

---

IST769: Advanced Big Data Management Systems

Instructor: Michael Fudge

Phone: 315-443-9686

Office: 314 Hinds Hall

Email: mafudge@syr.edu

Office Hours: TBA

---

**Description:**

An analysis of relational and non-relational 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:** IST659

**Audience:** Graduate

**Credits:** 3.0

**Learning Objectives:**

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

1. Demonstrate comprehension of 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, graph 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 before the live sessions.

**Course Requirements and Expectations:**

Type of Activity	Quantity	Points	Learning Outcomes	Total
A. Class Participation	10	1	1,2,3,5	10
B. Homework Assignments	10	3	1,2,3,4,5	30
C. Midterm/ Final Exam	2	20	1,2,3,4,5	40
D. Research Paper and Slide Deck	1	20	2,3,4,5	20
<b>Total Points</b>				<b>100</b>

- A. **Class Participation.** You are expected to come to class prepared, which assumes you will complete the assigned readings and coursework prior to the class session.
- All students start with 10 participation points. You only lose a point when you are absent, fail a participation diagnostic, or noticeably do not participate in class.
  - In class we will have discussions and group activities pertinent to the coursework and assigned readings. Your participation will be recorded by the instructor and student who do not participate will be warned or notified of a point deduction.
  - Attendance is required and there are no excused absences other than medical illness documented by the university. You cannot participate without being in attendance.
  - In some classes a pop-quiz (also known as a participation diagnostic) will measure your study of coursework and a passing grade will count as your participation grade for that week.
- B. **Homework Assignments** Homework assignments must be completed each week. These assignments are technical activities which enforce asynchronous concepts through practice. Homework must be completed before the week's class so that students are prepared to discuss the outcomes of the assignment and ask questions about it. Students will be assigned a grade of:
- High Pass (3 points) the assignment is complete and correct with no errors,
  - Pass (2 points) the assignment is complete on time and mostly correct with 1-2 errors,
  - Low pass (1 point) the assignment is incomplete or has 3 or more errors, or not on time
  - Fail (0 points) the assignment was not turned in by the due date.
- C. **Exams.** This course has a mid-term and final exam, which will be issued in class. Both exams will cover coursework completed up to the date of the exam. The exam format is a mix of multiple choice, short answer, and essay questions. Examinations are closed-book, individual assessments and students will be given approximately 60 minutes to complete the exam. Students must take the exam on the assigned date and there are no make-ups.

D. **Research Paper and Slide Deck.** This individual culminating assignment is your chance to demonstrate you can apply the principles learned in the course within a new context. You will be required to research a database system of your choosing, then critically evaluate the system based on the following defined criteria: 1. Key features 2. Storage requirements, 3. How the product fits within the CAP theorem, 4. Appropriate / inappropriate use cases, 5. Query language used, 6. Supported data models, 7. Technical limitations 8. Competing products and 9. A comparison of those competing products including a justification as to why they are competitors.

- **The whitepaper.** Should cover this material in depth and include in-text citations, and a list of sources cited. There should be at least 3 sources used to reference the product being researched, and this should exclude sources for competing products, which would include additional references.
- **The Slide Deck.** The slide deck is a summary of your whitepaper. It should include slides which outline the criteria of the whitepaper. In addition, it should include a title slide list of references.

Students will be evaluated based on quality and accuracy of their evaluation as it relates to the lessons learned in the course, quality and appropriateness of sources cited, spelling and grammar, coherence of writing.

**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

**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.
- Make-up work is not permitted. The only exception being university-approved extenuating circumstances.
- Final grades will not be rounded up. 94/100 is an A-, please don't ask.
- Sanctions for violations of academic integrity are an F in the course.

**Course Schedule:**

Week/ lecture, topic for the week/lecture, and required readings are in the columns below.

Week	Topic / Reading List	Due
1	<p><b>Course Introductions, Review and Lab Prep</b></p> <ul style="list-style-type: none"> <li>• Go over the syllabus, meet and greet, class expectations</li> <li>• IST659 Review <a href="https://www.youtube.com/embed/videoseries?list=PLA1261A156B60856E">https://www.youtube.com/embed/videoseries?list=PLA1261A156B60856E</a></li> <li>• Let's learn Docker and Docker-compose <a href="https://docker-curriculum.com/">https://docker-curriculum.com/</a> <a href="https://www.katacoda.com/courses/docker">https://www.katacoda.com/courses/docker</a></li> </ul>	
2	<p><b>Introduction</b></p> <ul style="list-style-type: none"> <li>• The world's most valuable resource is no longer oil, but data <a href="https://www.economist.com/leaders/2017/05/06/the-worlds-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</a></li> <li>• The Internet of Things Won't Be Big, It'll be Huge <a href="https://www.forbes.com/sites/markpmills/2016/09/28/the-internet-of-things-wont-be-big-itll-be-huge/#2d7589bb9915">https://www.forbes.com/sites/markpmills/2016/09/28/the-internet-of-things-wont-be-big-itll-be-huge/#2d7589bb9915</a></li> <li>• What is Polyglot Persistence? <a href="https://www.jamesserra.com/archive/2015/07/what-is-polyglot-persistence/">https://www.jamesserra.com/archive/2015/07/what-is-polyglot-persistence/</a></li> <li>• Digital Transformations Part 1: Rise of Microservices <a href="https://www.linkedin.com/pulse/part-1-rise-microservices-manoj-bhardwaj">https://www.linkedin.com/pulse/part-1-rise-microservices-manoj-bhardwaj</a></li> </ul>	HW1
3	<p><b>Relational: Programming</b></p> <ul style="list-style-type: none"> <li>• Why I avoid stored procedures and you should too <a href="https://kevinlawry.wordpress.com/2012/08/07/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/</a></li> <li>• It's time to get over that stored procedure aversion you have <a href="https://rob.conery.io/2015/02/20/its-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/</a></li> </ul>	HW2
4	<p><b>Relational: Transactions, Concurrency Control</b></p> <ul style="list-style-type: none"> <li>• Transactions and Concurrency Control <a href="https://gradeup.co/transactions-and-concurrency-control-i-4c5d9b27-c5a7-11e5-bcc4-bc86a005f7ba">https://gradeup.co/transactions-and-concurrency-control-i-4c5d9b27-c5a7-11e5-bcc4-bc86a005f7ba</a></li> <li>• Use Temporal tables for easy point-in-time analysis <a href="https://hackernoon.com/how-to-use-sql-temporal-tables-for-easy-point-in-time-analysis-38d43e4ee557">https://hackernoon.com/how-to-use-sql-temporal-tables-for-easy-point-in-time-analysis-38d43e4ee557</a></li> </ul>	HW3
5	<p><b>Relational: Performance, Security, noSQL</b></p> <ul style="list-style-type: none"> <li>• Working with different SQL Server index types: <a href="https://www.sqlshack.com/working-with-different-sql-server-indexes-types/">https://www.sqlshack.com/working-with-different-sql-server-indexes-types/</a></li> <li>• Combining Relational and noSQL concepts in SQL Server <a href="https://blogs.msdn.microsoft.com/sqlserverstorageengine/2015/09/01/combining">https://blogs.msdn.microsoft.com/sqlserverstorageengine/2015/09/01/combining</a></li> </ul>	HW4

	<a href="#">-relational-and-nosql-concepts-in-sql-server/</a>	
6	<b>Midterm Exam</b>	
7	<b>Hadoop: HDFS, MapReduce and YARN</b> <ul style="list-style-type: none"> <li>Don't use Hadoop – your data isn't that big! <a href="https://www.chrisstucchio.com/blog/2013/hadoop_hatred.html">https://www.chrisstucchio.com/blog/2013/hadoop_hatred.html</a></li> <li>Sqoop vs. Flume Battle of the Hadoop ETL tools. <a href="https://www.dezyre.com/article/sqoop-vs-flume-battle-of-the-hadoop-etl-tools-/176">https://www.dezyre.com/article/sqoop-vs-flume-battle-of-the-hadoop-etl-tools-/176</a></li> </ul>	HW5
8	<b>Hadoop: Pig, Hive, HCatalog</b> <ul style="list-style-type: none"> <li>Comparing Pig Latin and SQL for Constructing Data Processing Pipelines <a href="http://yahooohadoop.tumblr.com/post/982944444546/comparing-pig-latin-and-sql-for-constructing-data">http://yahooohadoop.tumblr.com/post/982944444546/comparing-pig-latin-and-sql-for-constructing-data</a></li> <li>Apache Pig Overview <a href="http://hadooptutorial.info/apache-pig-overview/">http://hadooptutorial.info/apache-pig-overview/</a></li> </ul>	HW6
9	<b>Hadoop: Hbase, Impala, and More</b> <ul style="list-style-type: none"> <li>Hbase – Overview of Architecture and Data Model <a href="https://www.netwoven.com/2013/10/10/hbase-overview-of-architecture-and-data-model/">https://www.netwoven.com/2013/10/10/hbase-overview-of-architecture-and-data-model/</a></li> <li>Impala vs. Hive: Difference between SQL on Hadoop <a href="https://www.dezyre.com/article/impala-vs-hive-difference-between-sql-on-hadoop-components/180">https://www.dezyre.com/article/impala-vs-hive-difference-between-sql-on-hadoop-components/180</a></li> </ul>	HW7
10	<b>noSQL: MongoDB and Redis</b> <ul style="list-style-type: none"> <li>The Little MongoDB Book (Chapters 1-4) <a href="https://github.com/karlseguin/the-little-mongodb-book/blob/master/en/mongodb.markdown">https://github.com/karlseguin/the-little-mongodb-book/blob/master/en/mongodb.markdown</a></li> <li>The Little Redis Book (Chapters 1-4) <a href="https://github.com/karlseguin/the-little-redis-book/blob/master/en/redis.md">https://github.com/karlseguin/the-little-redis-book/blob/master/en/redis.md</a></li> </ul>	HW8
11	<b>noSQL: Cassandra</b> <ul style="list-style-type: none"> <li>How to make MongoDB not suck for analytics <a href="https://www.scaleapi.com/blog/athena#asdf">https://www.scaleapi.com/blog/athena#asdf</a></li> <li>Introduction to Apache Cassandra's Architecture <a href="https://dzone.com/articles/introduction-apache-cassandras">https://dzone.com/articles/introduction-apache-cassandras</a></li> <li>Best Practices for Cassandra Data Modeling <a href="https://dzone.com/articles/best-practices-for-cassandra-data-modeling">https://dzone.com/articles/best-practices-for-cassandra-data-modeling</a></li> </ul>	HW9
12	<b>noSQL: Kafka and KSQL</b> <ul style="list-style-type: none"> <li>What Is Apache Kafka? Why Is It So Popular? Should I Use It? <a href="https://techbeacon.com/what-apache-kafka-why-it-so-popular-should-you-use-it">https://techbeacon.com/what-apache-kafka-why-it-so-popular-should-you-use-it</a></li> <li>Introduction to Apache Kafka <a href="https://kafka.apache.org/intro">https://kafka.apache.org/intro</a></li> </ul>	HW10

13	<p><b>noSQL: Graph Databases</b></p> <ul style="list-style-type: none"> <li>• Sql Server 2017 Graph Database <a href="https://www.sqlshack.com/introduction-sql-server-2017-graph-database/">https://www.sqlshack.com/introduction-sql-server-2017-graph-database/</a></li> <li>• Neo4j Getting Started <a href="https://neo4j.com/developer/get-started/">https://neo4j.com/developer/get-started/</a></li> </ul>	
14	<b>Final Exam</b>	Paper / Slide Deck

**Syracuse University Policies**

Syracuse University has a variety of other policies designed to guarantee that students live and study in a community respectful of their needs and those of fellow students. Some of the most important of these concerns:

**Diversity and Disability** (ensuring that students are aware of their rights and responsibilities in a diverse, inclusive, accessible, bias-free campus community) can be found here, at <https://www.syracuse.edu/life/accessibilitydiversity/>

**Religious Observances Notification and Policy** (steps to follow to request accommodations for the observance of religious holidays) can be found here, at: [http://supolicies.syr.edu/studs/religious\\_observance.htm](http://supolicies.syr.edu/studs/religious_observance.htm)

**Orange Success** (tools to access a variety of SU resources, including ways to communicate with advisors and faculty members) can be found here, at: <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. There may be aspects of the instruction or design of this course that result in barriers to your inclusion and full participation in this course. I invite any student to meet with me to discuss strategies and/or accommodations (academic adjustments) that may be essential to your success and to collaborate with the Office of Disability Services (ODS) in this process.

If you would like to discuss disability-accommodations or register with ODS, please visit their website at <http://disabilityservices.syr.edu> Please call (315) 443-4498 or email [disabilityservices@syr.edu](mailto:disabilityservices@syr.edu) for more detailed information.

ODS is responsible for coordinating disability-related academic accommodations and will work with the student to develop an access plan. Since academic accommodations may require early planning and generally are not provided retroactively, please contact ODS as soon as possible to begin this process.

**University Attendance Policy:**

Attendance in classes is expected in all courses at Syracuse University. Students are expected to arrive on campus 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. Faculty should use “ESPR” and “MSPR” in Orange Success to alert the Office of the Registrar and the Office of Financial Aid. A grade of NA is posted to any student for whom the Never Attended flag is raised in Orange SUccess. More information regarding Orange SUccess can be found here, at: <http://orangesuccess.syr.edu/getting-started-2/>

Students should also review the University’s religious observance policy and make the required arrangements at the beginning of each semester

**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 non-grade 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 pre-term check-in on MySlice.

**Discrimination or Harassment:**

The University does not discriminate and prohibits harassment or discrimination related to any protected category including creed, ethnicity, citizenship, sexual orientation, national origin, sex, gender, pregnancy, disability, marital status, age, race, color, veteran status, military status, religion, sexual orientation, domestic violence status, genetic information, gender identity, gender expression or perceived gender.

Any complaint of discrimination or harassment related to any of these protected bases should be reported to Sheila Johnson-Willis, the University’s Chief Equal Opportunity & Title IX Officer. She is responsible for coordinating compliance efforts under various laws including Titles VI, VII, IX and Section 504 of the Rehabilitation Act. She can be contacted at Equal Opportunity, Inclusion, and Resolution Services, 005 Steele Hall, Syracuse University, Syracuse, NY 13244-1120; by email: [titleix@syr.edu](mailto:titleix@syr.edu); or by telephone: 315-443-0211.

**Course evaluations:**

---

There will be an end of course evaluation for you to complete this term. This evaluation will be conducted online and is entirely anonymous. You will receive an official notification in your email account with the evaluation website link and your passcode. Please take the time and fill out this evaluation as your feedback and support of this assessment effort is very much appreciated. The school carefully reviews ratings and comments that you submit, and these factor into decisions about course, program and instructor development.

**Use of Blackboard:**

This course involves the use of Syracuse University's Blackboard system as an online tool. The environment is composed of a number of elements that will help you be successful in both your current coursework and your lifelong learning opportunities. To access Blackboard, <http://blackboard.syr.edu> use your Syracuse University NetID & Password. This specific course will appear in your course list.

To search for answers to your Blackboard questions, visit the Answers self-help knowledge <https://answers.syr.edu/display/blackboard01/Blackboard>. If you have problems logging in or need assistance with Blackboard, contact the ITS Service Center at: [help@syr.edu](mailto:help@syr.edu) or 315.443.2677. The Syracuse University Blackboard support team will assist you.