NoSQL Database Management
MISM Course S16-95-737 A4
Spring 2016
Carnegie Mellon University

Instructor: Dan Costa
Office: Software Engineering Institute / CERT – CIC
Office hours: By Appointment Only
Phone: 412-268-8006
E-mail: dlcosta@andrew.cmu.edu

TA: Dilip Gillela
Office hours: TBD
E-mail: dgillela@andrew.cmu.edu

Building / Room: HBH 236
Time: 6:00 PM - 8:50 PM, Thursday
Web site: http://www.cmu.edu/blackboard

Textbooks

Prerequisites and Requirements
Prerequisite: 95-703, Database Management
Requirement: Students MUST have a laptop with the ability to install multiple databases on it.
Note: This course will include in-class labs that involve the installation, configuration and programming of multiple databases. It is recommended, though not required, that students have some experience with programming languages (e.g. Java, HTML, SQL).

Course Description
The widespread emergence of big data storage needs has driven the development and adoption of a new class of non-relational databases commonly referred to as NoSQL databases. This course will explore the origins of NoSQL databases and the characteristics that distinguish them from traditional relational database management systems. Core concepts of NoSQL databases will be presented, followed by an exploration of how different database technologies implement these core concepts. We will take a closer look at 1-2 databases from each of the four main NoSQL data models (key-value, column family, document, and graph), highlighting the business needs that drive the development and use of each database. Finally, we will present criteria that decision makers should consider when choosing between relational and non-relational databases and techniques for selecting the NoSQL database that best addresses specific use cases.
Learning Objectives

<table>
<thead>
<tr>
<th>Learning Objective</th>
<th>How Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrate competency in designing NoSQL database management systems.</td>
<td>Final Exam, In-Class Labs, Assignments</td>
</tr>
<tr>
<td>Demonstrate competency in describing how NoSQL databases differ from relational databases from a theoretical perspective.</td>
<td>Final Exam, In-Class Labs, Assignments, Research Report</td>
</tr>
<tr>
<td>Demonstrate competency in selecting a particular NoSQL database for specific use cases.</td>
<td>Final Exam, In-Class Labs, Assignments</td>
</tr>
</tbody>
</table>

Schedule (tentative, subject to change during semester)

<table>
<thead>
<tr>
<th>Date</th>
<th>Lecture / Lab</th>
<th>Readings / References</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 17</td>
<td>Course Introduction / Differences from Relational Databases</td>
<td>PS: Ch. 1-2</td>
</tr>
<tr>
<td>March 24</td>
<td>Theory (CAP, Map-Reduce, Distributed Computing)</td>
<td>PS: Ch. 4-7</td>
</tr>
<tr>
<td>March 31</td>
<td>Key-Value Databases</td>
<td>PS: Ch. 8</td>
</tr>
<tr>
<td></td>
<td><em>Redis Lab</em></td>
<td>ER: Ch. 8</td>
</tr>
<tr>
<td>April 7</td>
<td>Document Stores</td>
<td>PS: Ch. 9</td>
</tr>
<tr>
<td></td>
<td><em>MongoDB Lab</em></td>
<td>ER: Ch. 5</td>
</tr>
<tr>
<td>April 14</td>
<td>Column Family Stores</td>
<td>PS: Ch. 10</td>
</tr>
<tr>
<td></td>
<td><em>Cassandra Lab</em></td>
<td>ER: Ch. 4</td>
</tr>
<tr>
<td>April 21</td>
<td>Graph Databases</td>
<td>PS: Ch. 11</td>
</tr>
<tr>
<td></td>
<td><em>Neo4j Lab</em></td>
<td>ER: Ch. 7</td>
</tr>
<tr>
<td>April 28</td>
<td>The Database Landscape / Choosing a NoSQL Database</td>
<td>PS: Ch. 13-15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ER: Ch. 9</td>
</tr>
<tr>
<td>May 5</td>
<td><strong>Final Exam</strong></td>
<td></td>
</tr>
</tbody>
</table>

* Remember to bring your laptop to class on lab days.

Assignments / Research Report

There will be 2 assignments based on topics covered in lectures and your work with the tools in the lab sessions. In addition, there will be a research report due on week 7. Students will select a topic germane to NoSQL databases for independent research, submit the topic for approval, and develop a 5-8 page report on their chosen topic. A draft version of each report will be due by week 5, and this draft version will be peer reviewed by another student in the class. Final versions of the research reports will be due by week 7. Following is a list of due dates for each assignment:

<table>
<thead>
<tr>
<th>Item</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework 1 - The CAP Theorem &amp; Map-Reduce</td>
<td>March 31 @ 6:00 PM</td>
</tr>
<tr>
<td>First Draft of Research Report</td>
<td>April 14 @ 6:00 PM</td>
</tr>
<tr>
<td>Peer Reviews of Research Reports</td>
<td>April 21 15 @ 6:00 PM</td>
</tr>
<tr>
<td>Homework 2 - Aggregate Oriented Design</td>
<td>April 21 @ 6:00 PM</td>
</tr>
<tr>
<td>Final Version of Research Report</td>
<td>April 28 @ 6:00 PM</td>
</tr>
</tbody>
</table>

Evaluation Method

Labs: 5%
Assignments: 25%
Research Report: 25% (90% final version, 5% peer review completion, 5% first draft submission)
Final Exam: 45%

Please note that class attendance is important. Although I don’t include attendance as part of the total percentage making up your grade, failure to attend class on a regular basis will have an adverse effect on your grade.
Students will only have 2 weeks after an assignment or exam is returned to question or challenge a grade. After the two week challenge period, the grade will not be changed. Please contact the instructor if you wish to question a grade.

**Grading Scale**

100 - 98 A+
97 - 92 A
91 - 90 A-
89 - 88 B+
87 - 82 B
81 - 80 C-
79 - 78 C+
77 - 72 C
71 - 70 C-

**Grade Distribution**

I plan on using the Heinz School guidelines in deciding on the overall grade distribution. Accordingly, the average grade will be an A-. However, I grade on an absolute scale. If every student does well in the class, each will get an A+ regardless of the recommended grading scale. The same holds true on the other end of the scale.

**Final Exam**

The final exam will cover material from the entire semester. The final exam is scheduled for May 5 from 6:00 pm – 8:50 pm. Please do not schedule anything that might conflict with the final exam. No one will be excused from it and there will be no make-up exam dates.

**Late assignment policy**

Homework is due at 6:00 pm on the assigned due date. I WILL NOT accept late homework unless the student has made arrangements with me prior to the assignment’s due date. PRIOR ARRANGEMENTS MUST BE MADE NO LATER THAN 12 PM ON THE DUE DATE.

**Policy on cheating and plagiarism**

Each student is responsible for handing in his/her own work. For any assignment found to be the partial or complete result of cheating or plagiarism, your grade for that assignment will be zero. Cheating is defined as inappropriate collaboration among students on an assignment. This can include copying someone else’s work with or without alteration. When students are found to be collaborating in this way, BOTH will pay the penalty regardless of who originated the work.

**Classroom Etiquette**

This is a Master’s level course taught as part of a professional degree program. Accordingly, you are expected to conduct yourself in a professional manner during the course and not engage in behavior in the class that would be considered unacceptable in the workplace.

I expect the following guidelines to be followed:

- Do not sleep in class
  - I realize you may occasionally have had a rough day, are unwell, or otherwise are too tired to stay awake. In those cases, it is acceptable to excuse yourself from that class.
  - If you are always too tired to stay awake during lectures, you should consider finding an alternative course offering.
  - If you find the material boring, please let me know. I will attempt to modify the content to better suit your interests and needs.
• Turn off your cell phones. You are not to answer calls while in class. If you have a need to be available during class, please let me know before the lecture begins.
• Please don’t browse the web, instant message, or check email during lectures.
• If you have a question about the content of the lecture, please direct it to me. If you are confused about an issue, chances are your classmates are confused as well. Please do not ask for clarification from your classmate during lecture.