

# IMPLEMENTATION OF DATABASE SYSTEMS

Fall 2021

---

<b>Instructor:</b>	Amir Hossein Rouhani Seraji	<b>Code:</b>	13755
<b>Email:</b>	<a href="mailto:ah.rouhaniseraji@gmail.com">ah.rouhaniseraji@gmail.com</a>	<b>Unit:</b>	3

---

**Course Pages:** Please check this website two or three times a week

- <https://rouhani-class.github.io/idbms>

**Discussion Group:** Every student has to be a member of our Telegram Channel and Telegram Group for our discussion and daily announcements.

- [Telegram Channel](#).
- [Telegram Group](#).

**Teaching Assistant:**

- [Mehran Moeini Jam](#).
- [Mohammad Hossein Khoshechin](#).

**Prerequisites:** An undergraduate-level understanding of Relational Databases Concepts, Data Structures, and Algorithms Design & Analysis.

**Objectives:** This is an upper-level course on the internals of database management systems. This course is on the design and implementation of database management systems. Topics include storage models (n-ary, decomposition), 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). Upon successful completion of this course, the student should be able to: 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.

- Introduction
  - Review Relational Theory Concepts
  - Review Relational Algebra
  - Review SQL Languages
  - History of Data Models and Database Systems
- Physical Storage Systems
  - Overview of Physical Storage Media
  - Storage Interfaces
  - Magnetic Disks

- Flash Memory
- Data Storage Structures
  - File Organization
  - Organization of Records in Files
  - Data-Dictionary Storage
  - Database Buffer
- Indexing
  - Ordered Indices
  - B<sup>+</sup>-Tree Index Files
  - B<sup>+</sup>-Tree Extensions
  - Hash Indices
- Query Processing
  - Measures of Query Cost
  - Selection Operation
  - Sorting
  - Join Operation
- Query Optimization
  - Transformation of RelationalExpressions
  - Estimating Statistics of ExpressionResults
  - Choice of Evaluation Plans
  - Materialized Views
  - Advanced Topics in QueryOptimization
- Transactions
  - Transaction Concept
  - Transaction Atomicity and Durability
  - Transaction Isolation
  - Serializability
  - Implementation of Isolation Levels
- Concurrency Control
  - Lock-Based Protocols
  - Deadlock Handling
  - Multiple Granularity
  - Timestamp-Based Protocols
  - Validation-Based Protocols
  - Multiversion Schemes
  - Snapshot Isolation
  - Weak Levels of Consistency inPractice

- Recovery from deadlock
- Recovery System
  - Failure Classification
  - Recovery and Atomicity
  - Recovery Algorithms
  - Buffer Management
  - Failure with Loss of Non-Volatile Storage
  - High Availability Using Remote Backup Systems
  - Early Lock Release and Logical Undo Operations
  - ARIES
- Database-System Architectures
  - Parallel Systems
  - Distributed Systems
  - Cloud-Based Services
- Parallel and Distributed Storage
  - Data Partitioning
  - Dealing with Skew in Partitioning
  - Replication
  - Parallel Indexing
  - Distributed File Systems
  - Parallel Key-Value Stores
- Parallel and Distributed Query Processing
  - Parallel Sort
  - Parallel Join
  - Parallel Evaluation of Query Plans
  - Query Processing on Shared-Memory Architectures
  - Query Optimization for Parallel Execution
  - Parallel Processing of Streaming Data
  - Distributed Query Processing
- Parallel and Distributed Transaction Processing
  - Distributed Transactions
  - Commit Protocols
  - Concurrency Control in Distributed Databases

**Main References:** Our Lectuers are mainly based-on this book.

- Abraham Silberschatz, Henry F. Korth, S. Sudarshan, *Database System Concepts*, McGraw-Hill, 7th ed, 2020.

**Supplementary References:** We also use these references for complementary discussion.

- Goetz Graefe, *On Transactional Concurrency Control*, Morgan & Claypool, 2019.
- M. Tamer Ozsu, Patrick Valduriez, *Principles of Distributed Database Systems*, Springer, 4th ed, 2020.
- Gerhard Weikum, Gottfried Vossen, *Transactional Information Systems Theory, Algorithms, and the Practice of Concurrency Control and Recovery*, Morgan Kaufmann, 2001.
- Jim Gray, Andreas Reuter, *Transaction Processing: Concepts and Techniques*, Morgan Kaufmann, 1993.
- Philip A. Bernstein, Eric Newcomer, *Principles of Transaction Processing*, Morgan Kaufmann, 2nd ed, 2009.
- Sam S. Lightstone, Toby J. Teorey, Tom Nadeau, *Physical Database Design: The Database Professional's Guide to Exploiting Indexes, Views, Storage, and More*, Morgan Kaufmann, 4th ed, 2007.
- Nicolas Bruno, *Automated Physical Database Design and Tuning*, CRC Press, 2011.

**Useful Links:** Courses with Video Lectures.

- [CMU 15-445/645 Database Systems](#), Carnegie Mellon University, Fall 2021.
- [CMU 15-445/645 Database Systems](#), Carnegie Mellon University, Fall 2019.
- [CMU 15-445/645 Database Systems](#), Carnegie Mellon University, Fall 2018.
- [CMU 15-445/645 Database Systems](#), Carnegie Mellon University, Fall 2017.
- [CMU 15-721 Advanced Database Systems](#), Carnegie Mellon University, Spring 2020.
- [CMU 15-721 Advanced Database Systems](#), Carnegie Mellon University, Spring 2019.
- [CMU 15-721 Advanced Database Systems](#), Carnegie Mellon University, Spring 2018.
- [CMU 15-721 Advanced Database Systems](#), Carnegie Mellon University, Spring 2017.
- [CMU 15-721 Advanced Database Systems](#), Carnegie Mellon University, Spring 2016.

Seminars.

- [Vaccination Database Tech Talks Second Dose](#) - 2021.
- [Vaccination Database Tech Talks First Dose](#) - 2021.
- [Quarantine Database Tech Talks](#) - 2020.
- [Hardware Accelerated Database Lectures](#) - Fall 2018.
- [Time Series Database Lectures](#) - Fall 2017.
- [The Databaseology Lectures](#) - Fall 2015.
- [Seven Databases in Seven Weeks](#) - Fall 2014.
- [A Celebration of Mike Stonebraker](#) - 2014.

Research Groups in Database Systems.

- [Carnegie Mellon Database Group](#).
- [MIT Database Group](#).
- [MIT Data Systems Group](#).
- [UC Berkeley Data Systems and Foundations](#).
- [Stanford Infolab Group](#).
- [University of Washington's database group](#).
- [UW-Madison Database Group](#).
- [Database of Databases](#)

Books website.

- [Database System Concepts](#).

Our passed Semesters Audio Lectures.

- [7232 Principles of Database Design, IAU South Tehran Branch, Fall 2018](#).
- [7232 Principles of Database Design, IAU South Tehran Branch, Spring 2017](#).

### **Grading Policy:**

- Assignments (40%)
- Quiz (30%)
- Final Exam (30%)
- (Optional) Presentation (10%)