

CMPS 277:
officially
Relational Databases
but this quarter
Database Implementation

Notes 01: Introduction

Arthur Keller

Isn't Implementing a Database System Simple?

Relations \Rightarrow Statements \Rightarrow Results

Introducing the

MEGATRON 30000

Database Management System

- The latest from Megatron Labs
- Incorporates latest relational technology
- UNIX compatible

Megatron 3000

Implementation Details

! First sign non-disclosure agreement !

Megatron 3000

Implementation Details

- Relations stored in files (ASCII)
e.g., relation R is in /usr/db/R

```
Smith # 123 # CS  
Jones # 522 # EE  
⋮
```

Megatron 3000

Implementation Details

- Directory file (ASCII) in /usr/db/directory

```
R1 # A # INT # B # STR ...  
R2 # C # STR # A # INT ...  
⋮
```

Megatron 3000

Sample Sessions

```
% MEGATRON3000
  Welcome to MEGATRON 3000!
&
  :
& quit
%
```

Megatron 3000

Sample Sessions

```
& select *  
  from R #
```

Relation R

<u>A</u>	<u>B</u>	<u>C</u>
SMITH	123	CS

```
&
```

Megatron 3000

Sample Sessions

```
& select A,B  
  from R,S  
  where R.A = S.A and S.C > 100 #
```

<u>A</u>	<u>B</u>
123	CAR
522	CAT

```
&
```

Megatron 3000

Sample Sessions

```
& select *  
  from R | LPR #  
&
```

Result sent to LPR (printer).

Megatron 3000

Sample Sessions

```
& select *  
  from R  
  where R.A < 100 | T #  
&
```

New relation T created.

Megatron 3000

- To execute “**select * from R where *condition***”:
 - (1) Read dictionary to get R attributes
 - (2) Read R file, for each line:
 - (a) Check condition
 - (b) If OK, display

Megatron 3000

- To execute "`select * from R
 where condition | T`":
 - (1) Process select as before
 - (2) Write results to new file T
 - (3) Append new line to dictionary

Megatron 3000

- To execute “**select A,B from R,S where *condition***”:
 - (1) Read dictionary to get R,S attributes
 - (2) Read R file, for each line:
 - (a) Read S file, for each line:
 - (i) Create join tuple
 - (ii) Check condition
 - (iii) Display if OK

What's wrong with the Megatron 3000 DBMS?

What's wrong with the Megatron 3000 DBMS?

- Tuple layout on disk

e.g., - Change string from 'Cat' to 'Cats' and we have to rewrite file

- ASCII storage is expensive
- Deletions are expensive

What's wrong with the Megatron 3000 DBMS?

- Search expensive; no indexes
- e.g.,
- Cannot find tuple with given key quickly
 - Always have to read full relation

What's wrong with the Megatron 3000 DBMS?

- Brute force query processing

e.g., `select *`

`from R,S`

`where R.A = S.A and S.B > 1000`

- Do select first?

- More efficient join?

What's wrong with the Megatron 3000 DBMS?

- No buffer manager
e.g., Need caching

What's wrong with the Megatron 3000 DBMS?

- No concurrency control

What's wrong with the Megatron 3000 DBMS?

- No reliability
 - e.g., - Can lose data
 - Can leave operations half done

What's wrong with the Megatron 3000 DBMS?

- No security

- e.g.,
- File system insecure
 - File system security is coarse

What's wrong with the Megatron 3000 DBMS?

- No application program interface (API)
e.g., How can a payroll program get at the data?

What's wrong with the Megatron 3000 DBMS?

- Cannot interact with other DBMSs.

What's wrong with the Megatron 3000 DBMS?

- Poor dictionary facilities

What's wrong with the Megatron 3000 DBMS?

- No GUI

What's wrong with the Megatron 3000 DBMS?

- Lousy salesman!!

Course Overview

- **File & System Structure**

Records in blocks, dictionary, buffer management,...

- **Indexing & Hashing**

B-Trees, hashing,...

- **Query Processing**

Query costs, join strategies,...

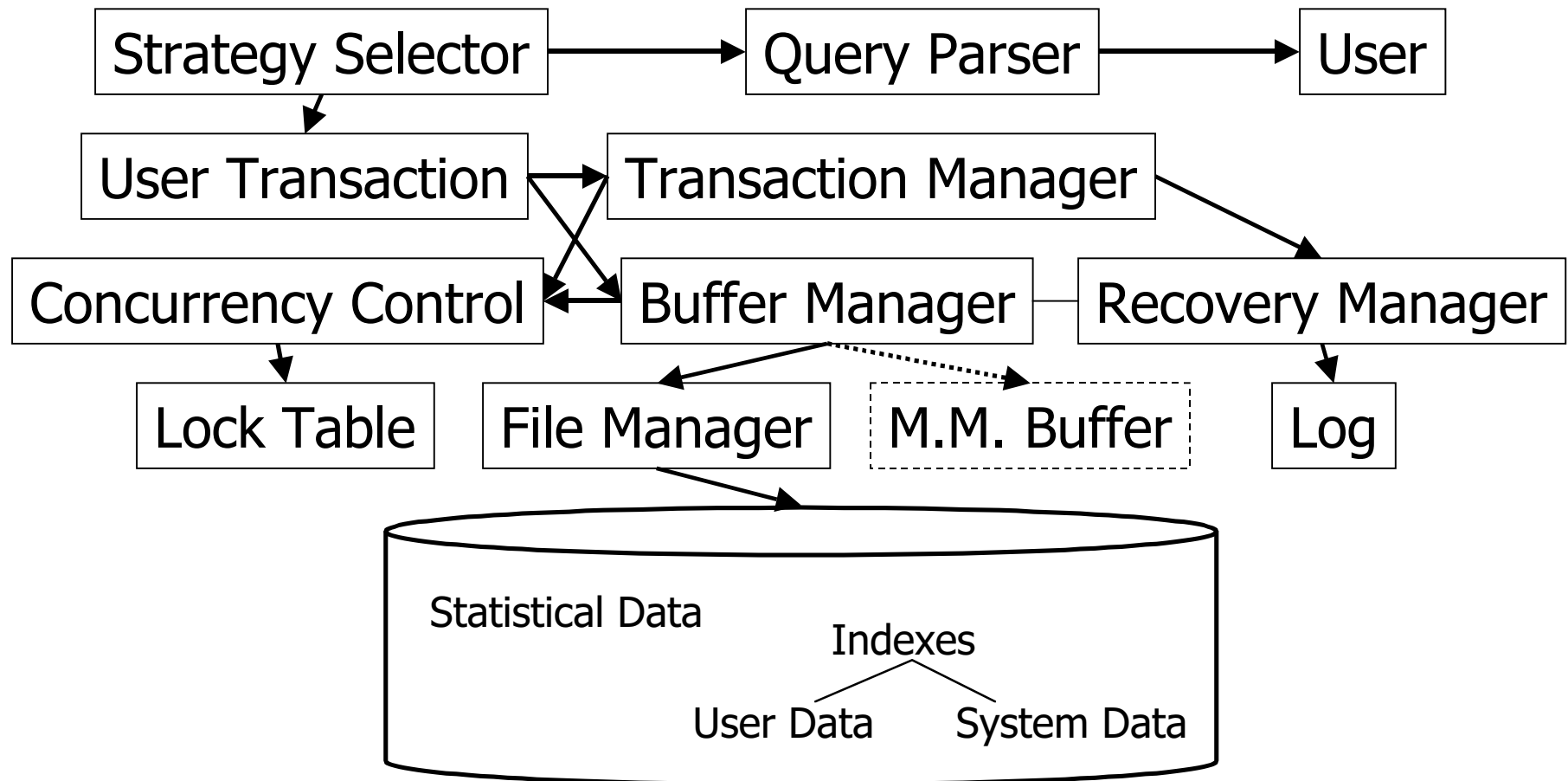
- **Crash Recovery**

Failures, stable storage,...

Course Overview

- **Concurrency Control**
Correctness, locks,...
- **Transaction Processing**
Logs, deadlocks,...
- **Security & Integrity**
Authorization, encryption,...
- **Distributed Databases**
Interoperation, distributed recovery,...

System Structure



Some Terms

- Database system
- Transaction processing system
- File access system
- Information retrieval system

Mechanics

<http://www.soe.ucsc.edu/classes/cs277/>

Coming soon

Prerequisite

- An introductory database course equivalent to CMPS180
- Knowledge of SQL (theory and practice)
- Algorithms and elementary analysis

- If you do not have the prerequisite, you may want to audit the class instead.

Staff

- INSTRUCTOR: Arthur Keller
- Office: Baskin Engineering 153A
- Email: ark@soe.ucsc.edu – a good way to reach me.
- Office Hours: Most Tuesdays, Some Thursdays 4:30-5:30pm, often for a few minutes after class, and by appointment.
- I'm coming from Palo Alto, so I may be late.
- TEACHING ASSISTANT: none
- GRADER: ?

Details

- LECTURES: Tuesday, Thursday 6-7:45pm, SS II 179
- TEXTBOOK: Garcia-Molina, Ullman, Widom: "DATABASE SYSTEMS, THE COMPLETE BOOK" (second half of book, first half was used for CMPS180).
- ASSIGNMENTS: Seven written homework assignments. No programming. Also readings in Textbook.
- GRADING: Homeworks: 21% (3% each), Survey paper: 19%, Midterm: 20%, Final: 40%.
- WEB SITE: All handouts and assignments will be posted on our Web site at <http://www.soe.ucsc.edu/classes/cs277/>
- Please check it periodically for last minute announcements.
- NEWSGROUP: ucsc.class.cmps277 is being set up.

Tentative Syllabus

	DATE	CHAPTER	TOPIC
•	Tue Mar 26		Introduction
•	Thu Mar 28		Class cancelled
•	Tue Apr 2	Ch. 11	Hardware
•	Thu Apr 4	Ch. 12	File and System Structure
•	Tue Apr 9	Ch. 12	File and System Structure
•	Thu Apr 11	Ch. 13	Indexing and Hashing
•	Tue Apr 16	Ch. 13	Indexing and Hashing
•	Thu Apr 18	Ch. 14	Indexing and Hashing
•	Tue Apr 23	Ch. 15	Query Processing
•	Thu Apr 25	Ch. 15	Query Processing
•	Tue Apr 30	Ch. 16	Query Processing
•	Thu May 2	Ch. 17	Crash Recovery
•	Tue May 7		Midterm
•	Thu May 9	Ch. 17	Crash Recovery
•	Tue May 14	Ch. 18	Concurrency Control
•	Thu May 16	Ch. 18	Concurrency Control
•	Tue May 21	Ch. 18	Concurrency Control
•	Thu May 23	Ch. 19	Transaction Processing
•	Tue May 28	Ch. 19	Transaction Processing
•	Thu May 30	Ch. 20	Information Integration
•	TBA		Review
•	Wed Jun 5		Final

Read: All Chapters

- Except following optional material:
 - Sections 11.7.4, 11.7.5
 - Sections 14.3.6, 14.3.7, 14.3.8
 - Sections 14.4.2, 14.4.3, 14.4.4
 - Sections 15.7, 15.8, 15.9
 - Sections 16.6, 16.7
 - In Chapters 15, 16: material on duplicate elimination operator, grouping, aggregation operators
 - Section 18.8
 - Sections 19.4, 19.5, 19.6, 19.7

Next time:

- Hardware
- Read chapter 11