CS8803: GPU HW/SW
Fall 2025

Delivery: 100% Web-Based, Asynchronous

Dates course will run: August 18 – December 11

Instructor: Prof. Hyesoon Kim

Head TA: Scott Madeira

Office Hours: will be announced at Ed

Course Description:

This course explores the software and hardware aspects of GPU development. Through hands-on projects, you'll gain basic CUDA programming skills, learn optimization techniques, and develop a solid understanding of GPU architecture. Additionally, you'll study compiler principles to comprehend software-related GPU issues and read research papers on hardware challenges. By the end, you'll have enhanced your knowledge of compilers, programming, and computer architecture for modern GPUs.

Course Objectives:

- Develop foundational CUDA programming skills.
- Optimize the performance of CUDA programs.
- Understand GPU architecture performance issues.
- Acquire static code analysis techniques to identify and address GPU performance issues.

Topics:

- GPU Programming
- Parallel Programming Fundamentals
- Compiler Backgrounds
- GPU Architecture
- CUDA program optimizations

Prerequisites:

- C/C++ Programming and Python Programming skills are necessary.
- Prior CUDA programming experience is not mandatory.

- Familiarity with basic computer organization, including topics like instruction sets and pipelining, as covered in undergraduate computer architecture courses, is expected.
- While not mandatory, taking CS6200 would be advantageous, especially for course projects #3, #4 & #5 involving large C++ projects.
- While not mandatory, CS6290 would be beneficial, particularly for course projects #3 & #4, as it includes topics like caching and aids in understanding architecture simulators.
- While not mandatory, CSE 6220 would be beneficial, especially for course projects #1 & #2 involving CUDA programming.
- While not mandatory, CS 6340 would be beneficial, especially for course projects #5 involving program analysis.

Textbooks:

None required. All reading materials will be provided on Canvas.

Grading Breakdown:

Assessment	Type	Weight	Description
Project 1: CUDA Programming	Programming Assignment	5%	Basic CUDA programming concepts. (C++)
Project 2: CUDA Programming II	Programming Assignment	20%	Performance optimization of CUDA programs. (C++)
Project 3: GPU Simulator	Programming Assignment	15%	Add GPU warp scheduler policies in a trace-driven GPU simulator (C++)
Project 4: GPU Simulator	Programming Assignment	15%	Add execution units in a trace-driven GPU simulator (C++)

Project 5: GPU Code Analysis	Programming Assignment	20%	Add GPU code analysis pass in a python based frame
Homework	Multiple-Choice Quizzes	15%	Test your understanding of lecture content and readings.
Final Exam	Exam	10%	Covers all course topics.
Participation points (optional)	Ed	3%	Ed participation points

Final Grade Algorithm:

- 90-100% & and at least 40% of the final exam score: A
- 80-89%: B
- 70-79%: C
- 60-69%: D
- Below 60%: F

Late Policy:

- Homework: No late submissions accepted.
- Projects: Each day late reduces the score by 10%.

Exams: During the first week of the course, we will administer a quiz on the syllabus which will also validate your ability to use the Honorlock proctoring system. The final exams is delivered using Honorlock. The Honorlock system will scan your room and work area. It will also record your exam session. For the exam, you are allowed **one 8.5x11 inch sheet of paper** (both sides) of notes for the exam. The exam will be comprehensive, and reviewing quiz questions will be the most helpful way to prepare.

Ed Participation points : When your final grade is within 3% of the cut-off point, we will use Ed contributions to re-evaluate the final grade. This means that good contributions to Ed can earn points.

The detailed formula will not be shared, but historically, good comments or answers receive 0.1 points per post.

Communication:

- Private: Use Ed private messaging for individual questions with the instructor and TAs.
- Public: Use Ed public posting for general questions where classmates can contribute.

Online Platforms:

- Canvas: For assignment submission and lecture note distribution.
- Ed : For class discussions and Q&A.

Deadlines and Extensions:

Projects and quizzes are due at the end of the day midnight of ET. Extensions will only be granted for projects/exams in the case of documented, excused absences. These absences must be verified and approved by either an instructor or the office of the dean of students.

Quiz:

In general, each quiz allows only one attempt. Please ensure you have selected the correct answers before submitting your responses. Some quizzes may allow two attempts such as Quizzes 6, 7, and 8 and in that case, we will announce it in advance.

Assignment Quality:
This is a graduate level course. You are expected to turn in graduate level work. This means properly formatted reports, adequate comments in code, and actual substance to your work. We reserve the right to deduct points for work we do not feel is graduate level.

Academic Integrity:

This course adheres to the Georgia Institute of Technology's academic honor code: https://policylibrary.gatech.edu/student-life/academic-honor-code

Accommodations:

Students with disabilities requiring accommodations should contact Prof. Kim within the first week. For more information, visit the Office of Disability Services: https://disabilityservices.gatech.edu/

Discrimination and Harassment:

Georgia Tech prohibits discrimination and harassment based on any protected class. This class welcomes diverse viewpoints but will not tolerate discriminatory or harassing behavior.

Subject to Change Statement: The syllabus and course schedule may be subject to change. Changes will be communicated via the Ed announcement tool. It is the responsibility of students to check email messages and course announcements to stay current in their online courses.