CS333: Storage Systems

Syllabus for Spring 2021

General Info

Instructor:	Bill Jannen
Email:	jannen@cs.williams.edu
Course Web Page:	`http://www.cs.williams.edu/~jannen/teaching/s21/cs333`
Department Remote Resources:	`http://www.cs.williams.edu/system`

Texts

We will be using the following textbook in this course:

- *Operating Systems: Three Easy Pieces* (version 1.0) by Remzi H. Arpaci-Dusseau and Andrea C. Arpaci-Dusseau.

This book is available for free online, and PDF links to all readings are provided on the course webpage. If you prefer reading from a physical book, the campus bookstore has new and used copies, and the book's webpage includes links for ordering low-cost versions as well.

Since CSCI 237 is a prerequisite for this course, course materials assume that you are proficient in writing and debugging C code. Therefore, a good C reference is highly recommended. Many copies of the following textbook are on reserve at the Schow Science Library:

- *The C Programming Language, 2nd edition.* by Brian Kernighan and Dennis Ritchie.

Additional readings will be assigned from various sources, including conference proceedings, research journals, magazines, and textbook excerpts. All required readings will be digitally accessible from the course website, although some may only be accessible while connected to the Williams network. If you are off-campus, you can access these readings using the library's proxy server or the college's VPN.

Course Description & Objectives

This course will examine topics in the design, implementation, and evaluation of storage systems. Topics include the memory hierarchy; ways that data is organized (both logically and physically); hardware characteristics and the ways that storage hardware influences storage software design; data structures and performance models; and system measurement/evaluation. The course emphasis will be on identifying and evaluating design trade-offs.

Upon the completion of this course, students will:

- be able to describe the performance profiles and behaviors of current and emerging storage technologies;
- when presented with a new storage technology, be able to describe how that technology will fit into or modify the "storage landscape";
- be able to critically read storage research texts and understand the texts' content and arguments;
- be able to describe and/or use common storage interfaces at the application, OS, and device layers;
- be able to read, write, document, debug, and test system code in C; and
- be able to analyze the performance of external memory algorithms and data structures.

Course Structure

Course meetings. There will be two synchronous conference meetings each week. To prepare for conference, you should complete all *required* readings and watch all *required* lecture videos. If there are lecture videos associated with a conference, the videos will be posted at least 24 hours before the conference meeting, and an announcement will be made on the course's GLOW page. For many units, *optional* readings will be posted to give more variety or depth. These readings are clearly marked and are truly optional.

During each conference, you will have a chance to ask questions about the course material. The remaining time in conference may be used for a combination of:

- group discussions,
- live coding exercises,
- simulation/demos designed to explore the material, and
- work on your programming projects

If you do not have access to a laptop or other computing device that you can use during conference, please contact me so that we can figure out a way to procure one for you to use.

Readings. There will be one or more required readings for each unit. In the second half of the semester, many of the readings are research papers; thus, there may be some assumed knowledge and "jargon" to sift through before we can understand a reading's core content. When this is the case, introductory recordings will be designed to help with that.

Labs. "Lab assignments" will be an important part of this course. Lab work is a way for us to explore select concepts at a deeper level. When completing labs, you should focus on completeness, correctness, and creativity. Throughout the course, we will build practices that will help us evaluate our work on these axes, with a focus on modular design and incremental testing.

Collaboration. Collaboration is important, and this course will permit a variety of collaboration styles. Each assignment specification will explain the types of collaboration that are permitted for that assignment.

Midterm Exam. There will be one midterm exam, and it will be a 24-hour open-book exam administered on GLOW. You will have a four-day window, starting on Thursday, 04/18, during which you can choose any contiguous 24-hour window that you would like to work on the exam.

Final Project. At the end of the course, you will either design your own project or select/customize a project proposal from a sample set that I will provide. In the first phase of your project, you will be required to submit a formal proposal that includes success criteria, methods for evaluation, and project deliverables. We will agree upon this proposal together. You will then complete your project, and submit your final write-up/artifact in place of a final exam.

Grading. This course will be ungraded, although the word "ungraded" is likely not the term I would have selected if I were the one to name the system. At a high level, you will propose your own letter grade, basing it on reflection and self-assessment. I will give you feedback on your submitted work, including your self-assessments, and when submitting your final letter grade to the registrar, I will strive to honor your assessment of your performance in the course. However, I reserve the right to adjust grades that I feel deviate from my own evaluation of your performance. If I feel that an adjustment is necessary, I will schedule a meeting with you to discuss it.

I suggest the following evaluation criteria as a starting point for your self-assessment of your performance:

Preparation and participation:	15%
Lab & written assignments:	45%
Midterm Exam:	20%
Final Project:	20%

You will be asked to submit mid-semester and end-of-term self assessments. In addition, you will be asked to submit small evaluations/reflections for each of your lab assignments; the specifics of these evaluations will be included in the assignments themselves. The purpose of these assignment-specific evaluations is to get you to think critically about what it means to write high-quality systems software—we should be able to design and write tests that verify that our code is complete (i.e., it satisfies all stated requirements); write documentation that makes our code readable and maintainable (i.e., interfaces are well described with clearly chosen, meaningful parameters); describe a system's performance in concrete terms (i.e., expected run-times based on appropriate models for computation and I/O); and describe our system's design/architecture in a way that others can reason about without manually inspecting your code.

Honor Code

Collaboration is strongly encouraged within the guidelines stated above. However, violations of course collaboration rules will be considered a violation of the honor code and will be forwarded to the honor committee. If in doubt of what is appropriate, do not hesitate to ask your instructor. For a full description of the Computer Science Honor Code, please see
`https://csci.williams.edu/the-cs-honor-code-and-computer-usage-policy/`.

Workload.

At Williams, we operate under the course unit system (rather than the credit hour system). You should expect to spend (on average) at least 13 hours per week on the academic and creative work related to class. This includes time spent meeting as a class and working on assignments. The Office of the Registrar explains the relationship of course units to credit hours in greater detail.

If you find yourself spending significantly more or significantly less time on this course, please let me know.

Course Calendar and Schedule.

The course calendar and schedule can be found on the course webpage, including links to the readings that should be completed prior to each course meeting (all required textbook materials are available as free PDF documents). The course webpage will be regularly updated after each course meeting to make all lecture materials and examples available for review. Links to materials will also be posted to GLOW.

Computing Resources.

You will need an Internet-connected digital device that can either run Linux, a VMware virtualization product (educational licenses can be provided by the department), or use `ssh` to connect to a machine in the department's computer labs.

I recognize that some students are unable to afford the cost of purchasing digital devices and that other students rely on older, more problem-prone devices that frequently break down or become unusable. I also recognize that those technology problems can be a significant source of stress for students. If you do not have access to reliable Internet or a reliable, Internet-connected digital device for programming, I encourage you to contact the Dean of Students Office to request assistance. If you contact me as well, I can help you write a message to the Dean of Students and also work with you to develop a plan for completing the work in this course.

If you are experiencing problems with your device or Internet access, I encourage you to contact the Office of Information Technology for support.

Help!!!

Help. We all need it. There are many resources available when *you* need it. You are encouraged to discuss any questions, concerns, difficulties, or thoughts about the course with your instructor. If you find yourself facing

challenges beyond the typical, we encourage you to reach out. Talk to your instructor, a friendly face from the Dean's Office, or some of the many professionals across campus who stand ready to help, including:

- The Peer Academic Support Network: For details, go to:

  `https://academic-resources.williams.edu/peer-academic-support/`

- Math & Science Resource Center: Support is available for students grappling with the more quantitative aspects of their coursework. For details, go to:

  `https://academic-resources.williams.edu/peer-academic-support/math-science/`

- Accessible Education and Disability Support Center: Some students with documented disabilities may require accommodations in certain situations. If that's you, take advantage of the options available. Go to

  `https://academic-resources.williams.edu/disabilities/` for details.

- The Health Center: Sometimes your challenges are not course-related. The Health Center provides a range of medical, psychological, and health/wellness services. Go to `https://health.williams.edu` for details.