Course Description
This course introduces students to parallel database management systems that scale in a data center and across data centers, physical organization of data for enhanced performance and availability, NoSQL, cache augmented database systems and memory management, disaggregated database systems, transactional storage managers, and data analytics. The discussion sections introduce basics of transactions with ACID properties, a file system, SDM and relational data models, SQL, persistent index structures such as hashing and B+-tree, relational algebra, SQL, normal forms, fundamentals of a magnetic disk drive and solid state drives including their working details, algorithms such as LRU, concurrency control protocols such as locking and time-stamp, and crash recovery protocols such as logging. We review an actor model named FLOW for expedited development of multi-node database systems and study how it is used in FoundationDB.

Guest lecturers include leading database researchers and practitioners. Confirmed presenters include Dr. Haoyu Huang of Google (Week 4), Dr. Hieu Nguyen of eBay (Week 5), Dr. Doug Terry of Amazon (Week 6), and Dr. Michael Carey of Couchbase (Week 12).

Learning Objectives
By the end of the semester, students should have a grasp of fundamental concepts, design decisions, protocols and algorithms to build a database system using off-the-shelf software and hardware components. They should understand ACID semantics and their application to both storage managers, database systems, and applications. Finally, they should be able to provide an abstraction of today’s hardware in support of database systems.

Prerequisite(s): Principles of Software Development (CSCI 201).

Recommended Preparation: A good understanding of in-memory data structures such as record and their organizations in arrays, trees, hash tables, etc. Knowledge of relational databases, SQL, relational algebra and physical database design is required. We will cover the relational data model+algebra, SQL, physical database design in the discussion sections.

Course Notes:
- All lecture material will be posted on the USC Brightspace.
- Grading breakdown
  - Exam 1, Oct 18, 2024: 40%
  - Exam 2, Dec 6, 2024: 40%
  - Class Participation: 20%
- This course has no comprehensive final exam. Exam 2 is held during the last lecture and covers material presented since Exam 1.
Communication
This course uses the USC Brightspace system to disseminate lectures and assigned readings, and make announcements.

Technological Proficiency and Hardware/Software Required

Required Readings and Supplementary Materials
This course assumes students are Internet savvy and capable of using it as a digital library. This includes use of ACM/IEEE/Springer digital libraries. All USC students have access to these digital collections automatically.

Grading Breakdown
Based on prior experience teaching this course, students who attend lectures do very well. In specific, students who skip lectures do poorly on Exam 1 and spend the second half of the semester concerned with their grade. To prevent this, classroom participation is a requirement for this course. It is worth 20% of the final grade.

<table>
<thead>
<tr>
<th>Assessment Tool (assignments)</th>
<th>Points</th>
<th>% of Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam 1, Oct 18, 2024</td>
<td>100</td>
<td>40%</td>
</tr>
<tr>
<td>Exam 2, Dec 6, 2024</td>
<td>100</td>
<td>40%</td>
</tr>
<tr>
<td>Participation</td>
<td>100</td>
<td>20%</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

Class Participation: How it Works?
At the start of each lecture, you are provided with 1 to 3 questions. Write your answer on a piece of paper and give it to the TA at the end of the lecture. DEN students email your answers to the TA.

Grading Scale
The final letter grade is based on a curve.

Course-specific Policies (Assignment Submission, Grading Timeline, Late work, and Technology)
All deadlines are firm. No late submissions or e-mail submissions are accepted for class participation.

Attendance
Based on prior experience, students who do not attend lectures perform poorly on exams. CSCI 585 accommodates student athletes with approved Travel Request Letters and students who give advance notice of religious observation.

Assigned Reading

Week 1: Introductions, Semantic Data Model & Relational Data Model

Discussion
Week 2: Parallel Relational Database Management Systems

Discussion
FLOW (Yiming Gao)

Week 3: Physical Data Design

Discussion
Transactions: Concurrency Control Protocols

Week 4: Separation of Storage from Processing, Guest Lecture by Dr. Haoyu Huang of Google

Discussion
Google's AlloyDB

Week 5: Transactional Storage Managers, Guest Lecture by Dr. Hieu Nguyen of eBay
Discussion
Disaggregated DBMSs

Week 6: NoSQL Database Systems, Guest Lecture by Dr. Doug Terry of Amazon
https://www.usenix.org/conference/atc22/presentation/elhemali

Discussion
Joseph Idziorek and Alex Keyes and Colin Lazier and Somu Perianayagam and Prithvi Ramanathan and James Christopher Sorenson III and Doug Terry and Akshat Vig. 2023. Distributed Transactions at Scale in Amazon DynamoDB. 2023 USENIX Annual Technical Conference (USENIX ATC 23).
https://www.usenix.org/conference/atc23/presentation/idziorek

Week 7: Fall Recess

Week 8: Exam 1
Exam 1 covers assigned readings Week 1–6.

Week 9: Benchmarks

Discussion

Week 10: Memory Management
https://doi.org/10.1145/2663165.2663317
Week 11: Cache Augmented Database Management Systems

Discussion

Week 12: Data Analytics, Guest Lecture by Dr. Michael Carey of Couchbase
Assigned Reading: TBA

Week 13: MapReduce

Discussion

Week 14: Thanksgiving Break

Week 15: Exam 2
<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topics/Daily Activities</th>
<th>Readings/Preparation</th>
<th>Key Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>8/30</td>
<td>Introduction, SDM, Relational</td>
<td>SDM, Conceptual model. Discussion led by Nima Yazdani.</td>
<td>Conceptual data models</td>
</tr>
<tr>
<td>Week 2</td>
<td>9/6</td>
<td>Parallel DBMSs</td>
<td>Gamma. Actor/Flow discussion led by Yiming Gao.</td>
<td>Parallel query processing, crash recovery, shared-nothing architecture</td>
</tr>
<tr>
<td>Week 3</td>
<td>9/13</td>
<td>Data Declustering</td>
<td>Hybrid Range Partitioning Strategy</td>
<td>Sharding, Load Balancing</td>
</tr>
<tr>
<td>Week 4</td>
<td>9/20</td>
<td>Disaggregated DBMSs</td>
<td>Nova-LSM, Google’s AlloyDB. Guest lecture by Dr. Haoyu Huang, Google.</td>
<td>Separation of storage from processing, shared disk architecture.</td>
</tr>
<tr>
<td>Week 5</td>
<td>9/27</td>
<td>Xact Storage Managers</td>
<td>FoundationDB, Guest Lecture by Dr. Hieu Nguyen, eBay.</td>
<td>OCC, MVCC, Fat Clients, DBMS as a collection of microservices</td>
</tr>
<tr>
<td>Week 6</td>
<td>10/4</td>
<td>NoSQL + Xacts at Scale</td>
<td>NoSQL DBMSs. Guest lecture by Dr. Doug Terry, Amazon.</td>
<td>NoSQL</td>
</tr>
<tr>
<td>Week 7</td>
<td>10/11</td>
<td>Fall Recess</td>
<td>Fall Recess</td>
<td>Fall Recess</td>
</tr>
<tr>
<td>Week 8</td>
<td>10/18</td>
<td>Review for Exam 1 (1 hour) + 1 hour Exam 1 is on the same day</td>
<td>Exam 1</td>
<td>Review for Exam 1 + Exam 1.</td>
</tr>
<tr>
<td>Week 9</td>
<td>10/25</td>
<td>Benchmarks</td>
<td>BG Social Networking Benchmark</td>
<td>Performance metrics, Closed and Open Simulation Models</td>
</tr>
<tr>
<td>Week 10</td>
<td>11/1</td>
<td>Memory Mgmt</td>
<td>CAMP and LRU</td>
<td>Low complexity memory mgmt algorithms that consider size and cost.</td>
</tr>
<tr>
<td>Week 11</td>
<td>11/8</td>
<td>Cache Augmented Database Mgmt Systems</td>
<td>IQ Framework</td>
<td>Impact of caches on transactions</td>
</tr>
<tr>
<td>Week 12</td>
<td>11/15</td>
<td>Data Analytics</td>
<td>Columnar stores. Guest Lecture by Dr. Michael Carey, Couchbase.</td>
<td>In-the-cloud JSON data analytics.</td>
</tr>
<tr>
<td>Week 13</td>
<td>11/22</td>
<td>MapReduce</td>
<td>Programming paradigm for parallelism.</td>
<td>Divide-and-Conquer, Resilient to failures.</td>
</tr>
<tr>
<td>Week 14</td>
<td>11/29</td>
<td>Thanksgiving Break</td>
<td>Thanksgiving Break</td>
<td>Thanksgiving Break</td>
</tr>
<tr>
<td>Week 15</td>
<td>12/6</td>
<td>Review for Exam 2 &amp; Exam 2</td>
<td></td>
<td>Exam 2 12/6</td>
</tr>
</tbody>
</table>
Statement on Academic Conduct and Support Systems

Academic Conduct:

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on Research and Scholarship Misconduct.

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University’s educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems:

Counseling and Mental Health - (213) 740-9355 – 24/7 on call studenthealth.usc.edu/counseling
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call suicidepreventionlifeline.org
Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL), press “0” after hours – 24/7 on call studenthealth.usc.edu/sexual-assault
Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086 eeoix.usc.edu
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298 usc-advocate.symplicity.com/care_report
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.
The Office of Student Accessibility Services (OSAS) - (213) 740-0776
osas.usc.edu
OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention - (213) 821-4710
campussupport.usc.edu
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion - (213) 740-2101
diversity.usc.edu
Information on events, programs and training, the Provost’s Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
dps.usc.edu, emergency.usc.edu
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call
dps.usc.edu
Non-emergency assistance or information.

Office of the Ombuds - (213) 821-9556 (UPC) / (323-442-0382 (HSC)
ombuds.usc.edu
A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice - (323) 442-3340 or otfp@med.usc.edu
chan.usc.edu/otfp
Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.