SYLLABUS

INSTRUCTOR

Dr. George Ray

COURSE INFORMATION

Course Title: CMSC 491, NoSQL Database Concepts

Time & Place: TuTh 6:00 to 8:20; room TBA.

Section: 02

TEXTBOOKS

There is no required textbook for this course. We will be reading articles from the computer science literature, which your instructor will cover in lectures. A link to a variety of software engineering resources will be available on the course web site. Additionally, we will be referencing material from the following textbooks:

- *NoSQL and SQL Data Modeling* by Hills
- *Graph Databases 2nd Edition* by Robinson, Webber and Eifrem
- *MongoDB in Action* by Banker, Bakkum, Verch, Garret, Hawkins
- *NoSQL* by Sullivan
- *Database System Concepts, 6th Edition* by Silberschatz, Korth, and Sudarshan

COURSE DESCRIPTION

A sudden change is taking place in database technology and well established relational systems are losing ground to NoSQL alternatives, especially in new information services. Job postings reflect this emerging trend: demand for traditional SQL database analysts and administrators has dropped over the past four years by 25 to 50 percent depending on the DBMS. At the same time, demand for NoSQL database analysts and administrators has grown over 250%.

This course is a comprehensive introduction into this new field. It presents the basic concepts of NoSQL databases that have propelled them into the vanguard, as the demands of big data are

surpassing the capabilities of relational databases. Students will get a thorough understanding of the advantages and disadvantages of the leading NoSQL database systems, as well as hands-on experience installing, administering and developing programs that use them. A data model will be introduced that represents the entities and principles in the problem space, can be used for any logical data design, as well as specific NoSQL document, key-value, columnar, and graph database implementations

PREREQUISITES

This course assumes a basic familiarity with programming and database concepts such as that provided in CMSC 201 and CMSC 461. A refresher on database concepts needed in the course as well as Python will be provided.

OBJECTIVES

After completion of this course, students will:

- Know the use cases for each of the four main models of NoSQL database and guiding principles for what problems they can best solve
- Understand model driven design for key-value, document, columnar, and graphical database designs, for developing both the logical and physical models
- Have hands-on installation and administration experience with NoSQL databases
- Have hands-on application development experience with NoSQL databases
- Understand BASE and how it contrasts with ACID
- Understand CAP and Brewer's theorem
- Understand scalability with sharding, parallelism with Map-Reduce, as well as reliability and availability through replication
- Understand approaches to eventual consistency
- Describe the basic ways distributed systems are synchronized, two-phase commit with quorum, vector clocks in both state transfer and operational transfer
- Understand key-based range partitioning and consistent hashing

GRADING

The final grade will be computed from the following components:

2 Individual Projects	30%
1 Group Project	30%
2 Exams	40%
Total =	100%

The initial plan calls for 2 individual projects. Each individual project will be worth 15% of your grade. A group project will be worth 30%. There will be 1 midterm and a cumulative final, each worth 20%.

Your final letter grades will be based on the standard formula:

$0 \leq F < 60, \quad 60 \leq D < 70, \quad 70 \leq C < 80, \quad 80 \leq B < 90, \quad 90 \leq A \leq 100$

Depending upon the distribution of grades in the class, there may be adjustments in the students' favor, but under no circumstances will the letter grades be lower than in the standard formula. Grades will not be "curved" in the sense that the percentages of A's, B's and C's are not fixed.

Grades are given for work done *during* the semester; incomplete grades will only be given for medical illness or other such dire circumstances.

ATTENDANCE

You are expected to attend all classes. If you miss a class, you are responsible for getting the notes and any verbal information given during class from a fellow classmate.

PROJECTS

You will not acquire NoSQL skills by watching someone else work with the various models. You must budget enough time to think about the assignments and then design solutions to them.

If you cannot complete an assignment, you should still submit your work. Partial credit will be given for reasonable effort. Late work will not be accepted.

You will be submitting your assignments and project electronically. Details will be explained in class before you need to submit your first project.

ACADEMIC INTEGRITY

By enrolling in this course, each student assumes the responsibilities of an active participant in UMBC's scholarly community in which everyone's academic work and behavior are held to the highest standards of honesty. Cheating, fabricating, plagiarism, and helping others to commit these acts are all forms of academic dishonesty and they are wrong. Academic misconduct could result in disciplinary action that may include, but is not limited to, suspension or dismissal. Full policies on academic integrity should be available in the UMBC Student Handbook, Faculty Handbook, or the UMBC Directory.

These policies recognize that students can learn productively from many sources including from other students in the class. Thus, policy allows small amounts of help but prohibits outright copying. Although, this leaves a gray area between "small amounts of help" and "outright copying", it is better that we live with some ambiguity than to have a clear-cut policy that deprives the students of productive learning opportunities. Students who have doubts about the propriety of an activity should consult the instructor.

Students who violate this academic integrity policy will receive a grade of 0 for that assignment, as

well as a reduction of one full letter grade in the student's final course grade. A second violation will result in very dire consequences. In the case where one student copies the work of another student, both students are considered to have violated this policy. Here, copying includes not just verbatim copies, but also work that is substantially similar and could not have been produced independently. Furthermore, all parties concerned will have their prior homework and classwork checked.

Violations of this policy may be reported to the University's Academic Conduct Committee for further action. Egregious cases of cheating will be written up as a "more serious" infraction. In this case, you will not be allowed to drop the course. Also, a "more serious" infraction would appear as a permanent part of your student record and would be seen by potential employers when they ask for an official copy of your transcript.

EMAIL

In order to facilitate email communication, please observe the following guidelines for email sent to the instructor.

Make sure that the subject line of the email message clearly identifies its content (e.g., mention CMSC 491).

- Use your UMBC email account. (I really shouldn't discuss your grade with some random person on the internet just because he has an email address that resembles your name.)
- Use your full real name.
- Submit your work instead of attaching it to your message.