UNIT 1

Lecture 2

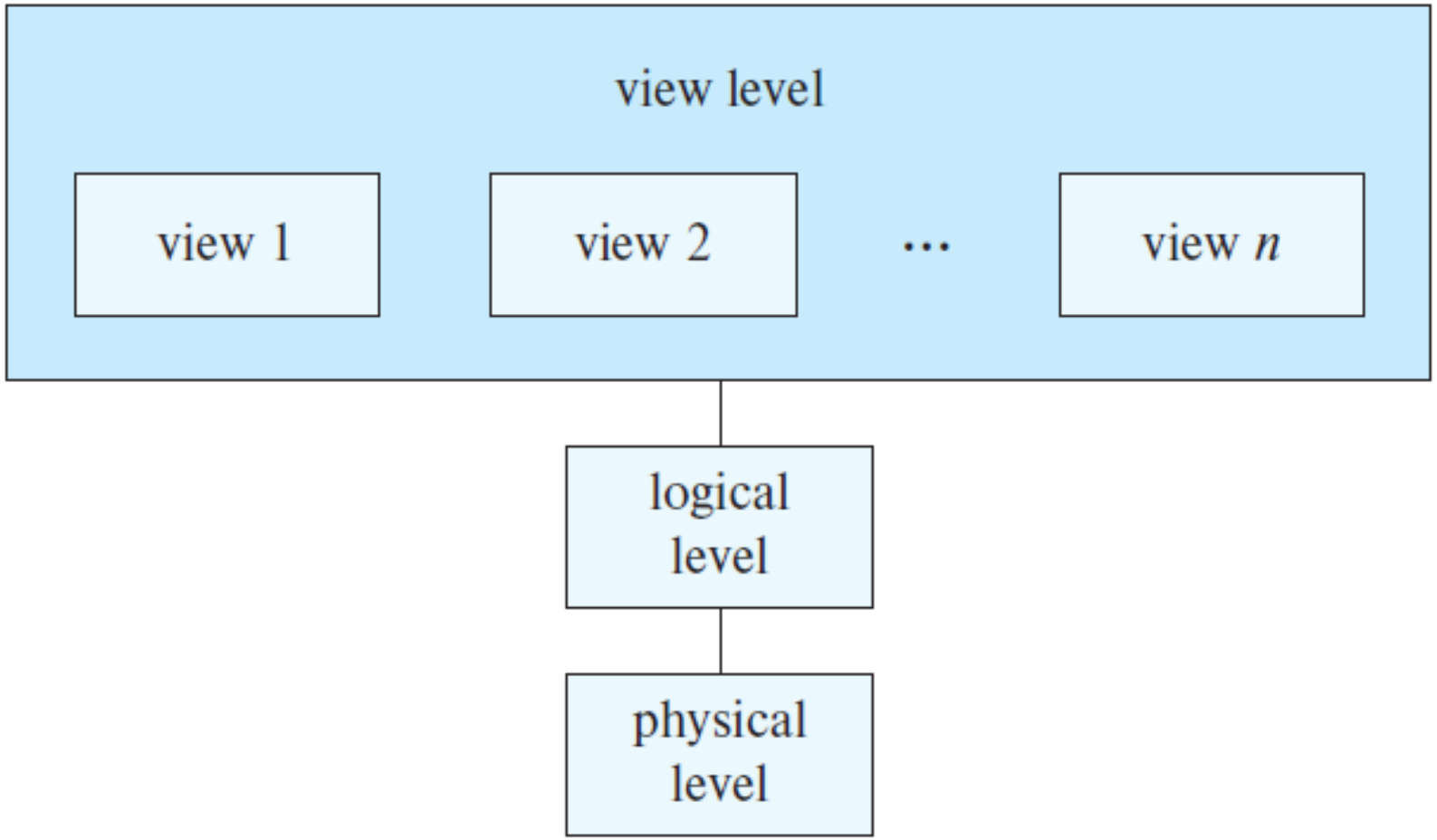
Data Abstraction

- It can be summed up as follows.
  - When the DBMS hides certain details of how data is stored and maintained, it provides what is called as the abstract view of data.
  - This is to simplify user-interaction with the system.
  - Complexity (of data and data structure) is hidden from users through several levels of abstraction.

Data Abstraction

- Data abstraction is used for following purposes:
  - To provide abstract view of data.
  - To hide complexity from user.
  - To simplify user interaction with DBMS.

Levels of Data Abstraction



Levels of Data Abstraction

- **Physical level:** It describes ***how*** a record (e.g., customer) is stored.
- **Features:**

  a) Lowest level of abstraction.

  b) It describes how data are actually stored.

  c) It describes low-level complex data structures in detail.

  d) At this level, efficient algorithms to access data are defined.

Levels of Data Abstraction

- **Logical level:** It describes ***what*** data stored in database, and the ***relationships*** among the data.
- **Features:**

a) It is next-higher level of abstraction. Here whole Database is divided into small simple structures.

b) Users at this level need not be aware of the physical-level complexity used to implement the simple structures.

c) Here the aim is ease of use.

d) Generally, database administrators (DBAs) work at logical level of abstraction..

Levels of Data Abstraction

- **View level:** Application programs hide details of data types. Views can also hide information (e.g., salary) for security purposes.
- **Features:**

a) It is the highest level of abstraction.

b) It describes only a part of the whole Database for particular group of users.

c) This view hides all complexity.

d) It exists only to simplify user interaction with system.

e) The system may provide many views for the whole system.

Instances and Schemas

- Instances and Schemas are similar to types and variables in programming languages.
- **Schema:**
- The overall design of a database is called ***database schema***.
- E.g., the database consists of information about a set of customers and accounts and the relationship between them.
- It is analogous to variable along with its type information in a program.

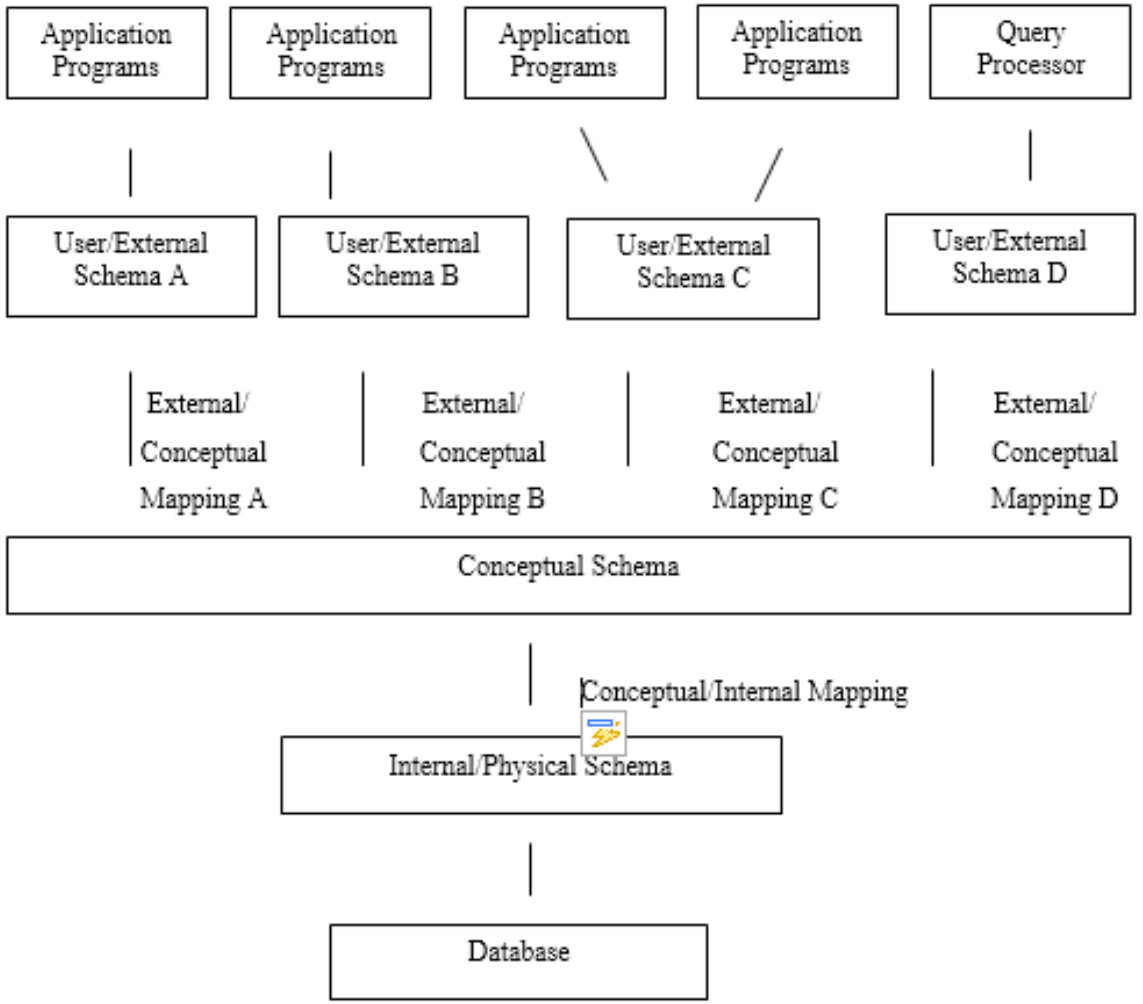
Instances and Schemas

- Types of Schemas (partitioned according to levels of abstraction):
  - **Physical schema**: It is database design at the physical level. It is hidden below logical schema, and can be changed easily without affecting application programs.
  - **Logical schema**: It is database design at the logical level. Programmers construct applications using logical schema. It is by far the most important schema, in terms of its effect on application programs.
  - **Subschema**: It is schema at view level.

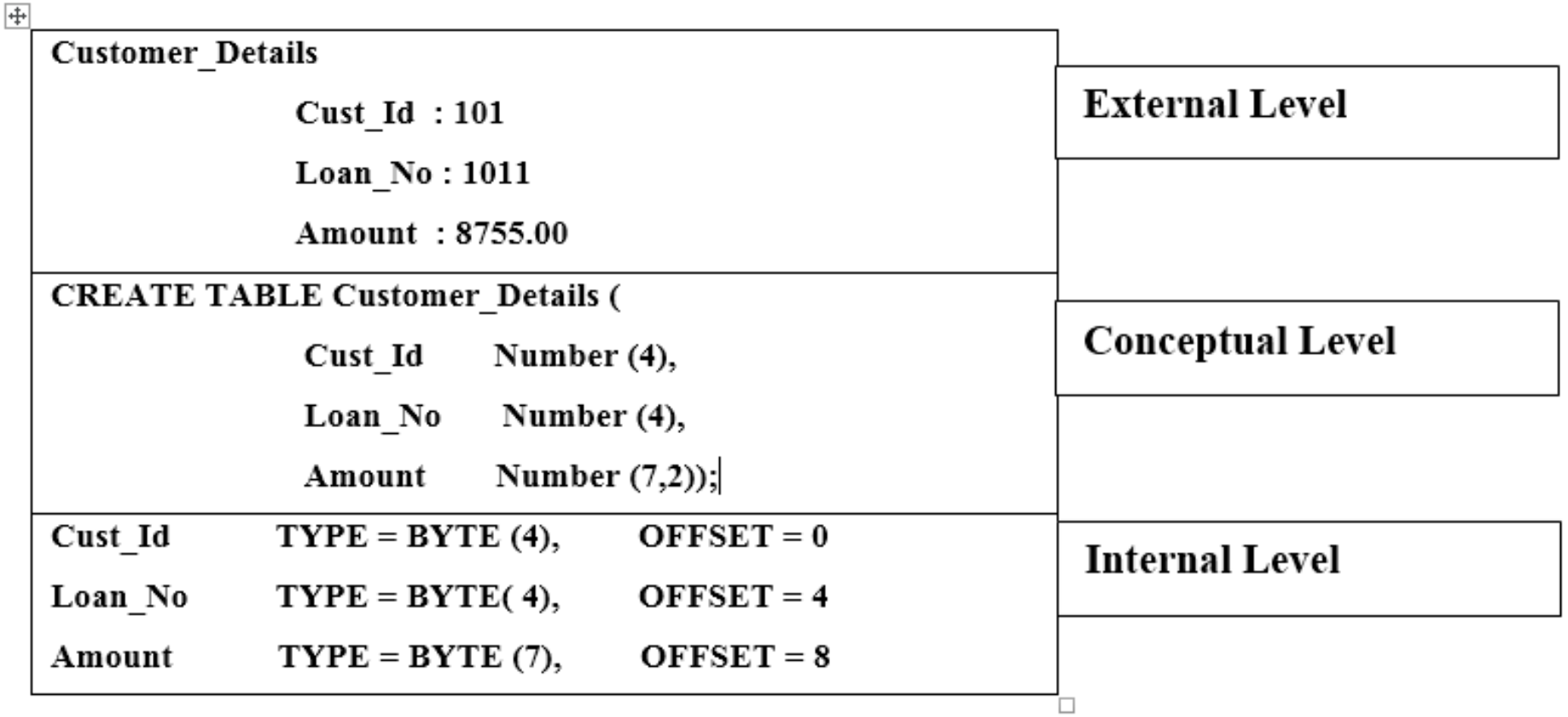
Instances and Schemas

- **Instance:**
- It is the actual content of the database at a particular point in time. It is analogous to the value of a variable.

ANSI / SPARC 3 level Architecture



ANSI / SPARC 3 level Architecture



Data Independence

- Data independence is the ability to modify a schema definition in one level without affecting a schema definition in a higher level is called data independence.

Data Independence

- There are two types of ‘data independence’:
- **Physical data independence**
  - It is the ability to modify the physical scheme without causing application programs to be rewritten.
  - Modifications at this level are usually to improve performance.
- **Logical data independence**
  - It is the ability to modify the conceptual scheme without causing application programs to be rewritten
  - It is usually done when logical structure of database is altered.
- Logical data independence is harder to achieve as the application programs are usually heavily dependent on the logical structure of the data. An analogy is made to abstract data types in programming languages.

University Questions based on lecture

1. Explain ANSI/ SPARC 3 level architecture of DBMS. Why the mapping between different levels are done?
2. Explain the level of data abstraction.
3. Write short notes on schema and instance.
4. What is data independence?

For Video lecture on this topic please subscribe to my youtube channel.

The link for my youtube channel is

https://www.youtube.com/channel/UCRWGtE76JlTp1iim6aOTRuw?sub_confirmation=1