

Project Ideas

CS 764, Fall 2022

This document contains a list of potential project topics. You are encouraged to explore your own ideas that interest you, which can be related to your own research.

1. Develop course projects using SQLite for the undergraduate database class (CS 564) at UW Madison. Please contact the instructor if you are interested in this project.

- <https://www.sqlite.org/index.html>
- Gaffney, Kevin P., et al. "SQLite: Past, Present, and Future." VDLB 2022
- Potential project ideas: (1) implement hash index, (2) implement hash join, (3) implement fine-grained concurrency control (e.g., record-level locking)

2. Conventional buffer management assumes a two-level storage hierarchy: DRAM as memory, and Disk or SSD as storage. Modern storage systems are more complex and some have more than two levels. For a storage hierarchy with multiple levels (or several different devices within the same level), can you develop better buffer management policies? You can consider OLAP and OLTP systems separately.

- Justin DeBrabant, et al., *Anti-Caching: A New Approach to Database Management System Architecture*. VLDB, 2013
- Ahmed Eldawy, et al., *Trekking Through Siberia: Managing Cold Data in a Memory-Optimized Database*. VLDB 2014
- Look for papers on “memory disaggregation”

3. Modern cloud-native databases can automatically scale the cluster (i.e., autoscaling) based on the features of the workloads. Please conduct a survey of existing solutions, identify their limitations, and propose a better design.

- Consider starting with AWS Aurora, and CockroachDB.

4. Conduct a survey on distributed query optimization techniques. What are the new challenges compared to single-node query optimization?

5. Some column store databases support query execution over compressed data without first decompressing the columns. These techniques can reduce CPU overhead by avoiding the expensive decompression operations. Please conduct a survey of such techniques and propose optimizations.

- Daniel Abadi, et al. *Integrating Compression and Execution in Column-Oriented Database Systems*, SIGMOD 2006
- Vijayshankar Raman, et al., *DB2 with BLU Acceleration: So Much More than Just a Column Store*, VLDB 2013

6. Early lock release (ELR) is a technique that allows transactions to release locks before they commit. ELR can reduce the time that transactions hold locks and thereby improve performance especially for contentious workloads. Please do a survey of ELR in the context of distributed systems and identify limitations or opportunities for further improvement.

- Hua Guo, *Lock Violation for Fault-tolerant Distributed Database System*, ICDE 2021
- Goetz Graefe, et al. *Controlled lock violation*. SIGMOD 2013
- Eljas Soisalon-Soininen and Tatu Ylönen. *Partial strictness in two-phase locking*. ICDT 1995