CS-UY 3254, Introduction to Parallel and Distributed Systems
Syllabus — Fall 2020

Jeff Epstein

Basic information

Course description   This course is an introduction to the practical and theoretical aspects of concurrent, parallel, and distributed systems. Students will learn about the algorithmic underpinnings and engineering concerns arising in building highly-reliable systems, such as modern cloud services. Topics will include a brief introduction to networking, approaches to distributed consistency and storage, handling of big data, and distributed peer-to-peer networks such as BitTorrent and Bitcoin.

Logistics   Meeting times are as follows:

Lecture	4:00pm - 5:50pm MW	online and JCBS 774

Contacting me   The best way to get in touch with me is by email. My address is jeff.epstein@nyu.edu. I encourage you to contact me with any question, at any time.

Our teaching assistants are:

Name	email	office hours
Nitin Patil	nap493@nyu.edu	Fri 2:30pm-3:30pm
Yash Shah	yjs310@nyu.edu	Fri 12:30-1:30pm

When contacting a member of the teaching staff by email, please use the course code ("CS-UY 3254"), so that we know the context.

Prerequisites   This course assumes a background in data structures and algorithms equivalent to CS-UY 2134. Students should be very comfortable reading, writing, and debugging programs in an imperative programming language.

Familiarity with network programming will be helpful but is not required.

Textbook   This textbook is optional. It may provide useful background information but is not required for this course. It is Distributed Systems Concepts and Design, Fifth Edition by George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair, ISBN 0-13-214301-1

Topics   A tentative list of topics to be covered this semester:

- Moore's law

- CAP theorem

- Network time protocol

- Go language

- Synchronization and concurrency control

- Logical clocks

- Distributed file systems

- Locking algorithms

- Consensus algorithms

- Broadcast algorithms

- Elixir language and actor-model concurrency

- Blockchains

- Distributed hash tables

- Consistency

- Replication

Coursework

Grading rubric   Your final grade for this class will be calculated as follows:

- Programming assignments: 90%

- Participation: 10%

Assignments   Assignments will be distributed on NYU Classes and should be submitted on Gradescope. Only work submitted on Gradescope will be accepted.

Credit is given for finding correct solutions to the problems posed in the programming assignments. Partial credit may be given for partially correct solutions. The cumulative homework represents a significant component of your final grade, and therefore students should complete the work carefully. I encourage you to start the homework early, as it may be more time-consuming than you expect.

Participation   Lectures are not intended to be an exercise in passive listening. Students should be engaged with the material, prepared to ask questions, and ready to respond to learning activities. With active learning, you'll feel more engaged, you'll remember the material better, and you'll be able to shape the direction of the discussion.

Of course, it's more difficult to be active during a lecture when most of us are at home, staring at a screen. To help motivate student engagement, we'll be using a student response system (SRS), Poll Everywhere. This tool allows the instructor to present polls to students during lecture, giving each student a chance to respond. Students can respond via their computer or smart phone. Responses to the poll can then form the basis of in-class discussions.

Participation with SRS contributes to your grade. In-class polls will typically be evaluated on the basis of good-faith effort: that is, if you demonstrate a sincere attempt to answer correctly, you will get credit, even if your answer is wrong. Some polls may be evaluated as "pop quizzes," i.e. your credit depends on getting the right answer, in which case you will be notified in advance.

Getting help   This is a challenging class. If you are struggling with the material, it is your responsibility to take action to improve your understanding. There are a variety of ways to get help:

- Ask questions in class. A classroom discussion is the most direct way to understand the material thoroughly.

- Meet with a TA. The teaching assistants offer regular office hours, and can meet with you one-on-one or in small groups.

- Meet with the professor. The professor has regular office hours, as well as office hours by appointment.

- Send an email. You can contact your professor or a TA by email to ask questions about the homework or course material.

- Read the book. You may prefer the presentation of material in a textual format.

Any case of any doubts or other situations, please discuss the situation with your professor.

Tools

Gradescope   You will need a Gradescope account for this class. You should receive an automatic email notifying you of your registration on Gradescope. If not, please create your Gradescope account according to the following guidelines:

- For your *name*, please give the customary name you use to mark your work. Give both first and last name.

- For your *student ID*, give your NYU NetID. Your NetID is the identifier based on your initials, followed by a number, such as `ab123`.

- For your *email address*, give your NYU email address. This is normally the address based on your NetID (for example, `ab123@nyu.edu`), unless you've changed it.

You may re-submit your work on Gradescope as many times as you like until the deadline. Submit early to ensure you don't miss the deadline. After the deadline, you will not be able to re-submit. Once you've submitted, please verify that you uploaded the correct file, as only those files that you've uploaded will be graded: it is your responsibility to check that the uploaded materials correspond to the work that you want evaluated. Graders have access only to your last submission. If you are submitting work on behalf of a group, make sure that you use the group submission option to include all members of your group. If you are using a Mac, please be aware that the operating system creates invisible files named `__MACOSX` and `.DS_Store`; please do not upload these files to Gradescope. Uploading them makes grading more difficult and may negatively impact your grade.

Your submitted homework will be evaluated and given a grade. The evaluation includes detailed feedback from the grader. Please be sure to read the comments provided to you in the evaluation, as they will help you understand your mistakes. Click on the question number in Gradescope to show the rubric items that were applied against that question, and view your submitted file in Gradescope to see additional comments provided by the grader.

If you have a question about the grading of an assignment, contact a teaching assistant. If you feel that your assignment has been graded incorrectly, submit a regrade request on Gradescope. Be sure to submit your regrade request promptly, as regrade requests are typically accepted only within one week from the time that grades are published.

Programming environment A substantial component of this class is programming work to be completed independently. You need to make sure that you are able to use the software tools necessary to accomplish this.

You should develop your programming assignments using tools as similar as possible to the environment under which they will be evaluated, to ensure that you don't lose points due to a difference in versions or different behaviors under different operating systems.

- I recommend that students use NYU Vital (`https://vital.engineering.nyu.edu/vital/`), a cloud-based virtual machine service. With Vital, you will have access to virtual machines, accessible anywhere, pre-configured with all the languages and tools you'll need for this class.

  Registration instructions for Vital are provided on Classes. The Vital registration code for this class is `8pcjw5xu`.

- You may install Linux into a virtual machine running on your own computer. This has the advantage that you won't need internet connectivity to use it, but on the other hand you are responsible for configuring the machine and installing the necessary software. Fortunately, most of the software you'll need can be easily installed through the repositories.

  Instructions for installing and configuring VirtualBox are provided on Classes.

- Windows 10 supports the Windows Subsystem for Linux, which allows you to run a complete Ubuntu installation without a virtual machine.

- If you have Linux installed on your own computer, you may choose to develop software without the use of a virtual machine. Please make sure that you're using the right version of any necessary software.

Poll Everywhere   We use Poll Everywhere as a student response system (SRS).

- To use it on your mobile device, download the app at `https://www.polleverywhere.com/mobile`.

- To use it from your web browser, go to `https://pollev.com/`.

In either case, you will need to enter my "presenter ID," which is `jeffepsteinnyu`. When prompted, log in to the NYU SSO with your NYU NetID. You may be asked to confirm a registration for the NYU College of Nursing; that's okay, just go with it.

You can consult the Poll Everywhere student guide for help.

Additional policies

Attendance   Let's define some terms we will use in the context of attendance:

- *in-person* — The standard approach to attending classes. You are physically present in the classroom. You can see the professor, and the professor can see you.

- *remote* — Remote attendance means that you watching the professor speak via Zoom. You can still ask questions and participate in the class.

- *asynchronous* — In this case, you are watching a recording of the presentation, rather than a live transmission. You won't be able to participate directly (although you can still ask questions by email).

Another factor to consider this semester is *where* students are located. Some of you are in New York, some elsewhere in the US, and some abroad. Everyone should be able to get a good experience from this course. I don't think it's

reasonable that some students have to stay up until midnight just to attend the lecture. Let's say that students located in one of the North American time zones (UTC-4:00 to UTC-8:00) are *local students* and those outside of those time zones are *non-local students*.

Now that we've defined our terms, let's discuss options for attendance:

- Lectures are hybrid. This means that students may freely choose to participate in person or remotely.

- Attendance at lecture is mandatory. SRS participation represents part of your grade.

- If you are non-local, and if you therefore cannot attend lectures synchronously during normal waking hours, you may be exempt from the attendance requirement. Contact your professor within the first week of class so that your exemption will be noted.

Please note that the choice to attend lecture in person or remotely is entirely yours. You don't need to request permission to attend remotely. If you feel sick, stay home. If you are worried about getting sick, stay home. I trust you to make the decision that is best for you.

Class attendance (either in-person or remote) is mandatory, unless you request an exemption. If this is a problem for you, please let me know.

Class expectations Health is of paramount concern. You have the option to attend lecture in person or remotely. Please make a healthy, responsible choice.

You are expected to take notes during lecture, based on our discussions and lectures. Please be prepared to do so.

Active participation in class discussions is strongly encouraged. This is the best time for students to ask questions or clarify any confusing concepts. In addition, you are responsible for any material covered in class, even if it isn't in the textbook. If you miss a class, you should contact a classmate to recover the missed content and assignments.

To participate in person, just raise your hand. To participate via Zoom, use the "raise hand" icon available on the Participants pane.

While attending a lecture remotely, please make sure that your microphone is muted when you're not speaking. You are not required to enable your camera, but it is appreciated, as it makes your professor feel that he is talking to real people. When attending in person, please disable or silence any device that may audibly disrupt the class.

While attending a lecture in person, you must obey the NYU COVID-19 policy:

- You must wear a face covering over your nose and mouth at all times.

- You must perform social distancing.

- You must select a seat in accordance with the Seating Assignments page on NYU Classes. On the first day of class, register your seat. Thereafter you should always use the same seat.

If you have registered for this course as a "remote student," you may not attend in person. Contact your professor if you have a problem.

Please note that lectures will be recorded and can be reviewed later on the Zoom page on NYU Classes.

Please bring your phone or computer to lecture, so that you can participate in Poll Everywhere. Make sure that you are logged in to the app during lecture.

During class time, please do not use your phone or computer for activities unrelated to class. Please do not eat or drink during class.

Students are expected to arrive to every class promptly. You should be actively engaged in the learning process during the duration of the class time.

Communication I may occasionally use email to make class announcements. It is your responsibility to check your NYU email account regularly.

Assignments will typically be posted on NYU Classes. It is your responsibility to check NYU Classes for assignments, and to submit your work in a timely manner.

Office hours All office hours are held on Zoom.

- TA office hours are scheduled at a regular time each week. You can find the schedule on the Zoom tab in NYU Classes.

- Professor office hours are made by appointment. Please email your professor to set up a time.

Late policy and make-up policy All homework should be submitted before the due date. Homework may be submitted without penalty up to 24 hours after the due date. After that time, absolutely no homework will be accepted.

In general, I do not accept late assignments (except as indicated above) and do not offer substitute times for exams.

Consideration may be given in case of a documented special circumstance, such as a medical condition or a family emergency, in accordance with university policy. If that case applies to you and you need additional time to complete an assignment or need a rescheduled exam, your first point of contact should be Tandon's Student Affairs office. Please provide to them written documentation of your circumstance, and they will in turn make a recommendation to me about how to proceed.

Academic integrity This class is bound by the Student Code of Conduct. We aim to ensure that your grade reflects your understanding of the material, and therefore we require that the work that you submit for grading be a result of

your own effort. This applies to all your work for this class, including homework, lab work, and exams.

Why do we take academic integrity seriously?

- If the work that you submit does not represent your own effort and understanding, you are doing yourself a grave disservice, by cheating yourself out of an opportunity to learn. Your grade has value only if it accurately represents what you've accomplished to get it. In the end, your diploma (should you achieve it) is worthless unless it is supported by personal growth and knowledge.

- Furthermore, by submitting work that does not represent your own effort and understanding, you are behaving dishonestly to the class's teaching staff. The professor and the TAs who work hard to create assignments, provide instruction, and evaluate students' work, do so in order to provide learning. Students who submit others' work as their own are therefore wasting the time of the teaching staff.

- Finally, misrepresenting yourself in your submitted work does real damage to your classmates. If you achieve a grade through dishonest work, students who complete the assignments honestly are at a disadvantage.

We realize that it's not always clear what activities are acceptable in the context of working on your assignments. This is especially true in Computer Science, where the web can provide excellent learning resources (good!) as well as solutions to assignments (bad!). To help you understand the boundaries of what is and is not acceptable, we provide the following guidelines, which are based on the Stanford CS107 collaboration rules. We classify activities into three categories: those that are always acceptable, those that are acceptable but require a citation, and those that are not acceptable.

- These activities are always acceptable:

    - **Discussion of general course topics.**

      You may freely discuss the course material, outside of the context of a specific assignment. This means you can ask and answer questions about the rules of the programming languages and tools that we use, as well as theoretical matters that were covered in lecture or in the textbook. However, the discussion may not refer to solutions for a specific assignment.

      For example: "What does the `elem` function do? What is the syntax for a function definition?"

    - **Discussion of assignment requirements.**

      You may freely discuss the requirements for a specific assignment. However, the discussion may not refer to solutions.

      For example: "Are we allowed to use `reverse`? Is performance important?"

– **Use of public resources for background information.**

Web sites, books, and other public resources provide lots of great information. You are expected and encouraged to refer to documentation for the languages and tools that we use, as long as they don't refer to the specific task for your assignment.

For example, the official Python documentation is okay, but "How to write Space Invaders in Python" is not okay (assuming the assignment is Space Invaders).

– **Discussion with teaching staff.**

This is the best way to get help! You can always turn to your professor and TAs with any questions. You can ask questions about any assignment during class, during office hours, or by email.

- These activities are acceptable but require a citation:

    – **Re-using your own work.**

    If you are submitting your own work, including code, that you originally completed before you took this class, you must provide a note stating this fact, including when you originally did the work and under what conditions.

    If you want to submit work that you have previously completed for an earlier instance of this class, or for any other class, you should discuss the matter with your professor before proceeding.

    – **Discussion of testing.**

    If you aren't sure if your solution is correct, it can be helpful to develop a strategy to test it. You may propose and discuss tests collaboratively, as long as you provide a note describing the incident. For example, you may remind a friend that a particular function needs to work with empty lists. However, you may not see each other's code or discuss it in detail.

    – **Discussion of debugging.**

    You may discuss particular error messages and propose possible solutions, as long as you provide a note describing the incident. For example, if a friend says they are getting a "segmentation fault," you may suggest that they might be dereferencing a null pointer. However, you may not see each other's code or discuss it in detail.

What is a citation? If you participate in one of the activities in this section, you must state so in the assignment that it pertains to. If the assignment is a coding exercise, you must write a comment in your code describing the activity. If the assignment is non-code work, you must write a preface describing the activity.

Every citation must describe (a) the specific source of the information that you used (a person, a book, or a web site), and (b) how you used

that information, and to what extent it influenced your work. Failing to provide a citation when participating in one of these activities renders the activity unacceptable, and therefore a violation of academic integrity.

- These activities are never acceptable, and always constitute a violation of academic integrity:

  - **Copying code or answers from others.**

    The work you submit must be written entirely by you. This means not only that you physically wrote it, but that the ideas behind it are yours. You may not copy code, read code, or consult code from another person, directly or indirectly. If you are in an environment where you are able to see another person's solution, it is your responsibility to take yourself out of that environment and actively prevent the possibility of being influenced by their work.

  - **Collaborative development.**

    You may not work together with another person, or "walk through" your code with another person. Even if you are not directly sharing code, any detailed sharing of design, pseudocode, structure, or substance, even in partial or incomplete form, is not acceptable. Solving bugs collaboratively or discussing particular code around debugging issues is prohibited. You must work on your assignments alone.

  - **Using public resources for assignment-specific code.**

    Web sites, books, and other public resources can provide valuable information, but any resource that refers to the task that you are to complete for the assignment, or to a task that is substantially similar to it, is not allowed. The design and implementation of your solution may not be influenced by an external resource. Taking code, in any quantity, from a web site is prohibited. If you find yourself reading a web site that addresses the design or implementation of a program similar to the task at hand, you are required to stop using it.

    The boundary between "general background resources" and "assignment-specific resources" may be fuzzy. If you have any doubt that a resource you are using is acceptable, ask your professor first.

  - **Allowing others to use your code or answers.**

    At no point should anyone else have access to your work. This means that you may not keep your work on a shared or public computer. You may not post your work on a public web site (including Github and similar sites). You may not show your work to others, even if they promise not to use it in their own work. Your work should remain private, and if it any time it isn't, then you can have contributed to academic dishonesty.

    You may use Github (and similar sites) as version control storage, but it is your responsibility to ensure that your repository is private.

If your code is publicly visible at any time, it constitutes a violation of this policy.

The prohibition on sharing your work remains in effect even after you have submitted your work, and after the course has ended. Students in future semesters should not have access to your work.

If you are in doubt about whether a particular activity is acceptable, it's best to consult with your professor *before* doing it.

If you are struggling in the class, please discuss your situation with your professor. Plagiarism is not a solution.

Violations of academic integrity will be dealt with strictly. Typically, this means receiving a failing grade for the course and a referral to Student Affairs.

Religious Holiday Observance   The instructor supports NYU's policy with regard to absence due to religious holidays. If you have a conflict due to observance of a religious holiday, please alert the instructor in advance so that a plan for makeup can be arranged.

Academic Accommodations   New York University provides reasonable accommodations to qualified students who disclose their disability to the Moses Center. Reasonable accommodations are adjustments to policy, practice, and programs that level the playing field for students with disabilities and provide equal access to NYU's programs and activities. Accommodations and other related services are determined on a case-by-case basis, taking into consideration each student's disability-related needs and NYU program requirements.

If you are 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 `http://www.nyu.edu/csd`. The Moses Center is located at 726 Broadway on the 2nd floor.

If you are authorized to receive accommodation for an exam, it is your responsibility to schedule the exam with the Moses Center in advance.

ABET learning goals

1. an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics

7. an ability to acquire and apply new knowledge as needed, using appropriate learning strategies