

## IT420: Database Management and Organization

---

Dr. Crăiniceanu  
Capt. Balazs

[www.cs.usna.edu/~adina/teaching/it420/spring2007](http://www.cs.usna.edu/~adina/teaching/it420/spring2007)

## Outline

---

- Class Survey
- Why Databases (DB)?
  - A Problem
  - DB Benefits
- In This Class?
- Admin
  - Syllabus
  - Policy

## Database Management and Organization

---

- How does Wal-Mart manage its 200 TB data warehouse?
- What is the database technology behind eBay's website?
- How do you build an Oracle 9i, MySQL or Microsoft SQL Server database?

## Database Management Systems (DBMS)

---

- Information is one of the most valuable resources in this information age
- How do we effectively and efficiently manage this information?
  - Relational database management systems
    - Dominant data management paradigm today
- 6 billion dollars a year industry!

## ICE: The Mid Store

---

- Create a system to keep track of inventory

## Problems

---

- Changes to data - Data model
- “on the fly” queries
- Data inconsistencies
- Security of information (views)
- Performance
- Partial processing
- Concurrency

## Why Database Management Systems?

---

- Benefits
  - High-level abstractions for data modeling, access, manipulation, and administration
  - Data integrity and security
  - Performance and scalability
  - Transactions (concurrent data access, recovery from system crashes)

## Data Model

---

- Entity-Relationship model
- Relational model
- Object-oriented model
- Object-relational model
- XML

## The Relational Data Model

---

- Turing Award for Codd in 1980
- Tables
- Database

cid	name	byear	state
1	Jones	1960	NY
2	Smith	1974	CA
3	Smith	1950	NY

## The Object-Oriented Data Model

---

- Richer data model. Goal: Bridge mismatch between programming languages and the database system.
- Example components of the data model:
  - Relationships between objects directly as pointers.
- Result: Can store abstract data types directly in the DBMS
  - Pictures
  - Geographic coordinates
  - Movies
  - CAD objects

## Object-Oriented DBMS

---

- Advantages:
  - Engineering applications (CAD and CAM and CASE computer aided software engineering), multimedia applications.
- Disadvantages:
  - Querying is much harder

## Object-Relational DBMS

---

- Mixture between the object-oriented and the object-relational data model
  - Combines ease of querying with ability to store abstract data types
  - Conceptually, the relational model, but every column type is a class
- All major relational vendors are currently extending their relational DBMS to the object-relational model

## XML

Computer Table

Id	Speed	RAM	HD
101	800Mhz	256MB	40GB
102	933Mhz	512MB	40GB

```

<Table>
  <Computer Id='101'>
    <Speed>800Mhz</Speed>
    <RAM>256MB</RAM>
    <HD>40GB</HD>
  </Computer>
  <Computer Id='102'>
    <Speed>933Mhz</Speed>
    <RAM>512MB</RAM>
    <HD>40GB</HD>
  </Computer>
</Table>

```

## Query Languages

- We need a high-level language to describe and manipulate the data
- Requirements:
  - Precise semantics
  - Easy integration into applications written in C++/Java/Visual Basic/etc.
  - Easy to learn
  - DBMS needs to be able to efficiently evaluate queries written in the language

## SQL: Structured Query Language

- IBM (System R) in the 1970s
- ANSI standard since 1986
- Example:

**SELECT** \*

**FROM** Customers

**WHERE** Customers.cid = 3

cid	name	byear	state
1	Jones	1960	NY
2	Smith	1974	CA
3	Smith	1950	NY

cid	name	byear	state
3	Smith	1950	NY

## Why Database Management Systems?

- Benefits
  - High-level abstractions for data modeling, access, manipulation, and administration
  - **Data integrity and security**
  - Performance and scalability
  - Transactions (concurrent data access, recovery from system crashes)

## Integrity Constraints

---

- Integrity Constraints (ICs): Condition that must be true for any instance of the database.
- ICs are specified when schema is defined.
- ICs are checked when tables are modified.
- A legal instance of a table is one that satisfies all specified ICs.
- DBMS should only allow legal instances.
- Example: Domain constraints.

## Security

---

- **Secrecy**: Users should not be able to see things they are not supposed to.
  - E.g., A student can't see other students' grades.
- **Integrity**: Users should not be able to modify things they are not supposed to.
  - E.g., Only instructors can assign grades.
- **Availability**: Users should be able to see and modify things they are allowed to.

## Why Database Management Systems?

---

- Benefits
  - High-level abstractions for data modeling, access, manipulation, and administration
  - Data integrity and security
  - **Performance and scalability**
  - Transactions (concurrent data access, recovery from system crashes)

## DBMS and Performance

---

- Efficient implementation of all database operations
- Indexes
- Query optimization
- Automatic high-performance concurrent query execution, query parallelization

## Why Database Management Systems?

---

- Benefits
  - High-level abstractions for data modeling, access, manipulation, and administration
  - Data integrity and security
  - Performance and scalability
  - Transactions (concurrent data access, recovery from system crashes)

## What is a Transaction?

---

The execution of a program that performs a function by accessing a database.

- Examples:
  - Buy an airline ticket.
  - Withdraw money from an ATM.
  - Verify a credit card sale.
  - Order an item from an Internet retailer.

## Transactions

---

- A transaction is an atomic sequence of actions
- Each transaction must leave the system in a consistent state
- The ACID Properties

## Example Transaction: Online Store

---

Your purchase transaction:

- **Atomicity**: Either the complete purchase happens, or nothing
- **Consistency**: The inventory and internal accounts are updated correctly
- **Isolation**: It does not matter whether other customers are also currently making a purchase
- **Durability**: Once you have received the order confirmation number, your order information is permanent, even if the site crashes

## What Makes Transaction Processing Hard?

- Reliability
- Availability
- Response time
- Throughput
- Scalability
- Security
- Configurability
- Atomicity
- Durability
- Distribution

## What Makes TP Important?

- It is at the core of electronic commerce
- Most medium-to-large businesses use TP for their production systems.
- It is a huge slice of the computer system market – over \$50 B/year

## Summary Of DBMS Benefits

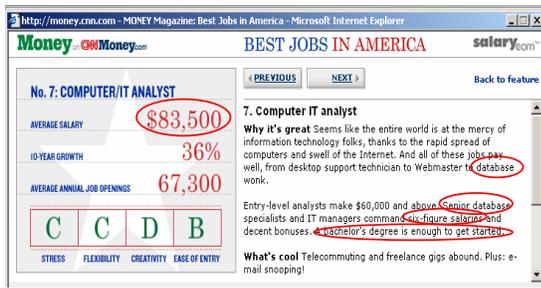
- High-level abstractions for data access
  - Data models
- Data integrity and security
  - Key constraints, integrity constraints, access control
- Performance and scalability
  - Parallel DBMS, distributed DBMS, performance tuning
- Transactions
  - ACID properties, concurrency control, recovery

## Best Jobs!

The screenshot shows the CNN Money.com website. At the top, there is a navigation bar with links for HOME, NEWS, MARKETS, TECHNOLOGY, JOBS & ECONOMY, PERSONAL FINANCE, and AUTO. Below the navigation bar, the main heading is "BEST JOBS IN AMERICA" in red. Underneath, it says "MONEY Magazine and Salary.com rate careers on salary and job prospects." and "By Tara Kalkwinski, Daphne Mosher, Janet Paskin and Donna Rosato". On the right side, there is a "GET QUOTES" button and a "Money salary.com" logo. The main content area features a large "10" in a blue box with stars, indicating the top 10 best jobs. Below this, the text reads "Top 10 best jobs" and "MONEY Magazine and Salary.com researched hundreds of jobs, considering their growth, pay, stress-levels and other factors. These careers ranked highest. (more)". The list of jobs is as follows:

1. Software Engineer	6. Market research analyst
2. College professor	7. Computer IT analyst
3. Financial adviser	8. Real Estate Appraiser
4. Human Resources Manager	9. Pharmacist
5. Physician assistant	10. Psychologist

## IT Analyst



## Course Topics

- Database design
- Relational model
- SQL
- Normalization
- Database administration
- PHP
- MySQL

## Course Goals

- Explain the main advantages of modern database management systems over file systems.
  - Design, create, and query relational databases to satisfy user requirements.
  - Design, build and deploy database-backed applications with dynamic website front-end.
  - Implement data access control mechanisms for database and application security.
  - Analyze the ethical issues and responsibilities related to records management
- Create applications that USE a Database Management System

## Things We Will NOT Cover

- Relational algebra and calculus
  - Implementation of index structures
  - Query evaluation and optimization
- How to BUILD a Database Management System

## Success in IT420

---

- Lecture – stay engaged
  - Take notes – provided slides are not enough!
  - Exams closed-book – but open-note!
  - Ask & answer questions
- Make the most of in-class lab time
  - Read lab in advance
  - Think before you start typing
  - Don't stay stuck!
- Don't fall behind
  - Finish lab early and leave time for reading
  - See me for help and/or talk to friends
  - Course material builds on itself and gets more complex

## Academic Integrity - Honor

---

- Collaboration on labs/ hws is allowed, but submitted work should be your own
  - Cite any assistance, from any sources
- **Collaboration on projects, quizzes and exams is prohibited**
- <http://www.cs.usna.edu/academics/honor.htm>

## Resources

---

- Lecture slides / your notes
- Textbook: Database Processing by David Kroenke
- Database Management Systems by R. Ramakrishnan and J. Gehrke
- PHP and MySQL Web Development by L. Welling and L. Thomson