




# CS 525: Advanced Database Organization



## 01: Introduction

Boris Glavic



Slides: adapted from a [course](#) taught by [Hector Garcia-Molina](#), Stanford InfoLab

CS 525

Notes 1 - Introduction
1


1

## Advanced Database Organization?



- =Database Implementation
- =How to implement a database system
- ... and have fun doing it ;-)

CS 525

Notes 1 - Introduction
2


2

## Isn't Implementing a Database System Simple?

Relations  $\implies$  Statements  $\implies$  Results



CS 525

Notes 1 - Introduction
3


3

## Introducing the MEGATRON 3000

Database Management System



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

CS 525

Notes 1 - Introduction
4


4

## Megatron 3000 Implementation Details

! First sign non-disclosure agreement !



CS 525

Notes 1 - Introduction
5


5

## Megatron 3000 Implementation Details

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

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

CS 525

Notes 1 - Introduction
6


6

## Megatron 3000

### Implementation Details

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

```

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

CS 525
Notes 1 - Introduction
7
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

7

## Megatron 3000

### Sample Sessions

```

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

CS 525
Notes 1 - Introduction
8
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

8

## Megatron 3000

### Sample Sessions

```

& select *
from R #

  Relation R
  A      B      C
SMITH  123   CS
&
```

CS 525
Notes 1 - Introduction
9
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

9

## Megatron 3000

### Sample Sessions

```

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

  A      B
123   CAR
522   CAT
&
```

CS 525
Notes 1 - Introduction
10
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

10

## Megatron 3000

### Sample Sessions

```

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

Result sent to LPR (printer).

CS 525
Notes 1 - Introduction
11
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

11

## Megatron 3000

### Sample Sessions

```

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

New relation T created.

CS 525
Notes 1 - Introduction
12
IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

12

## 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

CS 525



Notes 1 - Introduction

13

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

13

## 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

CS 525



Notes 1 - Introduction

14

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

14

## 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

CS 525



Notes 1 - Introduction

15

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

15

## What's wrong with the Megatron 3000 DBMS?

CS 525



Notes 1 - Introduction

16

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

16

## 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

CS 525



Notes 1 - Introduction

17

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

17

## 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

CS 525



Notes 1 - Introduction

18

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY



18

## 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?



CS 525  Notes 1 - Introduction 19  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

19

## What's wrong with the Megatron 3000 DBMS?

- No buffer manager



e.g., Need caching

CS 525  Notes 1 - Introduction 20  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

20

## What's wrong with the Megatron 3000 DBMS?

- No concurrency control



CS 525  Notes 1 - Introduction 21  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

21

## What's wrong with the Megatron 3000 DBMS?

- No reliability

e.g., - Can lose data  
- Can leave operations half done



CS 525  Notes 1 - Introduction 22  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

22

## What's wrong with the Megatron 3000 DBMS?

- No security

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



CS 525  Notes 1 - Introduction 23  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

23

## What's wrong with the Megatron 3000 DBMS?

- No application program interface (API)


e.g., How can a payroll program get at the data?

CS 525  Notes 1 - Introduction 24  IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

24

## What's wrong with the Megatron 3000 DBMS?


- Cannot interact with other DBMSs.

CS 525  Notes 1 - Introduction 25 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

25

## What's wrong with the Megatron 3000 DBMS?


- Poor dictionary facilities

CS 525  Notes 1 - Introduction 26 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

26

## What's wrong with the Megatron 3000 DBMS?


- No GUI

CS 525  Notes 1 - Introduction 27 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

27

## What's wrong with the Megatron 3000 DBMS?


- Lousy salesman!!

CS 525  Notes 1 - Introduction 28 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

28

## 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,...

CS 525  Notes 1 - Introduction 29 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

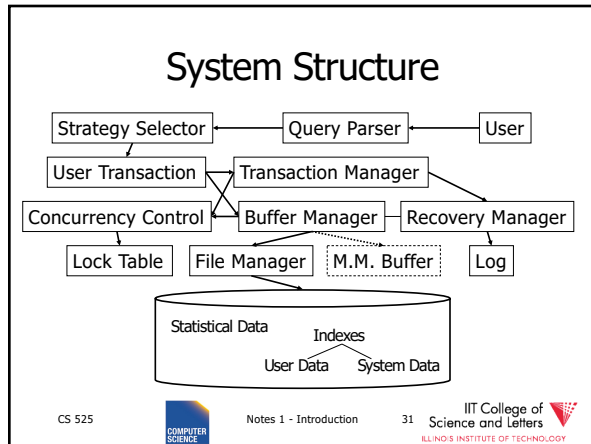
29

## Course Overview

- Concurrency Control  
Correctness, locks,...
- Transaction Processing  
Logs, deadlocks,...
- Security & Integrity  
Authorization, encryption,...
- Advanced Topics  
Distribution, More Fancy Optimizations, ...

CS 525  Notes 1 - Introduction 30 IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY


30



31

## Some Terms


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

CS 525        Notes 1 - Introduction    32    IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

32

## Instructor Info


- **Webpage:** <http://www.cs.iit.edu/~glavic/cs525/>
- **Instructor:** Boris Glavic
  - <http://www.cs.iit.edu/~glavic/>
  - **DBGGroup:** <http://www.cs.iit.edu/~dbggroup/>
  - **Office Hours: Wednesdays, 1pm-2pm**
    - <https://meet.google.com/gvy-qnyd-kqf>
- **TA: TBA**

CS 525        Notes 1 - Introduction    33    IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

33

## Course Information


- **Time:** Mon + Wed 5:10pm – 8:30pm
- **Live Video Call:**
  - <https://meet.google.com/siw-youu-kme>
- **Live Lectures will be recorded and uploaded to blackboard and to**
  - <https://drive.google.com/drive/folders/1CxZv2sXe8y1dOb4WrNuf30uMkwbjiiKRd?usp=sharing>

CS 525        Notes 1 - Introduction    34    IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

34

## Piazza


- <https://piazza.com/class/kdapgq7az2vg>
- Announcements
- Discussion forum
  - Student - Instructor/TA
  - Student – Student
- ->please join piazza to keep up to date

CS 525        Notes 1 - Introduction    35    IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

35

## Workload and Grading

- **Schedule and Important Dates**
  - On webpage & updated there
- **Programming Assignments (50%)**
  - 4 Assignments
  - Groups of 3 students
  - Plagiarism -> 0 points and administrative action
- **Quizzes (10%)**
- **Mid Term (20%) and Final Exam (20%)**

CS 525        Notes 1 - Introduction    36    IIT College of Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

36

## Textbooks

- Elmasri and Navathe , **Fundamentals of Database Systems**, 6th Edition , Addison-Wesley , 2003
- Garcia-Molina, Ullman, and Widom, **Database Systems: The Complete Book**, 2nd Edition, Prentice Hall, 2008
- Ramakrishnan and Gehrke , **Database Management Systems**, 3rd Edition , McGraw-Hill , 2002
- Silberschatz, Korth, and Sudarshan , **Database System Concepts**, 6th Edition , McGraw Hill , 2010

CS 525



Notes 1 - Introduction

37

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

37

## Programming Assignments

- 4 assignments one on-top of the other
- Optional 5<sup>th</sup> assignment for extra credit
- Code has to compile & run on server account
  - [Email-ID@fourier.cs.iit.edu](mailto:Email-ID@fourier.cs.iit.edu)
  - Linux machine
  - SSH with X-forwarding
- Source code managed in **git** repository on Bitbucket.org
  - Handing in assignments = submit (push) to repository
  - One repository per student
  - You should have gotten an invitation (if not, contact me/TA)
  - Git tutorials linked on course webpage!

CS 525



Notes 1 - Introduction

38

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

38

## Next:

- Hardware

CS 525



Notes 1 - Introduction

39

IIT College of  
Science and Letters  
ILLINOIS INSTITUTE OF TECHNOLOGY

39