Electrical & Computer ENGINEERING
Carnegie Mellon

Course Syllabus

18-742: *Computer Architectures and Systems - Spring*, **2021**
http://www.ece.cmu.edu/~ece742/

Instructor: Brandon Lucia
Office Location: CIC 4110
Email Address: blucia@ece.cmu.edu
Course Location: Zoom
Course Time: T/R 1040AM-1200PM

Course Description:
Historically, the performance and efficiency of computers has scaled favorably (according to "Moore's Law") with improvements at the transistor level that followed a steady trend ("Dennard scaling"). Unfortunately, device scaling has hit a limit on performance and power improvements dictated by physical device properties. To continue to make systems capable, fast, energy efficient, programmable, and reliable in this "post-Dennard" era, computer architects must be creative and innovate across the layers of the system stack. This course begins with a recap of conventional, sequential computer architecture concepts.

We will then discuss the end of convention, brought about by the end of Dennard Scaling and Moore's Law, and several trends that these changes precipitated. The first trend is the wholesale shift to parallel computer architectures and systems, covering parallel hardware and software execution models, cache coherence, memory consistency, synchronization, transactional memory, and architecture support for programming, debugging, and failure avoidance. The second trend is the shift to incorporating specialized, heterogeneous components into parallel computer architectures. Topics will include reconfigurable architectures, FPGAs in the datacenter, special purpose accelerators, GPU architectures, and architectural accelerators for machine learning, inference, and deep neural networks. The third trend is the emergence of newly capable hardware and software systems and new models of computation. Topics will include approximate and neuromorphic computing, intermittent computing, emerging non-volatile memory and logic technologies, and analog and asynchronous architectures, and may include future emerging topics.

Number of Units: 12

Pre-requisites: 18-447, 18-613, or permission of instructors

Pre-requisite for: -

Breadth Area: Hardware Systems

Undergraduate Course Designation: -

Undergraduate Course Area: -

Class Schedule:

- **Lecture**: T/R 1040-1200pm

- **Labs/Recitation**: Friday 9:10am-10am

Required Textbook:

Suggested:
Computer Architecture, Fifth Edition: A Quantitative Approach (The "Hennessy and Patterson" book)

Brief List of Topics Covered:
The precise set of topics covered may vary depending on the interest of the students in the class. Possible topics include:

- **Parallel Architecture:** programmability; coherence; consistency; hardware support for programming, debugging, and avoiding failures; parallel memory system design; synchronization and transactional hardware; manycore designs; interconnect networks; dark silicon
- **Heterogeneous Architecture:** FPGA/CPU hybrid systems; ASIC offloading and specialization; big/little heterogeneous multicores; CPU/GPU hybrid systems; heterogeneous memory consistency issues; debugging heterogeneous systems; machine learning and inference acceleration
- **Emerging Architecture Concepts:** intermittent and energy-harvesting computation (e.g., RF-based); neuromorphic hardware and software acceleration; approximate hardware, software, and programming; delay-based computation & race logic; novel memory technologies (e.g. STT, FRAM, PCM); 3D-stacking architectures
- **Any other, related topics of particular interest to the class participants**

Course Blackboard: To access the course blackboard from an Andrew Machine, go to the login page at: http://www.cmu.edu/blackboard. You should check the course blackboard daily for announcements and handouts.

Course Wiki:
Students are encouraged to use the ECE wiki to provide feedback about the course at: http://wiki.ece.cmu.edu/index.php.

Course Mechanics & Assignments:
You will be **reading one to two papers per class**. You will be **writing a short critique of the paper(s)**, highlighting the main contribution, contrasting with other work when

appropriate, and discussing (especially) the merit and (less so) the failings of the work. I appreciate how busy you all are, so **you will have license to skip critiques for up to two class sessions**. The goal is to get you all generating opinions on the papers we read so the in-class discussion is the best it can be. **The class will be as interesting as we all make it through our discussion of the work.** If we all put in the effort, our discussion has the potential to expand the known world of computer architecture knowledge. A student will also be required to **prepare a presentation on the paper(s) for each class session**, essentially leading the discussion for that class with their critique of the work. Paper leaders will be assigned in the first week based on student interest, after we've finalized the reading list. **The bulk of your grade is the project you do for the course.** You are expected to find a topic that is particularly interesting to you in the first few weeks of the course. You and one-ish other person (3 max, 2 preferred) will do background reading on the topic you choose. Then, in the first few weeks (exact dates TBD) of the course, you will write a project proposal. Projects could range in topic very widely from strictly engineering to strictly theory, and in between. The course website will describe the project scope, project goals, and a selection of project suggestions to help you get started. At the end of the course, your group will give a presentation on your project, which will contribute non-trivially to your project grade.

Grading:
This course is not about making grades on an endless churn of problem sets. Instead, your grade will be based on your discussion participation, your term paper, zero or one homework assignments depending on scheduling, and your project. I want the projects to be genuinely interesting and students genuinely invested. I want discussions to be interesting, thoughtful, and informed by the week's readings. **This course will not accept late work, except in extenuating circumstances, at the discretion of course staff.**

50%	Project
25%	Discussion/Presentation
25%	Paper critiques & Other Homeworks and writing

For Course Calendar and Reading List, see course website

Education Objectives (Relationship of Course to Program Outcomes)

(a) an ability to apply knowledge of mathematics, science, and engineering: Projects require expertise in all these areas.

(b) an ability to design and conduct experiments, as well as to analyze and interpret data: students will design, conduct, and analyze experiments on a novel or existing research topic for their project.

(c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability: student projects will be constrained by time in the semester, team compatibility, available resources (design tools, etc), and the research-focus constraints of the course.

(e) an ability to identify, formulate, and solve engineering problems: Students' discussions require them to identify and articulate the contributions of the research of others. Their projects require them to solve engineering problems associated with building and evaluating their projects.

(h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context: Students will study the cutting edge of computer architecture. The field is now striving to continue to scale device performance and efficiency, supporting the economy of the industry and the environmental issues surrounding energy-efficient computation .

(i) a recognition of the need for, and an ability to engage in life-long learning: emphasis on research, critiques, and presentations give students the skills to continue learning and doing research in the future.

(j) a knowledge of contemporary issues: The core content of the course if from contemporary research papers.

(k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice: Students will use modern simulators and research engineering tools to build their projects and run their experiments.

Academic Integrity Policy (http://www.ece.cmu.edu/student/integrity.html**):**

The Department of Electrical and Computer Engineering adheres to the academic integrity policies set forth by Carnegie Mellon University and by the College of Engineering. ECE students should review fully and carefully Carnegie Mellon University's policies regarding Cheating and Plagiarism; Undergraduate Academic Discipline; and Graduate Academic Discipline. ECE graduate student should further review the Penalties for Graduate Student Academic Integrity Violations in CIT outlined in the CIT Policy on Graduate Student Academic Integrity Violations. In addition to the above university and college-level policies, it is ECE's policy that an ECE graduate student may not drop a course in which a disciplinary action is assessed or pending without the course instructor's explicit approval. Further, an ECE course instructor may set his/her own course-specific academic integrity policies that do not conflict with university and college-level policies; course-specific policies should be made available to the students in writing in the first week of class.

This policy applies, in all respects, to this course.

Carnegie Mellon University's Policy on Cheating and Plagiarism
(http://www.cmu.edu/policies/documents/Cheating.html) states the following,

> Students at Carnegie Mellon are engaged in preparation for professional activity of the highest standards. Each profession constrains its members with both ethical responsibilities and disciplinary limits. To assure the validity of the learning experience a university establishes clear standards for student work.
>
> In any presentation, creative, artistic, or research, it is the ethical responsibility of each student to identify the conceptual sources of the work submitted. Failure to do so is dishonest and is the basis for a charge of cheating or plagiarism, which is subject to disciplinary action.
>
> *Cheating* includes but is not necessarily limited to:
> 1. Plagiarism, explained below.
> 2. Submission of work that is not the student's own for papers, assignments or exams.
> 3. Submission or use of falsified data.
> 4. Theft of or unauthorized access to an exam.
> 5. Use of an alternate, stand-in or proxy during an examination.
> 6. Use of unauthorized material including textbooks, notes or computer programs in the preparation of an assignment or during an examination.
> 7. Supplying or communicating in any way unauthorized information to another student for the preparation of an assignment or during an examination.
> 8. Collaboration in the preparation of an assignment. Unless specifically permitted or required by the instructor, collaboration will usually be viewed by the university as cheating. Each student, therefore, is responsible for understanding the policies of the department offering any course as they refer to the amount of help and collaboration permitted in preparation of assignments.
> 9. Submission of the same work for credit in two courses without obtaining the permission of the instructors beforehand.
>
> *Plagiarism* includes, but is not limited to, failure to indicate the source with quotation marks or footnotes where appropriate if any of the following are reproduced in the work submitted by a student:
> 1. A phrase, written or musical.
> 2. A graphic element.
> 3. A proof.
> 4. Specific language.
> 5. An idea derived from the work, published or unpublished, of another person.

This policy applies, in all respects, to 18-742