Unit : 1 Introductory Concepts of DBMS

Date
Page

* Write the application of DBMS :

→ Enterprise information :
Sales : customers, products, purchases
Accounting : payments, receipts, assets
Human resources : information about employees, salaries, payroll taxes

→ Manufacturing :
Management of production, inventory, orders, supply chain

→ Banking and finance :
customer information, accounts, loans and banking transactions
credit card transactions.
Finance : sales and purchase of financial instruments (e.g. stocks and bonds, storing real time market data)

→ Universities : registration, grades
→ Airlines : reservations, schedules.
→ Document database
→ Tele-communication :
records of calls, texts and data usage, generating monthly bills, maintaining balances on prepaid calling cards.
→ Web based services :
Online retailers : order tracking, customized recommendations
Online advertisements.
→ Navigation systems : For maintaining the locations of various places of interest along with the exact routes of roads, train systems, buses etc.

* Purpose of database / differentiate between file system and DBMS.
(disadvantages of file system)

1) Data redundancy : when the duplicate data exists in multiple places in the database.

2) Data inconsistency : when the duplicate data exists in different formats in multiple tables.

3) Data isolation : because data are scattered in various files and files may be in different formats. Therefore, writing new applications programs to retrieve the appropriate data is Difficult.

4) Difficulty in accessing the data : The conventional file processing environments do not allow needed data to be retrieved in a convenient and efficient manner. More responsive data retrieval systems are required for general use.

5) Integrity problems : There may be cases when some constraints need to be applied on the data before inserting it in database.
→ The file system does not provide any procedure to check these constraints automatically.
~~e. The database (DBMS) provides~~
→ Whereas the DBMS maintains data integrity by enforcing user defined constraints on data by itself.

6) Backup problems : The file system does not provide backup whereas DBMS provides.

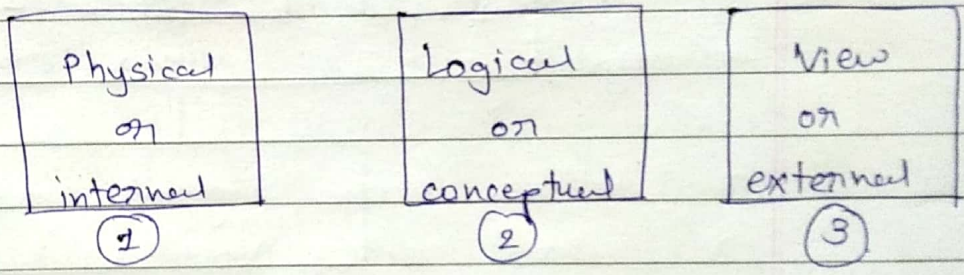
7) Data recovery : The file system doesn't provide data recovery if it is lost whereas DBMS provides.

8) Security issues : The file system provides less security in comparison to DBMS.

* Data Abstraction :

→ It is a process of hiding unwanted or irrelevant details from the end users for providing the feature of simple and easy interface.

→ 3 levels :



1) Physical or Internal level :

→ It is the lowest level of abstraction for DBMS which defines how the data is actually stored, it defines data-structures to store data and access methods used by the database.

2) Logical or Conceptual level :

→ It is the intermediate level or next higher level which describes what data is stored in the database and what relationship exists among those data.

→ Overall, the logical level contains fields and

attributes and relationships among table attributes.

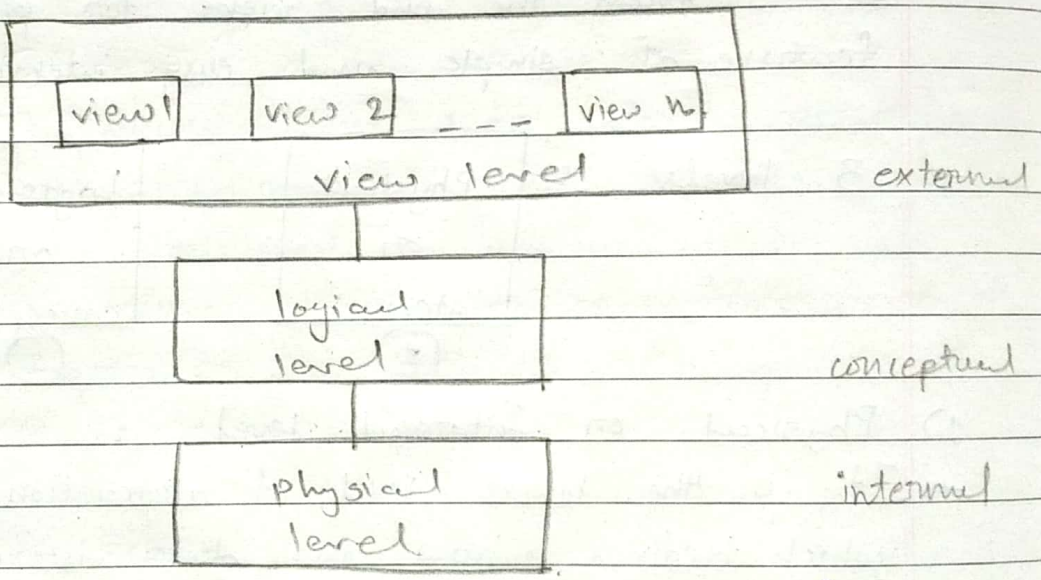
3) External or View level :

→ It is the highest level in which there are different levels of views and every views only defines a part of the entire data.

→ It is the least complex and easy to understand.

→ Ex :- a user can interact with a system using GUI that is view level and can enter details at GUI or screen.

→ Diagram :-



(Hierarchy of abstraction)

* Concept of mapping :

1) External conceptual mapping :

→ it defines the correspondence between a particular external view and conceptual view for representing the data in efficient manner.

→ it tells the DBMS which objects on the

conceptual level correspond to the objects requested on a particular user's external view for the information representation in the same data file.

2) Conceptual - internal mapping :
→ it defines the correspondence between the conceptual view and internal view.
→ it tells the DBMS that how the conceptual records are physically represented.

* Data independence :

1) Physical data independence :
→ the ability to change the physical level without affecting the logical level schemas
→ Hence, modification in the physical level should not result in any changes in the logical or view levels.

2) Logical data independence :
→ the ability to change the logical level without affecting the external level schemas.
→ Hence, modification in the logical level should not result in any changes in the view / external level.

* relation → table
* tuple → Row / Record
* Attribute → Column

* Different types of keys in DBMS :-

1. Super key :

→ A super key is a set of attributes that can identify each tuple uniquely in the given relation.
→ A super key may consist of any number of attributes.
→ All the attributes in a super key are definitely sufficient to identify each tuple uniquely in the given relation but all of them may not be necessary.

→ Ex :- Student (rollno, name, age, address, class, section)

→ super keys :

1) (rollno, name, age, address, class, section)
2) (class, section, rollno)
3) (name, address)

2. Candidate key :

→ A minimal superkey is called as a candidate key.

→ A set of minimal attributes that can identify each tuple uniquely in the given relation is called as a candidate key.

→ All the attributes in a candidate key are

sufficient as well as necessary to identify each tuple uniquely. Removing any attribute from the candidate key fails in identifying each tuple uniquely.

→ It is possible to have multiple candidate keys in a relation.

→ Those attributes which appears in some candidate key are called as prime attributes/ key attributes (ಮಹತ್ವದ non-prime)

→ Ex :- Student (rollno, name, age, address, class, section)

→ Candidate keys :

1) (class, section, roll no)
2) (name, address)

3. Primary key :

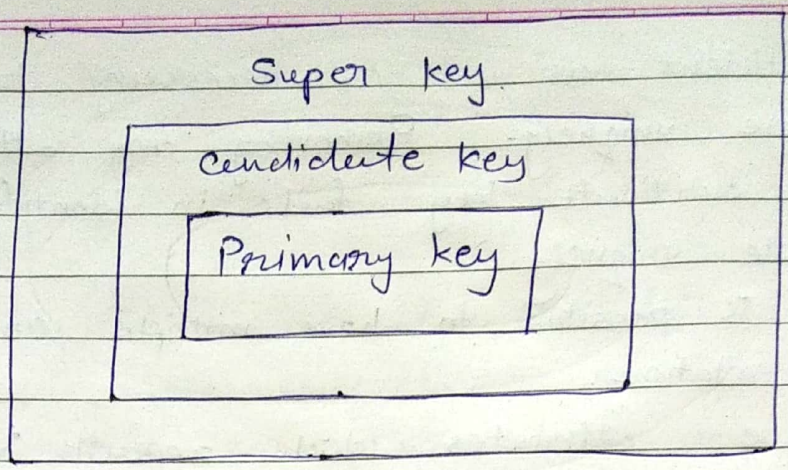
→ An attribute or a group of attributes those are able to identify two tuples uniquely is called a primary key.

→ The value of primary key can never be NULL.

→ The value of primary key must always be unique and can never be changed (no updation is possible.)

→ The value of primary key must be assigned while inserting.

→ A relation is allowed to have only one primary key.



☆ → two constraints :
1) It is always not null → You can't leave it blank
2) It is always unique → It must not be repetitive.

4. Alternate key :

→ Unimplemented candidate keys are called as alternate keys.

OR

→ Candidate keys that are left unimplemented or unused after implementing the primary key are called as alternate keys.

* Example :

empid	ename	edob	eaddress	deptno
e0001	Ramesh	13/11/80	17 MG Road	D001
e0002	Naresh	27/11/78	13 LBS Marg	D002
e0003	Mahesh	11/11/86	125 V Road	D001
e0004	Ramesh	13/11/80	27 MI Road	D002
e0005	Dinesh	31/5/81	17 MG Road	D001

should be in capital

→ Employee_schema = (empid, ename, edob, eaddress, deptno)

→ Super key : empid
empid + ename
empid + ename + edob
empid + eaddress
empid + deptno
ename + edob + eaddress + deptno
empid + ename + edob + eaddress + deptno
and so on...

→ Candidate key : empid
~~empid~~ ≠ ename + edob + deptno
ename + eaddress

→ Primary key : ename + eaddress // can be any of the candidate key

→ Alternate key : empid
ename + edob + deptno.

* Database schema : The logical design for a database is called database schema.
OR The scheme designed for a database is the database schema.

* Instances : "The data in a database at a particular time moment is called instance."

→ Since, the values or contents of a database may change time to time (more frequently or less frequently), it is quite obvious to have

different instances of database at different time.

* Basic building components of database system :-

1. People who interact with database :

1) Database users :

*→ Naive users / unsophisticated users / parametric users :

→ Naive users are the users who use database without seeing technical details and internal implementation of how they are using a database.

Ex :- bank customer using ATM.

*→ Sophisticated users :

→ Sophisticated users are those who use the database by writing queries.

Ex :- engineers, scientists use MYSQL, SQL. by writing queries having through knowledge of database management system

*→ Application programmers :

→ Application programmers are the people who develop application programs.

→ These people also have knowledge of database that how the application will provide the way of interfacing between users and database.

2) Database administrator : (DBA)

→ A DBA is a person who monitors and controls the database and is responsible for deciding the levels of access (read, read and write, read write and edit) to a particular person in an organization.

→ Responsibilities :

1) Schema definition : defining a structure of a database.

2) Modification of logical level and physical level schema : A DBA has to modify the structure of a database by executing queries.

3) Routine maintenance :

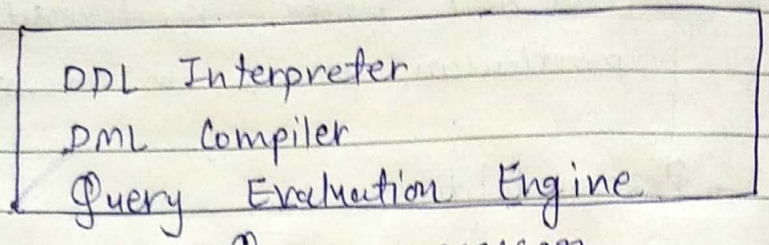
→ The activities which are done on a regular or on a fixed time interval are called routine activities.

- taking backup of data
- monitoring the resource consumption.
- availability of enough disk storage and prevention of data by unauthorized parties is also done.

4) Granting different level of authorization for a specific person : A DBA has to assign different levels of access to different people as every person should not be able to access all the data.

2. Query processor :

→ A query processor processes a query coming from one of the database users.



Query processor

1) DDL interpreter :

→ It interpretes DDL statements.
(Data Definition Language statements).

→ DDL is used to define the structure of a database and they put the definitions in the data dictionary.

it contains metadata

the data about data.

→ Ex :-

name char(24)

metadata → datatype is 'char' and length is '24'

which is put in the data dictionary by the DDL interpreter.

2) DML Compiler :

→ It translates DML statements.
(Data manipulation Language Statements)
(in short queries).

DML statements → [DML Compiler] → Low level instructions.

→ DML compiler provides feature of query optimization.

3) Query Evaluation Engine :

Low level instructions → [Query Evaluation Engine] → Evaluated query.

→ which is ultimately ready for fetching data from storage units.

3. Storage manager :

→ It is responsible for providing interface between the query coming from query processor and the data that is stored in disk storage.

** File manager :
→ It is responsible for managing the space on disk and for deciding the storage structure and the data structure to store the data.

** Buffer manager :
→ A buffer manager decides the data those have to be in cache memory. So, if a user ask for the same data it can be fetched from the cache and the time is reduced.

** Transaction manager :
→ It monitors the transaction that goes on a particular time.
→ If system gets crashed during the transaction then during the recovery, the sum of rupees should be rolled back to the user.

** Authorization and integrity manager :
→ It ensures the integrity to be maintained in the database and the authorization manager checks the user weather it is authorized for the specific access or not.

4. Disk storage :

Disk storage includes → Data files, Data dictionary

→ Data files are the main storage where actual data reside.

→ Data dictionary stores the meta data like structure of a database.

→ As in the textbook, we find the topics with the help of index, same like that a database has the indices to find the data fast and efficiently.

— x —