EIU
TRƯỜNG ĐẠI HỌC
QUỐC TẾ
MIỀN ĐÔNG
EASTERN
INTERNATIONAL
UNIVERSITY

CSE301 – DATABASE

EIU

CONTENT


Inamori Kazuo
CON ĐƯỜNG ĐI ĐẾN THÀNH CÔNG BẰNG SỰ TỬ TẾ
NHÀ XUẤT BẢN TRẺ
BLUMBERG
STARTUP CEO
A FIELD GUIDE TO SCALING UP YOUR BUSINESS
knaflic
storytelling with data
STEVE HILTON
Designing a World Where People Come First
MORE HUMAN
OWN YOUR FUTURE
STRESS TEST
TIMOTHY F. GEITHNER

Course Description

- This course is a core course for SE & CNDC program.
- This course aims to introduce the core principles and techniques required in the design and implementation of database systems.
- The course highlights the basic concepts of databases, database system concepts and architecture, data modeling using ER diagram, relational model, SQL, relational algebra and calculus, normalization, transaction processing, concurrency control, and database recovery.

Course Objectives

- Understand the basic database concepts, applications, data models, schemas and instances.
- Familiarize the student with Entity Relationship model for a database and demonstrate the use of constraints and relational algebra operations.
- **Understand the basics of SQL commands and construct queries using SQL**.
- Emphasize the importance of normalization in databases and the basic concepts of transaction processing and concurrency control.
- Familiarize the student with the concepts of database storage structures and identify the access techniques.

Course Learning Outcomes

Knowledge

- CLO1: To gain the basic knowledge of database design and database management system.
- CLO2: To understand application's data requirements using conceptual modeling tools like ER diagrams and apply it for database design.
- CLO3: To achieve the knowledge about relational algebra and relational calculus concepts and applying the same through SQL commands to interact with a relational DBMS.
- CLO4: To understand normalization concepts and apply it to database design in order to eliminate anomalies.
- CLO5: To understand the knowledge of database transaction management.

Course Learning Outcomes

Skills

CLO6: Design ER diagrams and design database schemas based on the conceptual model.

CLO7: Convert entity relationship diagrams into RDBMS and formulate SQL queries on the respect data into RDBMS and formulate SQL queries on the data.

Attitudes

CLO8: Having honest, hard- working attitude , awareness and responsibility when using laboratory equipment .

CLO9: Use communication skills, skills to work independently and as part of a team

Ability, responsibility and career

CLO10: Train and improve your self-study ability

Books and Teaching Materials

- **Teaching materials**
  - [1]. Abraham Silberschatz, Henry Korth, and S. Sudarshan. 2020. Database System Concepts (7th Edition). McGraw-Hill Publishers.
  - [2]. Ramez Elmasri and Shamkant B. Navathe. 2017. Fundamentals of Database Systems (7th Edition). Pearson Publishers.
- **References**
  - [3]. Raghu Ramakrishnan, Johannes Gehrke. 2003. Database management systems (Third Edition). BPB publications.
  - [4]. Ivan Bayross. 2009. SQL, PL/SQL the Programming Language of Oracle (4th Edition). McGraw-Hill Osborne Media.
  - [5]. Paul DuBois. 2013. MySQL: Developer's Library (5th Edition). Addison-Wesley Professional .

DATABASE

What is Database?

Collection of logically interrelated data and a description of this data, designed to meet the information needs of an organization.

Or

Collection of related data & storage area where we store it for retrieving information.

Examples:

Database : Employee

Data: Emp_Name, Emp_Phone, Emp_Address, Salary, Gender, DoB, Department, etc.

What is Data?

Raw facts which has some existence in the real world.

Examples:

Last Name	Class	Gender	IELTS
Nguyen	2023	M	6.5
Vo	2018	M	6.0
Le	2022	M	7
Nguyen	2021	M	7.5

Data is raw and unstructured

To derive some information from it.

What is Information?

When data is processed, organized, structured or presented in a given context to make it more useful, it is called as information.

Examples:

-

Last Name	Class	Gender	IELTS
Nguyen	2023	M	6.5
Vo	2018	M	5.0
Le	2022	M	7
Nguyen	2021	M	7.5

Information provides context and insights

For future analysis

Difference between Data & Information?

Roll	Name	Marks	Attendance
101	Jack	95	30
102	Bob	75	23
103	Tom	68	20
104	Dev	52	18
105	Sam	86	27

Data: Student's Attendance Data

Process: $\frac{Total\ no.of\ class\ attened\ X\ 100}{Total\ no.\ of\ class\ conducted}$

Information: % of attendance by students.

EIU

What is Database System?

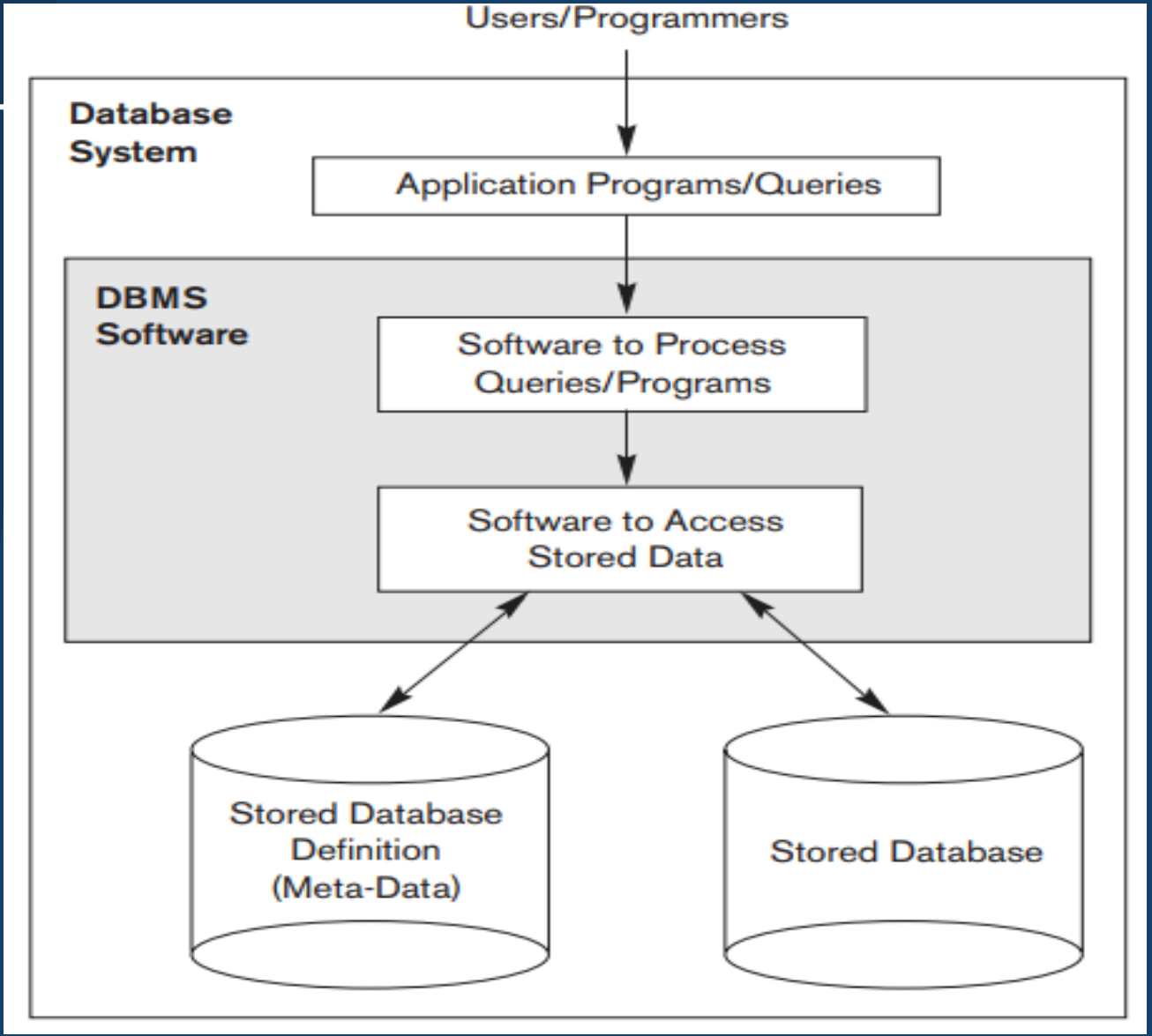
It provides a convenient way to store, delete, retrieve as well as manipulate the data.

What is Database Management System (DBMS)?

An application software that provides a convenient way to manage database systems.

Examples: MSSQL, MySQL, Oracle, IBMDB2, etc.

What is Database System?

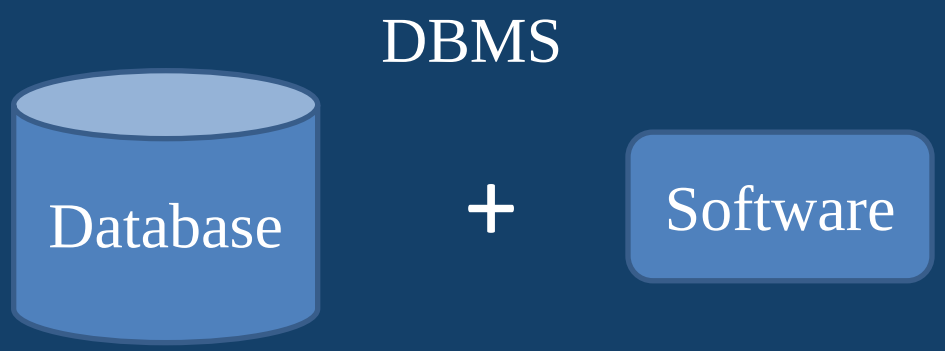


Definition: Database Management System

A database management system (DBMS) is a computerized system that enables users to create and maintain a database.

Or

The DBMS is a general-purpose software system that facilitates the processes of defining, constructing, manipulating, and sharing databases among various users and applications.



Processes of defining

Defining a database involves specifying the data types, structures, and constraints of the data to be stored in the database.

The database definition or descriptive information is also stored by the DBMS in the form of a database catalog or dictionary; it is called meta-data.

Processes of constructing, manipulating, and sharing

Constructing the database is the process of storing the data on some storage medium that is controlled by the DBMS.

Manipulating a database includes functions such as querying the database to retrieve specific data, updating the database to reflect changes, and generating reports from the data.

Sharing a database allows multiple users and programs to access the database simultaneously.

Traditional File System & Their Disadvantages

- Data Redundancy : Duplicating Data
- Data Inconsistency: Duplicate Data are not having same values.
- Lack of Data Integration: Data retrieval problem due to multiple files or Data may be required to satisfy some constraints.
- Program Dependency: The program should be changed if the file structure is changed.
- Data Dependency : modifications in the characteristics of data, such as changing a field from integer to decimal, require changes in all the programs that access the file.

Traditional File System & Their Disadvantages

- No / Less Data Security : Data Security problem.
- Data Integrity Issues: Wrong type data inserted into data fields.
- Data isolation: Because data are scattered in various files, and files may be in different formats.
- Data Manipulation capability is inadequate: No two data items in a database should present same given entity.
- Concurrent-access anomalies: Concurrent updates to same data by different transactions at the same time may result in inconsistent data.

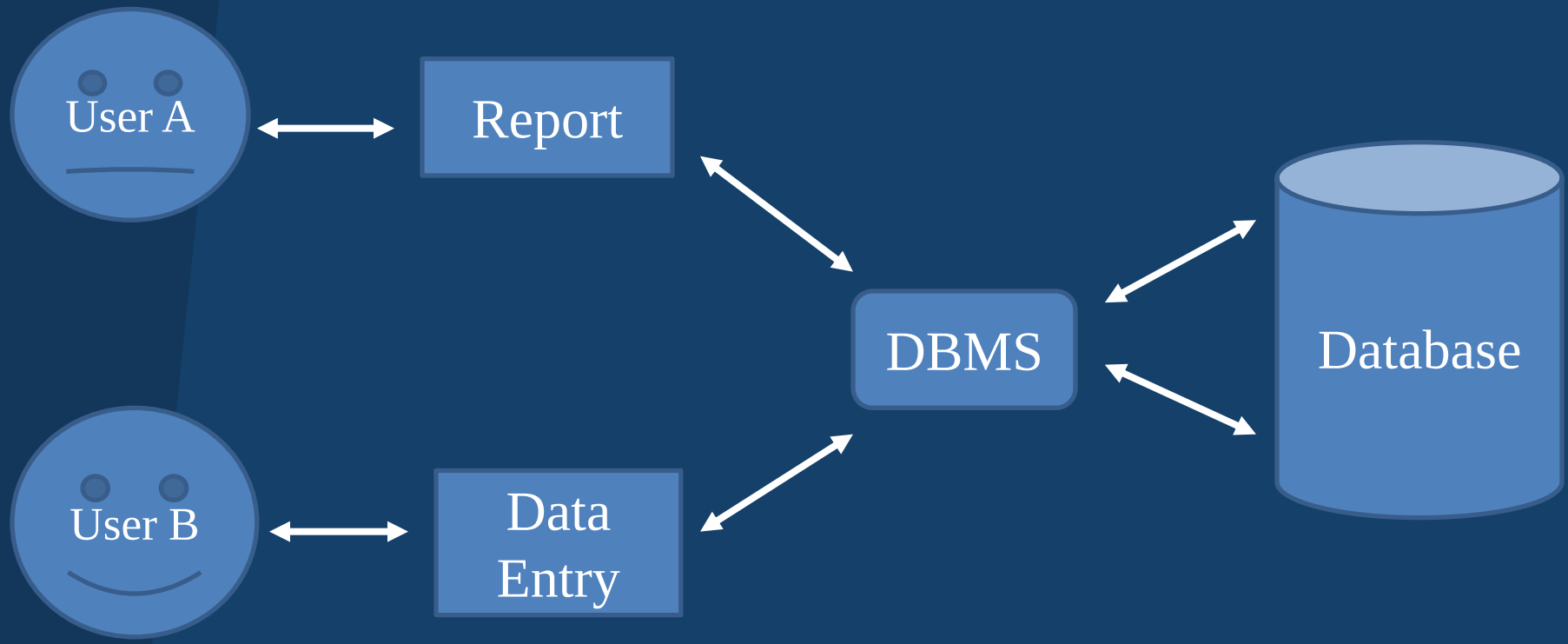
Advantages of DBMS

- Control Redundancy: Apply Normalization
- Data Consistency
- Program Data Independence
- Sharing of Data
- Improved Data Integrity
- Improved Security
- Efficient Data Access
- Improved Backup & Recovery Management
- Minimal Program Maintenance
- Removed Concurrent-access anomalies

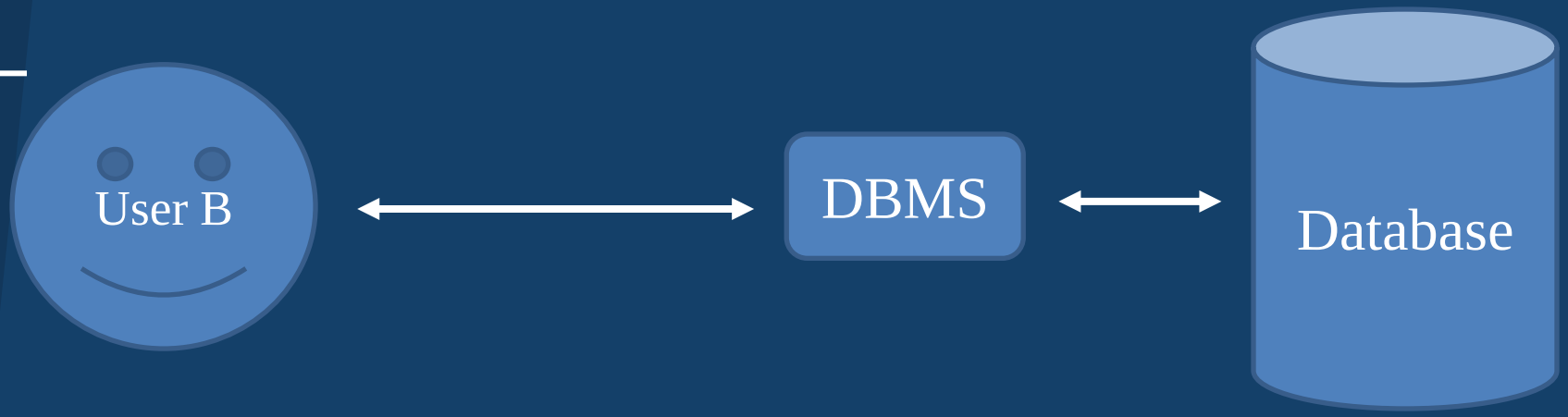
Disadvantages of DBMS

- Complexity Increases
- Requirements of Large Disk Space
- Addition Cost of Hardware
- Needs of Additional Specialized Manpower
- Need for Backup & Recovery Management
- More Installation & Maintaining Cost

Working of Database Architecture



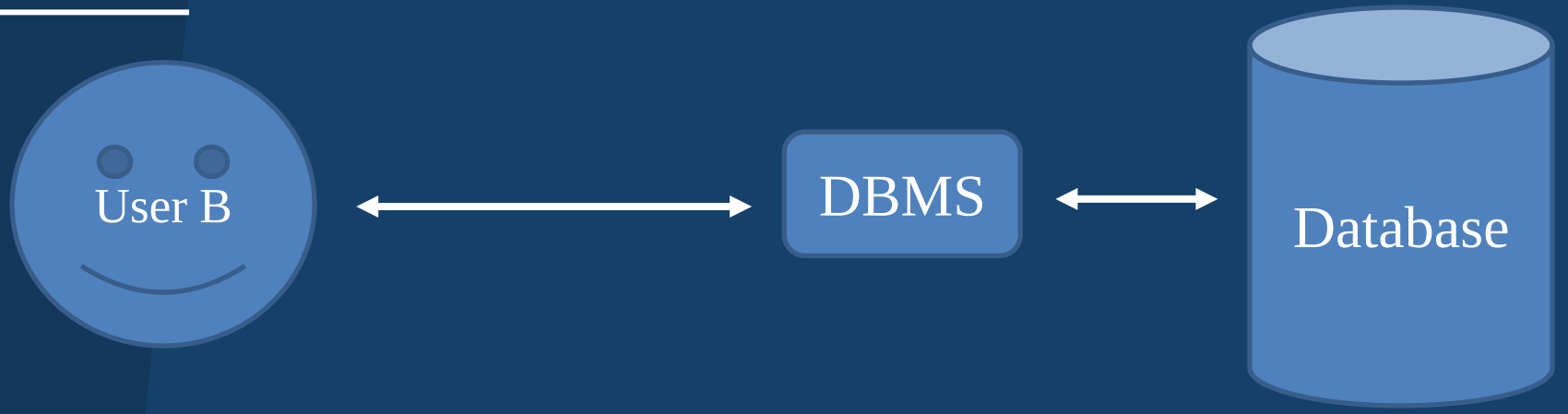
Database 1-Tier Architecture



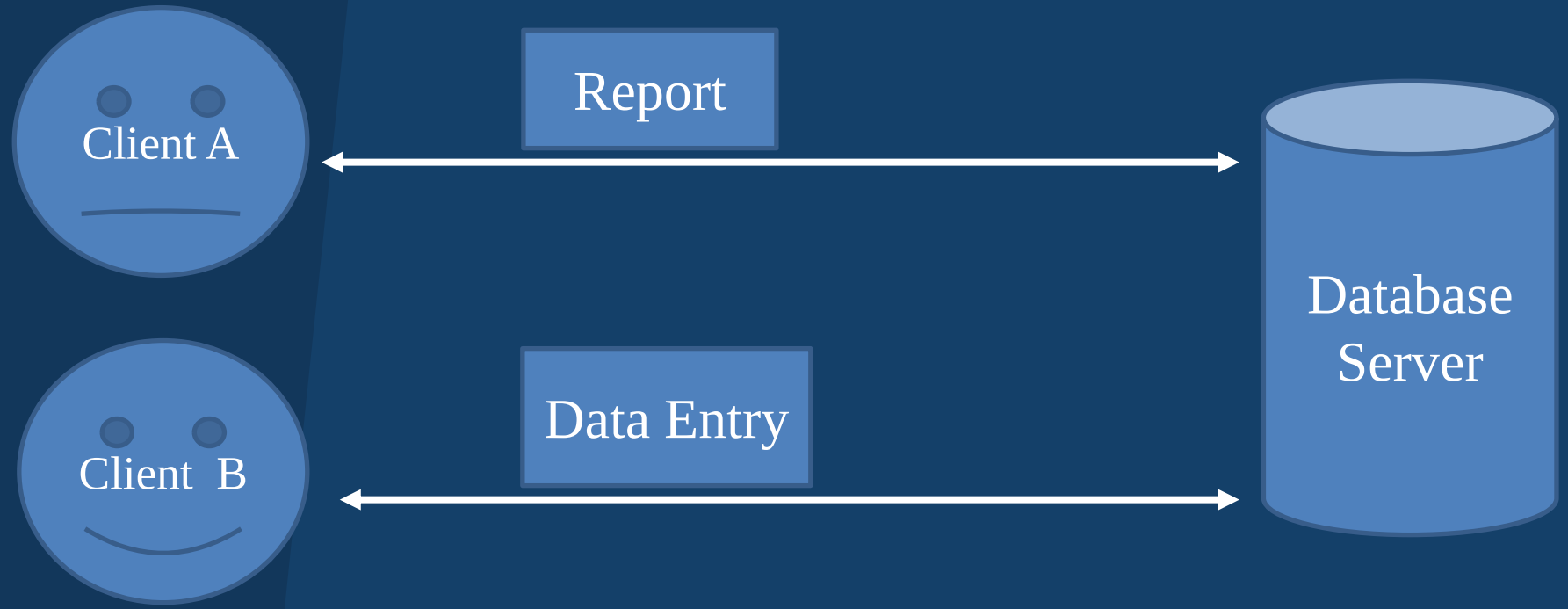
The database is directly available to the user. It means the user can directly sit on the DBMS and uses it.

Any changes done here will directly be done on the database itself. It doesn't provide a handy tool for end users.

Database 1-Tier Architecture

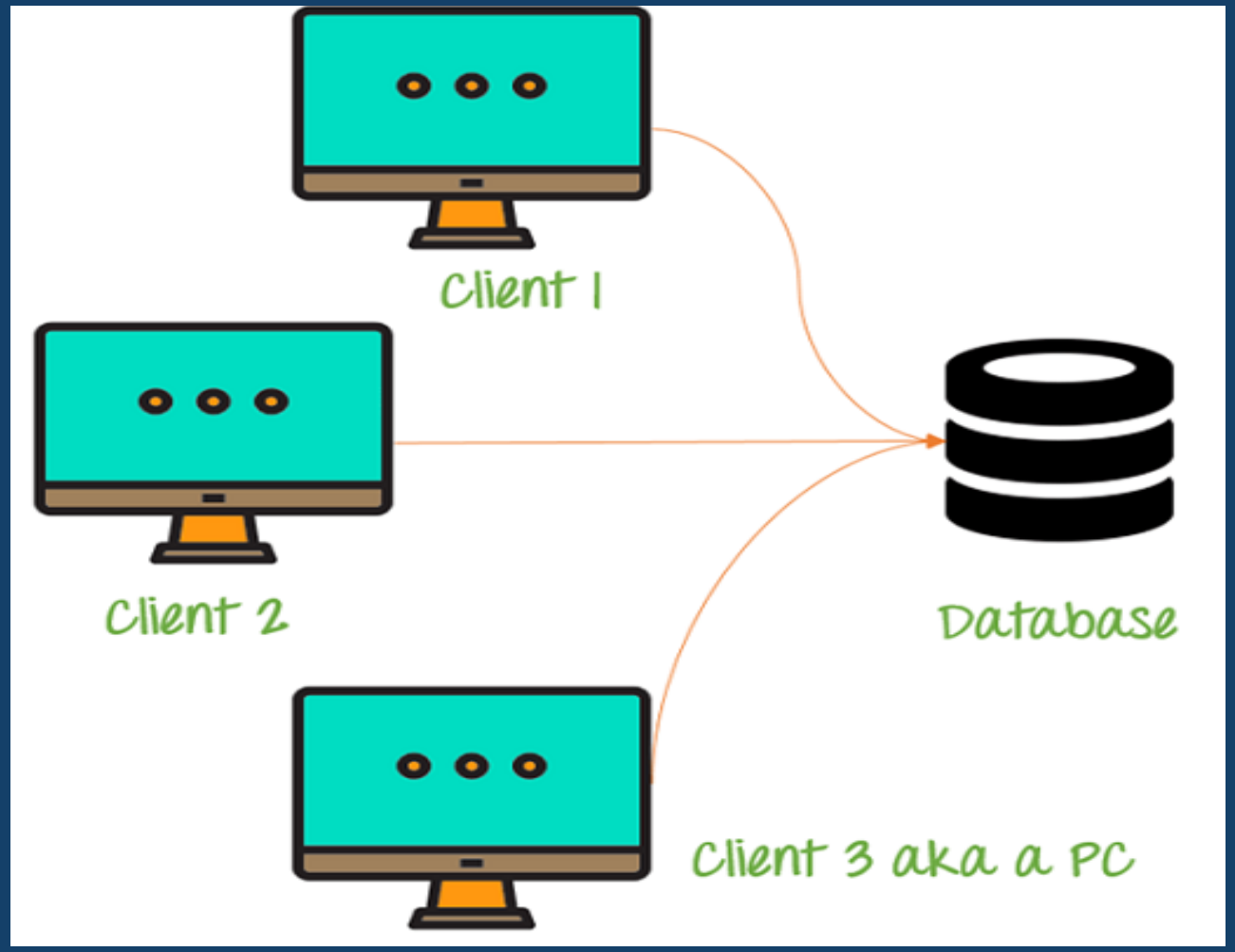


Database 2-Tier Architecture

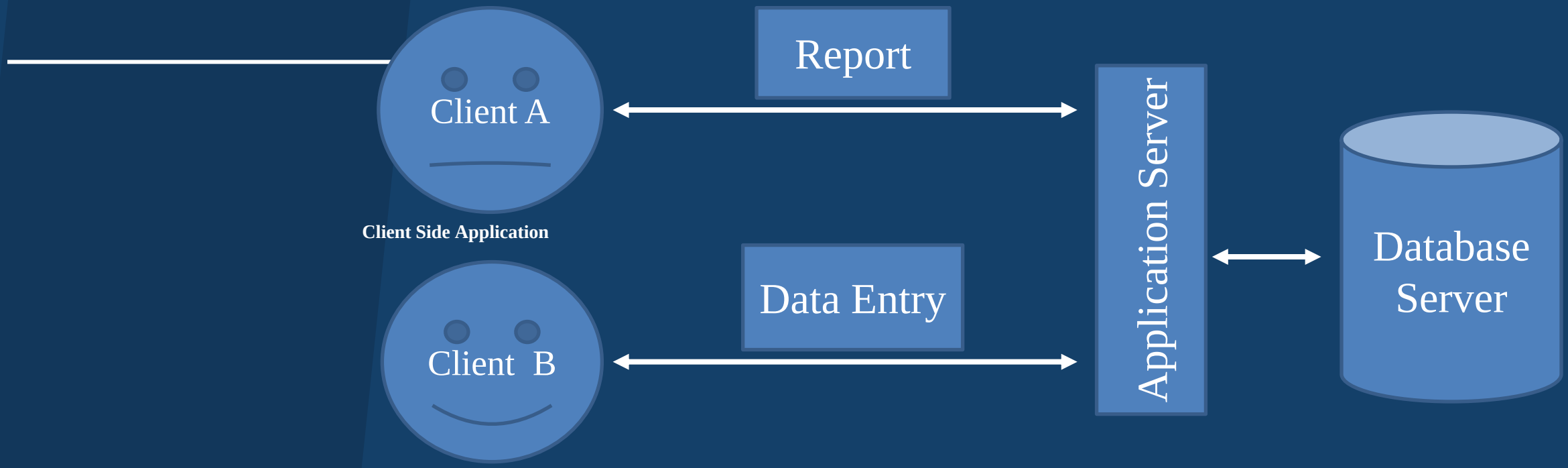


Application logic resides at the client machine and invokes database system functionality at the server machine through query language statements.

Database 2-Tier Architecture

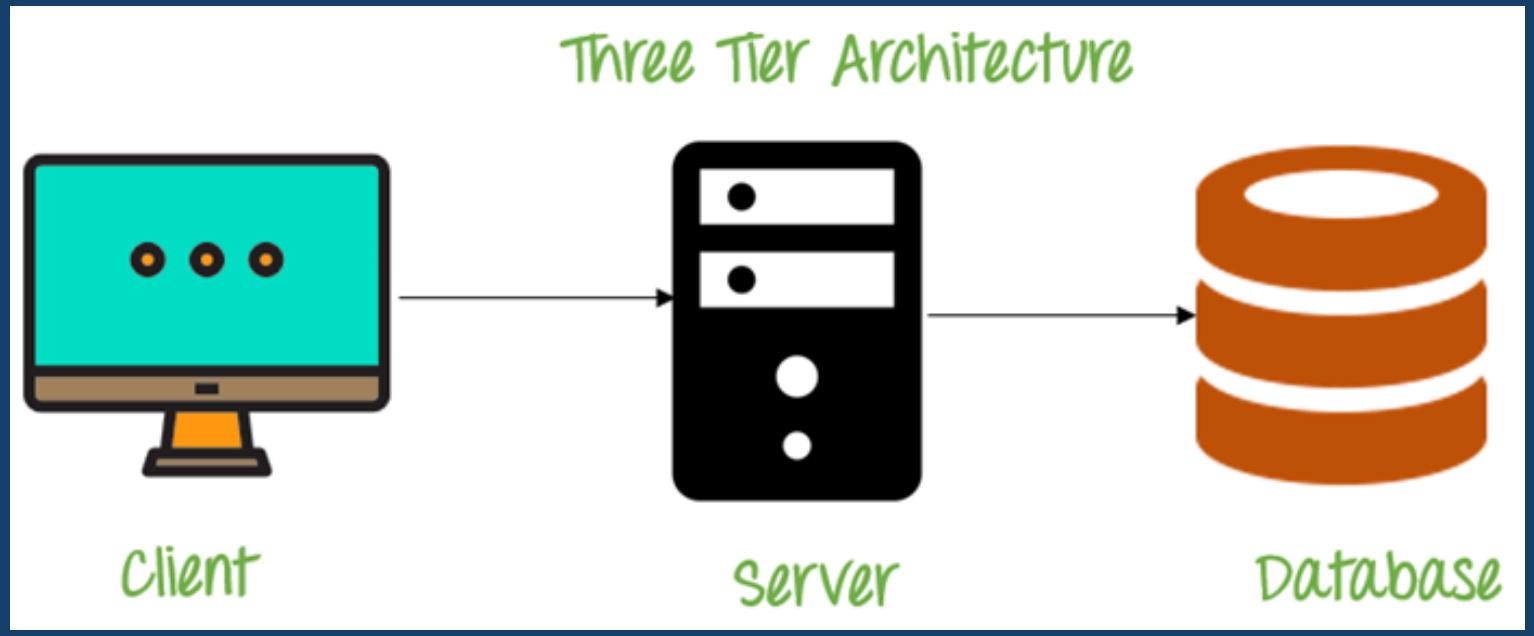


Database 3-Tier Architecture



Client machine acts as merely a front end and does not contain any direct database calls. The front end communicates with an application server. The application server, in turn, communicates with a database system to access data.

Database 3-Tier Architecture



DBMS Architecture: 2-Tire Vs 3-Tire

- Client Server Architecture
- Easy to build & maintain
- Slow Execution
- Less Secured

- Web based Application
- Complex to build & maintain
- Fast Execution
- More Secured

DBMS Components

- Hardware:
  - Processor / Main Memory: Program Execution
  - Secondary Storage: Physical Storage
- Data: Raw facts / Information which store in secondary storage.
- Software: layer between physical database and user.
- Users: Who are using the DBMS.
- Procedures: Set of rules to manage DBMS.

Types of Users in DBMS

- Native Users
  - are unsophisticated users who interact with the system by using predefined user interfaces, such as web or mobile applications.
- Sophisticated Users
  - interact with the system without writing programs, either using a database query language or by using tools such as data analysis software.
- Application programmers
  - are computer professionals who write application programs with many tools to develop user interfaces.
- DBA:
  - has such central control over the system

Database Schema, Instances and Sub Schema

- Schema
  - It is the overall description of the database.
  - Does not specify relationship among files.
  - Example,

    Employee(Emp_Name(varchar)(30), Roll (int){roll not null})

- Instance
  - Collection of information stored in the database at a particular moment is called as an instance.

- Sub-schema
  - It is an application programmer's or user's view of the data item types and record types, which he or she uses.

Database Schema Vs Database Instances

– Schema
  - Database Description
  - Changes rarely

Employee

EmpId	name	salary	did

Department

DeptId	DeptName

– Instance
  - Snapshot of a database at a specific moment.
  - Frequent changes
  - At t1 = 10 am, Instance 1
  - At t2 = 11 am, Instance 2

Employee

id	name	salary	did
1	A	1000	d1
2	B	2000	d2
3	C	1500	d3

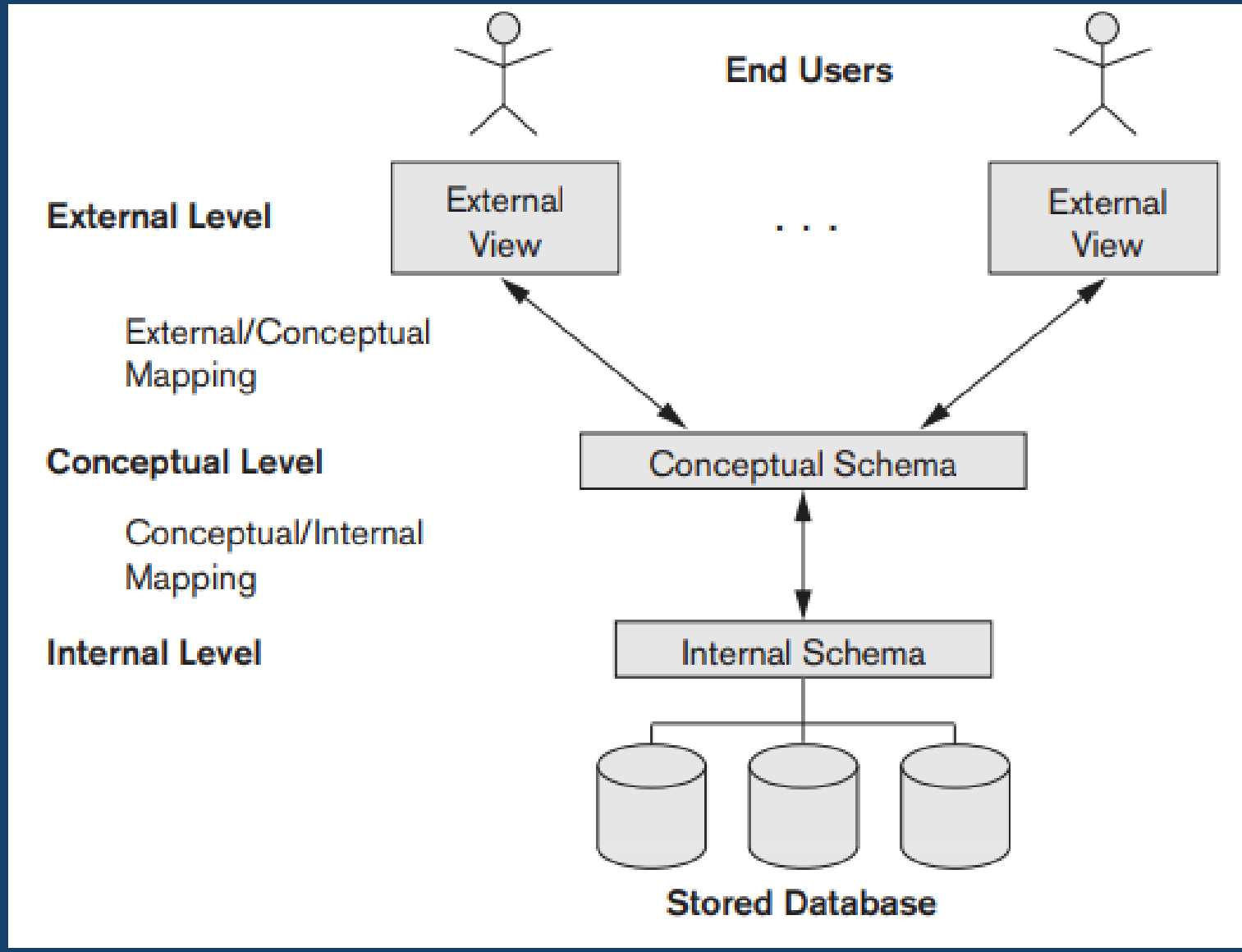
Instance 2

Instance 1

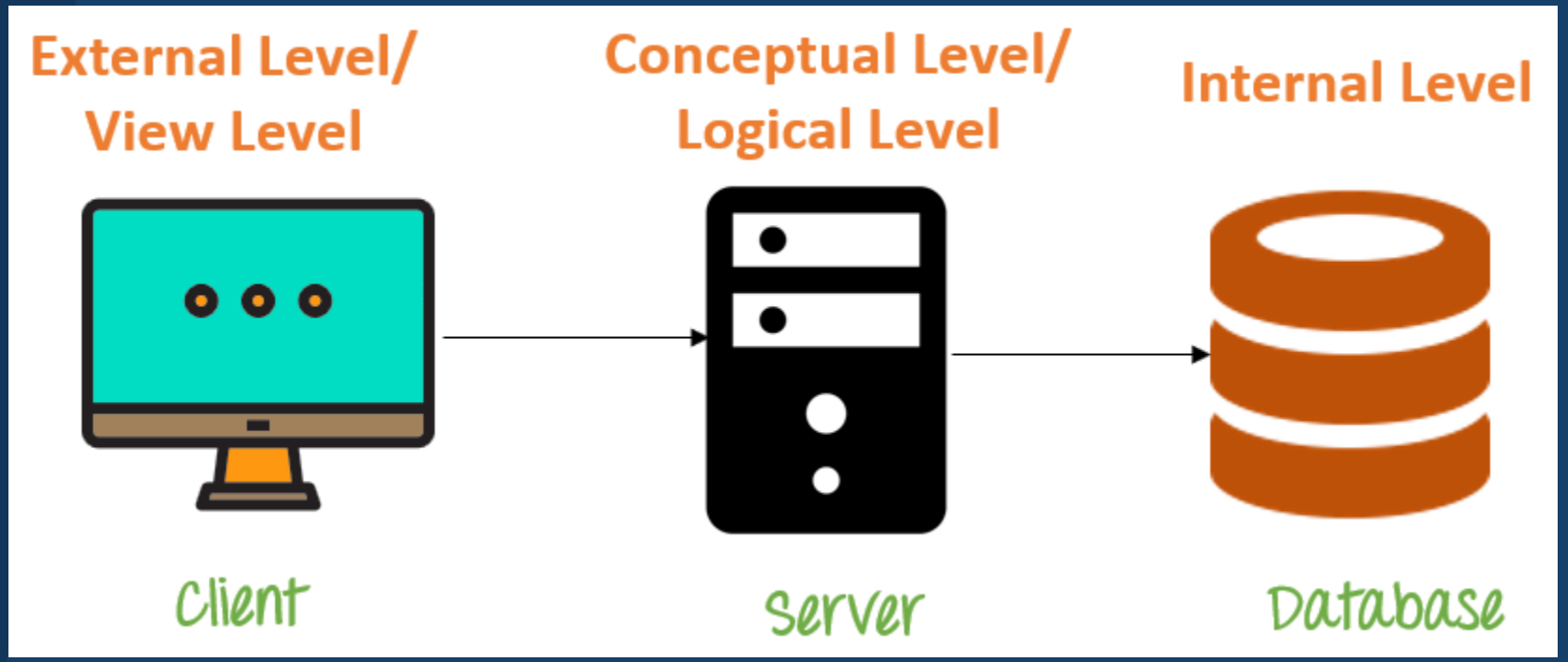
Three-Schema Architecture

- Internal Level / Internal Schema
  - describes the physical storage structure of the database
  - Such as space allocation, file system, data compression, etc.
  - How is data stored?
- Conceptual Level / Conceptual Schema
  - describes the structure of the whole database for a community of users
  - such as entities, data types, relationships, user operations, and constraints.
  - What data is to be stored?
- External or View Level / External Schema
  - includes a number of external schemas or user views.
  - What data or information would be displayed?

Three-Schema Architecture



Three-Schema Architecture



Data Independence

Data independence can be defined as the capacity to change the schema at one level of a database system without having to change the schema at the next higher level.

Or

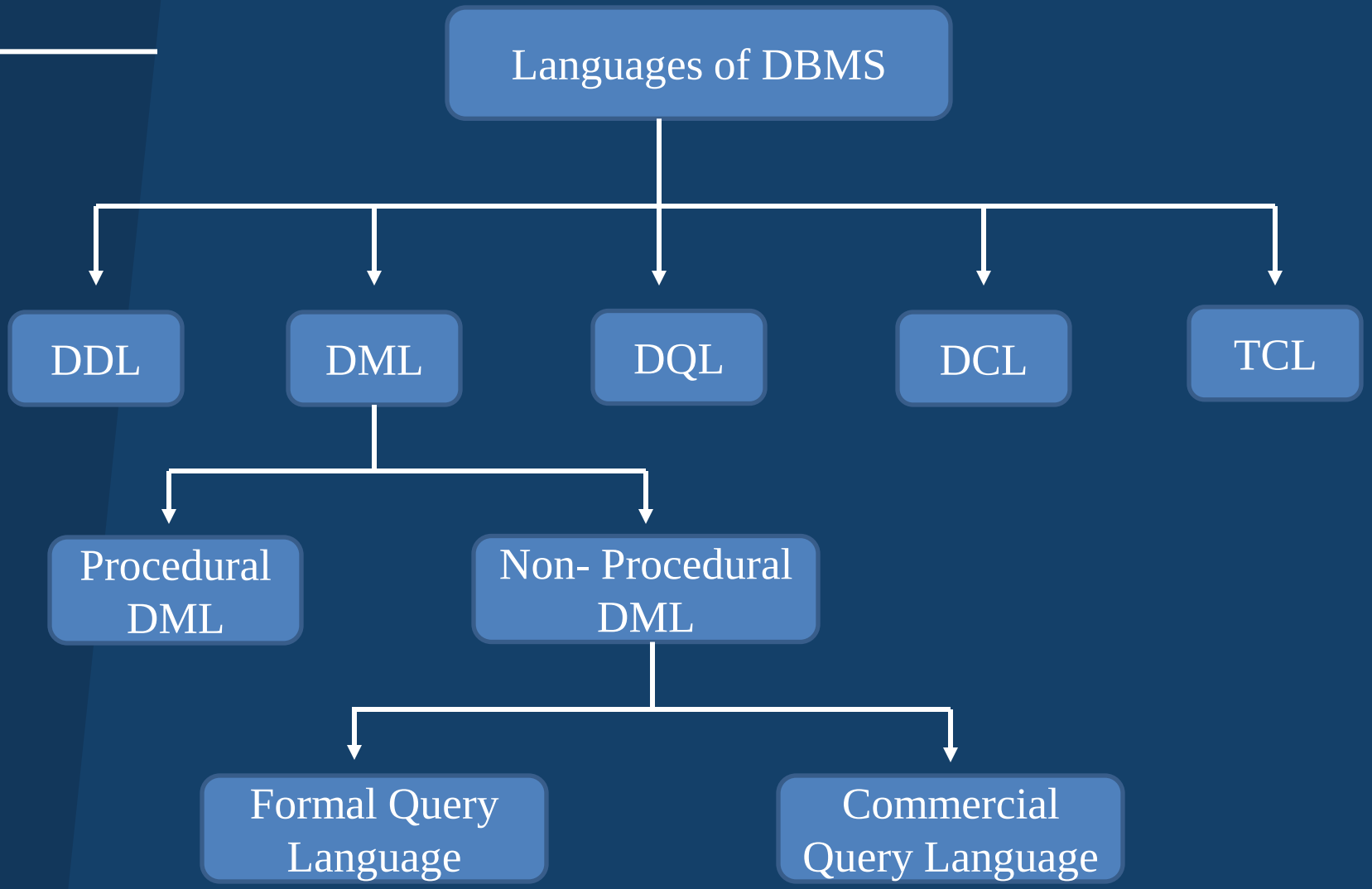
Upper levels are unaffected by changes in lower level.

– Two types of data independence:

- Logical data independence is the capacity to change the conceptual schema without having to change external schemas or application programs.

- Physical data independence is the capacity to change the internal schema without having to change the conceptual schema.

Database Languages



Data Definition Language (DDL)

- DDL is used to define database structure or pattern.
- It performs operations on schema, tables, indexes, constraints, etc. in the database.
- Commands:
  - Create: Create objects in the database.
  - Alter: Alter the structure of the database.
  - Drop: delete objects from the database.
  - Truncate: Remove all records from a table.
  - Rename: Rename an object.

Data Manipulation Language (DML)

- DML is used for manipulating data in a database.
- It handles user requests.
- Commands:
  - Insert: Insert data into a table.
  - Update: Update existing data within a table.
  - Delete: Delete records from a table.
  - etc.

Data Query/Select Language (DQL / DSL)

- DQL/DSL is used for retrieving or accessing data in a database.
- It handles user requests.
- Commands:
  - Select: Retrieve data from a database.

Data Control Language (DCL)

- DCL is used to retrieve the stored or saved data.
- Commands:
  - Grant: give user access privileges to a database.
  - Revoke: take back permissions from the user.

Transaction Control Language (TCL)

- TCL is used to run the changes made by the DML statement.
- TCL can be grouped into a logical transaction.
- Commands:
  - Commit: Save the transaction on the database.
  - Rollback: Restore the database to original since the last Commit.

EIU
TRƯỜNG ĐẠI HỌC
QUỐC TẾ
MIỀN ĐÔNG
EASTERN
INTERNATIONAL
UNIVERSITY

Q&A