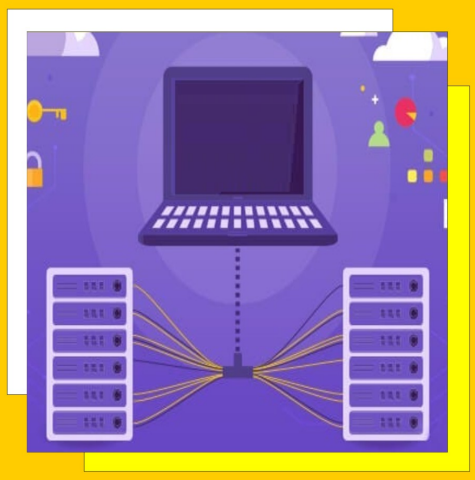
DATA BASE MANAGEMENT SYSTEMS

COMPUTER SCIENCE & IT

COMPUTER SCIENCE & IT



GATE / PSUs

STUDY MATERIAL

DATA BASE MANAGEMENT SYSTEMS

ENGINEERS INSTITUTE OF INDIA

ENGINEERS INSTITUTE OF INDIA

COMPUTER SCIENCE & IT

GATE & PSUs

STUDY MATERIAL

DATA BASE MANAGEMENT SYSTEMS

CONTENT

CHAPTER-1
FUNDAMENTALS OF DBMS

Database:

Database is a collection of related data while data is raw fact or figures or entity.

Information:

Processed data or meaningful data is called as information.

Example: 25 years then it is only data. If I say my age is 25 years then it will become information.

DBMS (database management system)

A Database – management system (DBMS) is a collection of interrelated data and a set of application programs used to access, update and manage those data. The goal of DBMS is to provide an environment that is both convenient and efficient to retrieve and store information into the database.

Disadvantage of File Processing System over Database Management System

Traditionally, the organization used files to store the data. In single organization, multiple files can be there. Each file is independent from one another. Files are manipulated using the program written in C, C++, and JAVA.

1. Data redundancy and inconsistency
2. Difficulty in accessing data
3. Data Isolation
4. Integrity problems
5. Atomicity problems
6. Concurrent access anomalies
7. Security Problems

Data Redundancy: The files are created in the file system as and when required by an enterprise over its growth path. So in that case the repetition of information about an entity cannot be avoided.

Example: The addresses of bank customers will be present in the two files makes wastage of storage:
a) One in savings account files and
b) Other in current account, files.

Data Inconsistency: Data redundancy leads to greater problem than just wasting the storage i.e. it may lead to inconsistent data. Same data which has been repeated at several places may not match after it has been updated at some places. Disadvantage of this is mismatch of data.

Example: If a customer requests to change the address for his account in the Bank and it has been updated at the saving bank account file only but his current bank account file is not updated. Afterwards the addresses of the same customer present in saving bank account file and current bank account file will not match. Finally there is no way to find out which address is latest out of these two.

Difficulty in Accessing Data: For analyzing the data the programs are not already present and only programmers will have to write a new program to generate requested report or will have to work manually. This is going to take impractical time and will be more expensive.

Data Isolation: Since the data files are created at different times by different people in different formats the structures of different files generally will not match. The data will be scattered in different files for a particular entity. So it will be difficult to obtain appropriate data.
Example: Suppose the Address in Saving Account file have fields: *Add line1, Add line2, City, State, Pin*
While the fields in address of Current account are: *House No., Street No., Locality, City, State, Pin.*
Administrator is asked to provide the list of customers living in a particular locality. Providing consolidated list of all the customers will require looking in both files. But they both have different way of storing the address. Writing a program to generate such a list will be difficult.

Security and Access Control: Database should be protected from unauthorized users for privacy and security purpose. Every user should not be allowed to access every data.
For example:

1. The Payroll Personnel in a bank should not be allowed to access accounts information of the customers
2. An employee can't see salary of other employees

View of Data

A major purpose of a database system is to provide users with an abstract view of the data. That is the system hides certain details how data are stored and maintained. Data abstraction is provided at different level of view.

1. **Physical View (internal view)**
   At physical level, we have very low level details. It provides the details about how data are actually stored.
2. **Logical View**
   At this level, we know about what data are stored and what relationship exist among those data. This level hides the details of physical level.
3. **View level (External View)**
   This level describes the part of the database that is relevant to each user. This level hides the details of logical level and physical level.

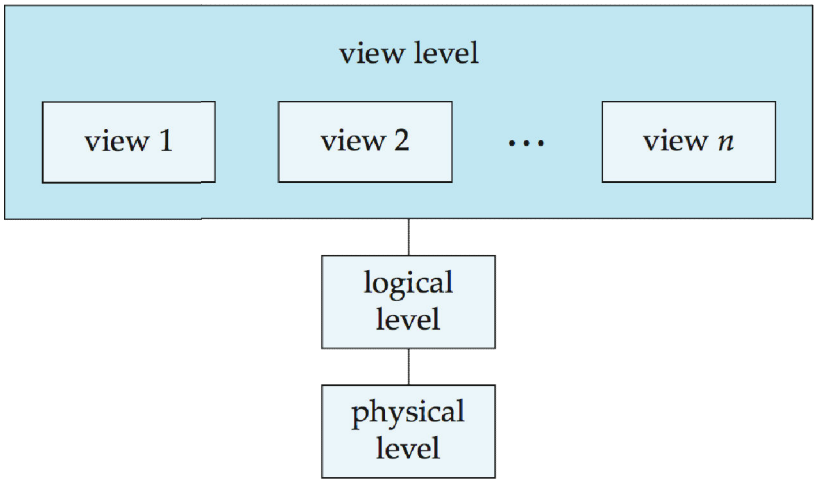


Figure: Abstraction level of Database

Instance

The collection of information stored in the database at a particular moment is called an instance of the database.

Schema

The overall design of the database is called schema.

a. **Physical schema: -** Physical view is represented by physical schema.
b. **Logical Schema: -**Logical view is represented by logical schema. Programmers construct the application using this schema.
c. **Sub Schema: -** External view is represented by sub schema. Schema at view level called subschema's

Database Users:-

There are five types of users that interact with database as follow:

1. **Naïve Users/unsophisticated user: -** indirectly interact with database using application program.

2. **Application programmers: -** application programmers are responsible for developing the application program and user interface. They use RAD (rapid application development) tools to construct forms and reports with minimal programming effort.

3. **Sophisticated user: -** Interact with system without writing program. Sophisticated user interacts to database by writing query.

4. **Specialized user: -** users Interact with system by writing application program.

5. **Database administrator:** These are the users who are the incharge for creating and maintaining the database. DBA can assign and remove access level to database user.DBA responsibility includes following :
    1. Schema definition

2. Storage structure and access-method definition
3. Schema and physical organization modification
4. Granting of authorization for data access
5. Routine maintenance(periodically backup, upgrade disk space, monitoring the job)

Functional Component of Database

Functional component of a database system can broadly divided into the storage manager and the query processor components. Functional component of a database system is shown in ***figure: system structure***.

A. Storage Manager

1. A storage manager is a program module that provides the interface between the low level Data stored in the database and the application program and query submitted to the system.
2. The storage manager is responsible for interaction with file manager. The storage manager translates the various DML statements into low level file system command.
3. The storage manager is responsible for storing, retrieving, and updating data in database.
4. The storage manager components include:
   a. Authorization and integrity manager
   b. Transaction manager
   c. File manager(storage on disk system)
   d. Buffer manager(use for fetching data from disk storage to memory)
5. Data structure implemented by storage manager includes:
   a. Data files, which stores the database itself.
   b. Data dictionary, which stores metadata and schema of database
   c. Indices are used to provide fast access to database.

B. The query processor

Query processor is used to facilitate access to data stored into database. The query processor components include:

1. **DDL interpreter**
   It interprets the DML statements and records the definition in the data dictionary.
2. **DML compiler**
   It converts DML statement into low level that query evaluation engine can understand and performquery optimization.
3. **Query evaluation engine**
   It is responsible for executing low level instruction.

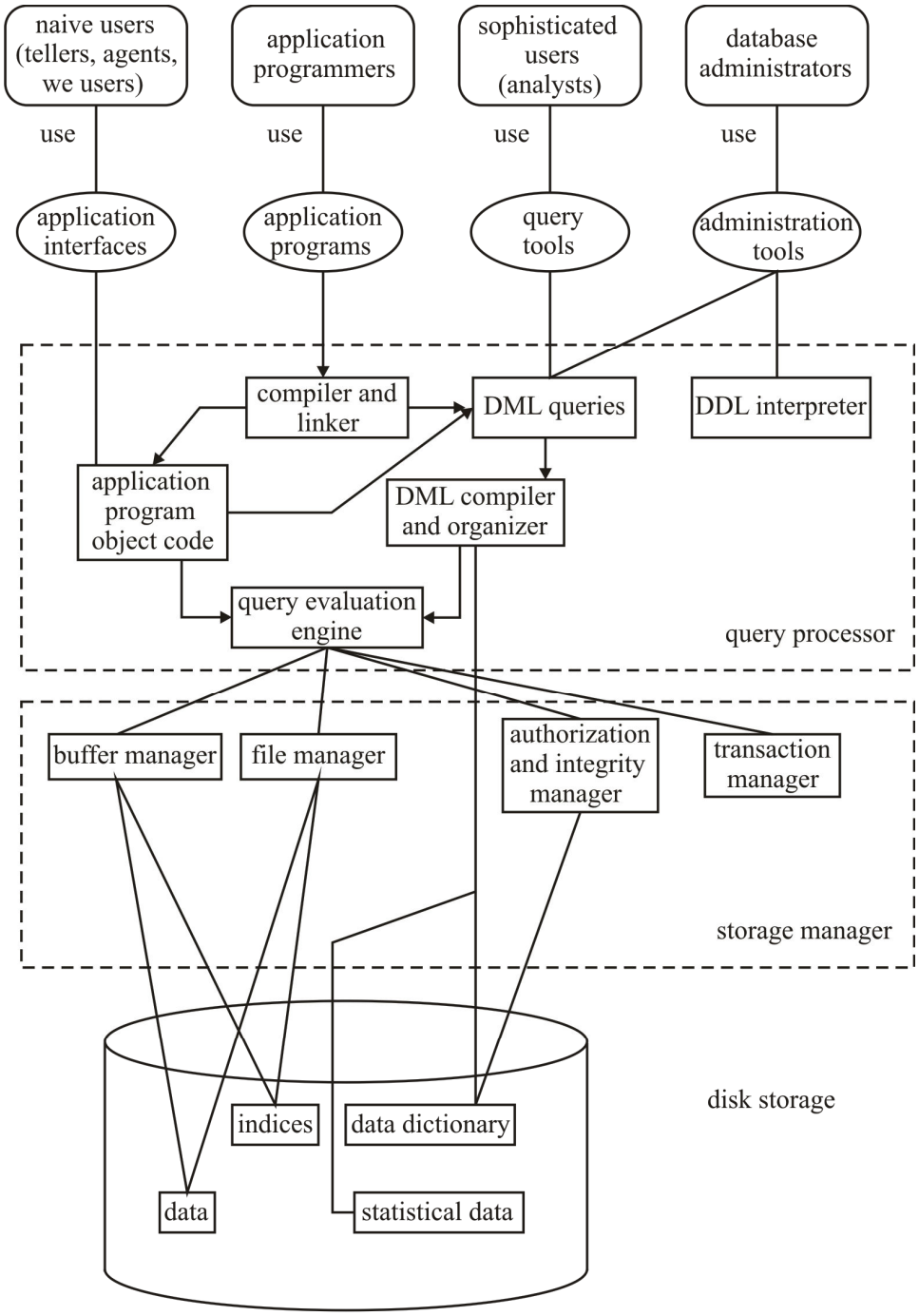
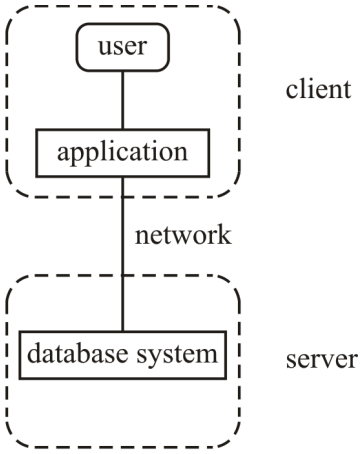


Figure: System structure

Database Architecture

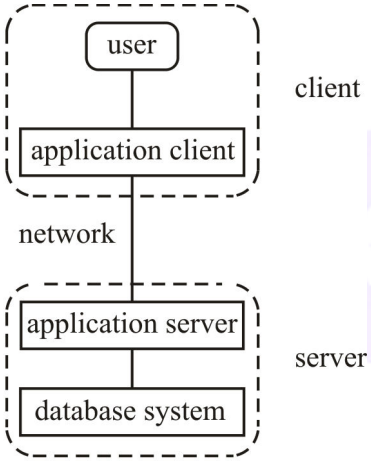
1. **Two tier architecture**

   In two tier architecture the application resides at client and database resides at server. Whenever client required accessing the database, it invokes the server for database access using query language statement.



(a) Two-tier architecture

2. **Three tier architecture**
   In the three tier architecture, the client machine works as a front end (user interface). Application logic (business logic resides) at application server machine and database resides at database server. Client machine interact with database server using application server.



(b) Three-tier architecture

3. **N-tier architecture**

Data Models

1. Relational Model
2. The Entity-Relationship Model
3. Object-Based Data Model
4. Semi structured Data Model(XML)

Relational model: The relational model uses a collection of tables. Each table is called as relation and each row is called as tuple and each column is called as attribute.

Entity Relationship Model: Entity relationship model uses a entity to show the object. Each object have attributes. This model also represents relationship between entities.

Object based data model: In this data model we combine the feature of object oriented data model and relation data model.

Semi structured data model: Semi structure data model permits the specification of data where individual data items of same type may have different attributes.

Database Language

There are two types of database language:

1. **DDL (Data Definition Language)**is used to specify database schema.DDL provides facility to specify the constraints on tables. It includes create, alter, drop, remove.
   Example:

   CREATE TABLE *student*
   (student_id ***int****,*
   student_name ***char(20)*****,**
   department_name ***char(20));***

   Constraints that can be specified by DDL include:

   I. Domain constraints
   Declaring an attribute to be of a particular domain (integer type, character type, and date/time types) acts as a constraints on the values that it can take. Domain constraints are tested by the system whenever a new data item is entered into the database.

   II. Referential integrity
   When we required to have same value for some attribute in two different table then the constraints on table is called as referential integrity constraints. These constraints sometimes called as foreign key constraints.
   Referential integrity constrain requires that the values appearing in specified attributes of any tuple in the referencing relation also appears in specified attributes of at least one tuple in the referenced relation.
   For example department name values in the student table should be the same as department name value in the department.

   III. Assertions
   An assertion is any condition that the database must always satisfy. For example minimum balance of saving account should be 1000.

   IV. Authorizations
   It is used to check the access of each user.

Example: The following table has two attributes A and C where A is the primary key and C is the foreign key referencing A with on – delete cascade.

A	C
2	4
3	4
4	3
5	2
7	2
8	5
9	4

The set of all tuples that must be additionally deleted to preserve referential integrity when the tuple (2, 4) is deleted is: **GATE – 2005**

(a.) (3, 4) and (6, 4)
(b.) (5, 2) and (7, 2)
(c.) (5, 2) (7, 2) and (9, 5)
(d.) 1

Solution: c

If (2, 4) is deleted then 2 is the primary key but in (5, 2) and (7, 2) 2 is the foreign key so these must be deleted. The primary key for (5, 2) and (7, 2) is and (7, 2) is 5 and 7 respectively but in (9, 5) 5 is the foreign key so it is also deleted.

Note: ***The output of DDL is placed in the data dictionary which contains meta data .Meta data is data about data.***

2. **DML (Data Manipulation Language)** is used to express database query and update to manipulate the data into database. It includes select, insert, delete, update.

   SELECT student.student_name
   FROM student
   WHERE student.department_name =’computer science’;

   There are two types of DML:
   a. **Procedural DML** require a user to specify what data are needed and how to get those data.
   b. **Declarative DML/nonprocedural DML** requires only to specify what data are needed.

3. **Data Control Language:** The Data Control Language (DCL) component of the SQL language is used to create privileges to allow users access to, and manipulation of, the database. There are two main commands:
   a. **GRANT** to grant a privilege to a user
   b. **REVOKE** to revoke (remove) a privilege from a user

Key Point

1. A data model is a collection of concepts that can be used to describe the structure of database.
2. A view can represent the subset of data stored in a table. View can join and simplify multiple tables into a single virtual table.
3. A query is a statement requesting the retrieval of information. The portion of DML that involves information retrieval is called a query language.

Level – 1

(1.) Which component of database system is responsible for interaction with file manager?
(a.) DDL interpreter (b.) compiler and linker
(c.) Storage manager (d.) buffer manager

(2.) Which of the following translate DML statement into the language understand by query evaluation engine.
(a.) DML interpreter (b.) DDL compiler
(c.) DML compiler (d.) DDL interpreter

(3.) Query optimization is performed by
(a.) DML compiler (b.) DML interpreter
(c.) DDL compiler (d.) Storage manager

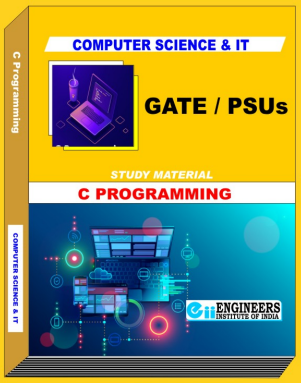
(4.) Which of the following is semi structured data model?
(a.) XML (b.) C++ (c.) HTML (d.) Relational model

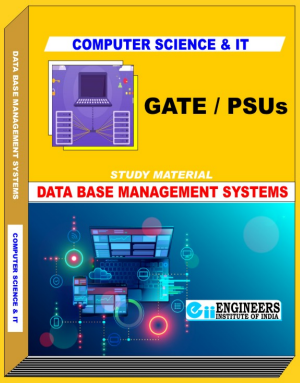
(5.) Data manipulation language used for
(a.) Describing physical structure of database
(b.) Modification of physical database structure of database
(c.) Defining the physical structure of database
(d.) Manipulation and processing of database

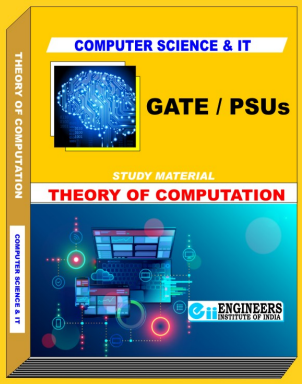
(6.) The database schema is represented in:
(a.) Data definition language (b.) Data manipulation language
(c.) Data control language (d.) Transaction control language

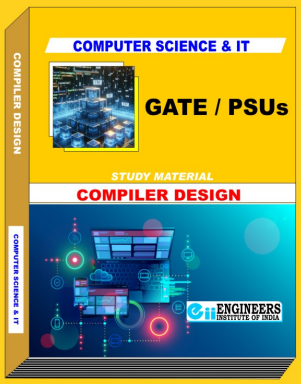
(7.) In the database architecture view level is also called as:
(c.) Conceptual level (b.) External level (c.) Internal level (d.) Logical level

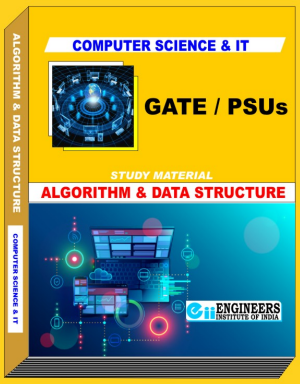
Published Books

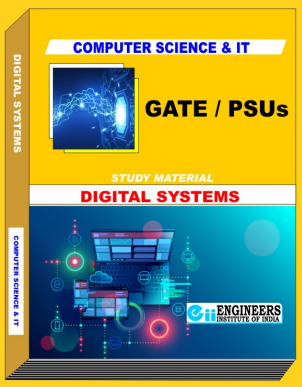


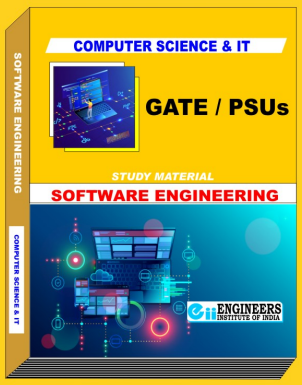


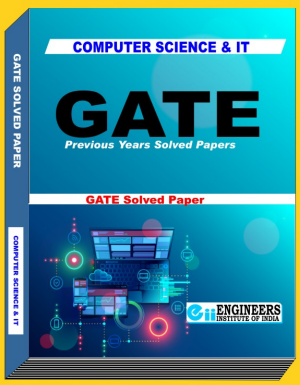


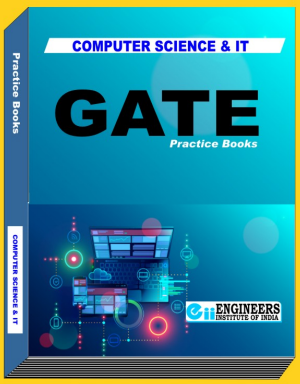












ENGINEERS INSTITUTE OF INDIA

Classroom Batches | Online Classes

Postal Course Classes | Online Test Series

Office: 58B, Kalu Sarai Near Hauz Khas Metro Station New Delhi-16

Helpline: 9990657855 , 9990357855

www.engineersinstitute.com