

DBMS Module - 1

Introduction

Database

- A database is a collection of related data.
- A database is a collection of related, logically coherent data used by the application programs in an organization

Database management System

- A DBMS is a collection of interrelated data and a set of programs to access those data.
- DBMS is hence a *general purpose software system* that facilitates the process of defining, constructing, manipulating and sharing databases among various users and applications.

DBMS Functions

- Defining : a database involves specifying the data types, structures, and constraints for the data to be stored in the database.
- Constructing : the database the process of storing data itself on some storage medium that is controlled by the DBMS.
- Manipulating : a database includes querying, updating the database.
- Sharing : a database allows multiple users and programs to access the database concurrently.

Database Management System

Database Management System (DBMS) provides....

... efficient, reliable, convenient, and safe multi-user storage of and access to massive amounts of persistent data.

Adjective of DBMS

- Massive
- Persistent
- Safe
- Multi-user
- Convenient
- Efficient
- Reliable

- Database applications may be programmed via “frameworks”
- DBMS may run in conjunction with “middleware”
- Data-intensive applications may not use DBMS at all

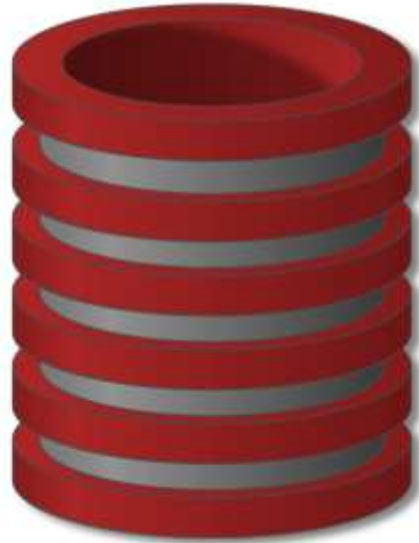
Key concepts

- Data model
- Schema versus data
- Data definition language (DDL)
- Data manipulation or query language (DML)

Key people

- DBMS implementer
- Database designer
- Database application developer
- Database administrator

**Whether you know it or not,
you're using a database every day**



Database system Vs File System

	File System <i>(like data in text file as in C programs)</i>	Database System
Data redundancy <i>(lead to higher storage and access cost)</i>	Same information is duplicated in many places eg. Mobile and email id of student in cse_student_detail records and in Btech_student_detail records as well.	Data redundancy can be deal by creating a separate table having student#, mob & email.
Difficulty in Accessing data <i>(Need to write a new program to carry out each new task)</i>	If user need all student having GPA>3.5 then we need to create such program or if this program exist then also file system can meet with changing needs as need student having GPA>3.5 and are from Delhi.	Searching required data is easy by writing small queries system get adapted to changing needs

	File System (<i>like data in text file as in C programs</i>)	Database System
Data Isolation	As data are scattered in different files and file may be of different formats as some in .doc or .txt or .xls need require coding for each of them.	Uniformity in the way data is stored
Integrity problems	<p>Data values must follow some consistency constraint such as no account should have less than \$25. in File System we need to code it and if in future we want to change it we need to recode it!</p> <ul style="list-style-type: none"> • <i>Integrity constraints (e.g. account balance > 0) become part of program code</i> • <i>Hard to add new constraints or change existing ones</i> 	New constraint can easily be add, modify & drop easily.
Atomicity Problem	<p>Computer systems are prone to failures. Suppose a program transferring \$50 from account A to B but in middle system crash then \$50 was removed from but not credited to B. This lead to <i>inconsistent state</i>. It is difficult to ensure atomicity in File system.</p> <ul style="list-style-type: none"> • <i>Transfer of funds from one account to another should either complete or not happen at all</i> 	Atomicity can easily be maintained, these system have recovery and back up tools.

	File System (like data in text file as in C programs)	Database System
Concurrent - access anomalies	Consider an account A holding \$500, if two customer C1 & C2 withdraw \$100 and \$50 from A simultaneously then initially C1 & C2 see \$500 now whichever write last it either show \$400 or \$450 while correct is \$350. Hence File system have greater challenge as many application programs access same data simultaneously in multi-user system.	Support multi-user system.
Security problem	Enforcing security constraint is difficult. As faculty can upload and see attendance of student, whereas only <i>class advisor</i> can modify already uploaded attendance and <i>students</i> can only see their own attendance.	Database has internal procedures and commands for this.

- Which one of the following refers to the copies of the same data (or information) occupying the memory space at multiple places.
- Data Repository
- Data Inconsistency
- Data Mining
- Data Redundancy

- Which one of the following refers to the copies of the same data (or information) occupying the memory space at multiple places.
- Data Repository
- Data Inconsistency
- Data Mining
- Data Redundancy

- The database management system can be considered as the collection of _____ that enables us to create and maintain the database.
- Translators
- Programs
- Keys
- Language activity

- The database management system can be considered as the collection of _____ that enables us to create and maintain the database.
- Translators
- Programs
- Keys
- Language activity

- Which of the following is a function of the DBMS?
 - a) Storing data
 - b) Providing multi-users access control
 - c) Data Integrity
 - d) All of the above

- What is information about data called?
 - a) Hyper data
 - b) Tera data
 - c) Meta data
 - d) Relations

- Which of the following is a function of the DBMS?
 - a) Storing data
 - b) Providing multi-users access control
 - c) Data Integrity
 - d) All of the above

- What is information about data called?
 - a) Hyper data
 - b) Tera data
 - c) Meta data
 - d) Relations

- _____ is a hardware component that is most important for the operation of a database management system.
 - a) Microphone
 - b) High speed, large capacity disk to store data
 - c) High-resolution video display
 - d) Printer

Which of the following is a feature of the database?

- a) No-backup for the data stored
- b) User interface provided
- c) Lack of Authentication
- d) Store data in multiple locations

- _____ is a hardware component that is most important for the operation of a database management system.
 - a) Microphone
 - b) High speed, large capacity disk to store data
 - c) High-resolution video display
 - d) Printer
- Which of the following is a feature of the database?
 - a) No-backup for the data stored
 - b) User interface provided
 - c) Lack of Authentication
 - d) Store data in multiple locations



Question

➤ **Which of the following is generally used for performing tasks like creating the structure of the relations, deleting relation?**

1. DML(Data Manipulation Language)
2. Query
3. Relational Schema
4. DDL(Data Definition Language)

➤ **Which of the following provides the ability to query information from the database and insert tuples into, delete tuples from, and modify tuples in the database?**

1. DML(Data Manipulation Language)
2. DDL(Data Definition Language)
3. Query
4. Relational Schema



Question

➤ **Which of the following is generally used for performing tasks like creating the structure of the relations, deleting relation?**

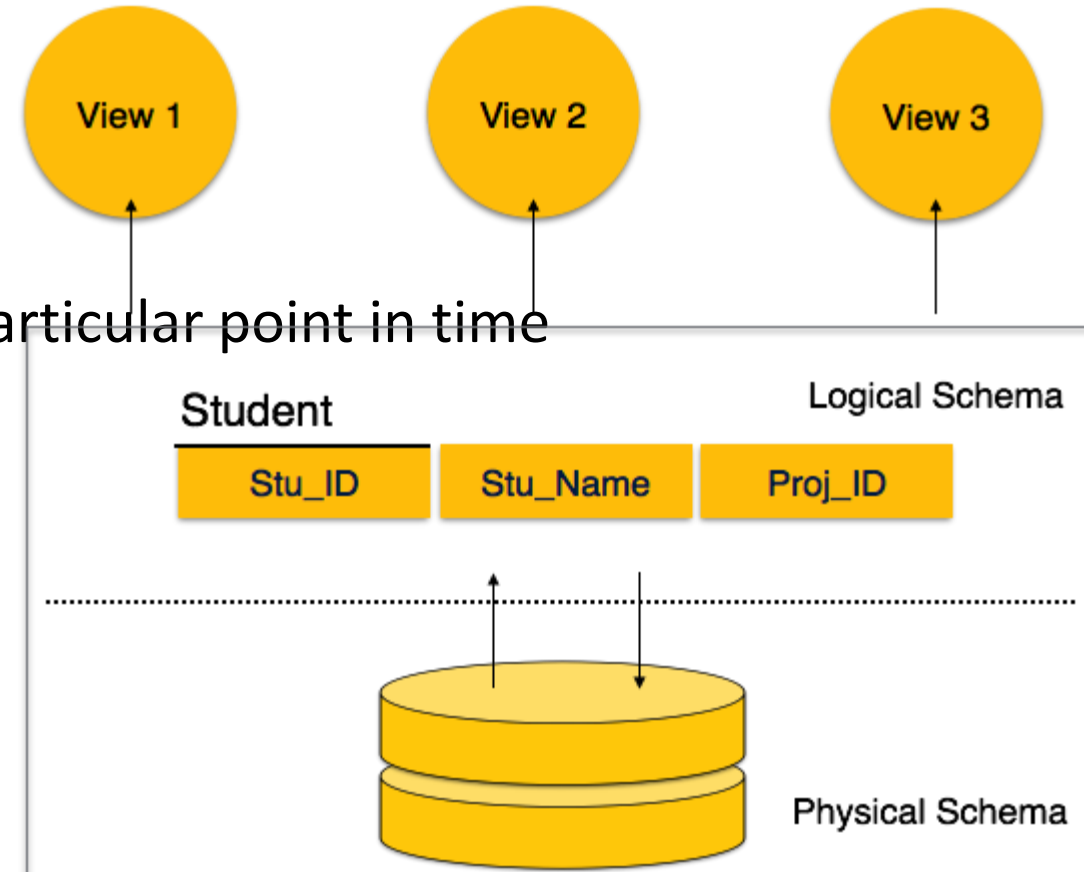
1. DML(Data Manipulation Language)
2. Query
3. Relational Schema
4. DDL(Data Definition Language)

➤ **Which of the following provides the ability to query information from the database and insert tuples into, delete tuples from, and modify tuples in the database?**

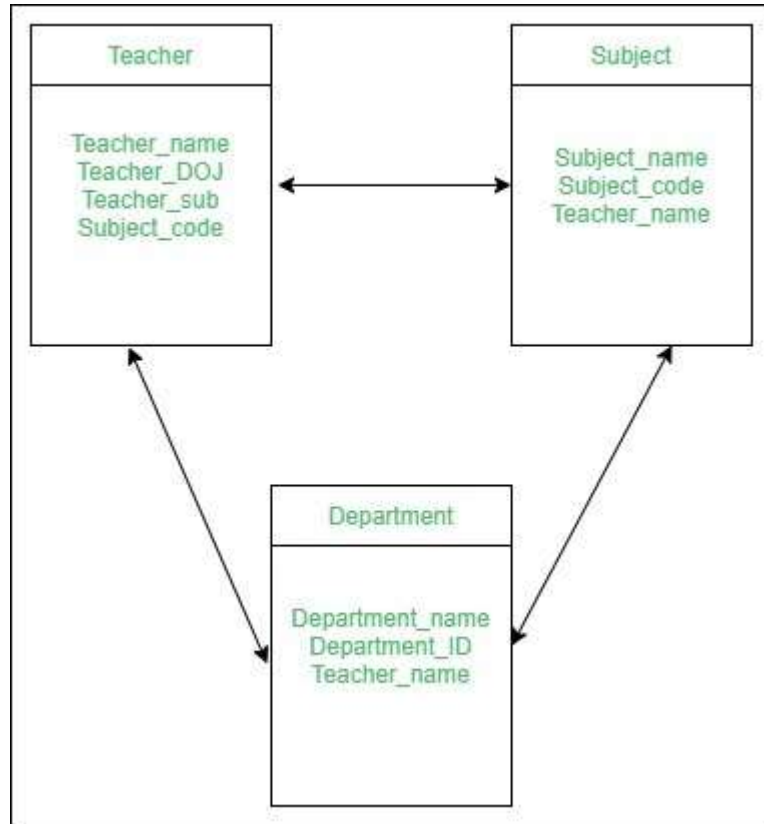
1. DML(Data Manipulation Language)
2. DDL(Data Definition Language)
3. Query
4. Relational Schema

Schema and Instance

- Similar to types and variables in programming languages
- **Schema** – the logical structure of the database
 - the overall design of the database
 - e.g., the database consists of information about a set of customers and accounts and the relationship between them)
 - Analogous to type information of a variable in a program
 - **Physical schema**: database design at the physical level
 - **Logical schema**: database design at the logical level
- **Instance** – the actual content of the database at a particular point in time
 - Analogous to the value of a variable



Schema



Collection of information stored in database at particular instance of time is called as _____.

Database Schema

Data Structure

Instance of Database

Collection of information stored in database at particular instance of time is called as _____.

Database Schema

Data Structure

Instance of Database

Overall design of the database is called as _____.

- a) Database Abstraction
- b) Database Schema
- c) Database Instance

Overall design of the database is called as _____.

a) Database Abstraction

b) Database Schema

c) Database Instance

Three Schema Architecture

Forms, Login Pages,
Command prompt

END User

External View

External View

...

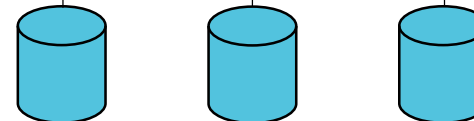
External View

What data are
stored?

CONCEPTUAL SCHEMA
OR
LOGICAL LEVEL

How data actually
stored?

INTERNAL SCHEMA
OR
PHYSICAL LEVEL



RollNo	Name	CPI

- Physical level
 - Has an internal schema
 - Describes the physical storage structure of the database
 - Data structure used to store data
 - Access paths for database
- Conceptual level
 - What data stored in database
 - Relationship among those data
 - DBA uses this level
 - Hides details of physical storage structures
- External View or View level
 - Application programs
 - Describe part of the database that a particular user group is interested in
 - Hides rest of database from that user group
 - hide information (e.g., salary) for security purposes.

RollNo	Name	CPI
1	Ajay	7.8
2	Ram	8.1

RollNo	Mobile
1	9897270768
2	8982720000

Which of the following refers to the level of data abstraction that describes exactly how the data actually stored?

- Conceptual Level
- Physical Level
- File Level
- Logical Level

Which of the following refers to the level of data abstraction that describes exactly how the data actually stored?

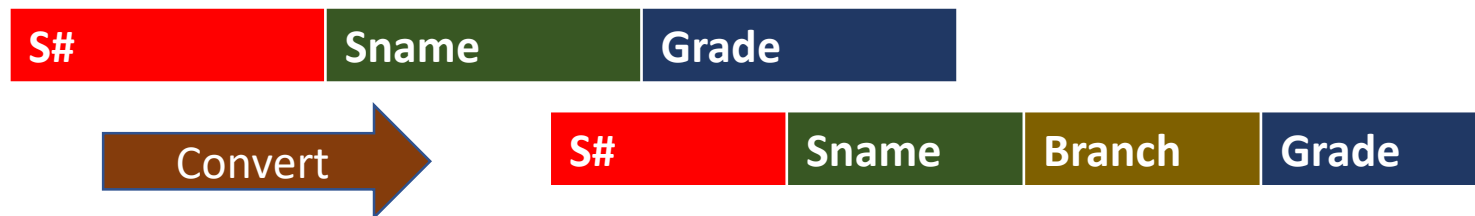
- Conceptual Level
- Physical Level
- File Level
- Logical Level

- Which of the following database abstraction levels is closest to the users?
 - a) Physical level
 - b) Conceptual level
 - c) Internal level
 - d) External level

- Which of the following database abstraction levels is closest to the users?
 - a) Physical level
 - b) Conceptual level
 - c) Internal level
 - **d) External level**

Data Independence

- **Physical Data Independence** – the ability to modify the physical schema without changing the logical schema
 - Applications depend on the logical schema
 - In general, the interfaces between the various levels and components should be well defined so that changes in some parts do not seriously influence others.
 - Numeric(7) to numeric(10)
- **Logical Data Independence**- the ability to modify the logical schema without having to change external schema or application programs.



- **Which of the following statements is/are TRUE regarding three tier DBMS architecture?**
- **(i) Logical Data Independence require structural and constraint changes without affecting application programs**
- **(ii) Physical data independence is required in most databases to hide compression, splitting and merging of records.**

1. Only (i)
2. Only (ii)
3. Both
4. None

- Which of the following database abstraction levels is closest to the users?
- a) Physical level
- b) Conceptual level
- c) Internal level
- d) External level

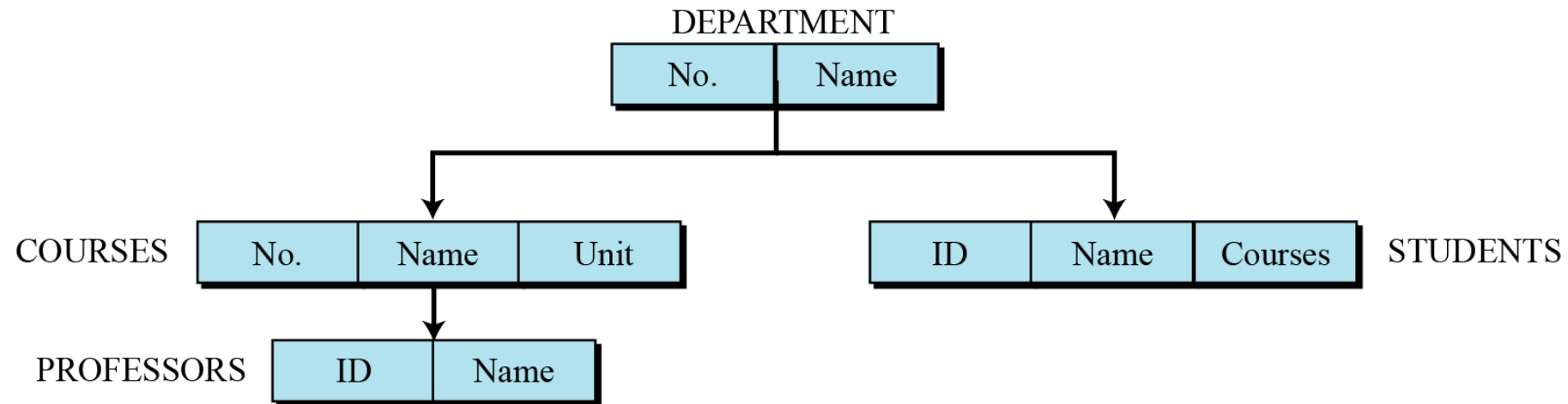
- Which of the following database abstraction levels is closest to the users?
- a) Physical level
- b) Conceptual level
- c) Internal level
- **d) External level**

Data Models

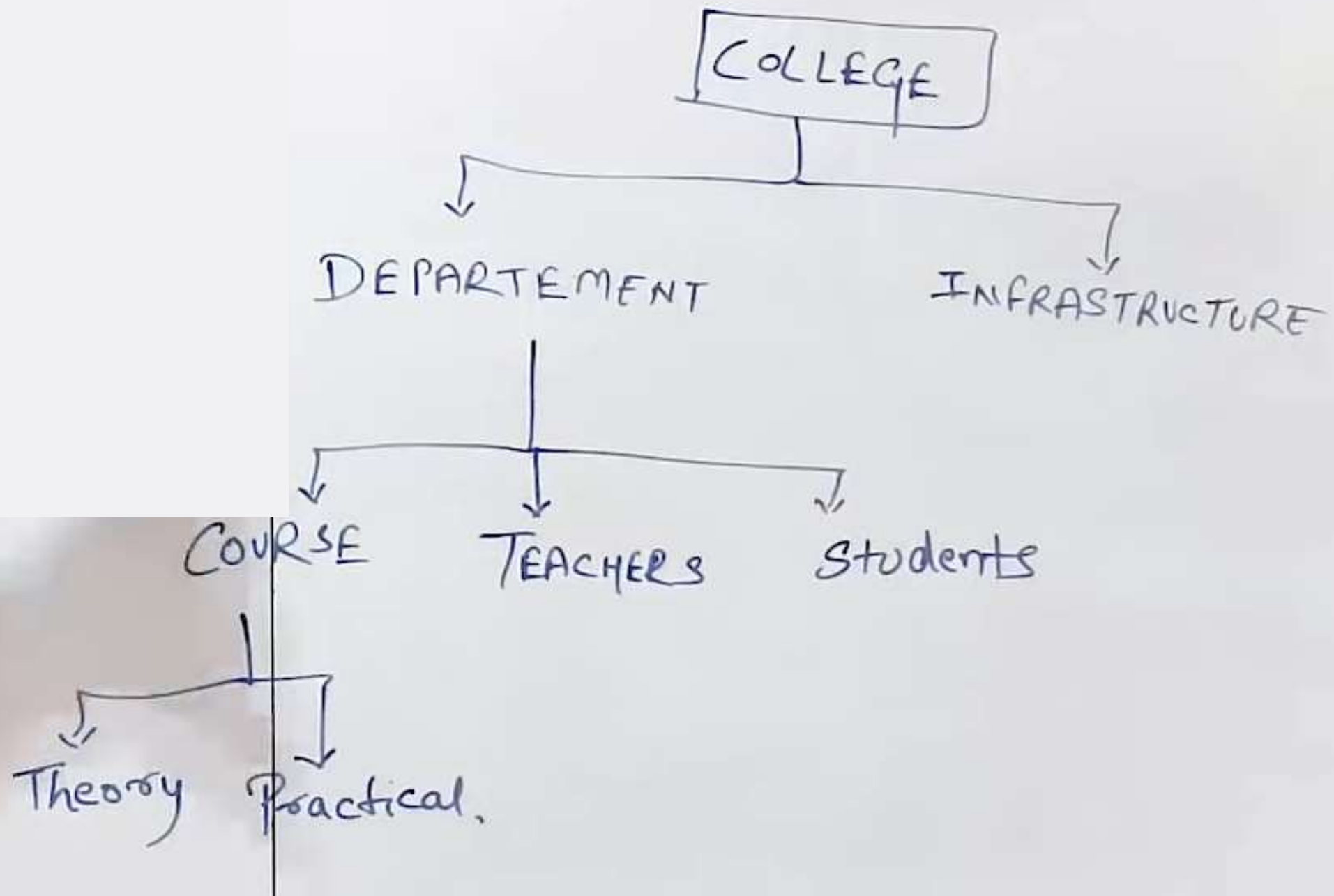
- A collection of conceptual tools for describing data, data relationships, data semantics, and consistency constraint.
- A data model provides a way to describe the design at physical, logical & view level.
- Various data models are:
 1. The entity- relationship model
 2. The object oriented model
 3. Relational model
 4. Network model and
 5. Hierarchical model

Hierarchical database model

In the hierarchical model, data is organized as an inverted tree. Each entity has only one parent but can have several children. At the top of the hierarchy, there is one entity, which is called the root.



An example of the hierarchical model representing a university



- In a Hierarchical database model, records are organized as _____ structure.
- a) Graph.
- b) List .
- c) Links.
- d) Tree.

- In a Hierarchical database model, records are organized as _____ structure.
- a) Graph.
- b) List .
- c) Links.
- **d) Tree.**

Network database model

In the network model, the entities are organized in a graph, in which some entities can be accessed through several paths (Figure).

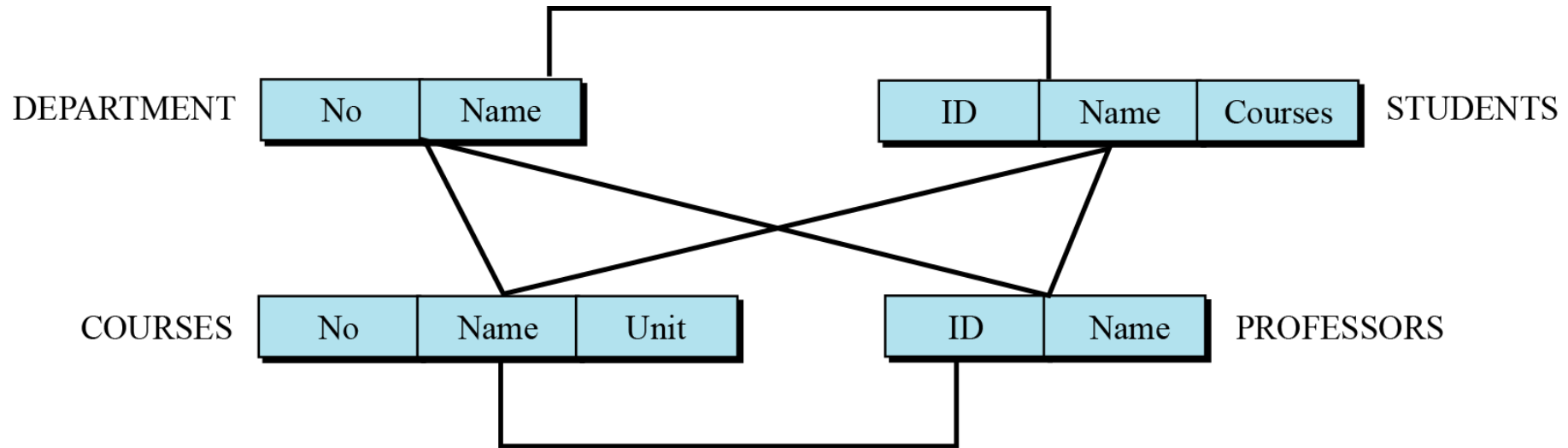


Figure An example of the network model representing a university

Relational database model

In the relational model, data is organized in two-dimensional tables called relations. The tables or relations are, however, related to each other.

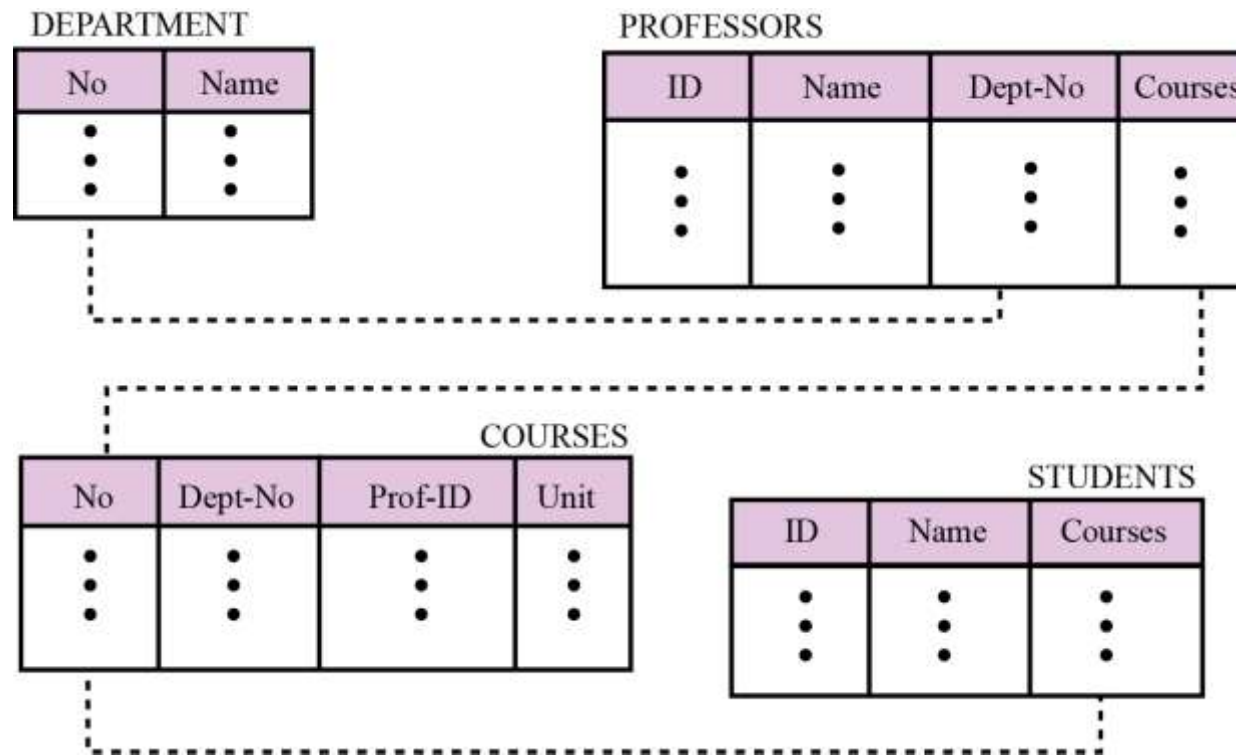


Figure 14.5 An example of the relational model representing a university

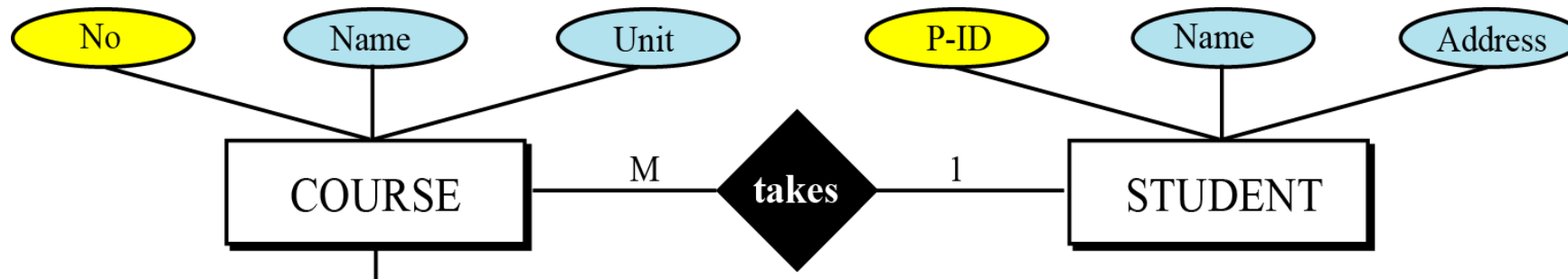
student_id	Name	age
1	AYUSH	17
2	PIHU	18
3	MOKSH	17
4	KIRAN	18

Subject_id	NAME	Teachers
1	Java	Mr. Jack
2	C#	PHILLIP
3	C++	Mr. Anand
4	PHP	Mr. patel

student_id	Subject_id	marks
1	1	52
1	2	89
2	1	66
3	2	78

The Entity-Relationship Model

- Entity is real world object or thing that is distinguishable from others
- The entities for which information needs to be stored and the relationship among these entities. E-R diagrams uses several geometric shapes, but we use only a few of them here:



Object Oriented Model

- This is extended ER model
- Having notion of encapsulation, methods (functions) & object identity.

Database languages

Data definition language(DDL)

- Use to define database structure & schema
- CREATE : to create object in DB
- ALTER: alters the structure of DB
- DROP: delete object from DB
- TRUNCATE: remove all record from table
- COMMENT: add comment to data dictionary
- RENAME: rename an object

Data Manipulation Languages (DML)

- To modify data
- Two types
 - Procedural DMLs: require a user to specify what data are needed and how to get those data.
 - Declarative DMLs: (non-procedural DMLs) require a user to specify what data are needed without specifying how to get those data.eg SQL
- SELECT: to retrieve data from DB
- INSERT: insert data to table
- UPDATE: update existing data within a table

- DELETE: delete all records from a table
- CALL: call a PL/SQL or java program.

DATA CONTROL LANGUAGE

- GRANT: gives users access privileges to DB
- REVOKE: withdraw access privileges given with the GRANT command

TRANSACTION CONTROL LANGUAGE

Use to manage changes by DMLs

- COMMIT: save work done
- ROLLBACK: restore DB to original state since last commit.
- SAVEPOINT: identify a point in transaction to which you can later rollback.

Which of the following are DDL commands?

- A. Select
- B. Delete
- C. Truncate
- D. Drop

Which of the following are DDL commands?

- A. Select
- B. Delete
- C. Truncate**
- D. Drop**

Which of the following are remove data + Schema?

- A. Select
- B. Delete
- C. Truncate
- D. Drop

Which of the following are remove data + Schema?

- A. Select
- B. Delete
- C. Truncate
- D. Drop**

Which of the following are remove data Only?

- A. Select
- B. Delete
- C. Truncate
- D. Drop

Which of the following are remove data Only?

- A. Select
- B. Delete**
- C. Truncate**
- D. Drop

Which of the following will be use to change mobile number of student in database?

- A. Update
- B. Insert
- C. Alter
- D. Create

Which of the following will be use to change mobile number of student in database?

- A. Update**
- B. Insert
- C. Alter
- D. Create

Which of the following will be use to change size of column mobile number from 10 to 11 to accommodate country code?

- A. Update
- B. Insert
- C. Alter
- D. Create

Which of the following will be use to change size of column mobile number from 10 to 11 to accommodate country code?

- A. Update
- B. Insert
- C. Alter**
- D. Create

If we want to undo some action during an ongoing transaction in database, which of the following could be used?

- A. Rollback
- B. Revoke
- C. Grant
- D. Commit

If we want to undo some action during an ongoing transaction in database, which of the following could be used?

- A. Rollback**
- B. Revoke
- C. Grant
- D. Commit

If we want to *finalize* some action during an ongoing transaction in database, which of the following could be used?

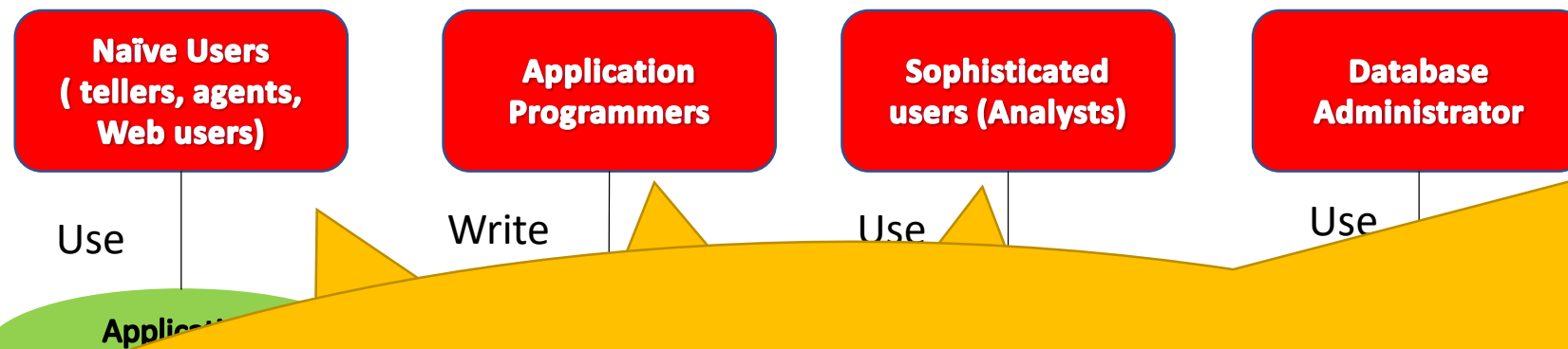
- A. Rollback
- B. Revoke
- C. Grant
- D. Commit

If we want to *finalize* some action during an ongoing transaction in database, which of the following could be used?

- A. Rollback
- B. Revoke
- C. Grant
- D. Commit**

Database Architecture

Overall database mgmt. structure



Database Administrator:

- Coordinates all the activities of the database system. The database administrator has a good understanding of the enterprise's information resources and needs.
- Database administrator's duties include:
 - **Schema definition:** The DBA creates the original database schema by executing a set of data definition statements in the DDL.
 - **Storage structure and access method definition.**
 - **Schema and physical organization modification:** The DBA carries out changes to the schema and physical organization to reflect the changing needs of the organization, or to alter the physical organization to improve performance.
 - **Granting user authority to access the database:** By granting different types of authorization, the database administrator can regulate which parts of the database various users can access.
 - **Specifying integrity constraints.**
 - **Monitoring performance and responding to changes in requirements.**

Naïve User

Database
Administrator

It interprets the requests (queries) received from end user via an application program into instructions. It also executes the user request which is received from the DML compiler.

DDL interpreter

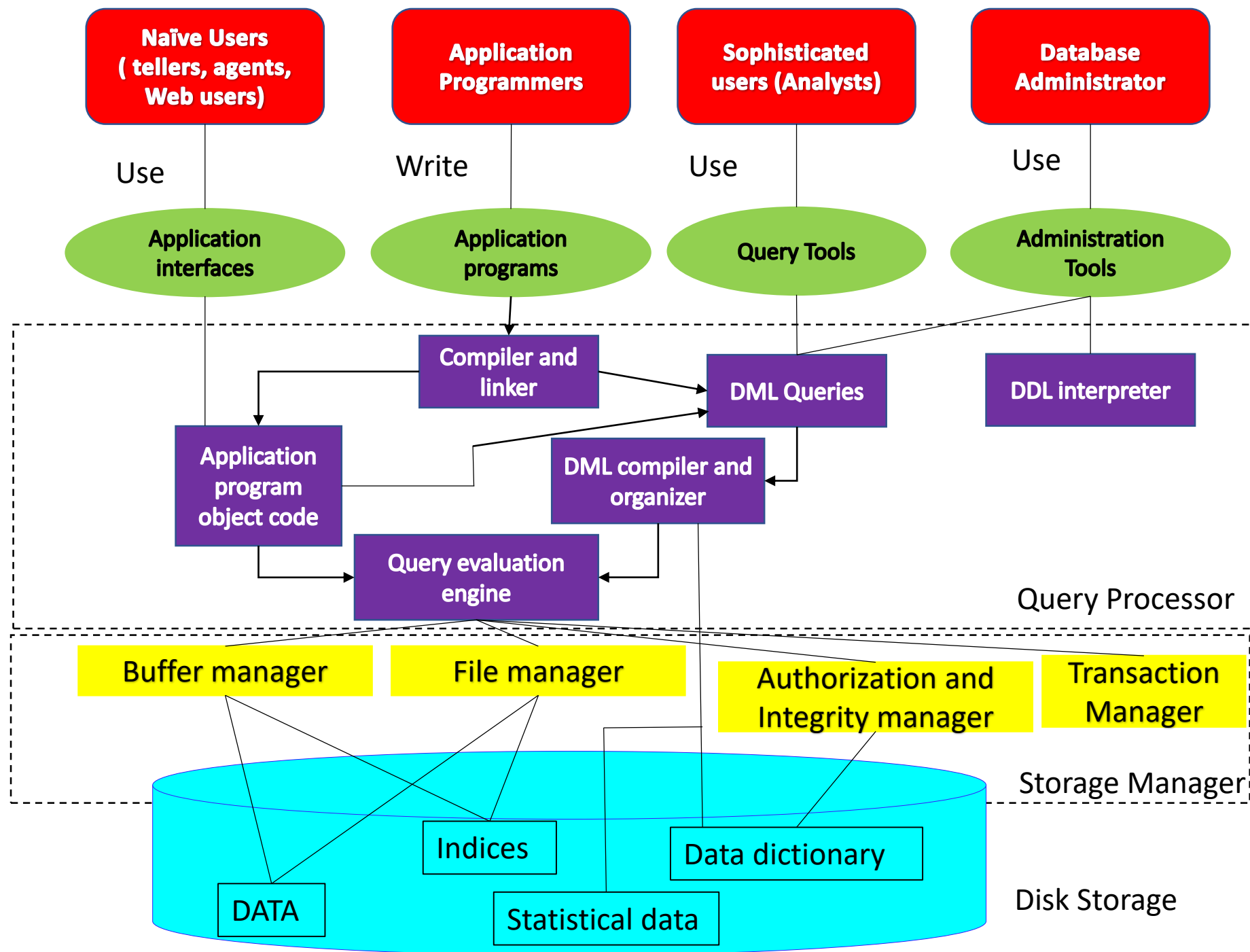
Application
program
object code

DML compiler and
organizer

Query evaluation
engine

Query Processor

It interprets the requests (queries) received from end user via an application program into instructions. It also executes the user request which is received from the DML compiler.



Database Users

Users are differentiated by the way they expect to interact with the system

- Application programmers – interact with system through DML calls, writes application programs
- Sophisticated users – form requests in a database query language eg analysts
- Specialized users – write specialized database applications that do not fit into the traditional data processing framework
- Naïve users – invoke one of the permanent application programs that have been written previously
 - E.g. people accessing database over the web, bank tellers, clerical staff

Which of following will be *your parent* when they check your result online using GLA University Portal?

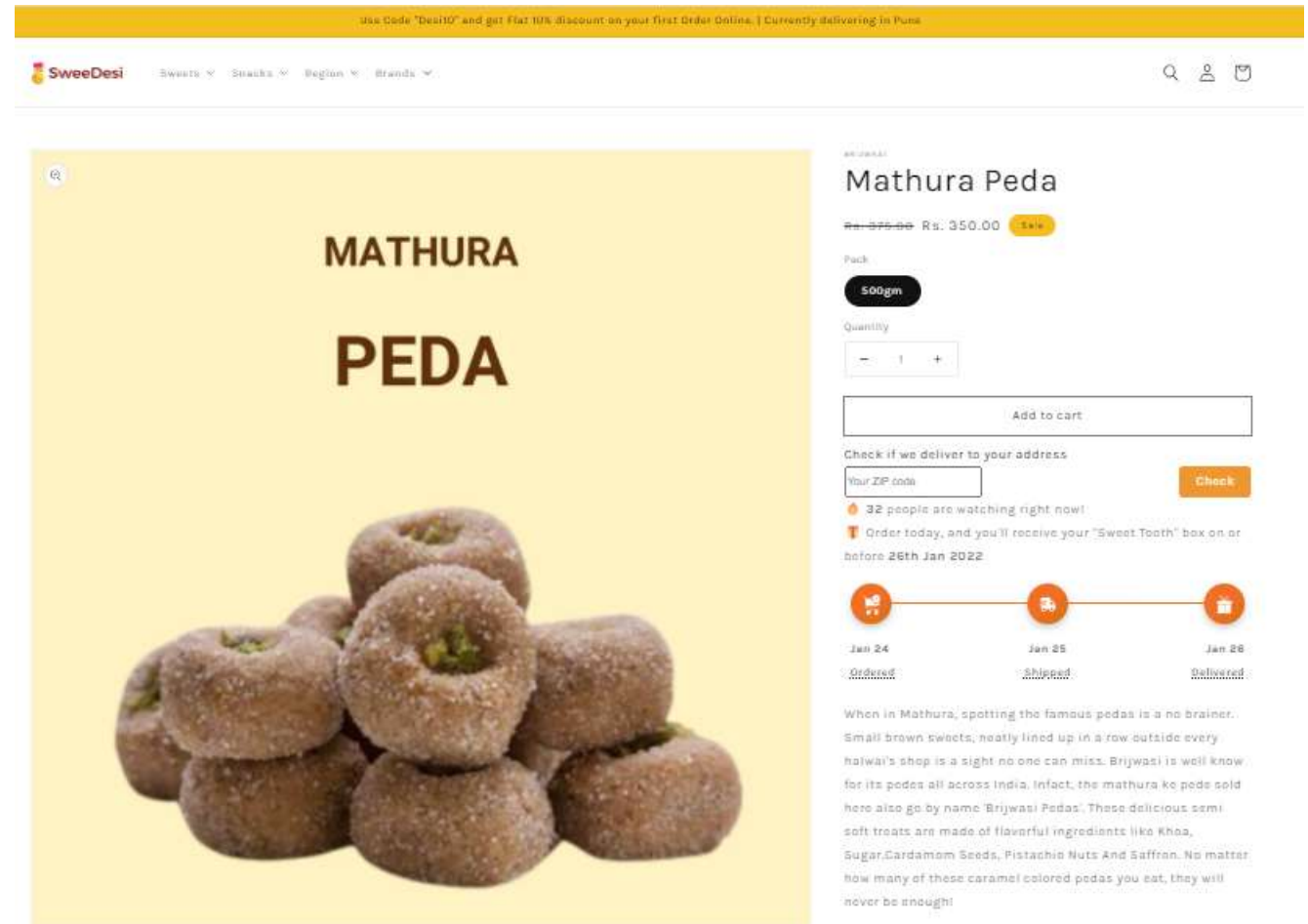
- A. Application programmers
- B. Sophisticated users
- C. Specialized users
- D. Naïve users

Which of following will be *your parent* when they check your result online using GLA University Portal?

- A. Application programmers
- B. Sophisticated users
- C. Specialized users
- D. **Naïve users**

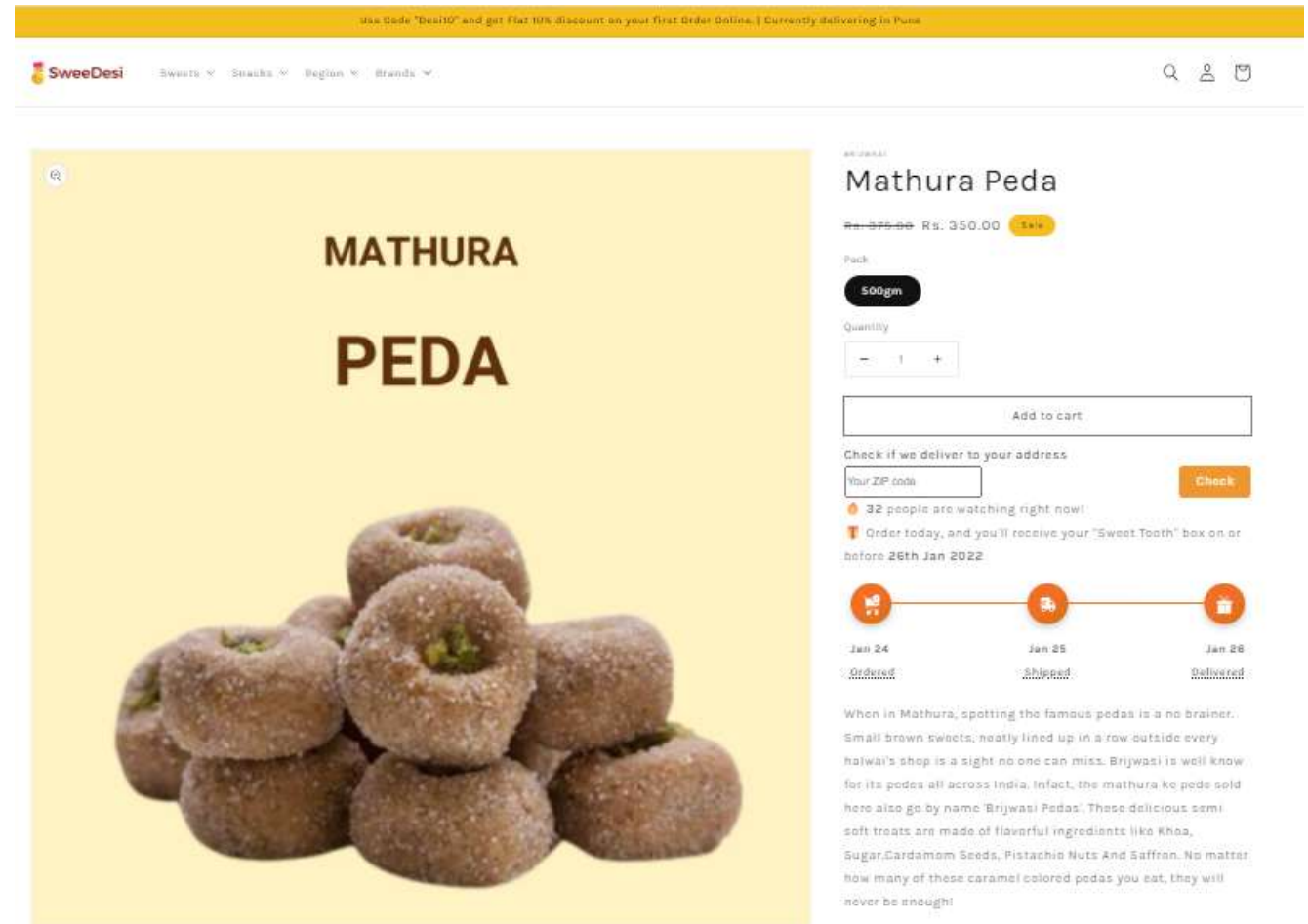
Which of following will be *your brother* when he creates a website for providing online presence of your *family business*?

- A. Application programmers
- B. Sophisticated users
- C. Specialized users
- D. Naïve users



Which of following will be *your brother* when he creates a website for providing online presence of your *family business*?

- A. Application programmers
- B. Sophisticated users
- C. Specialized users
- D. Naïve users



Suppose you work in Brokage firm they ask to query data certain stock and analyze the data on certain parameters, then u will fit in which of the following?

- A. Application programmers
- B. Sophisticated users
- C. Specialized users
- D. Naïve users

Suppose you work in Brokage firm they ask to query data certain stock and analyze the data on certain parameters, then u will fit in which of the following?

- A. Application programmers
- B. Sophisticated users**
- C. Specialized users
- D. Naïve users

Database Administrator (DBA)

- Coordinates all the activities of the database system; the database administrator has a good understanding of the enterprise's information resources and needs.
- Database administrator's duties include:
 - Schema definition- by executing statements in DDL
 - Storage structure and access method definition
 - Schema and physical organization modification
 - Granting user authority to access the database
 - Specifying integrity constraints eg. Primary key
 - Back up data
 - Ensure enough disk space for smooth performing of DB
 - Monitors jobs running on DB

Storage Manager

- Provide interface between low level data in database and application program.
- Components:
 - Authorization & integrity Manager: define role and responsibility for users, and provide Integrity checks.
 - Transaction manager: ensure DB consistency in concurrent access
 - File Manager: manage allocation of disk space and Data structure used for storing data
 - Buffer Manager: decide what data to *cache*

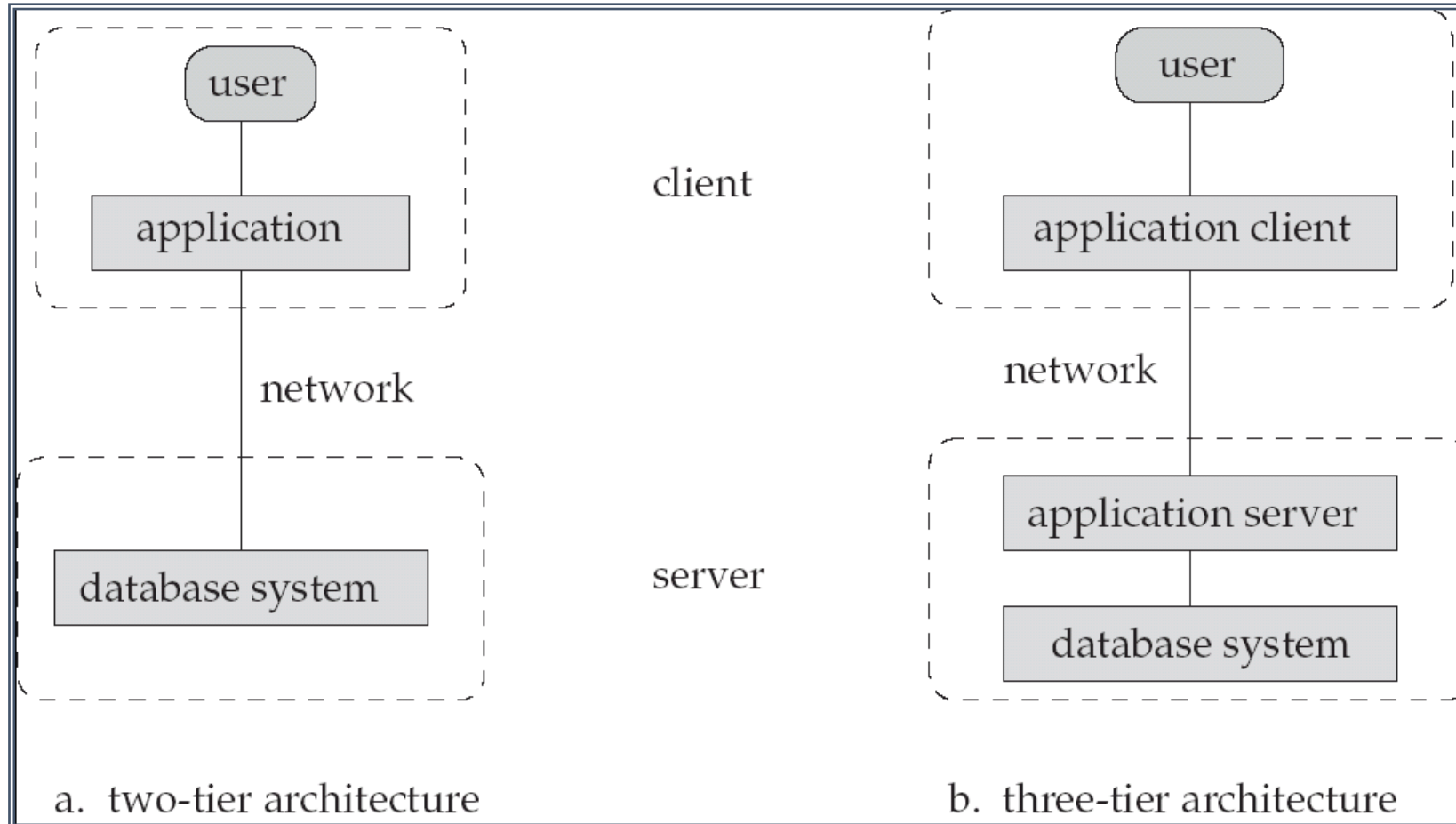
Disk Storage

- Data Files: which stores database itself
- Data Dictionary: stores metadata about the structure of the database
- Indices: provide fast access to data items

Query Processor

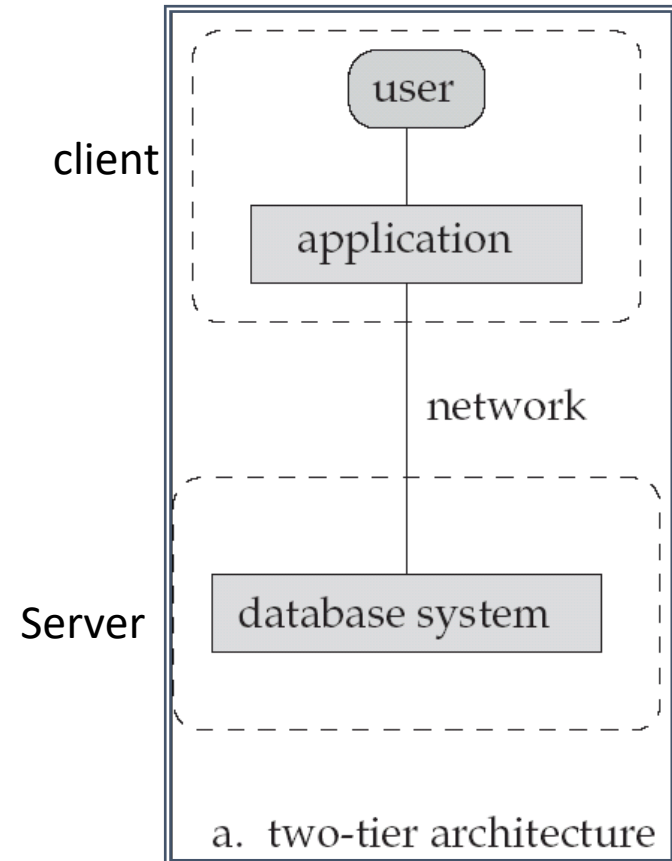
- DDL interpreter: which interprets DDL statements and records the definitions in the data dictionary
- DML Compiler: translate DML statements to evaluation plan. Also perform query optimization. (pick the lowest cost evaluation plan from among the alternatives)
- Query Evaluation Engine: executes low level instructions generated by DML compiler.

Application Architectures



Two – tier Architecture

- Application is partitioned into two components:
 - Server: runs query language statements
 - Clients: provide user interface and local processing
- JDBC & ODBC are used for interaction between the client and the server.



Three tier Architecture

- Client: merely a user interface (form interface) but no direct DB calls,
- Application Server: communicates with database system to access data. Had coding for what action to carried out under what conditions.
- Database Server: handle data access for query receive by Application Server

