Course: Advanced Database Organization - CS 525 - 01-02-03
Instructor: Yousef M Elmehdwi
Term: Spring 2018
Meetings: M/W 11:25 am-12:40 pm, Robert A. Pritzker Science Ctr 121

Office Location: 237D, Stuart Building
Email: yelmehdwi@iit.edu
Office Hours: F 11:50am-1:50pm or by appointment
Course Webpage: http://www.cs.iit.edu/~cs525/

Course Description: Databases management systems are a crucial part of most large-scale industry and open-source systems. This course provides comprehensive coverage of issues associated with database system development and an in-depth examination of structures and techniques used in contemporary database management systems (DBMSs). Students will learn about the inner workings of these exciting systems: Which algorithms are used? What are typical architectures used to build a system as complex as a DBMS? What are implementation strategies? These questions and more will be answered during the course.

The course is highly applied, emphasizing practical skills and habits through a series of programming assignments during which students will develop their own tiny DBMS like engine. We will cover the most important aspects/components of a DBMS: storage and buffer management, indexing, query optimization, query execution, and concurrency control and recovery.

This course is graduate-level introduction to the design and implementation of data management systems.

Prerequisite(s):
- Courses: CS425
- Programming experience in C, C++ or other low level languages
- Unix OS and file system knowledge is helpful
- Data structures (e.g., CS401)

Course Objectives:
- Understand the design decisions behind textbook DBMS architectures
- Know the trade-offs of various storage organization techniques
- Be able to build parts of a small-sized data processing system from scratch
- Understand the basics of query optimization
- Know standard implementations of relational operators such as join, aggregation, and set operations
- Be able to estimate the cost of executing an operator/query based on DB statistics
- Know standard database indexing techniques
- Understand concurrency control and recovery mechanisms

Suggested Texts, Readings, & Materials:
The following text books will be helpful for following the course and studying the presented material. All four textbooks have their merits, but any one should be sufficient as reading material.
Course Details: Major topics will be covered in the course:
- Disk Storage and Buffer Management
- Indexing and Hashing
- Query Optimization
- Query Execution
- Concurrency control and Recovery
- Advanced Topics (if time permits)

Workload and Grading Policies

Programming Assignments: There will be several programming assignments during the course. Starting from a storage manager you will be implementing your own tiny database-like system from scratch. You will explore how to implement the concepts and data structures discussed in the lectures and readings. The assignments will require the use of skills learned in this course as well as other skills you have developed throughout your program. Each assignment will build upon the code developed during the previous assignment. In the end there will be an optional assignment for extra credit. Each of the regular assignments will have optional parts that give extra credit. All assignments have to be implemented using C/C++.

- **Assignment 1 - Storage Manager:** Implement a storage manager that allows read/writing of blocks to/from a file on disk.
- **Assignment 2 - Buffer Manager:** Implement a buffer manager that manages a buffer of blocks in memory including reading/flushing to disk and block replacement (flushing blocks to disk to make space for reading new blocks from disk).
- **Assignment 3 - Record Manager:** Implement a simple record manager that allows navigation through records, and inserting and deleting records.
- **Assignment 4 - B⁺-Tree Index:** Implement a disk-based B⁺-tree index structure.
- **Potential Optional/extra Assignment:**
  - Implement a standard operator algorithm on top of your record manager, e.g., nested loop join, hash-aggregate, ...

Mid Term and Final Exam: There will be a mid term and a final exam covering the topics of the course.

Quizzes and Homework assignments: There will be quizzes during the course. The main objective of the quizzes is for you and the instructor to evaluate how well you internalized the topics covered in the course. In addition, five homework assignments turned in via Blackboard.

Grading Policies Grades will be posted on blackboard.

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming Assignments</td>
<td>45% (10% + 10% + 10% + 15% )</td>
</tr>
<tr>
<td>Homeworks</td>
<td>5%</td>
</tr>
<tr>
<td>Class participation and Quizzes</td>
<td>5%</td>
</tr>
<tr>
<td>Mid Term Exam</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>25%</td>
</tr>
</tbody>
</table>
### Letter Grade Distribution

<table>
<thead>
<tr>
<th>Points</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 - 100</td>
<td>A</td>
</tr>
<tr>
<td>80 - 89</td>
<td>B</td>
</tr>
<tr>
<td>70 - 79</td>
<td>C</td>
</tr>
<tr>
<td>60 - 69</td>
<td>D</td>
</tr>
<tr>
<td>0 - 59</td>
<td>E</td>
</tr>
</tbody>
</table>

### Attendance Policy
- Students are expected to attend all classes and are responsible for all material covered in class, even when absent.
- Students should understand that some material discussed in class is not covered in the textbook.
- Attendance is required.
- I realize that some absences are unavoidable, and you should inform your instructor prior to missing any classes.
- Missing more than 3 classes will decrease your overall grade by a letter grade.
- You will be advised to withdraw from the course if you miss more than 5 classes.

### Other Policies

#### Ethics:
- Any behavior on any assignment or exam that could be considered copying or cheating will result in an immediate zero on the assignment for all parties involved, failure in the class, and notification of the Undergraduate or Graduate Dean’s Office. Please see the IIT Code of Academic Honesty [https://web.iit.edu/student-affairs/handbook/fine-print/code-academic-honesty](https://web.iit.edu/student-affairs/handbook/fine-print/code-academic-honesty).
- We will check for plagiarism. Plagiarism will result in zero points for the assignment, potentially academic sanctions, and may result in an E grade.

#### Late policies
- 1-3 days late: -10% points
- 4-7 days late: -20% points
- > 7 days late: 0 points

#### Make-up Exams
- Only for officially proven health reasons.

#### Students with Disabilities:
- Reasonable accommodations will be made for students with documented disabilities.
- In order to receive accommodations, students must obtain a letter of accommodation from the Center for Disability Resources.
- The Center for Disability Resources (CDR) is located at 3424 S. State Street - 1C3-2, 312 567.5744 or disabilities@iit.edu
TENTATIVE COURSE The weekly coverage might change as it depends on the progress of the class.

OUTLINE

<table>
<thead>
<tr>
<th>Week</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Introduction/ Hardware</td>
</tr>
<tr>
<td>Week 2</td>
<td>File and System Structure</td>
</tr>
<tr>
<td>Weeks 3-4</td>
<td>Indexing and Hashing</td>
</tr>
<tr>
<td>Weeks 5-7</td>
<td>Query Processing</td>
</tr>
<tr>
<td>Week 8-9</td>
<td>Crash Recovery</td>
</tr>
<tr>
<td>Weeks 10-11</td>
<td>Concurrency Control</td>
</tr>
<tr>
<td>Weeks 12-13</td>
<td>Concurrency Control</td>
</tr>
<tr>
<td>Week 14</td>
<td>Transaction Processing</td>
</tr>
<tr>
<td>Week 15</td>
<td>Advanced topics</td>
</tr>
</tbody>
</table>