MAA OMWATI DEGREE COLLEGE HASSANPUR (PALWAL)

Notes

BBA 4th Sem

Data Base Management System

Unit - 1

* **Data :→** Data is defined as a collection of meaningfull facts which can be stored and processed by a computer or human. It is basically a collection of unorganised facts, concepts or instructions in a formalized manner suitable for communication, interpretation, processing by humans or some automatic means such as computer ATM's

Example of data are phone numbers, prices, costs, product name, addresses, imeges pictures, audio & video segments etc.

* **Information :→** Info. is the processed data that is meaningful and upon which some decision can be taken. It is the intelligent form of data. For example, paychecks, timetable, bills, profits, reports, receipts, comparision of sales figures, merit list, printed documents etc. The information is obtained by arranging data into meaningfull form. For example, the marks obtained by the students and their roll numbers form the data and their report card is the information which helps us to make decision that which student is pass and which is failed in the examination.

* **Diff. bet. Data & Information :→**

Data is a collection of raw facts, figures & other material on which computer programs work upon. It can be words, alphabets, numbers or some special symbols. They individually have no meaning of their own.

Data when organised can be used to represent usefull information. This information helps people in making decisions.

A set of words would be data when these are arranged in a proper manner to form some lines of text, it then represent information.

The processed information can be further used as input to some other sources. For example, the merit list of student of a particular class which is prepared from the output received form the individual marks of single students.

* **Qualities of Info.:→**

1) It must be meaningful.
2) It must be accurate and clean.
3) It should convey to the point and relevant information.
4) It should confirm and correct the previous knowledge.
5) It should reach at proper time so that proper action can be taken with it.

* Record: → A group of interselated field is called a record. A record consist of values for each field.

A meaningful group of characters is referred to as field. It is the smallest unit of data that has meaning to its user It is also called data item or data element.

Each record must contain atleast one fields that uniquely identifies it. The records are of two types:-
(i) Fixed length Records (ii) Variable length Records.

© Fixed Length Records — Each and every record, in fixed length record has exactly the same size (in bytes). These are simpler and easy to implement but it does have some problems associated with it.

Advantage — 1. These are simpler and easy to implement.
2. It is easy to insert record if a list of deleted records be kept.

Disadvantage- 1. Deletion is difficult.

© Variable Length Record — Each & Every record, in variable length records is not of same size i.e. the record size may very depending on the requirement. A record may be variable length record due to variable length fields, repeating fields or optional fields.

Advantage of variable length record :-
1) Size of record may be increased as per requirement
2) Optional use of memory.

Disadvantage of variable length record:-
1) Insertion of record is difficult.
2) Complex & difficult to implement.

24/7/11

* Files : →

A file is collection of related records. It can also be defined as collection of records of the same type grouped together

* Data Dictionary : →

A data dictionary keep track of data i.e. data about such as relation names, attribute names and types, storage information, integrity constraint and user information. It is also called system catalog and is an integral part of the database management system. Data dictionary helps the database administrator in the management of database. A data dictionary stores and manage the following info. 1) Description of database schema.
2) Description of database user.
3) Detailed info. on physical database design, such as storage structures, access path and file and record size.
4) Description of database transaction and data item referenced by them.
5) Statistical and ~~description~~ descriptive data.

relationship set.

Data Dictionary is a usw module of a DBMS which provides the definition of data item that they relate to other entities in the databases.

* Types of Data Dictionaries :→

(i) Active Data Dictionary — An active data dictionary is managed by the database management system automatically.

(ii) Passive Data Dictionary — These types of dictionaries are used for documentation purpose only.

* Components of Data Dictionaries:→

(i) Entities — An entity is an item about which the information is stored. A collection of entities of same type is called an entity set e.g. A row in a student file is an entity and whole student file is an entity set.

(ii) Attributes — The property of an entity is called an attribute. In the student file, Roll No, class and sections are all attributes of a student.

(iii) Relationships — The various ways in which the different entities relate to each other is called relationship. There are three type — One to one relationship ; one to many relationship, Many to many

(iv) Key — A key is a single attribute or combination of attribute of a table that is used to uniquely identify the various record of a table. (23/7/11)

- Primary Key:— It is used to uniquely identify a record.
- Concatenated Key:- When a combination of more than one field is used to identify a record, it is called a concatenated key.
- Secondary Key:— It is used to identify all those records that have a certain property.
- Super Key:- It contains any no. of attributes to uniquely identify a record.

* Database Administrator (DBA) :→

DBA is a person or group of persons who is responsible for management of database. DBA is a highly skilled person with strong technical background to monitor various operations such as creating, modifying and maintaining the database. DBA has most of the power such as defining schemes, storage structure and access method strategies, physical organisation, authorization and integrity constraint etc. So the DBA has all the power that system can give on all database objects. DBA are top level authority among all persons connected to the database.

(25/7/11)

* Duties or functions or Responsibilities of DBA :→

(i) Deciding the info. content of the database — It is job of the DBA to identify what kind of information is to be stored in the database.

(ii) Define the storage structure and access strategies — It is the DBA's job to decide the representation of data in the database and for this purpose he writes the storage structure definition which is translated by the Data Definition Language Compiler. He also specifies the mapping bet. physical schema (storage structure definition) and conceptual schema.

(iii) Define external schemes for the users — It is the job of DBA to write necessary external schemas and ensures that required data should be made available in the required format. He also specifies the mapping bet. the external schemas and conceptual schemas

(iv) Defines Integrity constraints — DBA is responsible for accuracy of data in the database by specifying integrity constraints and checks.

(v) Granting of authorization for data access — It is the job of DBA to decide which part of the database should be accessible by which types of users.

(vi) Define strategy for backup and recovery — It is the responsibility of DBA to recover the data which is lost due to the failure in the h/w or operating system or due to human error. DBA also provides backup of data and instructions so that they can be recalled when they are lost.

(vii) Monitoring and optimizing the performance of the database.

(viii) Plans the future storage requirements for the database system

(ix) Installing and upgrading the DBMS and application tools.

(26/7/11)

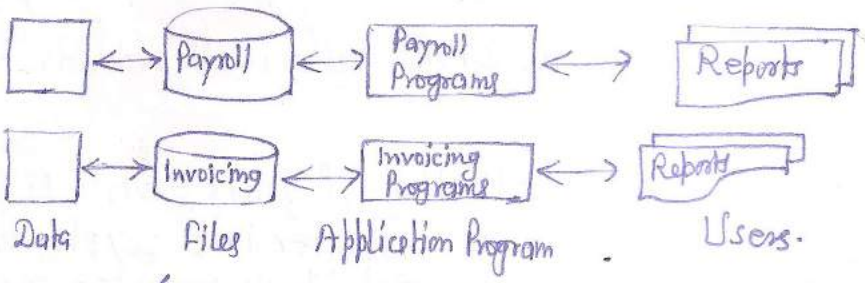
* Traditional/conventional file-based system Vs Database Approach. :→

With the growth of database and advancement in the technology related to computer, there was need to computerize the manual way of storing the data in the database and the approach used for this purpose used for this purpose was the File System Approach.

Limitations of the File System Approach :→

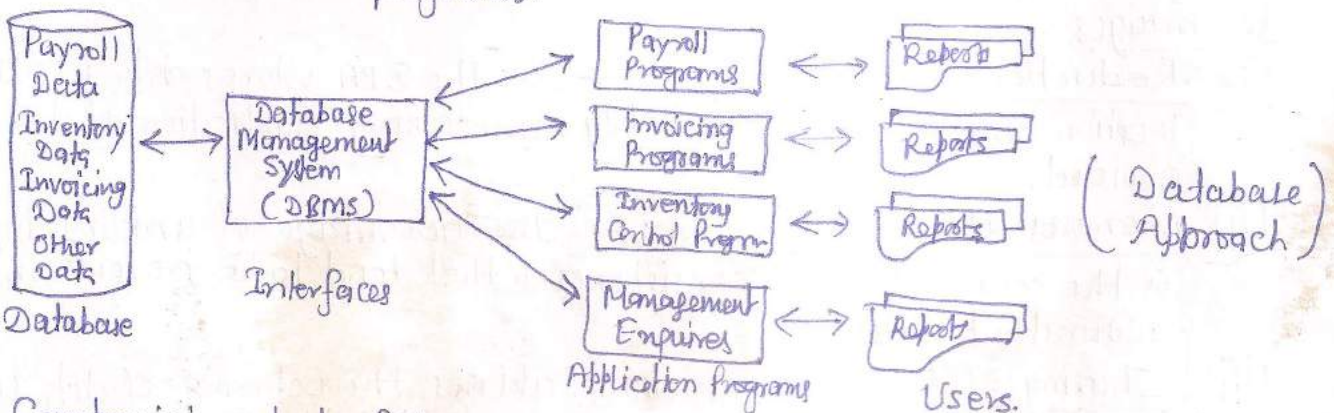
(i) Data redundancy and Inconsistency.
(ii) Data Independence.
(iii) Difficulty to get quick answers to simple queries.
(iv) Lack of Data Integrity and Quality
(v) Difficult to share data
(vi) Transactional Problem
(vii) Lack of Security
(viii) Concurrency problems

D.4



(File System Approach)

In order to overcome the problem of a conventional file-oriented data processing system. The concept of database was introduced. To use the database approach a database management system (DBMS) is required. A DBMS consist of program that can be used as an interface bet. the database and the application programs.



* Comparision bet. File Vs Database Approach

File System Approach	Database Approach
(i) Each application has its own private data files	(i) A database system has a collection of interrelated files and set of application program to access and modify these files.
(ii) It is used in small systems	(ii) It is used in large systems.
(iii) Cheaper	(iii) Costly.
(iv) Data dependency.	(iv) Data Independency.
(v) Data inconsistency.	(v) Data consistency.
(vi) Data Redundancy.	(vi) Controlled data redundancy.
(vii) Supports only fixed queries.	(vii) Support fixed & unforeseen queries.
(viii) User Interface only provided by application program	(viii) User Interface provided by application program, SQL and menu & wizards.
(ix) Cannot recover data from failures	(ix) Recover data from h/w or s/w failures
(x) Multiple users cannot access information.	(x) Allow multiple users to access the database.
(xi) Its examples are Cobol, C++ etc.	(xi) Its examples Oracle, SQL Server, Sybase etc.

Q. What is database ? what is DBMS also write down the advantages of DBMS ?

Ans. A database is an organised collection of facts (data). A database is a collection of records stored in a computer in a systematic way, so that a computer program can contact it to answer queries for better retrieval and sorting. Each record is usually organised as a set of data elements.

DBMS is a s/w that facilitates the user to design access and manipulate database. A DBMS is able to manage large, shared, collection of data while ensuring reliability and privacy. e.g. MS-Access, MS-SQL Server, MySQL and DB2.

Advantages: →

(i) Reduction of Redundancies: — As the DBA stores data in a central location and control it centrally, unnecessary duplication of data is avoided.

(ii) Removes the Inconsistency: — The advantage of avoiding duplication is the removal of inconsistencies that tend to be present in redundent files.

(iii) Sharing of Data: — A DBMS allows the sharing of data under its control to any number of users or application programs

(iv) Data Integrity: — Data integrity means that data stored in a database is accurate and consistent. Therefore data values being entered for storage could be checked to ensure that they fall with in specified range and is in correct format

(v) Data Security: — A Data base management system ensures security so that the unauthorized users can not access such confidential data.

(vi) Conflict Resolution: —
Since the database is under the control of database administrator (DBA), the DBA can resolve conflicting requirements of various users and applications. समाधान परस्पर विरोधी

(vii) Provides Quick Information: —
DBMS provides tools to convert the data into useful information, which can be used for decision-making purpose. For example, the sale database contains the records for sales made by all the sales persons in different zones. By using tools / commands provided by DBMS, the total sales for each zone could be produced very quickly and easily.

D-6

Q. What are the facilities provided by DBMS ?

Ans.
- Creation of database files.
- Addition / insertion, deletion and modification of data in a databases.
- Retrieving data from the database.
- Generating report from the database.
- Performing mathematical calculations on data stored in the database.
- Maintaining the data integrity
- Handling all access to the database.
- Applying the authorization checks and validation procedures.

Q. What are various types of files in which the records are collected and maintained ?

Ans
- Master file :— In business application, these are considered to be very important because they contain the essential records for the maintenance of organisation's business.
- Transaction file: — It is a temp file, used to collect data about the event as they occur. Second, it help in updating master files to reflect the result of current transaction.
- Table file — Table files are permanent files containing reference data used in processing transactions, updating master file or producing output.
- Report file — Report files are collected contents of individual output reports or documents produced by the system.
- Backup file — It is copy of master, transaction or table file that is made to ensure a copy is available if anything happens to original.
- Archival file — These files are copies of various files made for long term storage of data that may be required on a later time.
- Dump file — This is a copy of data stored in computer at a particular point of time. This may be a copy of master file to be retained to help recovery in the event of possible corruption of master file or it may be part of a program in which error is being traced.
- Library file — Library file generally contains applications programs, utility programs and system s/w packages.

DBMS 5

* Entity Vs Attributes :→

An Entity is a person, place, thing, event or concept about which information is recorded. In banking environment, examples of entities are Customers, Bank Accounts. In a warehouse the entities are suppliers, parts, etc.

Whereas

Attributes gives the characteristics of the entity. In other words, every entity has some basic attributes that characterize it e.g. (a) A house can be described by its size, colour, age & surroundings.

(b) A customer of bank may be described by such attributes as name, address & possible a customer identification number.

(c) A bank account can be represented by an account type, an account number & an account balance.

An attribute is often called a data element, a data field, a field, data item, or an elementary item.

D-8

Unit-2

* **Schema:→** The plan(scheme) of the database is known as schema. The schema gives the name of entities & attributes. It also specifies the relationship among them. The schema remains same, but the values stored into this changes from time to time.
Therefore, the schema includes the definition of database name, record type & components that makeup those records.

Types of Schema:→ There are two types of schema.

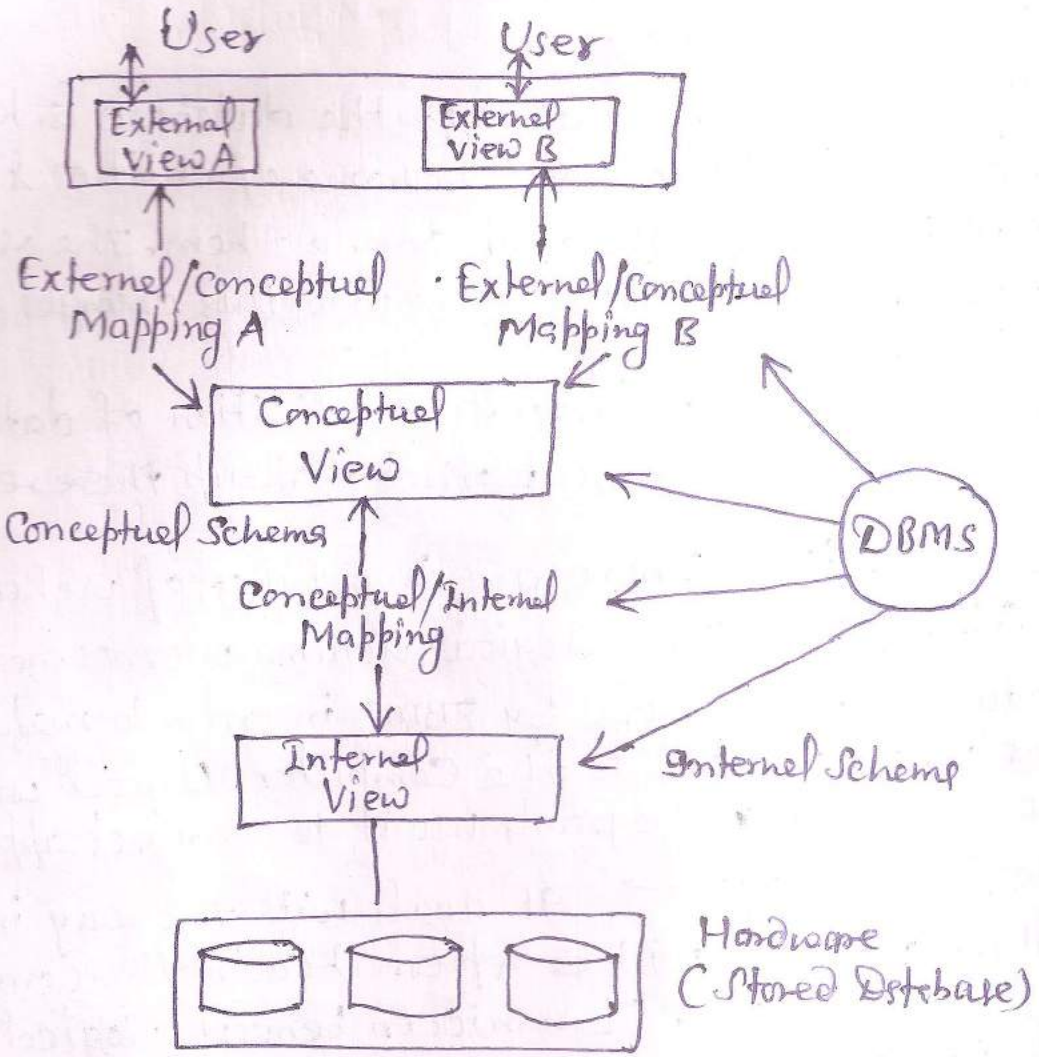
a) **Logical Schema:→** The logical schema is concerned with the data structure provided by DBMS in order to make the scheme understandable to the computer. Logical schema is more important as programs use it to construct applications.

b) **The Physical Schema:→** It deals with the way in which conceptual database will be represented in the computer as a stored database. It is hidden beneath logical schema. The physical scheme can be changed without affecting application program.

Subschema:→ It is subset of schema and inherits the same property that a schema has. The plan for view is generally called subschema.

* **Instances:→** The current state of database at any point of time, is called an instance. It is also called state of database.
The major difference bet. the database schema & database state is that database schema is specified to DBMS when new database is defined & at this point of time, database state is empty.

* **Three Level of database architecture:→**
The three level architecture was suggested by ANSI/SPARC. In this architecture, the database is considered as containg data about an orgn. The three levels of architecture are the three different views of data.



Three Level Architecture

(i) The External / User View — It is highest level of database architecture, where only the restricted portion of database is available to end users or application programmer. It is restricted view of database and same database may provide a no. of diff. view of for diff. classes of users.

(ii) Conceptual Level / View — It describes the structure of whole database for a community of users i.e. global view of data. It is represented by middle view/level in three level architecture. It is defined by conceptual ~~view~~ schema which includes definitions for each type of data. There is only single conceptual schema per database. It represents all the database entities, their relationship, constraints on data, security and integrity information. It is independent of both h/w & s/w

(iii) Internal / Physical view level : — It is the view about the actual physical storage of data. It tells how data is stored physically in the database.

The internal level is concerned with the following aspects:-
- Storage space allocation
- Access paths
- Data compression & encryption techniques
- Record placements etc.

* **Data Independance** :→ It is the characteristic of database system to change the schema at one level without having to change the schema at the next higher level. There are two types of data independence :—

- **Physical Data Independence** :→ It allow the changes in the physical storage devices or orgsnisation of files to be made without requiring changes in conceptual view or any of external views in the application program using database.
- **Logical Data Independance** :→ Logical data independe indicates that the conceptual schema can be changed without affecting the existing external schemas. It is more difficult to achive than physical data independance.

* **Mapping** :→ The process of transforming requests and results bet. the three levels are called mappings. DBMS is responsible for this mapping bet. internal, external and conceptual schemas. There are two types of Mapping :—

- **Conceptual / Internal Mapping** :→ It lies bet. the conceptual & internal levels and defines the correspondence bet. the records & fields of conceptual view & file and data structures of internal view. If the structure of stored database is changed, then the Conceptual / internal mapping must be changed acc. so that view from the conceptual level remains constt. This mapping provides physical data independance for the database.
- **External / Conceptual Mapping** :→ It lies bet. the external & conceptual levels and defines the correspondence bet. a particular external view & conceptual view. This mapping provides the logical data independance for the database.

* Data Models :—→

A data model is collection of concepts that can be used to describe the structure of the database including datatypes, relation ship and constraints that apply on data. A data model help in understanding the meaning of data and ensures that, we understand —

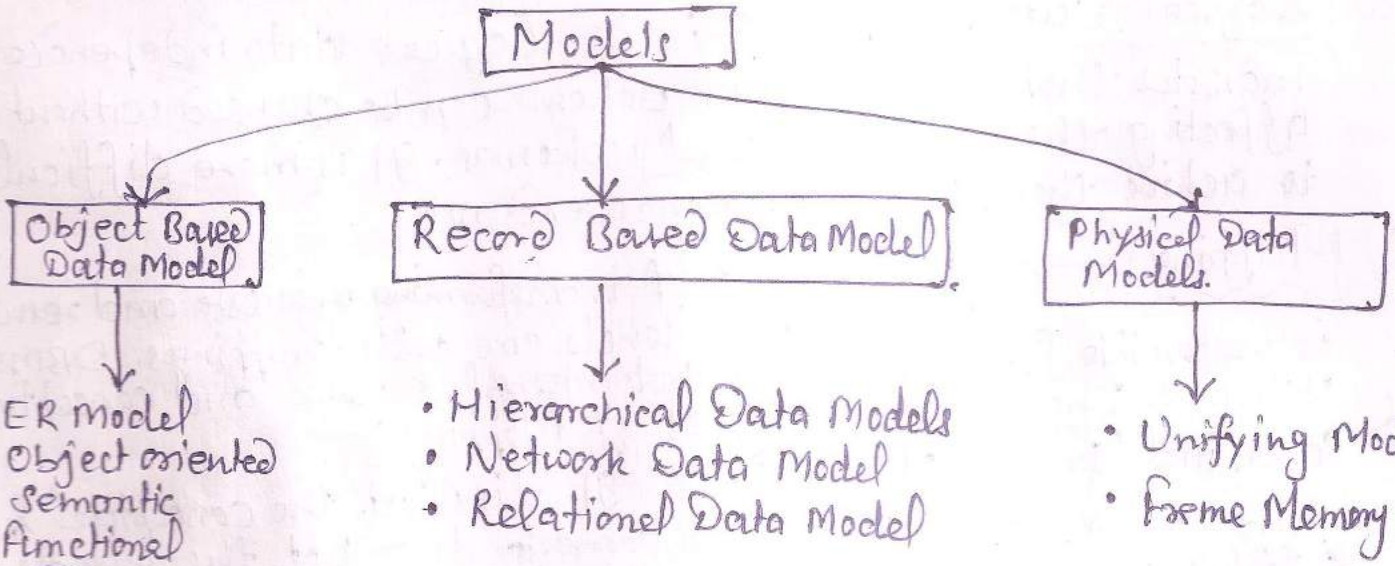
(i) The data requirement of each user.
(ii) The use of data across various applications
(iii) The nature of data independent of its physical repn.

A data model support communication bet. the users & database designers.

Characteristics of Data Models :—→

(i) Diagrammatic representation of data model
(ii) Simplicity in designing
(iii) Application independent
(iv) Data repn must be without duplication.
(v) Bottom-Up approach must be followed.

Types of Models :—→



Hierarchical Data Model :—→

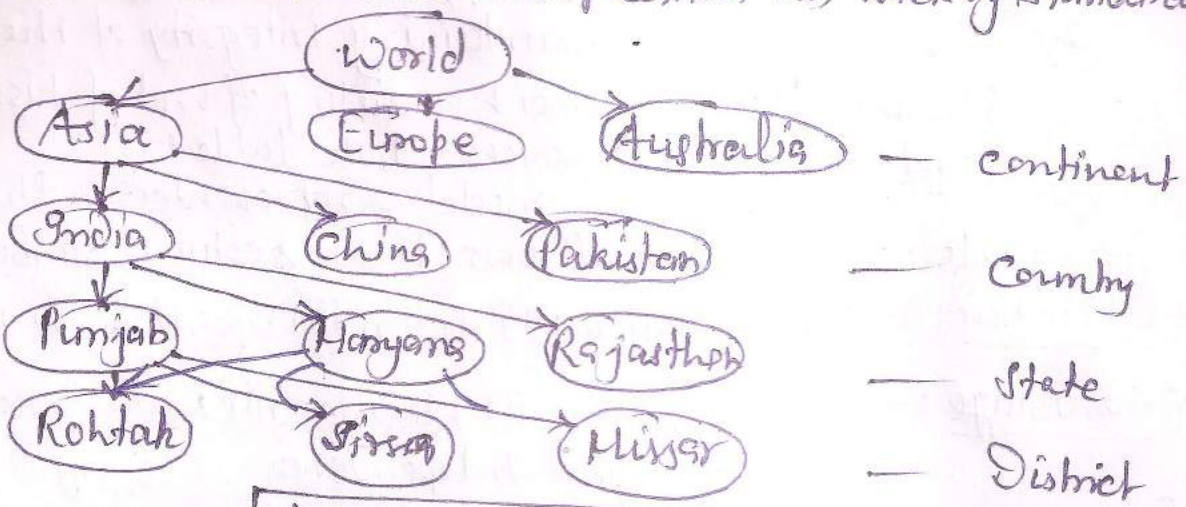
It is one of oldest models. It organizes records in a tree structure i.e hierarchy of parent & child record relationship. It employs two main concepts :— Record & Parent Child Relationship. A record is collection of field values that provide information of an entity.

A parent child Relationship type is a 1:N relationship bet. two records types. The record type of one side is called the parent record type and one on the N side is called the child record type. In term of tree data structure, a record type corresponds to node of tree and relationship type corresponds the edge of tree.

Advantage :— i) Simplicity ii) Integrity of data iii) Data Security iv) Efficiency (D-12)

Disadvantage —

i) Knowledge of physical level of data storage is required
ii) Complexity (iv) Lack of querying facilities
iii) Inflexibility (v) Database Management Problems.
vi) Problems with data manipulation vii) Lack of Stemdards.



A Hierarchical Model

Network Data Model :— This model overcome the limitation of hierarchical model. The ability of this model to handle many to many (N:N) relations bet its record is the main distinguishing feature from the hierarchical model. It permits a child record to have more then one parent. It uses the directed graphs instead of tree structure.

There are two basic data structures in this model— Record & Sets. The record containing the detailed information regarding the data which are classified into record types. A set type represents relationship bet record types and this model use linked list to represents there relationships. Each set type definition consist of three elements :— a name for set type, an owner record type (like parent) and a member record type (like child).

To represent Many to Many relationship in this model, the relationship is decomposed into two one to many (1:N) relationship by introducing an additional record type called an Intersection Record / Connection Record.

Operation :— 1) Insertion 2) Deletion 3) Updation 4.) Retrieval

Advantage :— 1. Eliminate Data Redendancy
2. Lesser storage requirement
3. Better Performance
4. Handle many types of Relationships.
5. Easy Access of data
6. Promotes Data Integrity

Disadvantage :— 1. Complexity
2. Difficulty in Quering data
3. Lack of Structural Independance.

D-13

Relational Model :→

The main significance of the model is absolute separation of logical view and physical view of data. It consists of three components –

1.) A structural components – A set of tables (Relation) & set of domain that defines the way that can be represented.

2.) A set of rules for maintaining integrity of the database.

3) A manipulative component consisting of set of high level operation which act upon & produces whole table.

This model is represented in the form of table which consist of rows (tuples) & columns (attributes).

Operation :→ i) Insertion ii) Deletion iii) Updation iv) Retrieval

Advantage :→ i) Simplicity ii) Flexible iii) Data Independence iv) Structural Independence v) Query Capability

Disadvantages :→

i) Mapping objects to Relational database can be difficult to learn

ii) Relational Model is not suitable for huge databases.

iii) Data integrity is difficult.

Comparision bet. diff. Record based Data Model

Hierarchical Model	Network Model	Relational Model
1. Organizes the records in a tree structure i.e. hierarchy a parent & child record relationship.	1. Organizes the collection of record connected to one another through link/pointer which is an association bet. two records	1. Records are represented in the form of table & relationship bet the tables are set using common fields.
2. 1:1, 1:N relationship can be implemented in this model	2. 1:1, 1:N, N:N relationships can also be implemented	2. All relationship can easily implemented.
3. It organizes the record in form of tree structure	3. It organizes records in form of directed graph.	3. It organize the record in the form of table.
4. Relationship among records are physically implemented using pointers	4. Relationship among record are represented physically using linked list	4. Relations among records are represented logically in the form of row & column.
5. Lack of declarative quering facilities	5. Lack of declarative quering facility	5. Provide declarative queriying facilities.
6. Less data independence	6. Partial data independence	6. Data independence.
7. Retrieval algorithm are Complex and asymmetric	7. Retrieval algo. are complex but asymmetric	7. Retrieval algorithms are simple & symmetric.
8. Deletion anomoly exists. It is difficult to delete parent records.	8. No deletion anomoly exists.	8. Free from deletion anomoly.

3) Object based data models: —> Object based data models are also known as conceptual models used for defining concepts including entries, attributes and relationships bet. them. These models are used in describing data at the logical and user view levels.

a) ER Model: —> The ER model is a high level conceptual data model developed by Chen in 1976 to facilitate design. It allow the representation of various constraints as well as their relationships.

The relationship bet. entity set is represented by a name. ER relationship is of 1:1, 1:N or N:N type which tells the mapping from one entity set to another.

Features: —> a) ER Model can be easily converted into relations (tables)
b) E-R Model is used for purpose of good database design by database developer.
c) It is helpful as a problem decomposition tool.
d) It is an iterative process.
e) It is very simple & easy to understand by various types of users.

The basic symbol used in E-R diagram are as —

 — represent entity set

 — represent relationship

 — represents attributes

 — represents link bet. entity set & attribute

Advantage: — i) Relation representation is straight forward.
ii) Better understanding due to graphical repn.
iii) Easy conversion to other data models.

Disadvantage: — i) No standard Notation
ii) It is only popular for high level database design.

b) Object Oriented Data Model: — It is logical data model that capture the semantics of objects supported in an object-oriented programming. It is based on collection of objects, attributes & relationships which together form the static properties. It also consists of integrity rules over objects & dynamic properties such as operations or rules defining new database states.

An object is a collection of data & methods. When different objects of same type are grouped together they form a class.

Advantages of object-oriented data models:—

(i) Capability to handle various data types.
(ii) Improved data access.
(iii) Improved productivity.
(iv) Integrated application development system.

Disadvantage of object-oriented Data Models:—

(i) Not suitable for all applications
(ii) No precise definition
(iii) Difficult to maintain

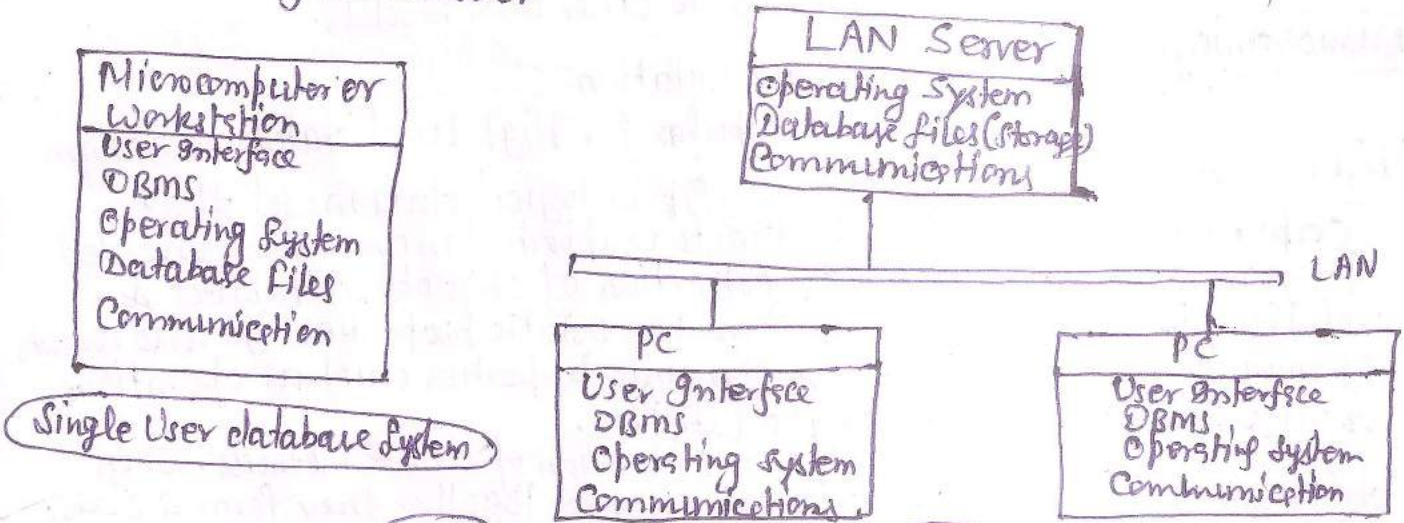
- Semantic Data Models — These models are used to express greater interdependencies among entities of interest.
- Functional Data Models — The functional data model describes those aspects of a system concerned with transformations of values functions, mapping, constraints and functional dependencies.

* Types of Database Systems: —

Acc. to no. of users, the database systems can be further subdivided into two categories.

i) Single-User database system: — In single user database system, the database reside on a PC- on the hard disk. All the applications run on the same PC & directly access the database. The application is the DBMS. A single user accesses the application and business rules are enforced in the application running on PC. The example is DBASE files on a PC.

ii) Multiuser database system: — In this system, many PC's are connected through a Local Area Network and a file server stores a copy of database files. The example is MS-Access / oracle files on a file server



Single User database System

Multiuser Database System

Merits of Multiuser Database System :→

(i) Ability to share data among various users.
(ii) Cost of storage is now divided among various users.
(iii) Low cost(since most components)

Demerits :→ It has limited data sharing ability.

Acc. to Type of Use :→

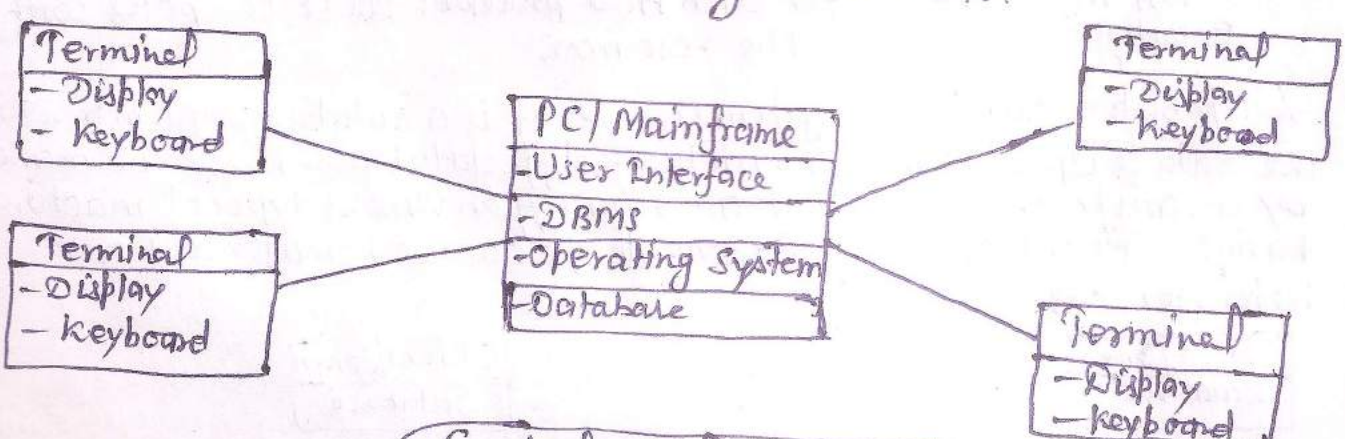
(i) Production or Transactional Database System :→ The production database systems are used for management of supply chain and for tracking production of items in factories, inventory of item in warehousing / stores and orders for items. The transactional database systems are used for purchase on credit cards & generation of monthly statements.

(ii) Decision Support System :→ Decision support system are interactive, computer based system that aid users in judgement & choice activities. They support framing, modelling and problem solving. The application area of DSS is are management and planning in business, healthcare, military and any area in which management will encounter complex decision situation.

(iii) Data Warehouse :→ A data warehouse is a relational database management system designed specifically to meet the transaction processing system. It can be loosely defined as any centralised data repository which may can be queried for bus. benefit.

Acc. to Database Site Locations :→

(i) Centralised database system :→ It consist of single processor together with its associated data storage devices & other peripherals. Database files resides on a personal computer or on a mainframe computer. The example of centralised database system is DB2 database & Cobol application programs running on IBM 390.



Centralised Database System

D-17

Advantage of Centralised Database System: →

i) The control over applications & security is excellent.
ii) The incremental cost per user is very low.
iii) Highly reliable
iv) Many functions such as query, backup, update etc are easier to accomplish

Demerits: →

(i) The failure of central computer blocks every user from using the system until the system comes back
(ii) The communication cost is high.

Parallel database system — A parallel database system can be defined as a database system implemented on tightly coupled multiprocessor or on a loosely coupled multiprocessor. These are applied where queries had done for large database. There are three main architectures for parallel database system. These are
(i) Shared Memory Architecture
(ii) Shared disk architecture
(iii) Shared nothing architecture.

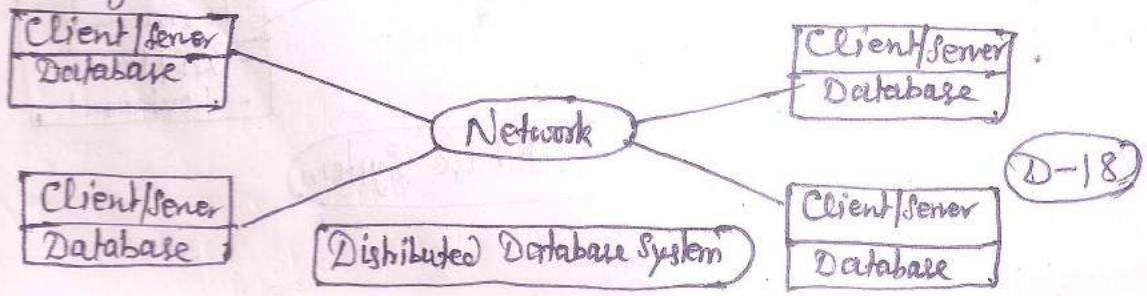
Advantage: →

(i) These are very useful in the applications where large databases have to be queried.
(ii) Response Time is very high.
(iii) Throughput is very high.
(iv) Speed is high
(v) Greater Scalability & reliability

Disadvantage: →

(i) Due to startup cost and startup time, the overall speed is adversely affected.
(ii) Due to processes executed in parallel, sharing the resources, a slow down may result offer each new processes as it competes with existing processes for the resources.

Distributed database systems: — It is a database system in which the data is spread across a variety of diff. databases. These are managed by a variety of DBMS's that are running on various types of machines having different operating system. It can be homogeneous or heterogeneous.



D-18

Advantages of Distributed Database System : →

1. Improved Sharing Ability
2. Local Autonomy
3. Availability
4. Reliability
5. Improved Performance

Disadvantage : →

1. Architectural Complexity
2. Lack of standards
3. Lack of professional support
4. Data integrity problem.
5. High Cost
6. Problem of Security

Client / Server database System : →

There are two logical components i.e. clients and server in the client / server database system. The clients are personal computers or workstations and the server is a large workstation or a mainframe computer. The clients run their own copy of an operating system. They also run one or more applications using their CPU's & memory. The server machines run their own copy of the operating system. They also run a DBMS that manage the database.

The database driver is a small portion of s/w that sits bet. clients & server. It also establishes a connection from the client to the server & passes commands bet. them.

Advantage : →

(i) The processing of entire database system is spread out over clients & servers.
(ii) The response time & throughput is very high.
(iii) It is more flexible as compared to centralised system.
(iv) It can take advantage of GUI
(v) DBMS can achieve high performance

Disadvantage : →

(i) The implementation is more complex.
(ii) Programming cost is high.

D-19

Client/Server Database System	Distributed Database System
1. Application is usually distributed across clients.	1. Application is distributed across sites.
2. Diff. platforms are often difficult to manage.	2. Diff. platform can be managed easily
3. Maintenance cost is low	3. Maintenance cost is much higher.
4. New sites can not be added easily	4. New sites can be added with little problem
5. Speed of database access is good.	5. Speed of database access is much better

Unit-3

* Q. What do you mean by data security? What is main difference bet. data security and data integrity.

Ans. Database security is the system, process and procedures that protect a database from unintended activity. Unintended activity can be categorised as an authenticated misuse, malicious attack or inadvertant ~~att~~ mistake by authorised individuals or processes. Database security involves allowing or disallowing users from performing actions on the database and objects with in it.

Database security refers to the mechanism that protect data stored in a database from unauthorised access.

Wheras Data integrity refers to the mechanism that protects a database from its accidental loss of consistency. It ensures that Data in the database is correct & consistent.

* Database Security Requirements :—>

- Confidentiality :— A secure system ensures the confidentiality of data. It allow individuals to see only the data that they are supposed to see.
- Privacy of communication :—
- Secure storage of sensitive data :— once data has been collected and entered into a database server, the database server must protect its integrity and privacy.
- Authenticated User :— Authentication method used should guarantee the identity of system user : that a person who say he is and not an imposer.
- Integrity :— A secure system must ensure that the data it contains is valid. Data integrity means that data is protected from deletion & corruption while it resides with in database & while it is being translated over the network.
- Availability :— A secure system make data available to authorised user without delay.

* Diff. bet. Authentication & Authorisation :—>

Authentication is the process by which a user's identity is checked.

Wheras Authorisation is the process by which the user's privileges are ensured.

* Issues/Level/Dimensions of database security :—>

- Network Level security :—> It is required in case of distributed database systems. Right security is another level of security provided by the network o/s. It controls the files, sub-directories and directories, a user or a group of users can access.
  Attribute security is also a level of security that determine whether a particular file or directory can be viewed, shared renamed, modified or deleted by a user or not.
- Operating System Level Security :— A weak operating system may serve as a means of unauthorised access of database file, so secure operating system is very important.
- Database System Level Security :— Some database users may be allowed to access only a portion of database and denied access to other portions.
- Program Level Security :— It prohibits unauthorised users from using particular program that may access the database. For example, bank clerk may be allowed to use a program that retrieves details of a customer account, but not a program that modifies the balance amount.
- Record Level Security :— It prohibits unathorised users from accessing or updating certain records such as record of managers in the Employee table.
- Field Level Security :— It prohibits unauthorised users from accessing or updating data in certain field such as Salary field.

* Threat :—>

A threat is any situation, event or personnel that can adversely affect the database security and smooth & efficient functioning of organisation. A threat may be caused by a situation or event involving a person, action or situation that is likely to bring some harm to the orgn.

- Excessive privilege abuse
- Legal privilege abuse :— Users may also misuse legal database privileges for unauthorised purpose. (D-22)
- Privilege elevation :— For example, with administrative privilege the rascal developer may turn off audit mechanism, create bogus accounts, transfer funds etc.

- Operating System Vulnerabilities:
- SQL Injection.
- Weak Audit Trail :— A database log that is used mainly for security purposes is called an audit trail. Weak database audit policy represents a serious organisational risk on many levels.
- Denial of Services.
- Weak Authentication :—> It allow attackers to assume identity of valid database users by stealing or otherwise obtaining login credentials (परिचय पत्र)

* Different Techniques for Database Security

- Authorisation :—> Authorisation is the granting of a right or privelage that enables a subject to have valid access to a system or system's object.
- Views :—> A view is a virtual table, which does not exist in the database but can be produced upon request by a particular user at the time of request. The view mechanism provides a powerful & flexible security mechanism by hiding parts of database from certain users. The users are not aware of existence of any columns or rows that are missing from table.
- Backup :—> Backup is the process of periodically taking a copy of database and log file onto offline storage media. To keep track of database transaction, DBMS maintains a special file called log file (journal) that contains information about all updates to the database.
- Integrity Constraints :—> Integrity constraints contribute to maintaing a secure database system by preventing data from becomming invalid & hence giving misleading or incorrect results. Data integrity refer to avoidance of accidental loss of consistency.
- Database Encryption :—> Encryption is the encoding of data by a special algorithm that makes the data unreadable by any program without decryption key.
- Audit Trail :—> A database log that is used mainly for security pupose is called an audit trail. Audit trail refers to the recorded history of operations performed on a database.

* Failures :→ System crashes can occur due to h/w or s/w errors resulting in loss of main memory. An example of user error is a user deleting a row or dropping a table Carelessness is the destruction of data or facilities by operators or users because they were not concentrating on the task at hands. Sabotage is the intentional corruption or destruction of data h/w or s/w facilities.

- Statement failure :→ It can be defined as the inability by the database to execute an SQL statement
- Application s/w errors :→ It include logical error in the programs that is accessing the database, which cause one or more transactions to fail.
- Network failures :→ It can occur while using a client server configuration or a distributed database system where multiple database servers are connected by communication network.
- Media failures :→ It is the most dangerous failures. A typical example of a media failure is a disk controller failure or diskhead crash, which causes all data database residing on that disk or disk to lost.
- Natural & Physical Disasters :→ are the damage caused to data hardware & s/w due to natural disasters like fires, floods, earthquakes, power failures etc

* Back-Up :→ A backup is a representative copy of data. This copy can include important parts of a database such as the control file, redo logs & data files. A backup protects data from application error and act as a safeguard against unexpected data loss by provide a way to restore original data

- Physical Backup → are copies of physical database file.
- logical Backup → It contains the data that is exported using SQL command & stored in binary file. These are used to supplement physical backups.

D-24

* Recovery of Database : —>

Database recovery is the process of restoring a database to the correct state in the event of failure. It is a service that is provided by DBMS to ensure that the database is reliable and remains in consistent state in case of a failure.

There are two types of Recovery : —>

i) Backward Recovery : —> In backward recovery, the uncommitted changes made by a transaction to a database are undone and the system is brought back to previous consistent state of database

ii) Forward Recovery : —> In forward recovery, changes made by a transaction are reapplied to an earlier copy of the database.

* Transaction : —> A transaction is a group of database operations combined into a logical unit of work that is either entirely committed or rolled back.

Properties : —

i) Atomicity :- A transaction is an atomic unit of processing it should be performed or not performed at all.

ii) Consistency Preservation :- A transaction is said consistency preserving if its complete execution takes the database from one consistent state to another

iii) Isolation :— A transaction should be executed seperately, that is, a transaction should be executed without interfering with any other concurrent executing transaction.

iv) Durability : — The changes made to database by a committed transaction must persist in the database. These changes must not be lost due to any failure.

* Technique of Database Recovery :—>

The idea behind deferred update technique is to defer

(a) **Defered Update:—>** or postpone any actual updates to the database until the transaction completes its execution successfully & reaches its commit point

One can state a typical deferred update protocol as follows :-

1) A transaction can not change the database on disk until it reaches its commit point.
2) A transaction does not reach its commit point until all its update operations are recorded in the log and log is force-written to disk.
3) In the second step of this protocol is write-ahead logging (WAL) protocol. Because the database is never updated on disk until after the transaction commits, there is never a need to UNDO any operations. Hence it is also known as the NO UNDO/REDO recovery algo. , REDO is needed in case the system fails after a transaction commits but before all its changes are recorded in the database on the disk. In this case the transaction operations are redone from the log entries

(b) **WAL :—>** In write-ahead Logging, the transaction logs (updates) are first written to the log before they are applied to the database.

(c) **Immediate Update:—>** In this, the database may be updated by some operation of transaction before the transaction reaches its commit point. These operations are recorded in the log on disk by force-writing before they are applied to the database. If a transaction fails after recording changes in the database but before reaching the commit point, the effect of transaction on database must be undone (rolled back).

So in the case of immediate update technique, both undo & redo operations are required during recovery. Hence the immediate update technique is also known as UNDO/REDO algo.

⑤ Log :—→ Log is the most widely used structure for recording database modifications. The log is the sequence of log records and maintains a history of all updates activities in the database.

The following steps are involved in the log-based recovery process:—

i) Find the first log record to be processed.
ii) Create the Redo Undo Lists.
iii) Undo Transactions
iv) Redo Transactions.

⑥ Checkpoint :—→ A checkpoint is an entry in the log and is represented as [Checkpoint]. A checkpoint operation is performed periodically by the system. It copies logging information onto stable storage. The [Checkpoint] record is written into log when all the changes up to this point are reflected in the database. The information & operations performed at each checkpoint consist of the following:—

- A start of checkpoint record giving the identification, that is checkpoint along with the date & time of the checkpoint.
- All log information from the buffer in the volatile storage is copied to the log on the stable storage.
- All database updates from the buffers in the volatile storage are propagated to the physical database.
- An end-of-checkpoint record is written and the address of checkpoint record is saved on a file accessible to the recovery routine on the start-up after system crash.

⑦ Shadow Paging :—→ is a technique where transaction logs are not required. Two directories for each database page are created during the life of a transaction - the current directory & the shadow directory. When the transaction start both the directories are same.

The shadow directory never changed during the duration of transaction and the current directory is updated when the transaction perform a write operation. All I/O operations use the current directory to locate the database page on the disk.

When the value in the database is updated, the old record is not updated but a new one is created and the link in the current directory is changed to point to the newly created record. When a transaction commits, the shadow directory is discarded and the current directory becomes the database page directory. If the transaction aborts, the current directory is discarded.

* **Distributed Database** :→ A distributed database is a network of database managed by multiple database servers on multiple sites that appears to as a user a single logical database. A distributed database is distributed into seperate partitions / fragments. Each partition/ fragment of a distributed database may be replicated.

◎ **Demerits** :→

i) Complexity :— Replication of data in a distributed database adds to the DDBMS.
ii) More costly
iii) Security
iv) Difficult to maintain integrity
v) Increased processing overheads
vi) Inexperience :— Distributed databases are difficult to work with and as the young field there is not much readily available experience on proper practice.
v) Lack of standards :— There are no tools or methodology yet to help users convert a centralised DBMS into a distributed DBMS.
vi) Database design more complex.
vii) Greater potential for bugs.

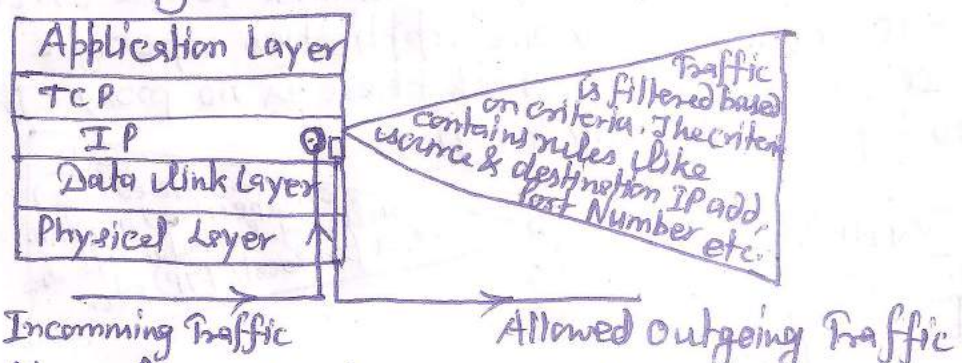
⊙ Merits :—>

i) Reflects Organisational Structure
ii) Local autonomy
iii) Improved availability of data
iv) Improved performance
v) Improved reliability
vi) Economic
vii) Faster Query Processing
viii) Easy to expand
ix) Modularity — System can be modified, added & removed from the distributed database without affecting other modules

* Firewall : —> A firewall is a software program or h/w device that filters the inbound and outbound traffic between user network or computer and the Internet.

Types of firewall : —

i) Packet Filtering Firewalls : —>

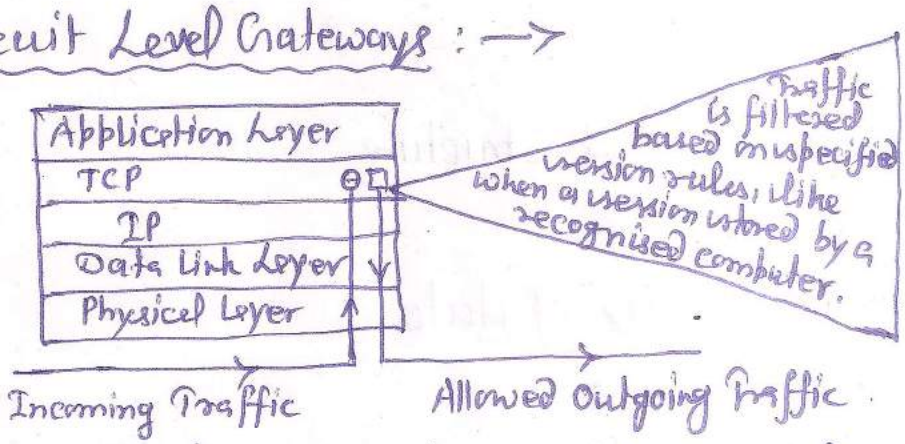


This firewall work at the Network layer of OSI model or the IP layer of the TCP/IP model. It compare each packet with a set of criteria before forwarding them. It can drop the packet, forward it to the network or send a message to originator, depending on the outcome of comparison of packet & criteria

- Merits :—>
  1. Low cost and low impact on Network performance
  2. Most routers support packet filtering
- Disadvantages :— Since it work at network layer, it does not support rule based models.

(2-29)

⑧ Circuit Level Gateways :→



This work at session layer of OSI model or TCP layer of the TCP/IP model. It monitor the TCP handshaking bet. packets to determine whether a requested session is legal. The information passed to remote computer through circuit level gateways appears to have originated from gateway. It is useful for hiding information about protected network.

Merits :→

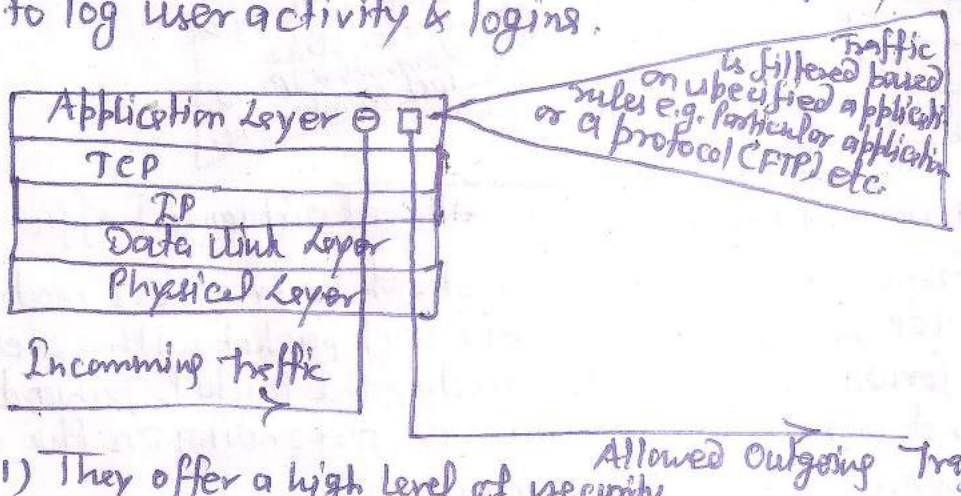
1. They are relatively inexpensive
2. They hide the information about the private network they protect.

Demerits :→

1. They do not filter individual packets.

⑨ Application Level Gateways :→

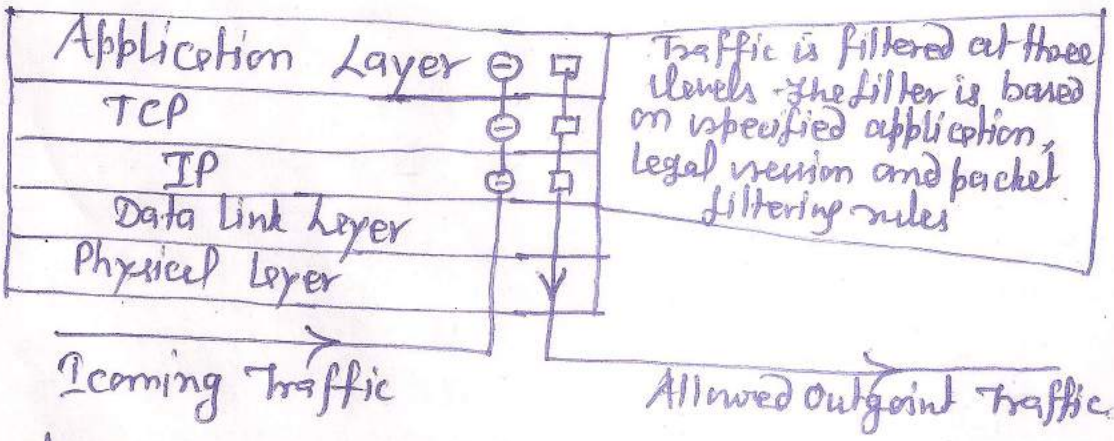
It work at the Application layer of OSI or TCP/IP model. They are similar to the circuit level gateway except that they are application specific. The packets cannot access services for which there is no proxy. There can also be used to log user activity & logins.



- Merits :— 1) They offer a high level of security.
  2) They can filter application specific command like http: post & get.
  3) They can log user activity and logins.
- Demerits :— 1) They have significant impact on network performance, since context switches slow down network access
  2) They are not transparent to end user.

⊙ Stateful Multilayer Inspection Firewalls :→

It is a combination of all the three previous firewall. These firewall filter packets at the network layer, determine whether version packets are legal and evaluate the contents of packets at the application layer. These firewalls allow a direct connection bet. client & host. They depend on the algorithm to recognize and process application layer data instead of running application specific proxies.



Advantages —

They provide a high level of security, good performance & transparency to the end users.

Disadvantages —

1. They are expensive.
2. They are less secure than other types if not administered by competent personnel.

* Dr. Codd's 12 Rules for a relational database model :→

① Information Rule :→ All info. in a relational database is represented explicitly as values in tables.

② Guaranteed Access :→ Every value in a relational database is guranteed to be accessible by using a combination of table name, primary key value and column name.

③ Systematic null value support :→ The DBMS provide systematic support for handling null values, distinct from default values and independent of any domain.

④ Active, online relational catalog :→ The description of database and its contents is represented at the logical level as tables and can therefore be queried using database language.

⑤ Comprehensive data sublanguage :→ At least one supported language must have a well-defined syntax and be comprehensive. It must support data definition, manipulation, integrity rules, authorisation and transactions.

⑥ View Updating rule :→ All views that are theoretically updatable can be updated through the system.

⑦ Set-level insertion, update and deletion :→ The DBMS support not only set-level retrievals but also set-levels inserts, updates and deletes.

⑧ Physical data independance :→ Application programs and ad hoc programs are logically unaffected when physical access method or storage structure are modified.

⑨ Logical data independence :→ Application programs and ad hoc programs are logically unaffected, to extent possible, when changes are made to the table structures.

⑩ Integrity independence :→ The database language must be able capable of the defining integrity rules. They must be stored in the online catalog and they can not be avoided.

⑪ Distribution Independence :→ Application programs and ad hoc request are logically uneffected when data is first distributed or when it is redistributed.

⑫ No Subversion :→ It must not be possible to ignore the integrity rules defined through the database language by using lower-level languages.

(24)

D-31-II

Relation :→ A table in relational database is called relation. Relations are unordered. The term relation is often used to refer to both schema (intension) and instance (state of that schema) or extension. Each cell of the relation contains exactly one atomic (single) value.

Relation Instance:→ A relation instance is a set of tupples (rows or records) that each conform to the schema of the relation.

Relation Schema: → A relation schema can be thought of as the basic info. describing a table or relation. This includes, a set of column names, the data types associated with each column and name associated with entire table.

Attribute : → Column of a table/relation is called attributes. Column or attributes have names. Attributes can be thought of a columns in a table. The order of attributes in a relation is immaterial. Each attribute in a relation is distinct.

Domain : → A set of permissible values for an attribute (column) is called domain. The set of all possible atomic values for one or more attributes is known as a domain.

Cardinality : → Number of tuples (records) in a relation is called cardinality of the relation.

Q. What is key ? Diff. bet. primary key and foreign key ?

A. A key is a collection of one or more columns (fields) in one table where values match corresponding column in other tables. Any column can be a key or multiple column can be grouped together to form a single key.

A primary key is the candidate key that has been chosen from group of candidate key to uniquly identify records in a particular table. A foreign key is a reference to key in another table. A relationship bet. two tables is created by creating a common field to the two tables. Foreign keys allow us to ensure refesential integrity. This means a foreign key that contains a value must refer to an existing record in the related table.

Candidate Key:→ is a field or combination of fields, that identifies a record uniquely. The fields of candidate key must contain unique values and can not contain null values.

Super key : → is a set of attributes that contains a key.

* **Tupple : →** A row in database table is called a tupple. All tupples have the same no. of attribute/columns.

* **Relation degree: →** Relation degree is the no. of fields (column) in the relation.

* Database Encryption :→ Encryption is the encoding of data by special algo. that makes the data unreadable by any program without the decryption key. If a database system hold particularly sensitive data, it may be deemed necessary to encode it as a precaution against possible external threats or attempts to access it. Some DBMS is provide encryption facility for this purpose. The DBMS can access the data (after decoding it), although there is degradation in performance because of time taken to decode it. Encryption also protect data transmitted over communication lines. There a no. of techniques for encoding data to conceal the information; some are termed irreversible and other reversible. Irreversible technique, as name implies, do not permit original data to be known. However, the data can be used to obtain valid statistical information. Reversible techniques are more commonly used. To transmit data securely over insecure network requires the use of a cryptosystem, which includes:—

1) An encryption key to encrypt data (plain text)
2) An encryption algo. that, with encryption key, transforms plaintext into ciphertext
3) A decryption key to decrypt the ciphertext.
4) A decryption algo that, with decryption key, transforms ciphertext into plaintext.

* Failure :→ when the system does not functions according to its specifications and fails to deliver the services for which it was intended to is known as failure.

* Fault :→ when an error is propagated from one component to another component of a system or when the component fails, the a fault is occured.

* Error :→ when a component of the system acquires an undesirable state then the error is said to occur. This undesirable state that a component achieves is known a erroneous state and further use of component will lead to failure.

D-33

Unit-4

* **Data Warehousing :→** A data warehousing is subject oriented integrated non-volatile, time varying collection of data in support of its decision making.
  - **Subject oriented :—** A data warehouse is oriented around the major business subjects of enterprise, such as customer, vendor, product or activity.
  - **Non-Volatile :—** Data warehouse are nonvolatile i.e. information in the data warehouse changes for less often.
  - **Time-variant :—** Data in the data warehouse contain an element of time, such as day, week, month or year.

⊙ **Characteristics of Data Warehousing :→**

1) Subject Oriented 2) Integrity 3) Time Variant
4) Non-Volatile

⊙ **Purpose of Data Warehouse**

1) To provide Business users with Access to Data.
2) To provide one version of the Truth.
3) To record the past Accurately.
4) To slice and Dice Through data.
5) To Separate Analytical & Operational Procesing
6) To support the Reengineering of Decisional Proces.

⊙ **Structure of Data Warehouse:—** It consists of following.

i) **Physical Data Warehouse :—** Physical database in which all the data for the data warehouse are stored, along with methods and procesing logic for organising, packaging & processing the detail data.

ii) **Logical Data Warehouse :—** The logical data warehouse contains metadata, including enterprise rules & procesing logic for organising, packaging & procesing the data, but does not contain actual data.

iii) **Data Mart :—** It is a subset of an entreprise-wide Data warehouse, which typically supports an enterprise element.

⊙ Data Warehouse Components :—>

(i) Data Warehouse Database
(ii) Sourcing, Acquisition, cleanup & Transformation Tools.
(iii) Meta Data.
(iv) Access Tools
(v) Data Marts
(vi) Data Warehouse Administration & Mgt.
(vii) Information Delivery System.

⊙ Advantage of a Data Warehouse :—>

i) More Cost-Effective Decision-making.
ii) Better Enterprise Intelligence
iii) Enhanced Customer service.
iv) Business Reengineering
v) Information System ReEngineering.

⊙ Use of Data Warehousing :—>

i) Standard Report & Queries.
ii) Queries against Summarised Data.
iii) Data Mining
iv) Interface with other Data Warehouse

⊙ Limitation of Data Warehouse :—>

i) Extracting, Cleaning and loading data is time consuming.
ii) The demand of resources & maintenance cost is very high.
iii) It is query-intensive.
iv) Performance tunning is very bad.
v) Scalability can be problem.

D-35

* Data Mining :—> Data Mining or Knowledge discovery in databases (KDD) as it is also known, is the nontrivial extraction of implicit, previously known and potentially usefull information from the data. This encompasses a no. of different technical approaches, such as clustering, data summarization, learning classification rules, finding dependency networks, analyzing changes and detecting anomalies."

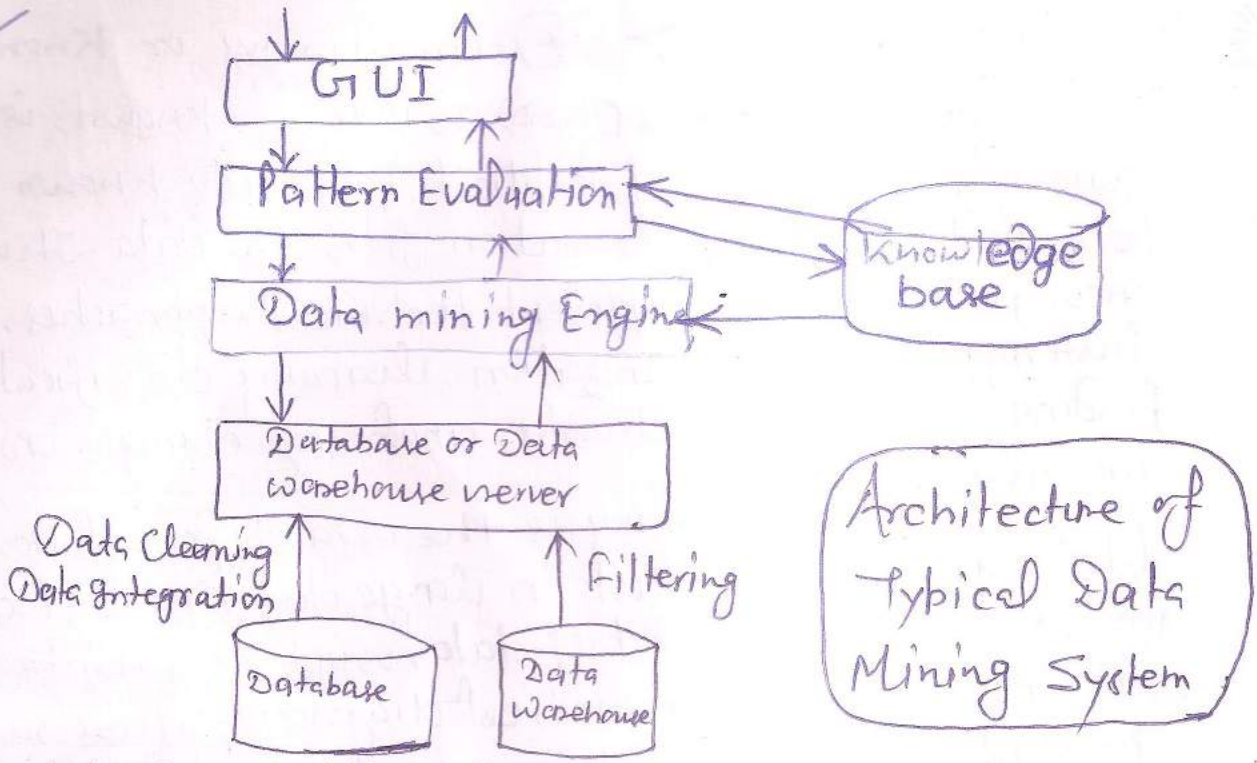
Data mining is the search for relationships & global pattern that exists in large database but are "hidden" among the vast amount of data, such as a relationship bet. patient data and their medical diagnosis. These relationship represents valuable knowledge about the database and the objects in the database.

Data mining refers to "using a variety of techniques to identify nuggets of information or decision-making knowledge in bodies of data and extracting these in such a way that they can be put to use in the areas such as decision support, prediction, forecasting and estimation. The data is often voluminous, but as it stands of low value as no direct use can be made of it, it is the hidden info. in the data that is useful"

⊙ Components of Data Mining :—>

The architecture of typical data mining system may have following components.

i) Database, Data Warehouse or other Information Repository
ii) Database or Data Warehouse Server
iii) knowledge base
iv) Data Mining Engine
v) Pattern Evaluation Module
vi) Graphical User Interface.



Architecture of Typical Data Mining System

⊙ Data Mining Applications: →

i) Retail / Marketing: →

a) Identify buying patterns from customers
b) Find associations among customer characteristics
c) Predict response to mailing campaigns.
d) Market basket analysis.

ii) Banking: →

a) Detect patterns of fraudulent credit use
b) Identify 'loyal' customers
c) Predict customers likely to change their credit card affiliation.
d) Determine credit card spending by customer groups.

iii) Insurance & Health care: →

a) Claims Analysis – i.e. which medical procedures are claimed together.
b) Predict which customers will buy new policies
c) Identify behavior patterns of risky customers.
d) Identify fraudulent behavior

iv) Transportation: →

a) Determine the distribution schedules among outlets.
b) Analyze loading pattern

D-37

v) Medicine: →

a) Characterize patient behavior to predict office visits
b) Identify successful medical therapies for different illness.

⊙ Advantage of Data Mining :—>

a) Automated Prediction of Trends & Behaviors,
b) Automated Discovery of Previously Unknown Patterns.
c) Databases can be larger in both Depth & Breadth.

✳ Internet Database :—> An Internet database is a database that one access by using a web browser (i.e Internet Explorer), a programming language & a web server. Internet database are an integral part of most company internets and the skills needed to support them are highly valued. Creating an internet database can solve many problems for users.

A Internet database is a wide term for managing data online. An internet database gives you the ability to build your own database/data storage without you being a database guru or even a technical person.

⊙ Internet Database Related Languages :—>

i) HTML (Hyper text markup language)
ii) XML (Extensible markup language).

✳ Web Databases :—> A web database is a database used online accessed through a website or some other interface. A web database is handy because any person can get access to this data through the internet from anywhere in the world. A web database is usually created in a way that is accessible to a script language such as Hypertext Pre-processor or PHP. Web database is basically a database used for internet.

⊙ Web database Tools :—>

(i) Common Gateway Interface (CGI) (ii) XML (Extended Markup Language)

⊙ Merits of Web database :—>

i) Good & user friendly graphical use interface.
ii) It is used by both developers & end-users, which is simple & easy to use language.
iii) The network access is transparent
iv) These support cross-platform.

D-38

* Disadvantage of Web Database: →

i) The communication medium is slow.
ii) The security is the major concern
iii) The cost of meeting the demands & expectation of customer is high
iv) Scalability is problem.
v) The functionality of HTML is limited.

* Digital Library: →

A digital library is a library in which collections are stored in digital formats and accessible by computers. The digital content may be stored locally or accessed remotely via computer networks.

⊙ Objectives: →

i) To collect, store, organise and access information in digital form.
ii) To preserve valuable and rare documents.
iii) To have large digitized database accessible to multiple user at the same time.
iv) To save time of Library staff by avoiding routine jobs.
v) To reduce cost involved in various library activities.

⊙ Types: →

i) Born-digital: → In this library, the information that has been converted from physical medium i.e. paper, by digitizing & contents that was created in a digital format exists.
ii) Hybrid Library: → is used for libraries that have both physical & digital collections. For example American Memory is a digital library with in the Library of Congress.

⊙ Components: →

i) Geographically distributed digital information collections
ii) Geographically distributed users.
iii) Information represented by a variety of digital objects
iv) Large & diverse collections
v) 'Seamless' access.

⊙ Services provided by Digital Library :—>

i) Searching facility through Online Public Access Catalogue (OPAC).
ii) Free Text Search Facility
iii) Document delivery
iv) Bibliographics
v) Discussion Groups, Forums, News.
vi) Marketing
vii) Regular E-mail updates on specific subjects

⊙ Merits of Digital Libraries :—>

i) No physical boundary.
ii) Round the clock availability.
iii) Multiple access.
iv) Information retrieval
v) Preservation & conservation
vi) Space
vii) Added value
viii) Access current information

⊙ Demerits of Digital Libraries :—>

i) Attitude of Library professionals.
ii) Lack of information policy & information culture.
iii) Lack of efficient & effective Library n/w.
iv) The cost involved in the creation & maintenance of digital Library environment is high.

⊙ Requirements for Digital Library :—>

i) Audio Visual ii) Computer iii) Network
iv) Printer v) Scanner vi) Storage Device vii) Software.

D-40

✳ **Multimedia Database :—>** A multimedia database is a database that can store one or more primary media file types. It is a database that hosts one or more primary media file types such as .txt (document), .jpg (images), .swf (videos), .mp3 (audio), etc. And loosely fall into three main categories :-

i) Static media (time-independent, i.e images & handwriting)
ii) Dynamic media (time-dependent, i.e. video & sound bytes)
iii) Dimensional media (i.e. 3D games or computer-aided drafting programs — CAD)

⊙ **Components :—>**

i) Text ii) Speech (iii) Graphics
iv) Image v) Animation (vi) Sound (vii) Video

⊙ **Application :—>**

i) Repository Application.
ii) Presentation Application.
iii) Collaborative work.
iv) Documents & record management
v) Knowledge disseminating. (GATE)
vi) Education & training
vii) Marketing, advertising, retailing, entertainment & travel.
viii) Real time control & monitoring.

⊙ **Advantages of Multimedia Database :—>**

i) Optimized Storage ii) Efficient Access
iii) Referential integrity of links.
iv) Transaction protected multi-user mode.
v) Recovery etc.

D-41

✳ Mobile Database :—> A mobile database is a database that can be connected to by a mobile computing device over a mobile network. The client & server have wireless connections. A cache is maintained to hold frequent data & transactions so that they are not lost due to connection failure. Mobile database allow the development & deployment of database applications for handheld devices. Thus, enabling relational database based applications in the hands of mobile workers. The database technology allows employees using handheld to link to their corporate networks, download data, work offline & then connect to the network again to synchronise with the corporate database.

⊙ Need :—>

i) Mobile users must be able to work without a wireless connection due to poor or even non-existent connections.
ii) Limited life of power supply
iii) The changing topology of network.
iv) Bandwidth must be conserved.

⊙ Merits :—>

(i) Provide access to data by using only a web browser.
(ii) Provide Internet access to data.
(iii) Significantly lower the cost of using a database.
(iv) Make a database available to more users.
(v) Provide users a familiar web interface.

⊙ Demerits :—>

i) Development time increases
ii) Skill-level requirements for the development team are higher.
iii) Maintenance needs are not decreased.
iv) The risk of data corruption may be increased.

⊙ Characteristics :—>

i) Users are not attached to a fixed geographical location.
ii) Mobile computing devices : low-power, low-cost, portable.
iii) Wireless networks.
iv) Mobile computing constraints.

* **Spatial Database:** —> A spatial database is a database that is optimized to store and query data that is related to objects in space, including points, lines & polygons. A spatial database system is a database system. It has following functions:—

i) It offers spatial data types in its data model & query language.

ii) It supports spatial data types in its implementation, providing at least spatial indexing & efficient algorithm for spatial join.

⊙ **Applications:** —> Spatial databases have applications in various areas as wide as geography, biology, CAD/CAM, Medicine etc.

⊙ **Merits:** —>

i) It has the capability to treat the spatial data like anything else in the database such as transactions, integrity checks, backups. Less data redundancy, multi-user support, security control, locking, fundamental organisation & operations handled by the database.

ii) It offsets the complicated tasks to the database server such as organisation & indexing.

iii) Spatial databases donot have to re-implement operators & functions.

iv) Spatial database use simple SQL expression to perform spatial operations such as area, length, union, intersection & buffer.

v) Spatial databases use simple SQL expression to determine spatial relationships such as distance, adjacency & containments.

⊙ **Demerits:** —>

i) The cost to implement spatial database is very high.

ii) The spatial database are less flexible.

iii) The spatial database are incompatible with some GIS (geographical information system) softwares.

iv) The spatial database are slower than local & specialized data structures.

v) The people that uses & manages the spatial databases are inexperienced.

D-43

RAM-RAM G