CS275 – Intro to Databases

How does a DBMS work? - Chap. 1.3 - 1.9

File Systems vs. DBMS

- We have to write special programs for queries
- We have to protect data from inconsistencies
- We have to handle crashes
- We have to control access w/ passwords

- What is the common theme?

Why is a DBMS so important?

- A DBMS is software designed to assist in maintaining and utilizing large collection of data.
  - A large amount of data,
  - Concurrent access by many users,
  - Fast access,
  - Consistent data update,
  - Role-based security,
  - Robust against hardware failures and OS crashes.
Why and When Do We Need DBMS?

• Advantages of DBMS
  – Data independence (you can still access the data using Windows 95).
    • Hide data representation
    • Hide data storage

• Advantages of DBMS
  – Efficient data access
    • High-quality data compression schemes,
    • Fast data retrieval and search algorithms.

• Advantages of DBMS
  – Data integrity and security
    • Prevent invalid queries from being executed,
    • Access control can be enforced.
Why and When Do We Need DBMS?

• Advantages of DBMS
  – Centralized data administration
    • The DBA can optimize the organization of the data to facilitate its uses.

• Advantages of DBMS
  – Concurrent access
  – Crash recovery
  – Reduced application development time

• Disadvantages of DBMS
  – High startup cost (time and effort)
  – Relatively high maintenance
When Not to Use a DBMS?

How Is Data Represented?

• A data model
  — A collection of high-level data descriptions,
  — Hides low-level storage details.
• A semantic data model
  — More abstract,
  — Serves as a startup point for the design,
  — Farther away from the physical storage than a data model.

How Is Data Represented?

• The relational model
  — Relation (records)
  — Schema
    • Data descriptions, such as name of the relation and individual field.
    Students(sid: string, name: string, login: string,
    age: integer, gpa: real)
  — Integrity Constraints
How Is Data Represented?

• Levels of Abstraction

How Is Data Represented?

• Conceptual Schema (logical schema)
  – Data Model/Relationships

Students(sid: string, name: string, login: string, age: integer, gpa: real)
Faculty(fid: string, fname: string, sal: real)
Courses(cid: string, cname: string, credits: integer)
Enrolled(sid: string, cid: string, grade: string)
Teaches(fid: string, cid: string)

How Is Data Represented?

• Physical Schema
  – Data Storage
  – Based on Access
How Is Data Represented?

• External Schema
  – Different Views
  – Defined by end user requirements

Courseinfo(cid: string, fname: string, enrollment: integer)

How Is Data Represented?

• Data Independence
  – Logical data independence

How Is Data Represented?

• Data Independence
  – Physical data independence
How Is Data Retrieved and Manipulated?

• Queries
  – Data manipulation language (DML)
    • Retrieval
    • Add
    • Delete
    • Update

How Does DBMS Support Concurrent Access & Crashes?

• Transaction
  – Any one execution of a user program in DBMS.
    • A single command
    • Group several commands/queries into one transaction
  – Locking protocol
    • Shared vs. Exclusive
    • Row level vs. Table level
    • Read vs. Write
  – Log
    • Finalizes transaction
    • Checkpoints

Who is Involved With Databases in Real Life?

• Database application programmers
• Database administrators
  – Conceptual and physical schemas
  – Security and authorization
  – Data availabilities and recoveries
  – Database tuning
Questions?