MODULE – I
INTRODUCTION TO DATABASE SYSTEMS, DBMS ARCHITECTURE INTRODUCTION TO DATA MODELING

Dr. Geetha Mary A
Associate Professor
SCOPE
VIT

□ Why to store data?

□ Data : Known facts that can be recorded and have an implicit meaning

□ Database: Collection of related data

□ DBMS: Data Base Management System

- A software package/ system to facilitate the creation and maintenance of a computerized database.

□ Database Systems:

- The DBMS software together with the data itself.

Some of the DBMS software

- Oracle
- Ms-Access
- Foxpro
- dbase

Database System Application

- Banking
- Airlines
- Universities

Need

- Need WRT **Paper work** (Hard copy of data)
  - Paper gets degraded over time
- Need WRT **File system**
  - Store the data in OS file types
    - Eg: Notepad(*.txt)
      MS word (*.doc)

File System Contn…

- To allow users to manipulate information the system application programs that manipulate the files, including

  - A program to debit or credit an account

  - A pgm to add a new account

  - A pgm to find the balance of an account

  - A pgm to generate monthly statements

Drawbacks of using File system

- **Data redundancy and inconsistency**
  - Multiple file formats, duplication of information in different files
- **Difficulty in accessing data**
  - Need to write a new program to carry out each new task
- **Data isolation** — multiple files and formats
- **Integrity problems**
  - Integrity constraints
  - Hard to add new constraints or change existing ones

Drawbacks of using File system

- **Atomicity of updates**
  - Failures may leave database in an inconsistent state with partial updates carried out
  - E.g. transfer of funds from one account to another should either complete or not happen at all
- **Concurrent access by multiple users**
  - Concurrent accessed needed for performance
  - Uncontrolled concurrent accesses can lead to inconsistencies
    - E.g. two people reading a balance and updating it at the same time
- **Security problems**

Advantages of using a DBMS

- Controlling Redundancy

- Restricting Un-authorized Access

- Providing persistent storage for pgm objects and data structures

- Providing multiple user interfaces

- Representing complex relationships among data

- Enforcing integrity constraints

- Providing Backup and recovery

Simplified database system environment
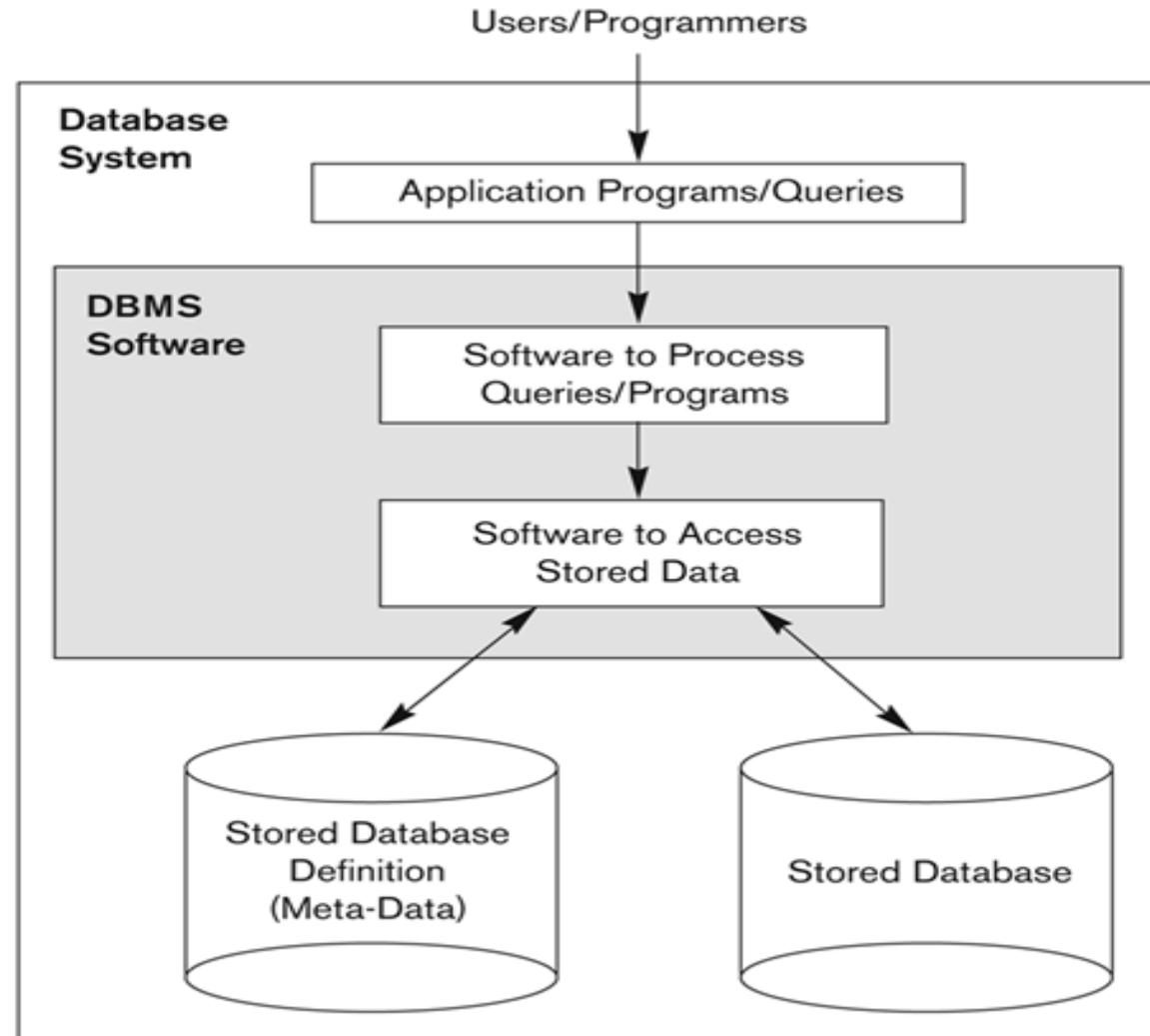
Typical DBMS Functionality

- *Define* a particular database in terms of its data types, structures, and constraints
- *Construct* or Load the initial database contents on a secondary storage medium
- *Manipulating* the database:
  - Retrieval: Querying, generating reports
  - Modification: Insertions, deletions and updates to its content
  - Accessing the database through Web applications
- *Processing* and *Sharing* by a set of concurrent users and application programs – yet, keeping all data valid and consistent

Example of a Database
(with a Conceptual Data Model)

- **Mini-world for the example:**
  - Part of a UNIVERSITY environment.
- **Some mini-world *entities*:**
  - STUDENTs
  - COURSEs
  - SECTIONs (of COURSEs)
  - (academic) DEPARTMENTs
  - INSTRUCTORs

Example of a simple database

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	04	King
92	CS1310	Fall	04	Anderson
102	CS3320	Spring	05	Knuth
112	MATH2410	Fall	05	Chang
119	CS1310	Fall	05	Anderson
135	CS3380	Fall	05	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Database Users

☐ Users may be divided into

  ❑ Those who actually use and control the database content, and those who design, develop and maintain database applications (called "**Actors on the Scene**"), and

  ❑ Those who design and develop the DBMS software and related tools, and the computer systems operators (called "**Workers Behind the Scene**").

Database Users

- Actors on the scene

  - **Database administrators:**

    - Responsible for authorizing access to the database, for coordinating and monitoring its use, acquiring software and hardware resources, controlling its use and monitoring efficiency of operations.

  - **Database Designers:**

    - Responsible to define the content, the structure, the constraints, and functions or transactions against the database. They must communicate with the end-users and understand their needs.

Categories of End-users

□ Actors on the scene (continued)

▫ **End-users:** They use the data for queries, reports and some of them update the database content. End-users can be categorized into:

- **Casual**: access database occasionally when needed
- **Naïve** or Parametric: they make up a large section of the end-user population.
  - They use previously well-defined functions in the form of "canned transactions" against the database.
  - Examples are bank-tellers or reservation clerks who do this activity for an entire shift of operations.

Categories of End-users (continued)

■ **Sophisticated:**

■ These include business analysts, scientists, engineers, others thoroughly familiar with the system capabilities.

■ Many use tools in the form of software packages that work closely with the stored database.

■ **Stand-alone:**

■ Mostly maintain personal databases using ready-to-use packaged applications.

■ An example is a tax program user that creates its own internal database.

■ Another example is a user that maintains an address book

Data Models

□ **Data Model:**

- ▣ A set of concepts to describe the ***structure*** of a database, the ***operations*** for manipulating these structures, and certain ***constraints*** that the database should obey.

□ **Data Model Structure and Constraints:**

- ▣ Constructs are used to define the database structure
- ▣ Constructs typically include ***elements*** (and their ***data types***) as well as groups of elements (e.g. ***entity, record, table***), and ***relationships*** among such groups
- ▣ Constraints specify some restrictions on valid data; these constraints must be enforced at all times

Data Models (continued)

□ **Data Model Operations:**

 ▫ These operations are used for specifying database *retrievals* and *updates* by referring to the constructs of the data model.

 ▫ Operations on the data model may include ***basic model operations*** (e.g. generic insert, delete, update) and ***user-defined operations*** (e.g. compute_student_gpa, update_inventory)

Categories of Data Models

- **Conceptual (high-level, semantic) data models:**
  - Provide concepts that are close to the way many users perceive data.
    - (Also called *entity-based* or *object-based* data models.)
- **Physical (low-level, internal) data models:**
  - Provide concepts that describe details of how data is stored in the computer. These are usually specified in an ad-hoc manner through DBMS design and administration manuals
- **Implementation (representational) data models:**
  - Provide concepts that fall between the above two, used by many commercial DBMS implementations (e.g. relational data models used in many commercial systems).

DBMS Component
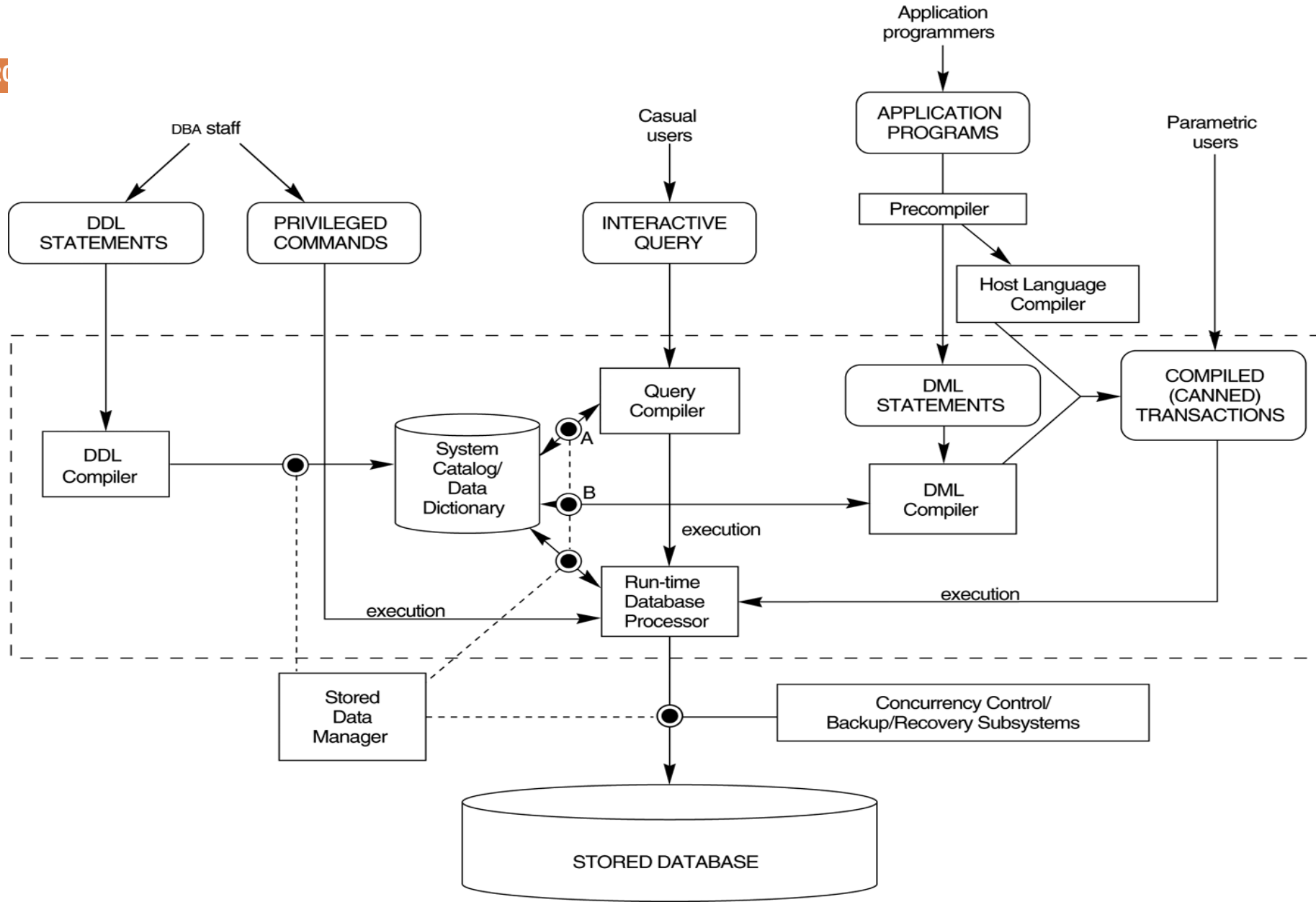
Schemas versus Instances

- Database Schema:
  - The ***description*** of a database.
  - Includes descriptions of the database structure, data types, and the constraints on the database.
- Schema Diagram:
  - An ***illustrative*** display of (most aspects of) a database schema.
- Schema Construct:
  - A ***component*** of the schema or an object within the schema, e.g., STUDENT, COURSE.

Schemas versus Instances

□ Database State:

  ◻ The actual data stored in a database at a ***particular moment in time***. This includes the collection of all the data in the database.

  ◻ Also called database instance (or occurrence or snapshot).

  ▪ The term *instance* is also applied to individual database components, e.g. *record instance, table instance, entity instance*

Database Schema vs. Database State

- ## Database State:

  - Refers to the ***content*** of a database at a moment in time.

- ## Initial Database State:

  - Refers to the database state when it is initially loaded into the system.

- ## Valid State:

  - A state that satisfies the structure and constraints of the database.

Database Schema
vs. Database State (continued)

□ Distinction

  ▫ The ***database schema*** changes very infrequently.

  ▫ The ***database state*** changes every time the database

  is updated.

□ **Schema** is also called **intension**.

□ **State** is also called **extension**.

Example of a Database Schema

STUDENT

Name	Student_number	Class	Major

COURSE

Course_name	Course_number	Credit_hours	Department

PREREQUISITE

Course_number	Prerequisite_number

SECTION

Section_identifier	Course_number	Semester	Year	Instructor

GRADE_REPORT

Student_number	Section_identifier	Grade

Example of a database state

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	04	King
92	CS1310	Fall	04	Anderson
102	CS3320	Spring	05	Knuth
112	MATH2410	Fall	05	Chang
119	CS1310	Fall	05	Anderson
135	CS3380	Fall	05	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Three schema architecture

- Important characteristics of database approach:

  - Insulation of programs and data

  - Support multiple user views

  - Use of a catalog to store the database description (Schema)

- An architecture of schema is proposed to achieve these characteristics

- Architecture goal: To separate the user application and the database

Three schema architecture

External
Level

**External
View**

**External
View**

Conceptual
Level

**Conceptual
Schema**

Internal
Level

Internal Schema

Stored Database
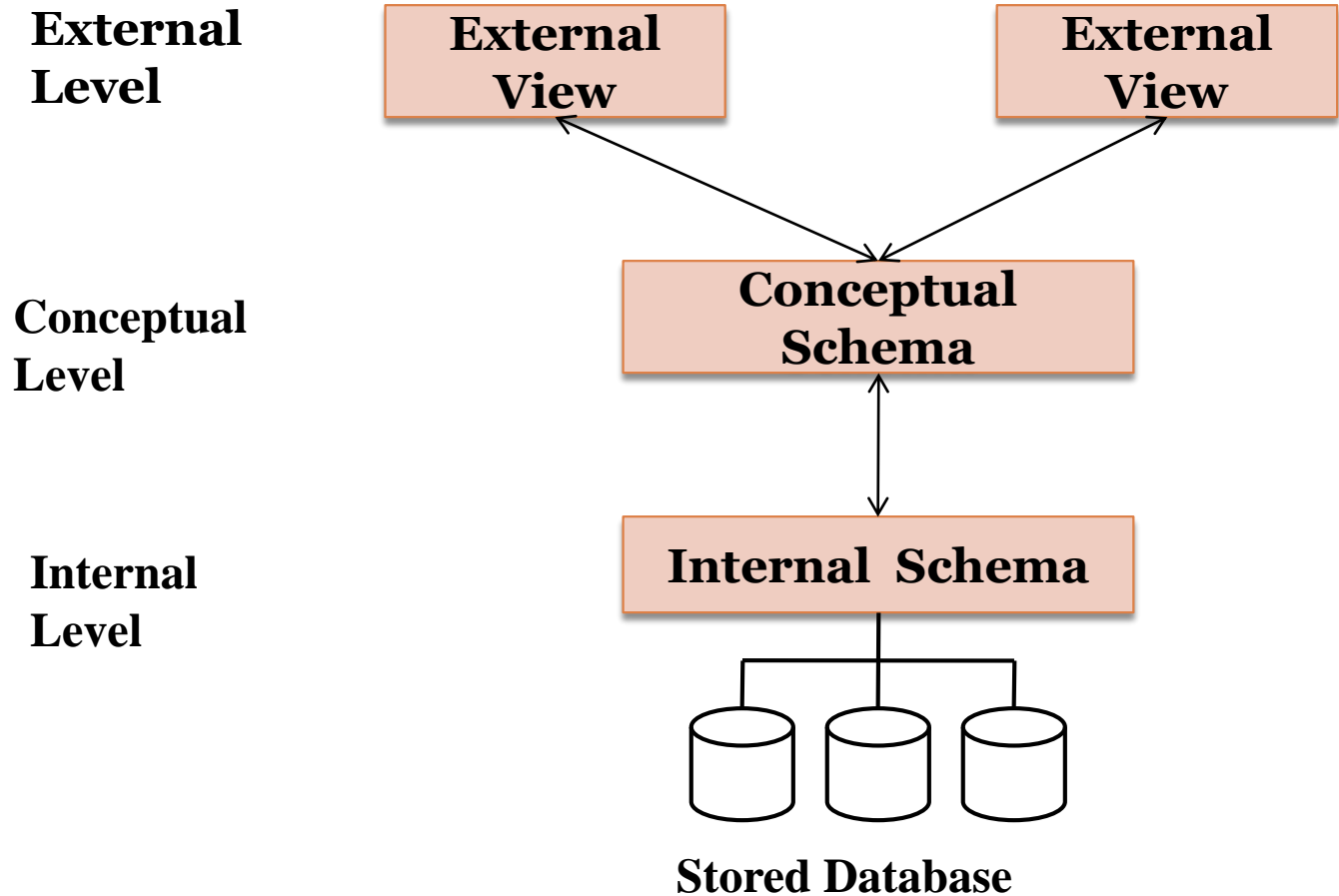
Three-Schema Architecture

Defines DBMS schemas at *three levels*:

- **Internal schema** at the internal level to describe physical storage structures and access paths. Typically uses a *physical* data model.

- **Conceptual schema** at the conceptual level to describe the structure and constraints for the *whole* database for a community of users. Uses a *conceptual* or an *implementation* data model.

- **External schemas** at the external level to describe the various user views. Usually uses the same data model as the conceptual level.

Three-Schema Architecture

Mappings among schema levels are needed to transform requests and data. Programs refer to an external schema, and are mapped by the DBMS to the internal schema for execution.

Access path

Structure that makes the search for particular database records efficient

select * from employee where sal>1000 and empid>20

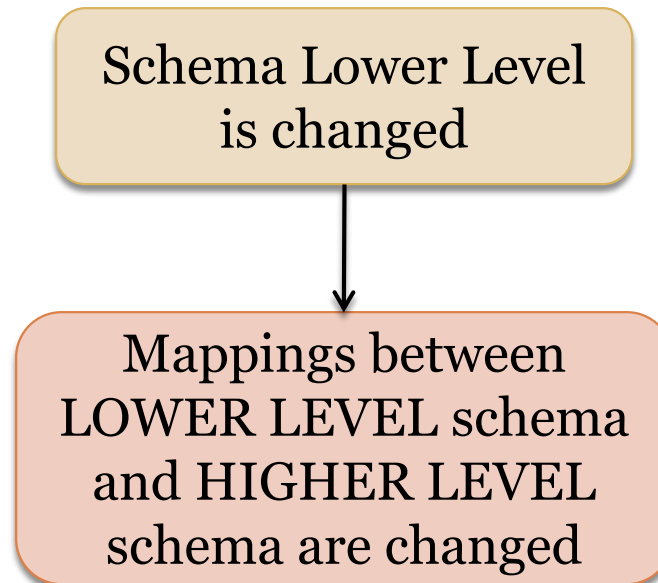| empid	Name	Sal	place
5	Pavithra	5000	Chennai
2	Kirthika	10000	Bangalore
55	Amrith	15000	Chennai
25	Jai	20000	Bangalore
33	Anu	26000	Chennai

- Which condition need to be applied first.
- Query optimizing

Data Independence

Capacity to change the schema at one level of database system without having to change the schema at the next higher level

- **Logical Data Independence**: The capacity to change the conceptual schema without having to change the external schemas and their application programs.

- **Physical Data Independence**: The capacity to change the internal schema without having to change the conceptual schema.

Data Independence

Schema Lower Level
is changed

Mappings between
LOWER LEVEL schema
and HIGHER LEVEL
schema are changed

Higher level schemas are unchanged.