San José State University
College of Science/Computer Science Department
CS 249, Distributed Systems, 01, Fall, 2021

Course and Contact Information

Instructor(s): Ben Reed
Office Location: MH 213
Telephone: (408) 924-517
Email: ben.reed@sjsu.edu
Office Hours:

Monday 3:30-5:30 in office
Tuesday: 10:00-11:00 in office
Wednesday: 3:30-5:30 in office
Thursday: 10:00-12:00 via zoom https://sjsu.zoom.us/j/4077267356

Class Days/Time: MW 2:15-3:30
Classroom: MH 223
Prerequisites: CS 149

Course Description

Current issues in operating systems, including multiprocessor systems and distributed computing, networks, security and performance. Case studies of current operating systems.

Course Format

This is an in-person course.

Course Learning Outcomes (CLO)

Upon successful completion of this course, students will be able to:

1. define the terminology and common ideas of distributed computing.
2. explain fundamental ideas of distributed computing.
3. demonstrate experience with distributed computing principles in modern applications.
4. explain the challenges present in a distributed environment.
5. explain the differences and trade-offs between various solutions for distributed systems problems.

Required Texts/Readings

Textbook

Designing Data-Intensive Applications by Martin Kleppmann publish by O'Reilly. Available online at the SJSU library.

Other Readings

The following papers will be studied as part of the class:

1. ZooKeeper: https://www.usenix.org/legacy/events/atc10/tech/full_papers/Hunt.pdf
2. Lamport clock: https://amturing.acm.org/p558-lamport.pdf
3. Vector clocks: https://www.vs.inf.ethz.ch/publ/papers/VirtTimeGlobStates.pdf

4. FLP: https://groups.csail.mit.edu/tds/papers/Lynch/jacm85.pdf
5. ADB: https://groups.csail.mit.edu/tds/papers/Attiya/TM-423.pdf
6. Paxos: https://lamport.azurewebsites.net/pubs/paxos-simple.pdf
7. BFT: https://www.cs.princeton.edu/courses/archive/fall17/cos418/papers/bft.pdf
8. order & agreement: http://www.cs.cornell.edu/lorenzo/papers/sosp03.pdf
9. RSM: https://www.cs.cornell.edu/fbs/publications/SMSurvey.pdf
10. primary/backup: https://www.cs.cornell.edu/fbs/publications/DSbook.c8.pdf
11. dynamo: https://www.allthingsdistributed.com/files/amazon-dynamo-sosp2007.pdf
12. chain replication: https://www.cs.cornell.edu/home/rvr/papers/OSDI04.pdf
13. wait free synchronization: http://dx.doi.org/10.1145/114005.102808
14. concurrent objects: https://lamport.azurewebsites.net/pubs/interprocess.pdf
15. witnesses: http://www2.cs.uh.edu/~paris/MYPAPERS/Icdcs86.pdf
16. GFS: https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf
17. big table: https://static.googleusercontent.com/media/research.google.com/en//archive/bigtable-osdi06.pdf
18. map/reduce: https://static.googleusercontent.com/media/research.google.com/en//archive/mapreduce-osdi04.pdf
19. raft: https://web.stanford.edu/~ouster/cgi-bin/papers/raft-atc14
20. DHT: https://www.cs.princeton.edu/courses/archive/fall17/cos418/papers/chord.pdf
21. ceph: https://www.ssrc.ucsc.edu/Papers/weil-osdi06.pdf

Course Requirements and Assignments

Discussion participation (5%)

Homework will be given, but will not be graded. It is intended for self-evaluation and will be the basis for future exams. I encourage students to work on homework in groups and discuss possible solutions together. We will take time at the beginning of each class to discuss any difficulties students have completing the homework. There will be discussion assignments associated with each set of homework questions to encourage participation and preparation before the in-class discussions.

Ethical assignment (10%)

Along with technical questions in the homework, we will also discuss ethical issues related to distributed systems. We want you to understand that along with technical choices come moral implications, and we want to be able to identify and reason about them. There will be 2 written (1 page) assignments to discuss contemporary ethical issues.

Programming assignments (20%)

We will be doing individual programming assignments. At the end of many of the assignments we will be doing a bakeoff in class to see if our implementations can interoperate. THERE ARE NO LATE ASSIGNMENTS. THE BAKE-OFF IS DONE IN CLASS. Individual programming assignments are not group projects. If students get help on assignments, even to resolve a stupid problem, it must be documented in the code with the name of the person rendering the help and a brief description of the help provided. Extensive help on a project will result in a reduced grade. Failure to document help, or any other forms of cheating will result in a failing grade on the assignment at a minimum and may result in failure of the course. All incidents will be reported to the Office of Student Conduct & Ethical Development. Even in open source, you cannot copy code from one open source project to another without attribution. Sharing solutions with other students, even if it is indirectly through public source repositories, falls under "aiding and abetting".

The University Policy S16-9, Course Syllabi (http://www.sjsu.edu/senate/docs/S16-9.pdf) requires the

following language to be included in the syllabus:
"Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the syllabus

Paper video (20%)

As a graduate student, you should be able to learn from original research papers. You will need this skill when you do your thesis. We will be covering important research papers in distributed systems. Each student will be assigned a research paper that they will be making a short video (more than 10 mins but less than 20 mins) about. The video will identify and explain the important contribution of the paper and show how the paper has impact on current systems. The due date for the video depends on the paper you are assigned.

Final Examination and exams (45% total)

All exams, including the final, will be taken in class using lockdown browser. This course will have a cumulative final exam given during exam week.

There will be three in-class exams given in the semester (the last being the final exam :)). The second exam will have two questions derived from the previous exam, and the final exam will have two questions derived from the first exam and two questions derived from the second exam.

I do not grade on a curve. The exams and projects measure what you are expected to have learned. There aren't many opportunities for extra credit, but there are bonus questions on exams.

Grading Information

Grades will be calculated based on the individual project grades, the two mid-semester exams, the final, discussion participation, the video explanation of a paper, and the ethic assignment. Thus, the grade distribution is

programming assignment	20%
Midterm 1	15%
Midterm 2	15%
Final	15%
Video explaining paper	20%
Ethics assignment	10%
Canvas discussion participation	5%

I do not curve grades. I do round the reported grade up. I don't feel that it is fair that only students who beg for their grade to be rounded up to get it, so I automatically round-up for everyone. So, if you get an 89.1, that will be rounded to an A-. However, do not expect a 68.9 to be rounded to a C-, even if you can go back and argue for some points to get it to a 69.1. The whole point of rounding is to get people that got close enough the grade they deserve. Getting a 68.9 is not close to a C-.

Grade	*Percentage*
A plus	*96 to 100%*
A	*93 to 95%*

Grade	*Percentage*
A minus	*90 to 92%*
B plus	*86 to 89 %*
B	*83 to 85%*
B minus	*80 to 82%*
C plus	*76 to 79%*
C	*73 to 75%*
C minus	*70 to 72%*
D plus	*66 to 69%*
D	*63 to 65%*
D minus	*60 to 62%*

Classroom Protocol

This is your class. Please ask questions. Please come prepared. Do not engage in activity that may distract other students.

I do not take attendance except for the first two classes. Students not attending either of the first two classes will be dropped to make room for students on the waiting list. Attempting to get marked as present (by having someone else attend in your place or using technological deceptions) will be considered academic dishonesty and at a minimum will result in you getting dropped from the course.

University Policies

Per University Policy S16-9 ,relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on Syllabus Information web page (https://www.sjsu.edu/curriculum/courses/syllabus-info.php). Make sure to visit this page to review and be aware of these university policies and resources.

Course Schedule

Week/Lesson /Module *(*Delete if not applicable*)*	**Date**	**Topics, Readings, Assignments, Deadlines** ***(If appropriate, add extra column(s) to meet your needs.)***
1	8/23	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch01.html In class assignment #1 due
1	8/25	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch01.html In class assignment #2 due
2	8/30	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch05.html ADB
2	9/1	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch05.html

Week/Lesson /Module *(*Delete if not applicable*)*	**Date**	**Topics, Readings, Assignments, Deadlines** ***(If appropriate, add extra column(s) to meet your needs.)***
3	9/8	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch05.html
4	9/13	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch06.html
4	9/15	ADB bakeoff
5	9/20	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch06.html
5	9/22	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch06.html
6	9/27	Exam 1
6	9/29	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch07.html GFS & ZooKeeper
7	10/4	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch07.html
7	10/6	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch07.html
8	10/11	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch07.html
8	10/13	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch08.html ZooKeeper mini-bake off
9	10/18	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch08.html
9	10/20	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch08.html
10	10/25	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch08.html
10	10/27	SGFS Bake off
11	11/1	Exam 2

Week/Lesson /Module *(*Delete if not applicable*)*	**Date**	**Topics, Readings, Assignments, Deadlines** ***(If appropriate, add extra column(s) to meet your needs.)***
11	11/3	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch09.html
12	11/8	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch09.html
12	11/10	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch09.html
13	11/15	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch09.html Map/Reduce
13	11/17	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch10.html
14	11/22	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch10.html
15	11/29	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch10.html
15	12/1	https://learning.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/ch10.html
16	12/6	Simple Map Bake off
Final Exam	12/9 @ 12:15	