Large-scale distributed systems are the core software infrastructure underlying cloud computing. These systems consist of tens of thousands of networked computers working together to provide unprecedented performance and fault-tolerance. Examples include distributed databases (e.g. Google's Spanner, Amazon's S3 and Dynamo), distributed computation frameworks (e.g. Hadoop/Spark and TensorFlow) or even decentralized currency systems (e.g. BitCoin). This class teaches the abstractions, design and implementation techniques that allow you to understand and build such distributed systems.

It will present abstractions and implementation techniques for engineering distributed systems. Major topics include multithreading, network programming, fault tolerance, replication, and consistency. We will also spend a lot of time studying several case studies of distributed systems.

Contact
Email: gustavo.sandoval@nyu.edu
2 Metrotech, 10th Floor, Office 10.040
Office hours: Wed: 2:30 – 4:00

Student Assistants & Office Hours:
- Adam Sanghera – as8513@nyu.edu
- Dov Salomon – dsalomon@nyu.edu

Prerequisites:
CS 6233 – Intro to Operating Systems.
Expertise in C/C++ programming.

Academic Honesty

All work submitted in this course must be your own. Cheating and plagiarism will not be tolerated. If you have any questions about a specific case, please ask me.
NYU Poly’s Policy on Academic Misconduct is here:
http://engineering.nyu.edu/academics/code-of-conduct/academic-misconduct

Course requirements
- This term we will be using Piazza for class discussion. The system is highly
catered to getting you help fast and efficiently from classmates, the TA and
myself. Rather than emailing questions to the teaching staff, I encourage you to
post your questions on Piazza. Here’s the link for our class:
https://piazza.com/nyu/fall2018/cs9223/home
- In class quizzes and participation in Piazza will count towards 10% of your grade.
So Participate!!

Course schedule (*Tentative and subject to change*)

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sept 6</td>
<td>Introduction to Distributed Systems and : MapReduce</td>
<td>MapReduce</td>
</tr>
<tr>
<td>2</td>
<td>Sept 13</td>
<td>Go Intro. Web Architectures and Cloud Computing</td>
<td>Online go Tutorial Guest Speaker: Andy Cheung (SignalSciences)</td>
</tr>
<tr>
<td>3</td>
<td>Sept 20</td>
<td>Distributed Communication: Sockets and RPC</td>
<td>Online go Tutorial</td>
</tr>
<tr>
<td>4</td>
<td>Sept 27</td>
<td>Time, Clocks and Distributed Systems</td>
<td>Logical Clocks</td>
</tr>
<tr>
<td>5</td>
<td>Oct 4</td>
<td>Fault Tolerance: Primary/backup Replication</td>
<td>GFS VR Revisited</td>
</tr>
<tr>
<td>6</td>
<td>Oct 11</td>
<td>CAP Theorem Consistency Models</td>
<td>Perspectives on the CAP Theorem</td>
</tr>
<tr>
<td>7</td>
<td>Oct 18</td>
<td>Consensus: Paxos</td>
<td>Paxos Made Simple</td>
</tr>
<tr>
<td>8</td>
<td>Oct 25</td>
<td>Consensus: Raft</td>
<td>Raft</td>
</tr>
<tr>
<td>9</td>
<td>Nov 1</td>
<td>Consensus in Practice: Chubby, and BigTable</td>
<td>Special Guests from Google to talk about: Chubby, BigTable</td>
</tr>
<tr>
<td>10</td>
<td>Nov 8</td>
<td>Relaxed Consistency</td>
<td>Dynamo Bayou</td>
</tr>
<tr>
<td>11</td>
<td>Nov 15</td>
<td>Distributed Transactions</td>
<td>Spanner</td>
</tr>
<tr>
<td></td>
<td>Nov 22</td>
<td>Thanksgiving</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Nov 29</td>
<td>Peer to peer: Bitcoin</td>
<td>Bitcoin Summary</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>-----------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>13</td>
<td>Dec 6</td>
<td>Project Presentations</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Dec 13</td>
<td>Project Presentations</td>
<td></td>
</tr>
</tbody>
</table>

**Textbook:** None.

**Useful Reference Books:**
- Designing Data Intensive Applications (Kleppmann)
- Distributed Systems (Tanenbaum and Steen)
- Distributed Systems (Colouris, et al)
- Distributed Systems for Fun and Profit

**Reading**
We will be reading published papers and discussing them in class. The papers will also be covered in the exams.

**Acknowledgements**
This course builds upon several existing distributed systems courses from other universities:
- MIT Distributed Systems course by Robert Morris.
- NYU Distributed Systems course by Jinyang Li
- Columbia Distributed Systems course by Roxana Geambasu

**Grading**
Grading will be based on the following weights.
- 30 % Assignments
- 30 % Quizzes
- 40 % Project

**Assignments:**
There will be several assignments involving programming projects. Every assignment has to be done individually by each student. Discussions between students and help in using the programming environment are permitted and encouraged, however no code or solutions may be copied. The assignments must be in Completed in Go. There is no late policy for assignments; any absence notifications should be communicated via [http://engineering.nyu.edu/life/student-affairs/advocacy-privacy-and-compliance](http://engineering.nyu.edu/life/student-affairs/advocacy-privacy-and-compliance).
Project:
The goal of this project is for the student to gain experience in implementing a Distributed System different to what we see in class. The project will be performed in teams, details to be announced. Students will be responsible for selecting a domain area problem (some example problems will be suggested). The grade for the project will be divided between: problem and background research, proposed solutions (short presentation), and implementation. More details will be provided in class.

Quizzes:
There will be a series of very short quizzes at the beginning of selected classes covering material to date. The lowest quiz mark will be dropped. Quiz marks will not be excused.

Other Grading notes:

Please take the following into consideration during and after the semester and save yourself one or many emails.

1) **I must grade every student EXACTLY the same way.** To this end, I cannot give you special consideration as a result of your academic status (probation or otherwise), scholarships, work status, family situation, visa status, race, color, creed, religious beliefs, past alien abductions, current moon cycle, location of the sun in the sky or anything other than your academic performance. **Your grade must be based on your academic performance in my class.**

2) **I cannot change your grade simply because you ask me to.** Your grade is calculated based on your performance from the first day of class to moment you turn in the final exam.

3) **I will not give you additional work.** Please remember that I must treat all students the same, so if I give you additional work, I would have to give it to the entire class. This is unfair to the students who complete their work on time.

4) **Your grade is a measure of your performance in my class.** If you receive an “F” it is because you have demonstrated that you do not understand the material in the course; if you receive an “A” it is because you have demonstrated that you fully understand the material covered in the course. Other grades are assigned accordingly.

Moses Center Statement of Disability
If you are a student with a disability who is requesting accommodations, please contact New York University’s Moses Center for Students with Disabilities at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations. Information about the Moses Center can be found at www.nyu.edu/csd. The Moses Center is located at 726 Broadway on the 2nd floor.