prior knowledge and goals of application.
2. Creating a target data set: Data Selection
3. Data cleaning and preprocessing.
4. Data reduction and transformation: Find useful features, dimensionality/variable reduction, and invariant representation.
5. Choosing functions of data mining: Summarization, Classification, Regression, Association, Clustering.
6. Choosing the mining algorithm.
7. Data mining: Search for patterns of interest.
8. Pattern evaluation and knowledge presentation: visualization, transformation, removing redundant patterns etc.
9. Use of discovered knowledge.

IMPORTANT QUESTIONS:-

SUMMER-2016 (18 Marks)

1. Enlist different components of DBMS. *(Listing any 2 components - 1 mark each) S-16.*
2. What is meant by data redundancy? *(Correct Definition - 2 marks)*
3. List any four DBMS software. *(Any four software - ½ mark each)*
4. List and explain any four functions of Database Administrator. *(Any four functions - 1 mark each)*
5. What is data warehousing and data mining? *(Data Warehousing - 2 marks; Data Mining - 2 marks)*
6. Draw and explain client server architecture. *(Diagram - 2 marks; Description - 2 marks)*

WINTER-2016 (26 Marks)

1. List any four advantages of DBMS. *(Any 4 Advantages: 1/2 mark each)*
2. List four DDL commands. *(Any 4 DDL commands: ½ mark each)*
3. List DCL commands any four. *(Any 4 DDL commands: ½ mark each)*
4. Explain Distributed Database with example.*(Definition of Distributed database:3 marks, example:1 mark (Any valid Example))*
5. Explain three levels of data abstraction with suitable diagram. *(Diagram – 2 mark; Description -2marks)*
6. Explain client-server architecture.
7. Explain data warehousing and data mining.
8. Draw overall structure of DBMS.

WINTER-2016 (28 Marks)

1. List any four advantages of DBMS. *(Any Four advantages – ½ Mark)*
2. What is data redundancy? *(Data Redundancy – 2 Marks)*
3. Explain three tier architecture with suitable diagram.*(Diagram – 2 Marks, Description – 2 Marks)*
4. Compare DBMS and RDBMS (any four points). *(Any four points - 1 Mark each)*
5. What are data independence? What are its types?*(Definition of data independence - 1 Mark, list of types - 1 Mark, definition of each type - 1Mark each)*
6. Explain the overall structure of DBMS with suitable diagram.*(Explanation- 2 Marks, Diagram – 2 Marks)*
7. Explain different applications of DBMS. *(Any four relevant applications – 1 Mark each)*
8. What is DBMS? Explain its functions.*(DBMS definition – 1 Mark, explanation of any three function – 1 Mark each (Any other functions of DBMS is considered))*