CS 437: Database System Implementation

Lecture: Wednesday 11:00am - 12:20pm, Friday 1:00pm - 2:20pm

Address: Online

Pre-requisites: (01:198:214 SYSTEMS PROGRAMMING) AND (01:198:336 PRIN INFO & DATA MGT)

Textbook: Database System Concepts (7th Edition)

Instructor: Dong Deng (dong.deng AT rutgers.edu)

TA: Chaoji Zuo (cz296@scarletmail.rutgers.edu)

Course Description

This course is on the design and implementation of database management systems. Topics include data models (relational, document, key/value), storage models (n-ary, decomposition), query languages (SQL, stored procedures), storage architectures (heaps, log-structured), indexing (order preserving trees, hash tables), transaction processing (ACID, concurrency control), recovery (logging, checkpoints), query processing (joins, sorting, aggregation, optimization), and parallel architectures (multi-core, distributed).

Educational Objectives

This is an upper-level course on the internals of database management systems. This course has a heavy emphasis on programming projects. Upon successful completion of this course, the student should be able to:

- Use relational algebra to express database queries.
- Use SQL to interact with database management systems.
- Design appropriate database tables, using functional dependencies and normal forms.
- Implement a disk-oriented database storage manager with table heaps and indexes.
- Understand, compare, and implement the fundamental concurrency control algorithms.
- Implement database recovery algorithms and verify their correctness.
- Identify trade-offs among database systems techniques and contrast distributed/parallel alternatives for both on-line transaction processing and on-line analytical workloads.
- Interpret and comparatively criticize database system architectures.

Grading

The final grade for the course will be based on the following weights:

- 20% — Homework
- 45% — Programming Projects
- 35% — Final Exam

Homework

Students will complete multiple homework assignments during the course. These homeworks are designed to reinforce the lectures and reading materials. The homework due dates are the ones that are posted on this website.

Each homework will be graded out of a total of 100 points and are counted equally when computing the homework portion of the final grade.

Programming Projects

Students will complete programming projects during the course. Each assignment is cumulative. That is, you need to successfully complete each assignment in order to complete the next one. We will not release solutions for the programming projects.

Each project will be graded out of a total of 100 points and are counted equally when computing the project portion of the final grade.

Exams

There will be a final exam during the University's final examination period at the end of the semester. The exams will be based on topics discussed in class.

Plagiarism Policy

All homeworks and projects are to be done individually. Whatever you turn in must be your own work.

- Students are allowed to discuss about homework and project problems with others.
- Students are not allowed to copy the contents of a white-board after a group meeting with other students.
- Students are not allowed to copy the solutions from another colleague.

Rutgers CS Diversity and Inclusion Statement

Rutgers Computer Science Department is committed to creating a consciously anti-racist, inclusive community that welcomes diversity in various dimensions (e.g., race, national origin, gender, sexuality, disability status, class, or religious beliefs). We will not tolerate micro-aggressions and discrimination that creates a hostile atmosphere in the class and/or threatens the well-being of our students. We will continuously strive to create a safe learning environment that allows for the open exchange of ideas and cherished freedom of speech, while also ensuring equitable opportunities and respect for all of us. Our goal is to maintain an environment where students, staff, and faculty can contribute without the fear of ridicule or intolerant or offensive language. If you witness or experience racism, discrimination micro-aggressions, or other offensive behavior, you are encouraged to bring it to the attention to the undergraduate program director and/or the department chair. You can also report it to the Bias Incident Reporting System http://inclusion.rutgers.edu/report-bias-incident/

Tentative Schedule

Date	Note	Topic
1-Sep		Introduction
3-Sep	Programming 0 (Due 9/17)	Relational Model
8-Sep		SQL Recap
10-Sep		Storage Management I - Physical Storage
15-Sep		Storage Management II - Data Storage Structure
17-Sep		Storage Management III - Buffer Pool
22-Sep	Programming 1 (Due 10/13)	Indexing I - Hash Table
24-Sep		Indexing II - B+ Tree
29-Sep	Homework 1 (Due 10/13)	Indexing III - Indexing Concurrency Control
1-Oct		Indexing IV - Advanced Indexes
6-Oct		Sorting and Aggregation
8-Oct		Join Algorithms
13-Oct	Programming 2 (Due 11/3)	Query Execution I
15-Oct		Query Execution II
20-Oct	Homework 2 (Due 11/3)	Query Optimization I
22-Oct		Query Optimization II
27-Oct		Concurrency Control
29-Oct		Two Phase Locking
3-Nov	Programming 3 (Due 11/24)	Timestamp Ordering Concurrency Control
5-Nov		Multiversion Concurrency Control
10-Nov	Homework 3 (Due 11/24)	Logging
12-Nov		Recovery
17-Nov		Introduction to Distributed Databases
19-Nov		Distributed Storage Model
24-Nov	No Class	Thanksgiving
26-Nov	No Class	Thanksgiving
1-Dec		Distributed Query Processing
3-Dec		Distributed Transaction Processing
8-Dec		Advanced Topics I - Big Data
10-Dec		Advanced Topics II - Information Retrieval
15-Dec	No Class	Reading Day
17-Dec	Final Exam	