PAGE No.
DATE / /

Unit 1

Overview of Database Management System

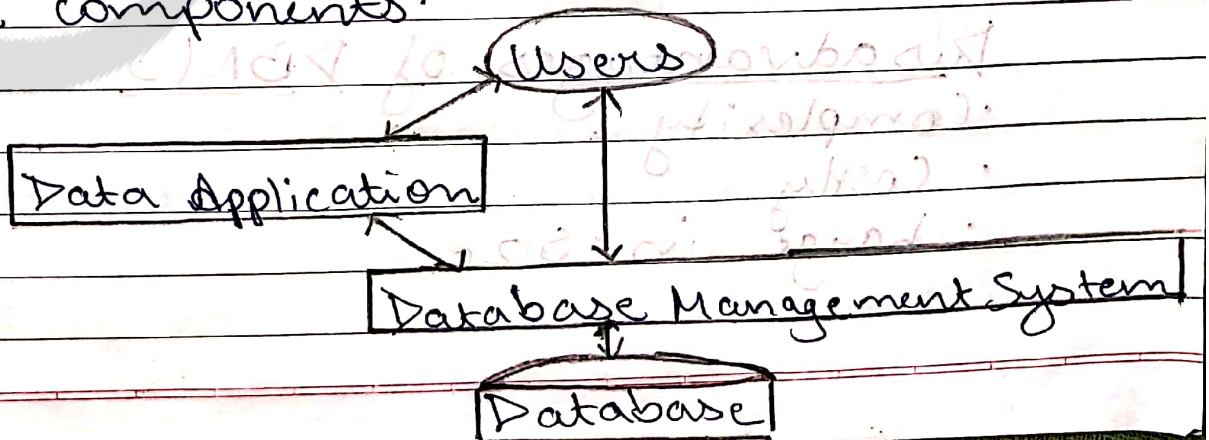
A Database is a collection of related data organised in a way that data can be easily accessed, managed and updated. Any piece of information can be a data, for example name of your school. Database is actually a place where related piece of information is stored and various operations can be performed on it.

DBMS

A DBMS is a software that allows creation, definition and manipulation of database. Dbms is actually a tool used to perform any kind of operation on data in database. Dbms also provides protection and security to database. It maintains data consistency in case of multiple users. Some examples of popular dbms are MySQl, Oracle, Sybase, Microsoft Access and IBM DB2 etc.

Components of Database System

The database system can be divided into four components.



- **Users:** Users may be of various type such as DB administrator, System developer and End users.
- **Database application:** Database application may be Personal, Departmental, Enterprise and Internal.
- **DBMS:** Software that allow users to define, create and manages database access, Ex: MySql, Oracle etc.
- **Database:** Collection of logical data.

Functions of DBMS

- Provides data independence
- Concurrency Control
- Provides Recovery services
- Provides Utility services
- Provides a clear and logical view of the process that manipulates data.

Advantages of DBMS

- Segregation of application program
- Minimal data duplicacy
- Easy retrieval of data.
- Reduced development time and maintainance need.

Disadvantages of DBMS

- Complexity
- Costly
- Large in Size

PAGE No.
DATE / /

Q] File systems versus DBMS

File system	DBMS
Software that manages the data files in a computer systems.	Software to create and manage databases
Helps to store a collection of raw data files into the hard disk.	Helps to easily store, retrieve and manipulate data in a database.
Tasks such as storing, retrieving and searching are done manually, so it is difficult to manage data	Operations such as updating, searching, selecting data is easier since it allows using SQL querying
Has data inconsistency	Provides higher data consistency using normalization
There is more redundant data	There is low data redundancy
Provides more security to data	Comparitively less data security
Handling is easy	Handling is complex

Backup and recovery process is not efficient because it is not possible to recover the loss lost data	Has a sophisticated backup and recovery
Appropriate to handle data of a small-scale organization or individual users	Suitable for medium to large organizations or multiple users
Eg:- NTFS and Ext	Eg: MySQL, MSSQL, Oracle, DB2

Q) Advantages of DBMS

1) Minimize Data Redundancy :- In file Processing System, duplicate data is created in many places because all the programs have their own files. This creates data redundancy which in turns wastes labor and space. In Database Management System, all the files are integrated in a single database. The whole data is stored only once at a single place so there is no chance of duplicate data.
For example: A student record in library or examination can contain

duplicate values, but when they are converted into a single database, all the duplicate values are removed. Complete redundancy can be removed because somehow we need duplicate value to relate tables with each other. But still DBMS controls data redundancy that saves lots of labor and time.

2) Sharing of Data:- In DBMS, Data can be shared in between authorized user of database. All the users have their own right to access the database up to a level. Database Administration has complete access of database. He can assign users to access the database. Other users are also authorized to access database and also they can share data between them. Many users have same authority to access the database.

3) Data Consistency:- DBMS controls data redundancy which in turns controls data consistency. Data consistency means if you want to update data in any files then all the files should not be updated again. As in DBMS, data is stored in a single database so data becomes more consistent in comparison to file processing system. Also updated values are available to

all the users immediately.

4) Data Integrity :- Data integrity means unification of so many files into a single file. In DBMS data is stored in different tables. A database contains different tables that are linked to each other. Many users feed entries in these tables so it is important to maintain data items and association between data items. DBMS allows data integrity that makes it easy to decrease data duplicity. Data integration reduces redundancy as well as data consistency.

5) Search Capability :- Users of database may require to fetch data from the database. There are numerous queries users may ask about the data. Search speed of the database must be fast to produce quick results. If users execute any query then it is required that he get fastest results from the database. It is an objective of database to maintain flexible search capability.

6) Security :- Data security means protecting your precious ~~data~~ data from unauthorized access. Data

in database should be kept secure and safe to unauthorized modifications. Only authorized users should have the grant to access the database. There is a username set for all the users who access the database with password so that no other guy can access these information. DBMS always keep database tamperproof, secure and theft free.

7) Privacy :- Privacy means up to what extent a user can access the data. It is predetermined by the DBA that who will access the data and up to what level he will be able to access it. Lets say when you make a facebook page then you have the power to give rights to other users that who will be the promoter, editor and admin.

8) Simplicity :- Simplicity means to represent the overall logical view of data in a simple and clear manner. DBMS is very simple for its users who use it. All the operations like insert, delete, creat and update are very easy to implement.

9) Backup and Recovery :- Data loss is a very big problem for all the organizations. In traditional file processing system, a user needs to backup

the database after a regular interval of time that wastes lot of time and resources. If the volume of data is large then this process may take a very long time.
DBMS solves this problem of taking backup again and again because it allows automatic backup and recovery of database. For example, if a system fails in the middle of any process then DBMS stores the values of that state in which database were before query execution.

10) Integrity Constraints :- Constraints are used to store accurate data because there are many users who feed data in database. Data stored in database should always be correct and accurate. DBMS provides the capability to enforce these constraints on database.
For example :- the maximum marks obtained by the students can never be more than 100. Also account balance of Banks like Axis should not be less than 2500 otherwise you will be penalized.

11) Data Atomicity :- Any complete transaction in database is called atomic unit. It is the duty of

DBMS to store a complete transaction in database. If any transaction is partially completed then it rolls back them.
For example, in railway reservation system, if user has completed the process of ticket reservation then his record will be stored and amount of money will be deducted from his account otherwise no amount will be deducted and if deduced it will be given back.

12) Development of new applications:- If a new application is required and data is available for creating the application then it is very easy to develop new application. No time will be consumed in creating stored data again and again.

13) Concurrency Control:- If two users are accessing data simultaneously and they both want to update values of same record then it may create concurrency. DBMS has the power to control concurrency so that no transactions are lost.

14) Data Migration:- Data migration means adjusting storage of data according to its popularity. In a

database, there is some kind of data that is accessed frequently and at the same time some data is accessed occasionally. So it is required to store frequently accessed data in a manner that it can be accessed quickly.

15) Tunability:- Tuning means adjusting something to get better performance. Same in the case of DBMS, as it provides tunability to improve performance DBA adjust database to get effective results.

16) Solves Enterprise and Individual Requirement:- A DBMS provides a wide range of user interfaces to use a database. There are many users working on the database having a different level of knowledge.

17) Powerful User language:- A DBMS permits end users to use database without having special training or expertise. Any untrained user can easily query, search and update data in Database. He can easily generate report or documents with less knowledge.

18) Standards can be enforced:-

PAGE No.
DATE / /

As DBMS have central control of database so a DBA can ensure that all the applications follow some standards such as format of data, document standards etc. These standards help in data migrations or in interchanging the data.

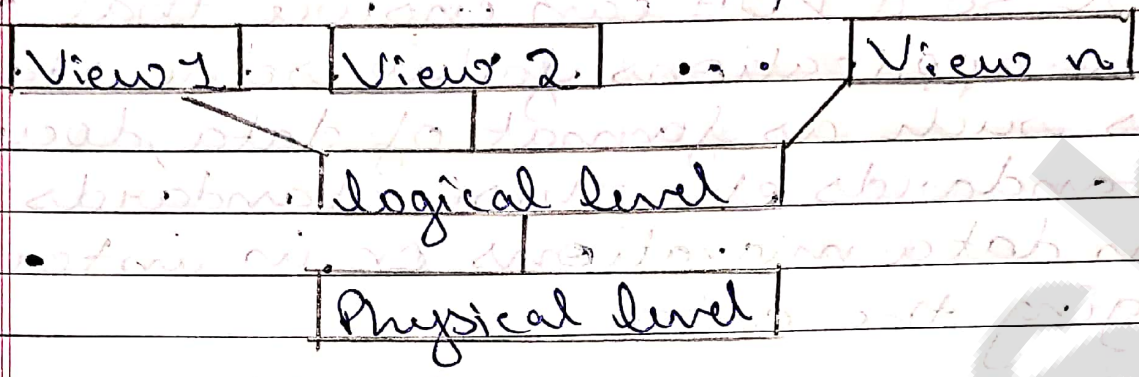
19) Maintaining Cost is lower :- DBMS systems are costly but after purchasing them their maintenance cost is very less. It can be maintained by few programmers that is not costly for an enterprise.

20) Very less chances of Data loss :- As there is lot of security constraint made on database so chances of data loss are minimum. One can store their precious data or many years in DBMS without loss of any information.

Q) Views of data :- Data Abstraction, Instances and Schemas

Data abstraction :- Database systems are made-up of complex data structures. To ease the user interaction with database, the developers hide internal irrelevant details from users. This process of hiding irrelevant details from user is called data

abstraction



Three levels of data abstraction

We have three levels of abstraction:-

Physical level: This is the lowest level of data abstraction. It describes how data is actually stored in database. You can get the complex data structures details at this level.

Logical level:- This is the middle level of 3-level data abstraction architecture. It describes what data is stored in database.

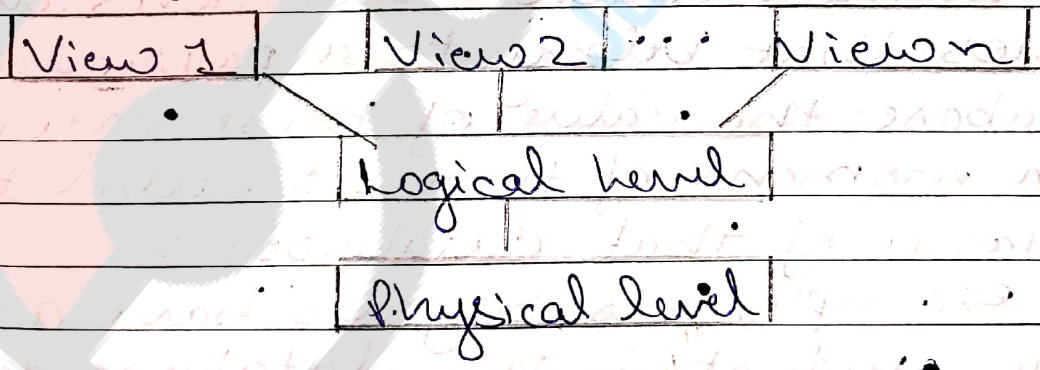
View level:- Highest level of data abstraction. This level describes the user interaction with database system.

Example:- Let's say we are storing customer information in a customer table. At physical level these records can be described as blocks of storage (bytes, gigabytes, terabytes etc) in memory. These details are often hidden from the

programmers.
At the logical level these records can be described as fields and attributes along with their data types, their relationship among each other can be logically implemented. The programmers generally work at this level because they are aware of such things about database systems.
At view level, user just interact with system with the help of GUI and enter the details at the screen, they are not aware of how the data is stored and what data is stored; such details are hidden from them.

Schema

Design of a database is called the schema. Schema is of three types: Physical schema, logical schema and view schema.



Three levels of Data Abstraction

The design of a database at physical level is called physical schema; how the data stored in blocks of storage is described at this level.

Design of database at logical level is called logical schema, programmers and database administrators work at this level, at this level data can be described as certain types of data records gets stored in data structures, however the internal details such as implementation of data structure is hidden at this level (available at physical level).

Design of database at view level is called view schema. This generally describes end user interaction with database systems.

Instance

The data stored in database at a particular moment of time is called instance of database. Database schema defines the variable declarations in tables that belong to a particular database; the value of these variables at a moment of time is called the instance of that database.

For example, lets say we have a single table student in the database, today the table has 100 records, so today the instance of the database has 100 records. lets say we are going to add another 100 records in this table by tomorrow so the instance of database tomorrow will have 200

records in table. In short, at a particular moment the data stored in database is called the instance, that changes every time when we add or delete data from the database.

Q] Database Models :-

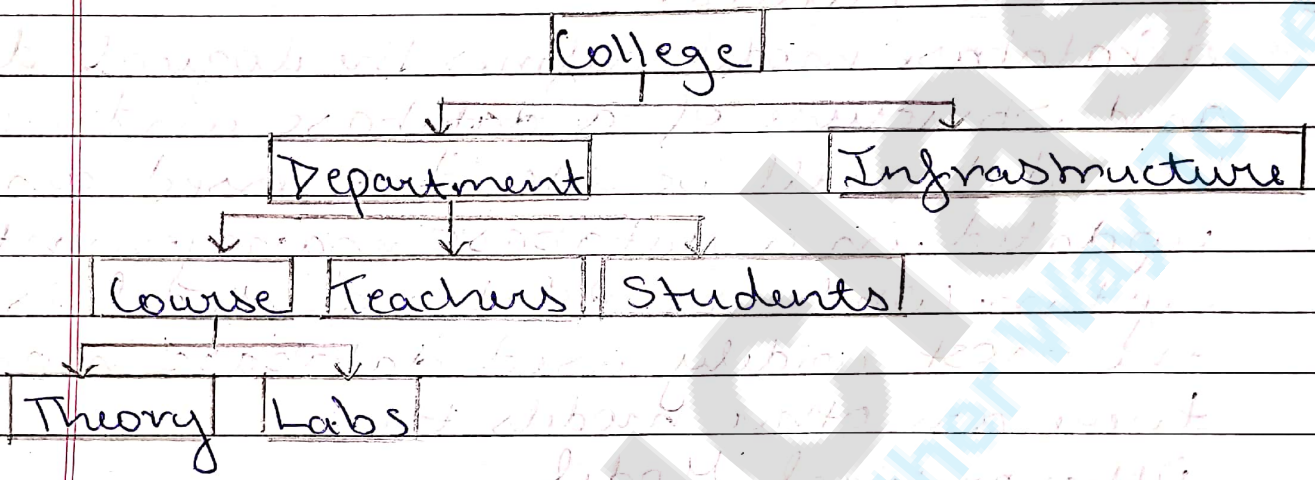
A Database model defines the logical design and structure of a database and defines how data will be stored, accessed and updated in a database management system. While the Relational Model is the most widely used database model, there are other models too:

1) Hierarchical Model
2) Network Model
3) Entity-relationship Model
4) Relational Model

1) Hierarchical Model :-

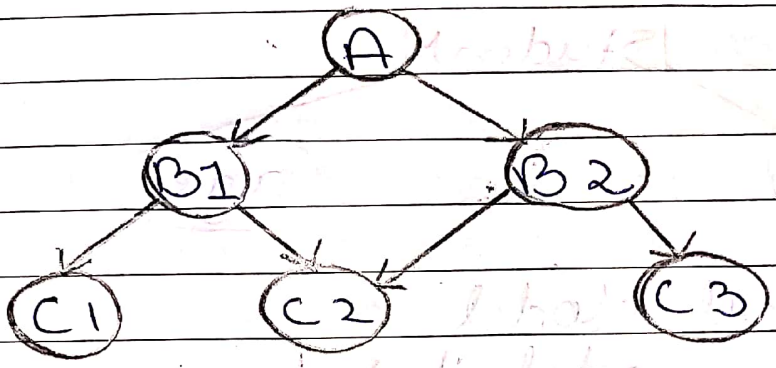
This database model organises data into a tree-like-structure, with a single root, to which all the other data is linked. The hierarchy starts from the root data, and expands like a tree, adding child nodes to the parent nodes. In this model, a child node will only have a single parent node. This model efficiently describes many real-world relationships like index of a book, recipes etc.

In hierarchical model, data is organised into tree-like structure with one-to-many relationship between two different types of data, for example, one department can have many courses, many professors and of course many students.

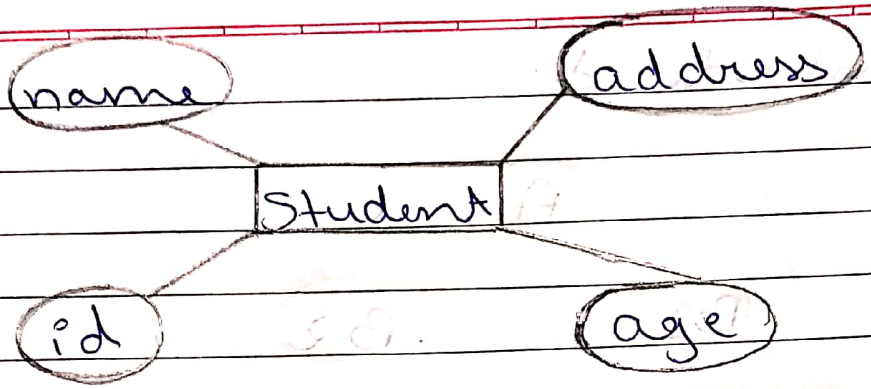


2) Network Model :-
This is an extension of the Hierarchical method. In this model data is organised more like a graph and are allowed to have more than one parent node.
In this database model data is more related as more relationships are established in this database model. Also, as the data is more related, hence accessing the data is also easier and fast. This database model was used to many many-to-many data relationships.
This was the most widely used database model, before Relational Model

was introduced.



3) Entity-relationship Model :- In this database model, relationships are created by dividing object of ~~intered~~ interest into entity and its characteristics into attributes.
Different entities are related using relationships.
E-R Models are defined to represent the relationships into pictorial form to make it easier for different stakeholders to understand.
this model is good to design a database, which can be turned into tables in relational model.
let's take an example, if we have to design a school Database, then Student will be an entity with attributes name, age, address etc. As Address is generally complex, it can be another entity with attributes street name, pincode, city etc and there will be a relationship between them.
Relationships can also be of different types.



4) Relational Model

In this model, data is organised in two-dimensional tables and the relationship is maintained by storing a common field.

The basic structure of data in the relational model is tables. All the information related to a particular type is stored in rows of that table. Hence, tables are also known as relations in relational model.

Student_id	name	age
1	AKon	17
2	BKon	18
3	CKon	17
4	DKon	18

subject_id	name	teacher
1	Java	Mr.J
2	C++	Mr.C
3	C#	Mr.C# (Hash)
4	Php	Mr.Php

student_id	subject_id	marks
1	1	98
1	2	78
2	1	76
3	2	88

5) Database languages

PAGE No.
DATE / /

Read, Update, Manipulate and Store data in a database using Database languages. The following are the database languages:-

1) Data Definition language
2) Data Manipulation language
3) Data Control language
4) Transaction Control language

Data Definition language

The language is used to create database, tables, alter them etc. With this, you can also rename the database, or drop them. It specifies the database schema.

The DDL statements include:

1) CREATE :- Create new database, table etc
2) ALTER :- Alter existing database, table etc.
3) DROP :- Drop the database
4) RENAME :- Set a new name for the table.

Data Manipulation language

The language used to manipulate the database like inserting data, updating table, retrieving record from a table etc is known as Data Manipulation language.

1) SELECT :- Retrieve data from the database
2) INSERT :- Insert data

3) UPDATE :- Updata data
4) DELETE :- Delete all records

Data Control language
Grant privilge to a user using the GRANT statement. In the same way, revoke the privilege using the REVOKE statement. Both of these statements come under the Data Control Language (DCL).
1) GRANT :- Give privilege to access the database.
2) REVOKE :- Take back the privilege to access the database.

Transaction Control language
Manage transactions in the Database using the Transaction Control Language:-
1) COMMIT :- Save the work.
2) SAVEPOINT :- Set a point in transaction to rollback later
3) ROLLBACK :- Restores since last commit

Q] Structure of DBMS :-
At very high level, a database is considered as shown in diagram. These are explained as follows:-

1) Applications :- It can be considered as a user friendly web page where the user enters the requests. Here he simply

enters the details that he needs and presses buttons to get the data.

2) End User :- They are the real users of the database. They can be developers, designers, administrator or the actual users of the database.

3) DDL :- Data Definition language is a query fired to create database, schema, tables, mappings etc in the database. These are the commands used to create the objects like tables, indexes in the database for the first time. In other words, they create structure of the database.

4) DDL Compiler :- This part of database is responsible for processing the DDL commands. That means these compiler actually breaks down the command into machine understandable codes. It is also responsible for storing the metadata information like table name, space used by it, number of columns in it, mapping information etc.

5) DML Compiler :- When the user inserts, deletes, updates or retrieves the record from the database, he will be sending request which he understands by pressing some buttons. But for the

database to work/understand the request, it should be broken down to object code. This is done by this compiler. One can imagine this as when a person is asked some question, how this is broken down into waves to reach the brain!

6.) Query Optimizer :- When user fires some request, he is least bothered how it will be fired on the database. He is not all aware of database or its way of performance. But whatever be the request, it should be efficient enough to fetch, insert, update or delete the data from the database. The query optimizer decides the best way to execute the user request which is received from the DML compiler. It is similar to selecting the best nerve to carry the waves to brain!

7.) Stored Data Manager :- This is also known as Database Control System. It is one of the main central system of the database. It is responsible for various tasks.

- It converts the requests received from query optimizer to machine understandable form. It makes actual request inside the database. It is

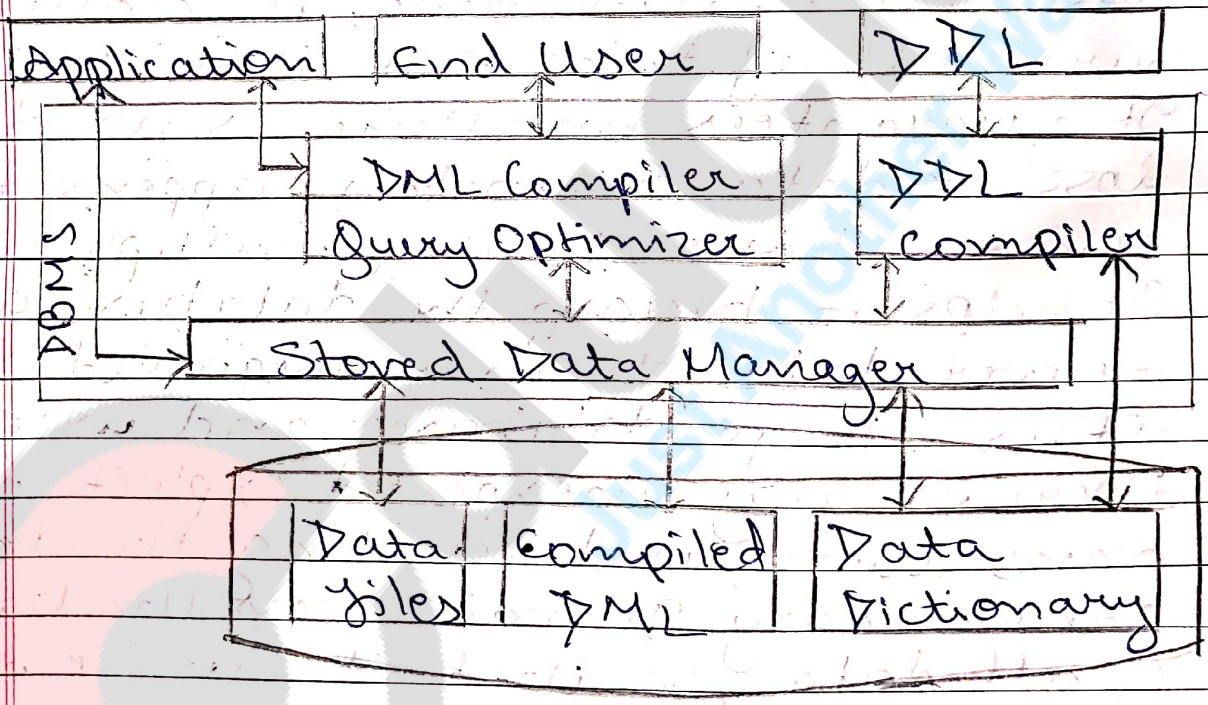
like fetching the exact part of the brain to answer.
It helps to maintain consistency and integrity by applying the constraints. That means, it does not allow inserting / updating / deleting any data if it has child entry. Similarly it does not allow entering any duplicate value into database tables.
It controls concurrent access. If there is multiple users accessing the database at the same time, it makes sure, all of them see correct data.
It guarantees that there is no data loss or data mismatch happens between the transactions of multiple users.
It helps to backup the database and recover data whenever required. Since it is a huge database and when there is any unexpected exploit of transaction, and reverting the changes are not easy. It maintains the backup of all data, so that it can be recovered.

8) Data files :- It has the real data stored in it. It can be stored as magnetic tapes, magnetic disks or optical disks.

9) Compiled DML :- Some of the processed DML statements (insert, update,

delete) are stored in it so that if there is similar requests, it will be re-used.

10) Data Dictionary:- It contains all the information about the database. As the name suggests, it is the dictionary of all the data items. It contains description of all the tables, view, materialized views, constraints, indexes, triggers etc.



Q) Role of DBA:-
The role of DBA are as follows:-
1) Deciding the hardware device:-
Depending upon the cost, performance and efficiency of the hardware, it is DBA who have the duty of deciding which hardware device will suit the company.

requirement. It is hardware that is an interface between end users and database, so it is needed to be of best quality.

2) Managing Data Integrity :-
Data integrity should be managed accurately because it protects the data from unauthorized use. DBA manages relationship between the data to maintain data consistency.

3) Decides Data Recovery and Backup method :-
If any company is having a big database, then it is likely to happen that database may fail at any instance. It is required that a DBA takes backup of entire database in regular time span. DBA has to decide that how much data should be backed up and how frequently the back up should be taken. Also the recovery of database is done by DBA if they have lost the database.

4) Tuning Database Performance :-
Database performance plays an important role for any business. If user is not able to fetch data speedily then it may be a loss for company's business. So by tuning and modifying sql commands a DBA can improve the

performance of database.

5) Capacity issues:-
All the databases have their limits of storing data in it and the physical memory also has some limitations. DBA has to decide the limit and capacity of database and all the issues related to it.

6) Database design:-
The logical design of the database is designed by the DBA. Also a DBA is responsible for physical design, external model design and integrity control.

7) Database accessibility:-
DBA writes subschema to decide the accessibility of database. He decides the users of the database and also which data is to be used by which user. No user has the power to access the entire database without the permission of DBA.

8) Decides validation checks on data:-
DBA has to decide which data should be used and what kind of data is accurate for the company. So, he always puts validation checks on data to make it more accurate and consistent.

PAGE No.
DATE / /

9) Monitoring performance :-
If database is working properly then it doesn't mean that there is no task for the DBA. Yes ofcourse, he has to monitor the performance of the database. A DBA monitors the CPU and memory usage.

10) Decides content of the database :-
A database system has many kind of content information in it. DBA decides fields, types of fields and range of values of the content in the database system. One can say that DBA decides the structure of database files.

11) Provides help and support to user :-
If any user needs help at any time then it is the duty of DBA to help him. Complete support is given to the users who are new to database by the DBA.

12) Database implementation :-
Database has to be implemented before anyone can start using it. So DBA implements the database system. DBA has to supervise the database loading at the time of its implementation.

13) Improve query processing performance

PAGE No.
DATE / /

Queries made by the users should be performed speedily. Users need fast retrieval of answers so DBA improves query processing by improving their performance.