



CSE 215: Database

Department of CSE
Bangladesh University of Engineering & Technology



Course Overview

- Class Schedule
 - Saturday (Sec A: 2 pm, Sec B: 11 am, Sec C: 9 am)
 - Sunday (Sec A: 10 am, Sec B: 8 am, Sec C: 9 am)
 - Monday (Sec A: 11 am, Sec B: 9 am, Sec C: 12 pm)
- Course Teacher (1st part)
 - Prof. Tanzima Hashem
 - Office: Room# 313
 - Email: tanzimahashem AT cse DOT buet DOT ac DOT bd
- Course page
 - <https://cse.buet.ac.bd/moodle/course/view.php?id=179>



Course Overview

- Text Books
 - Database System Concepts by Silberschatz, Korth, and Sudarsan
- Other Book
 - Database Systems: The Complete Book, by Garcia-Molina, Ullman and Widom



Course Topics (TH)

- Concepts of database systems; Data Models: Entity-Relationship model, Relational model; Query Languages: Relational algebra, SQL; Constraints and triggers;



Database and DBMS

- Database
 - Collection of interrelated data
- Database Management System (DBMS)
 - Collection of interrelated data
 - Set of programs that store, organize, and manage data, and allow users to efficiently add, update, retrieve, and protect information



Key Properties

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



Applications

- Enterprise Information
 - Sales, Accounting, Human Resources
- Banking and Finance
- Manufacturing
- Universities
- Web Based Services
 - Online retailers, online advertisement
- Telecommunication
- Document database
- Navigation systems
- ...



File Systems- Limitations

- Data redundancy and inconsistency
 - Information duplication in different files, inconsistent updates
- 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 (e.g. account balance > 0) become part of program code
 - Hard to add new constraints or change existing ones



File Systems- Limitations

- 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
 - Hard to provide user access to some, but not all, data



Data Models

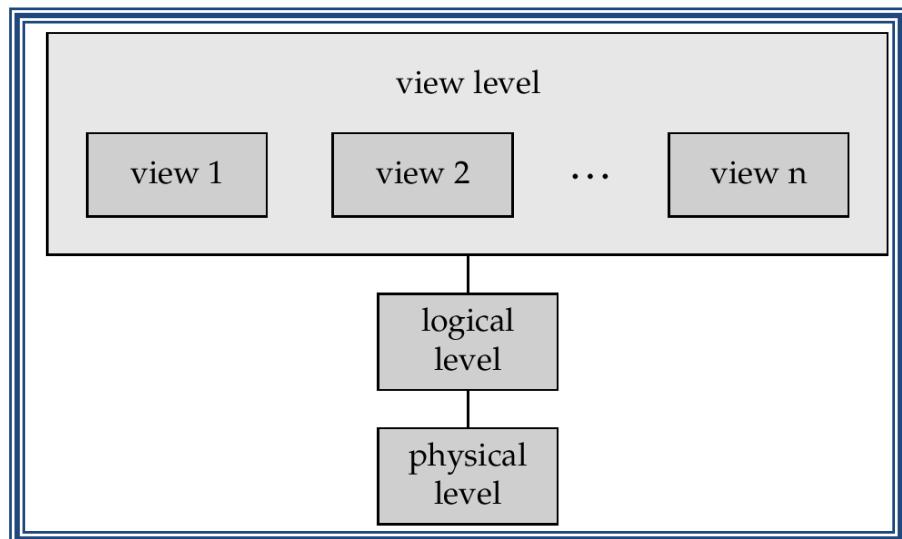
- A collection of tools for describing
 - Data
 - Data relationships
 - Data semantics
 - Data constraints
- Relational model
- Entity-Relationship data model
- Object-based data models
- Semi-structured data model (XML)
- Other older models:
 - Network model, Hierarchical model



Data Abstraction

- Physical level
 - Describes how data is stored
- Logical level
 - Describes data stored in database, and the relationships among the data
- View level
 - Describes only part of the database
 - May provide many views of the same database

Hide the complexity of data structures to represent data in the database from users through several levels of data abstraction





Instances and Schemas

- **Schema** – Overall design of the database
 - Analogous to type information of a variable in a program
- **Instance** – the actual content of the database at a particular point in time
 - Analogous to the value of a variable



Data Definition Language (DDL)

- Specification notation for defining the database schema
 - Example: `create table instructor (`
 `ID char(5),`
 `name varchar(20),`
 `dept_name varchar(20),`
 `salary numeric(8,2))`
- DDL compiler generates a set of table templates stored in a data dictionary



Data Definition Language (DDL)

- Data dictionary contains metadata (i.e., data about data)
 - Database schema
 - Constraints
 - Integrity constraints
 - Referential constraints
 - Domain constraints
 - Authorization



Data Manipulation Language (DML)

- Language for accessing and updating the data organized by the appropriate data model
- There are basically two types of data-manipulation language
 - **Procedural DML** -- require a user to specify what data are needed and how to get those data
 - **Declarative DML** -- require a user to specify what data are needed without specifying how to get those data
- Declarative DMLs are usually easier to learn and use than are procedural DMLs
- Declarative DMLs are also referred to as non-procedural DMLs
- The portion of a DML that involves information retrieval is called a query language



Database Design

- Initial phase -- characterize fully the data needs of the prospective database users
- Second phase -- choosing a data model
 - Applying the concepts of the chosen data model
 - Translating these requirements into a conceptual schema of the database
 - A fully developed conceptual schema indicates the functional requirements of the enterprise
 - Describe the kinds of operations (or transactions) that will be performed on the data



Database Design

- Final Phase -- Moving from an abstract data model to the implementation of the database
 - Logical Design – Deciding on the database schema
 - Database design requires that we find a “good” collection of relation schemas
 - Business decision – What attributes should we record in the database?
 - Computer Science decision – What relation schemas should we have and how should the attributes be distributed among the various relation schemas?
 - Physical Design – Deciding on the physical layout of the database



Database Design

- In designing a database schema, we must ensure that we avoid two major pitfalls:
 - **Redundancy:** a bad design may result in repeat information.
 - Redundant representation of information may lead to data inconsistency among the various copies of information
 - **Incompleteness:** a bad design may make certain aspects of the enterprise difficult or impossible to model

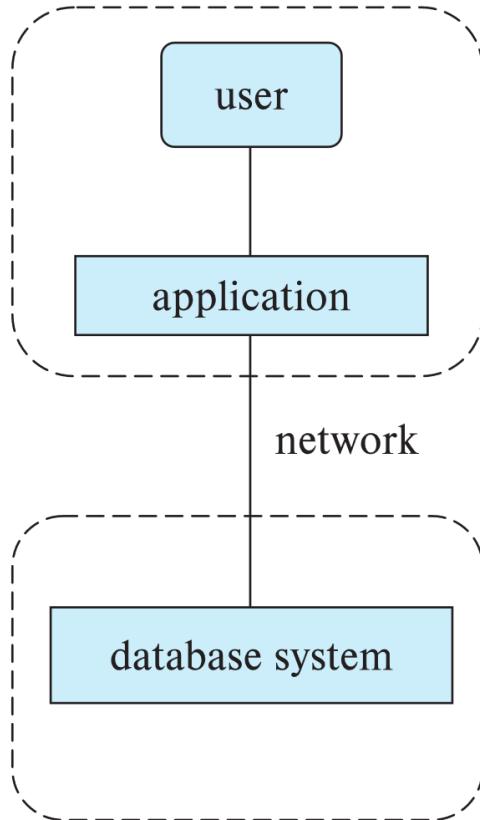


Database Engine

- A database system is partitioned into modules that deal with each of the responsibilities of the overall system
- The functional components of a database system:
 - The storage manager
 - Provides the interface between the low-level data and the application programs and user queries
 - The query processor component
 - Interprets, optimizes, and executes user queries by converting them into efficient low-level instructions
 - The transaction management component
 - Keeps database consistent even after system or transaction failures



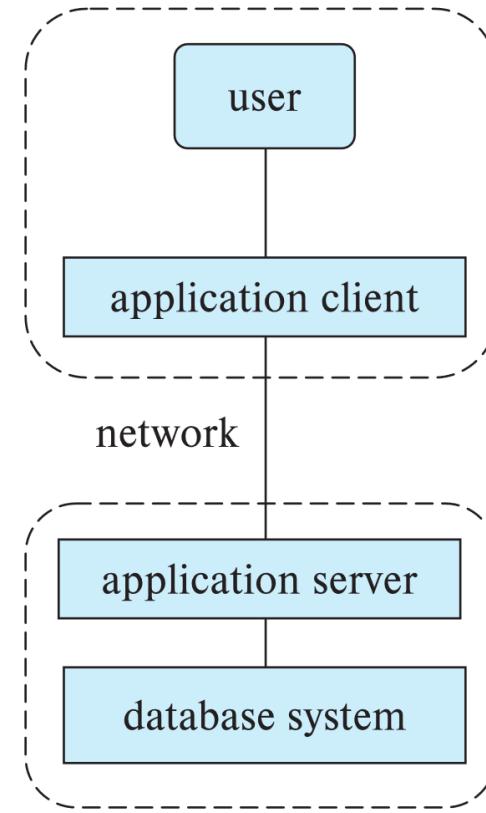
Database Architecture



(a) Two-tier architecture

client

server



(b) Three-tier architecture



Key People

- Database Users
 - Naïve Users: Interact with the system by using the predefined interface
 - Application Programmers: Write application programs
 - Analysts: Submit query to explore data without writing programs
- Database Administrator
 - Has central control over the system
 - Schema definition
 - Storage structure and access-method definition
 - Schema and physical organization modification
 - Granting authorization for data access
 - Routine maintenance



Acknowledgement

- Some slides and figures are collected from
 - <https://www.db-book.com/>