

Introduction to Database Systems CSE 444

Lecture #1
September 26, 2007

1

Staff

- Instructor: Hal Perkins
 - CSE 548, perkins@cs
 - Office hours: Mondays 3:40-4:30 and tbd, CSE ugrad labs. (Where would be best?)
- TAs:
 - Andrew Hemmaplardh, plardh@cs
 - Huei-hun Elizabeth Tseng, lachesis@cs
- Office hours tbd

2

Communications

- Web site: www.cs.washington.edu/444
 - Lectures will be available here (usually after class)
 - Homework will be posted here (HW1 is posted)
 - The project description will be here
- Discussion board
 - Will be linked from web site
 - Please watch, contribute
- Mailing list
 - Everyone automatically subscribed
 - Mainly for announcements from course staff, etc.

3

Textbook(s)

Main textbook, available at the bookstore:

- *Database Systems: The Complete Book*, Hector Garcia-Molina, Jeffrey Ullman, Jennifer Widom
 - Most chapters are good. Some are not (functional dependencies).
- **COME TO CLASS ! ASK QUESTIONS ! READ SLIDES !**

4

Other Texts

Available at the Engineering Library (not on reserve):

- *Database Management Systems*, Ramakrishnan
- *XQuery from the Experts*, Katz, Ed.
- *Fundamentals of Database Systems*, Elmasri, Navathe
- *Foundations of Databases*, Abiteboul, Hull, Vianu
- *Data on the Web*, Abiteboul, Buneman, Suciu

5

Outline of Today's Lecture

1. Overview of DBMS
2. DBMS through an example
3. Course outline
4. Assignment 1, Homework 1

6

Database

What is a database ?

Give examples of databases

7

Database

What is a database ?

- A collection of files storing related data

Give examples of databases

- Accounts database; payroll database; UW's students database; Amazon's products database; airline reservation database

8

Database Management System

What is a DBMS ?

Give examples of DBMS

9

Database Management System

What is a DBMS ?

- *A big C/C++ program written by someone else that allows us to manage efficiently a large database and allows it to persist over long periods of time*

Give examples of DBMS

- DB2 (IBM), SQL Server (MS), Oracle, Sybase
- MySQL, Postgres, ...

10

RDBMS Market Shares

From 2006, www.gartner.com

- Oracle: 47% market share, \$7.2BN in sales
- IBM: 21% market share with \$3.2BN in sales
- Microsoft: 17% market with \$2.6BN in sales

11

An Example

The Internet Movie Database

<http://www.imdb.com>

- Entities:
Actors (800k), Movies (400k), Directors, ...
- Relationships:
who played where, who directed what, ...

12

Tables

Directors:

id	fName	lName
15901	Francis Ford	Coppola
...		

Movie_Directors:

id	mid
15901	130128
...	

Movies:

mid	Title	Year
130128	The Godfather	1972
...		

13

What the Database Systems Does

1. Create/store large datasets
2. Search/query/update
3. Change the structure
4. Concurrent access to many user
5. Recover from crashes
6. Security

14

Possible Organizations

- Files
- Spreadsheets
- DBMS

15

1. Create/store Large Datasets

- Files
- Spreadsheets
- DBMS

Yes, but...

Not really...

Yes

16

2. Search/Query/Update

- Files
- Spreadsheets
- DBMS

Simple queries (grep);
Updates are difficult

Simple queries;
Simple updates

All

Updates: generally OK

17

3. Change the Structure

Add Address to each Actor

- Files
- Spreadsheets
- DBMS

Very hard

Yes

Yes

18

4. Concurrent Access

Multiple users access/update the data concurrently

Lost updates; inconsistent reads,...

- What can go wrong ?
- How do we protect against that in OS ? locks
- This is insufficient in databases; why ?

A logical action consists of multiple updates

19

5. Recover from crashes

- Transfer \$100 from account #4662 to #7199:

```
X = Read(Account, #4662);
X.amount = X.amount - 100;
Write(Account, #4662, X);

Y = Read(Account, #7199);
Y.amount = Y.amount + 100;
Write(Account, #7199, Y);
```

CRASH !

What is the problem ?

20

6. Security

- Files
- Spreadsheets
- DBMS

File-level access control

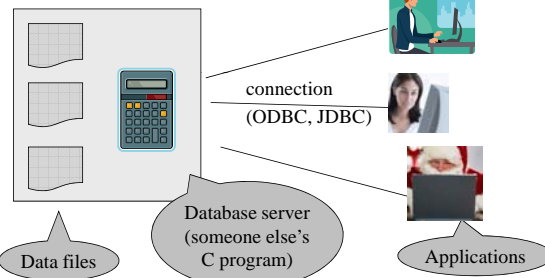
Same [?]

Table/attribute-level access control

21

Enters a DMBS

“Two tier system” or “client-server”



Data Independence

Logical view

Directors:

id	fName	lName
15901	Francis Ford	Coppola
...		

Movie_Directors:

id	mid
15901	130128
...	

Movies:

mid	Title	Year
130128	The Godfather	1972
...		

Directors_file

Moviews_title_index_file

Directors_fname_index_file

Movies_file

Physical view

What the Database Systems Does

1. Create/store large datasets SQL DML
2. Search/query/update
3. Change the structure SQL DDL
4. Concurrent access to many user
5. Recover from crashes Transactions ACID
6. Security Grant, Revoke, Roles

24

Course Outline - TENTATIVE !!

1. 9/26: SQL
 2. 10/1: SQL in C#; Database Design: E/R, NF
 3. 10/5: Views, Constraints, security
 4. 10/12: Database design
 5. 10/17: XML/XPath/XQuery
 6. 10/26: Transactions, concurrency
 7. 11/14: Database storage, indexes
 8. 11/26: Physical operators, optimization
- Calendar on web site – updated as we go

25

Grading (TENTATIVE)

- Homework 30%
- Project 25%
- Midterm 15%
- Final 25%
- Intangibles: 5%

Late days: Up to 4 total per quarter, at most 2 on any particular assignment/project phase. Otherwise **no late assignments accepted**

26

Reading Assignment

- Reading assignment for Fri, Sept.28
 - Introduction from **SQL for Web Nerds**, by Philip Greenspun, <http://philip.greenspun.com/sql/> (link on the course web)
- This is a one-time assignment, no grading, BUT *very* instructive and lots of fun reading

27

Homework

- Homework 0:
 - Due this Friday! (Don't panic – page with your name/picture/etc.)
- Homework 1:
 - SQL Queries
 - Due Friday, Oct. 5
 - It is posted already!
- Homework 2:
 - Conceptual design: E/R diagrams, Normal Forms
 - Due Friday, Oct. 17
- Homework 3:
 - XML/Xquery
 - Due Friday, Nov. 2
- Homework 4:
 - Transactions: concurrency control and recovery
 - Due Friday, Nov. 16

28

The Project: Boutique Online Store

- Phase 0:
 - Partner details; due middle of next week
- Phase 1:
 - Design a Database Schema, Build Related Data Logic
 - Due Friday, Oct. 12
- Phase 2:
 - Import data, Web Inventory Data Logic
 - Due Friday, Oct. 26
- Phase 3:
 - Checkout Logic
 - Due Friday, Nov. 9
- Phase 4:
 - Database Tuning
 - Due Friday, Nov. 30

29

Project

SQL Server, C#, ASP.NET

- Supported
- Will provide starter code in C#, ASP.NET
- The import data is in SQL/XML on SQL Server

Alternative technologies: MySQL, postgres, PHPs

- Not supported (you are on your own)
- Worry about the SQL/XML part...

30

Accessing SQL Server – Today!

SQL Server Management Studio

- Server Type = Database Engine
- Server Name = IISQSRV
- Authentication = SQL Server Authentication
 - Login = your UW email netid (*not* your CSE email, and *just* the netid without “u.washington.edu”)
 - Password = 2007#cse444
 - Change your password on first login (must be “secure”)
- Details on the course web

Then play with IMDB, start thinking about HW1