IST 769: Advanced Database Management

Description:

An analysis of relational and nonrelational databases and their corresponding database management system architectures. Learn to build complex database objects to support a variety of needs from both the big data and traditional perspectives. Data systems performance, scalability, and security.

Additional Course Description:

This course provides tour of relational, document, key-value, columnar, and streaming database systems through the lens of the CAP theorem. We will explore the strengths and weaknesses of various database systems in the relational, Hadoop, and noSQL spaces. Where possible you will experience these systems first-hand as to gain an understanding of how they can be used to address complex, big data challenges.

Prerequisite: IST 659

Audience: Graduate

Credits: 3.0

Learning Objectives:

After taking this course, the students will be able to:

1. Understand advanced issues with the relational database model, such as transactions, performance, and security, as to understand the need for other database models.
2. Explain the CAP theorem and describe how any given database system's architecture fits within the CAP context.
3. Compare different database models such as document, key-value, column-family, streaming, and relational.
4. Identify the most suitable database systems for a specific application's data storage requirements.
5. Evaluate relational, Hadoop, and noSQL database tooling as to understand their underlying similarities and necessary differences.

Bibliography/Texts/Supplies—Required:

- Readings are listed below in the course schedule and should be completed 24 hours before the live session.

Course Requirements and Expectations:

Type of Activity	Quantity	Points	Notes	Total
1. Asynchronous Participation	10	2	Lowest score dropped (9 count). Due 24 hours before the live session.	18
2. Synchronous Participation	10	2	Lowest score dropped (9 count). Participation in the live session	18
3. Lab Homework	10	3	All 10 assignments count. Due before the live session.	30
4. Final Exam	1	34	Comprehensive. Issued in the final live session.	34
Total Points				**100**

1. **Asynchronous Participation**. Measured through a student's participation in the asynchronous portion of the course, this includes assigned readings, watching lecture video, answering lecture questions, completing quizzes, and engaging in discussion questions posed in the lecture. Activity will be monitored weekly and students will be assigned feedback and grades of:
    a. High pass (2 points): 100% of the work was completed 24 hours before the live session,
    b. Low pass (1 point): at least 70% of the work as completed 24 hours before the live session, or
    c. Fail (0 points): less than 70% of the work was completed 24 hours before the live session.
2. **Synchronous Participation**. Measured through a student's participation in the course's synchronous components or "live session," this includes a discussion on course content (asynchronous content and homework), participation in live-session activities, group breakouts, project work, and report-out activities. Activity will be monitored weekly and students will be assigned feedback and grades of:
    a. High pass (2 points): noticeable contributions to live session,
    b. Low pass (1 point): passive engagement in live session, or
    c. Fail (0 points): no attendance in live session.
3. **Lab Homework** assignments are technical activities which enforce asynchronous concepts through practice and are based on the demos in the course videos. Homework must be completed before the week's live session, where students must be prepared to discuss the outcomes of the assignment. Students will be assigned a grade of:
    a. High pass (3 points): the assignment is complete and correct with no errors,
    b. Pass (2 points): the assignment is complete on time and mostly correct with one or two errors,
    c. Low pass (1 point): the assignment is incomplete or has three or more errors, or not on time, or
    d. Fail (0 points): the assignment was not turned in by the due date.
4. **Final Exam.** The final exam will be issued and must be completed individually during the last live session. The exam will be given online and students will have 60 minutes to complete the 34-question exam. The exam is cumulative and covers all topics in the course, including assigned readings, homework, asynchronous course work, and synchronous activities.

Grading:

Student Achievement	Percentage	Registrar Grade
Mastery	95–100	A
	90–94	A–
Satisfactory	85–89	B+
	80–84	B
Low Passing	75–79	B–
	70–74	C+
Unsatisfactory	65–69	C
	60–64	C–
	0–59	F

University Attendance Policy

Attendance in classes is expected in all courses at Syracuse University. Students are expected to arrive in time to attend the first meeting of all classes for which they are registered. Students who do not attend classes starting with the first scheduled meeting may be academically withdrawn as not making progress toward degree by failure to attend. Instructors set course-specific policies for absences from scheduled class meetings in their syllabi.

It is a federal requirement that students who do not attend or cease to attend a class to be reported at the time of determination by the faculty. Students should also review the university's religious observance policy and make the required arrangements at the beginning of each semester.

Course-Specific Policies:

- All work is due on the dates provided. No late work is accepted. The reasoning is the grading is participation-/effort-based and most of the content time sensitive.
- No makeup work. Your lowest synchronous and asynchronous session grades are dropped to account for what might otherwise be an excused absence.
- Final grades will not be rounded up. 94/100 is an A–; please don't ask.

Syracuse University Policies:

Students should review the University's policies regarding: Diversity and Disability, https://www.syracuse.edu/life/accessibilitydiversity/; the Religious Observances Notification and Policy, http://supolicies.syr.edu/studs/religious_observance.htm; and Orange Success, http://orangesuccess.syr.edu/getting-started-2/.

Disability-Related Accommodations

Syracuse University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. If you believe that you need academic adjustments (accommodations) for a disability, please contact the Office of Disability Services (ODS), visit the ODS website—http://disabilityservices.syr.edu, located in Room 309 of 804 University Avenue, or call (315) 443-4498 or TDD (315) 443-1371, for an appointment to discuss your needs and the process for requesting academic adjustments. ODS is responsible for coordinating disability-related academic adjustments and will issue students with documented Disabilities Accommodation Authorization Letters, as appropriate. Since academic adjustments may require early planning and generally are not provided retroactively, please contact ODS as soon as possible. Our goal at the iSchool is to create learning environments that are useable, equitable, inclusive, and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please meet with me to discuss additional strategies beyond official accommodations that may be helpful to your success.

Academic Integrity Policy

Syracuse University's Academic Integrity Policy reflects the high value that we, as a university community, place on honesty in academic work. The policy defines our expectations for academic honesty and holds students accountable for the integrity of all work they submit. Students should understand that it is their responsibility to learn about course-specific expectations, as well as about university-wide academic integrity expectations. The policy governs appropriate citation and use of sources, the integrity of work submitted in exams and assignments, and the veracity of signatures on attendance sheets and other verification of participation in class activities. The policy also prohibits students from submitting the same work in more than one class without receiving written authorization in advance from both instructors. Under the policy, students found in violation are subject to grade sanctions determined by the course instructor and nongrade sanctions determined by the School or College where the course is offered as described in the Violation and Sanction Classification Rubric. SU students are required to read an online summary of the University's academic integrity expectations and provide an electronic signature agreeing to abide by them twice a year during preterm check-in on MySlice. For more information about the policy, see http://class.syr.edu/academic-integrity/policy/.

Course Schedule:

Week/lecture, topic for the week/lecture, and required readings are in the columns below. These readings will be discussed in the live session. It is recommended that you first complete the asynchronous coursework prior to reading the items on this list.

Week	Topic/Reading List	Due
1	**Course Introduction: Setting the Stage** • The World's Most Valuable Resource Is No Longer Oil, but Data https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data • The Internet of Things Won't Be Big It'll be Huge https://www.forbes.com/sites/markpmills/2016/09/28/the-internet-of-things-wont-be-big-itll-be-huge/#2d7589bb9915 • What Is Polyglot Persistence? https://www.jamesserra.com/archive/2015/07/what-is-polyglot-persistence/ • Digital Transformations, Part 1: Rise of Microservices https://www.linkedin.com/pulse/part-1-rise-microservices-manoj-bhardwaj	HW1
2	**Relational Databases: Programming** • Why I Avoid Stored Procedures (and You Should Too) https://kevinlawry.wordpress.com/2012/08/07/why-i-avoid-stored-procedures-and-you-should-too/ • It's Time to Get Over That Stored Procedure Aversion You Have https://rob.conery.io/2015/02/20/its-time-to-get-over-that-stored-procedure-aversion-you-have/	HW2
3	**Relational Databases: Transactions, Concurrency, and Triggers** • Transactions and Concurrency Control https://gradeup.co/transactions-and-concurrency-control-i-4c5d9b27-c5a7-11e5-bcc4-bc86a005f7ba • How to Use SQL Temporal Tables for Easy Point-in-Time Analysis https://hackernoon.com/how-to-use-sql-temporal-tables-for-easy-point-in-time-analysis-38d43e4ee557	HW3
4	**Relational Databases: Performance, Security, and noSQL** • Working With Different SQL Server Indexes Types https://www.sqlshack.com/working-with-different-sql-server-indexes-types/ • Combining Relational and noSQL Concepts in SQL Server https://blogs.msdn.microsoft.com/sqlserverstorageengine/2015/09/01/combining-relational-and-nosql-concepts-in-sql-server/	HW4

5	**Hadoop: HDFS, MapReduce and YARN** • Don't Use Hadoop—Your Data Isn't That Big! https://www.chrisstucchio.com/blog/2013/hadoop_hatred.html • Sqoop vs. Flume Battle of the Hadoop ETL Tools https://www.dezyre.com/article/sqoop-vs-flume-battle-of-the-hadoop-etl-tools-/176	HW5
6	**Hadoop: Pig, Hive, and HCatalog** • Comparing Pig Latin and SQL for Constructing Data Processing Pipelines http://yahoohadoop.tumblr.com/post/98294444546/comparing-pig-latin-and-sql-for-constructing-data • Apache Pig Overview http://hadooptutorial.info/apache-pig-overview/	HW6
7	**Hadoop: Hbase, Impala, and Other Tools** • Hbase—Overview of Architecture and Data Model https://www.netwoven.com/2013/10/10/hbase-overview-of-architecture-and-data-model/ • Impala vs Hive: Difference Between SQL on Hadoop https://www.dezyre.com/article/impala-vs-hive-difference-between-sql-on-hadoop-components/180	HW7
8	**NoSQL: MongoDB and Redis** • The Little MongoDB Book (Chapters 1–4) https://github.com/karlseguin/the-little-mongodb-book/blob/master/en/mongodb.markdown • The Little Redis Book (Chapters–1-4) https://github.com/karlseguin/the-little-redis-book/blob/master/en/redis.md	HW8
9	**NoSQL: Cassandra** • How to Make MongoDB Not Suck for Analytics https://www.scaleapi.com/blog/athena#asdf • Introduction to Apache Cassandra's Architecture https://dzone.com/articles/introduction-apache-cassandras • Best Practices for Cassandra Data Modeling https://dzone.com/articles/best-practices-for-cassandra-data-modeling	HW9
10	**NoSQL: Kakfa and KSQL** • What Is Apache Kafka? Why Is It So Popular? Should I Use It? https://techbeacon.com/what-apache-kafka-why-it-so-popular-should-you-use-it • Introduction to Apache Kafka https://kafka.apache.org/intro	HW10
11	**Final Exam**	