CISC 432/832
Advanced Topics in DBMSs

Pat Martin
Goodwin 630
martin@cs.queensu.ca

Learning Objectives

- Gain an appreciation of DBMSs from a *systems* point of view
- Become familiar with main components of a large DBMS
- Become familiar with algorithms and concepts used to implement a relational DBMS
- Understand how DBMS performance is measured, how it is affected by the different components and how the system can be tuned

Reference Materials

Textbook:

Database Management Systems (3rd Edition) by Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill Higer Education

Other materials:

Supplementary papers on reserve and on Web.

TAs

- Lei Wu (wul@cs.queensu.ca)
- Hao Yu (yuh@cs.queensu.ca)
- Jared Zebedee (zebedee@cs.queensu.ca)

Marking Schemes

- **Undergraduate students**
  - 4 assignments (45 marks)
  - 2 term tests (25 marks
  - One of a final exam, programming project or survey paper (30 marks)

Marking Schemes (Cont.)

- **Graduate students**
  - 4 assignments (45 marks)
  - 2 term tests (25 marks)
  - 2 of a final exam, programming project or survey paper (60 marks)

  Total is converted to an equivalent mark out of 100

Marking Schemes (Cont.)

- A hard copy of assigned work is to be handed at the beginning of class on the date due. Do not hand-in assignments via email.
- Late assignments are subject to a 10% per day late penalty, with weekends counted as one day. Late assignments will not be accepted beyond 5 days past the date due.
- A list of all updated marks will be posted after each piece of work is returned. Be sure to check your marks and to report any problems to the professor.
- Students must send an email to the professor declaring their choice(s) of final exam, project or paper by September 30, 2003.

Plagiarism

- Plagiarism is defined in the Arts and Science calendar as the act of "presenting work done (in whole or in part) by someone else as if it were one's own". Plagiarism is different from cooperation or collaboration. Students may, where explicitly permitted, work in groups and present results collectively. Examples of plagiarism include
  - Submitting a paper or assignment prepared in whole or in part by someone else as one's own.
  - Copying a paper or assignment or knowingly allowing someone else to copy your paper or assignment.
  - Using direct quotations or large sections of paraphrased materials without acknowledgement.
  - Submitting the same piece of work in more than one course without the permission of the instructors.
  - Taking large sections of papers found on the Web and submitting as part of your own work.
- The minimum penalty for plagiarism will be a mark of zero for the specific piece of work where plagiarism occurred. A more severe penalty may be assessed when it is warranted by the situation.

Course Outline

- **DBMS storage and indexing**
  - 2 weeks
  - Chs 8 and 9
  - File organizations, indexing, buffer management, RAID, SAN

Course Outline (cont.)

- **Query evaluation and optimization**
  - 4 weeks
  - Chs 12, 14 and 15
  - System catalog, evaluating query optimizers, query optimization

Course Outline (cont.)

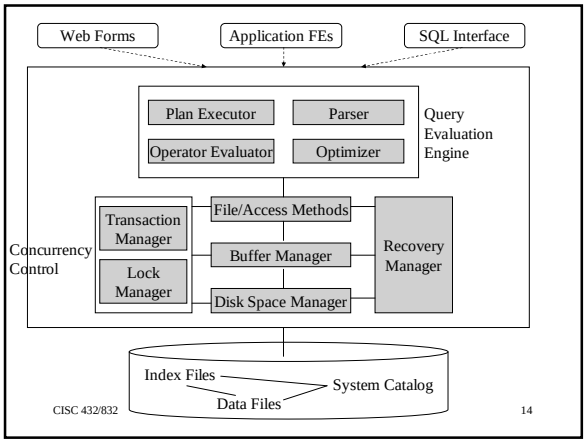
- **Transaction management**
  - 3 weeks
  - Chs 16, 17 and 18
  - Schedules, serializability, concurrency control, recovery, Aries recovery algorithm

Course Outline (cont.)

- **Database tuning**
  - 2 weeks
  - Chs 20 and 21
  - Index selection, database tuning, security and authorization

Assumed Background

- Relational model, SQL, relational algebra, schema design
- Java



DBMS Performance

- Common metrics
  - Throughput
  - Response time
- Key factors that influence performance
  - Number of disk accesses
  - Amount of data retrieved to answer query
  - Competition for DBMS resources