New York University Tandon School of Engineering

Electrical and Computer Engineering

Course Outline **ECE 6913** [Computing Systems Architecture], Fall 2024 **INET Section**

Instructor: Azeez Bhavnagarwala

Instructor Contact : ajb20@nyu.edu

Course Assistants & Graders contact:

1. Aishwarya Kandam, ak11049@nyu.edu
2. Aditya Krishna, ak11137@nyu.edu
3. Srinatha Sivareddy, ss18364@nyu.edu
4. Manoj Srinivasan, ms14845@nyu.edu
5. Aditya Ojha, ao2612@nyu.edu
6. Nitisha Shetty, ns6108@nyu.edu
7. Sarthak Gupta, sg8304@nyu.edu
8. Pratik Pattanaik, pp2861@nyu.edu
9. Sivaram Goriparthi, sg8099@nyu.edu

Instructor Weekly Discussion Hour:

Wednesdays 9AM – 10AM

Course Assistant Off. Hours:

Monday, Tuesday, Wednesday, Thursday & Friday **9:30 AM - 11:00 AM**

Instruction:

Prerecorded Videos uploaded to Brightspace weekly

Course Prerequisites: Basic knowledge of digital logic and Computer Architecture is assumed. If you have not taken an undergraduate level class on Computer Architecture, you will need to supplement course work with additional preparation – *please review introductory texts posted on Brightspace*.

Weekly Schedule:

ECE 6913	MON	TUE	WED	THU	FRI
9AM – 10AM			Discussion Hour (INET Sec only)		
9:30-11:00 AM	CA Office Hours (Zoom)	CA Office Hours (Zoom)	CA Office Hours (Zoom)	CA Office Hours (Zoom)	CA Office Hours (Zoom)
11 AM – 1:30 PM				Lecture Sec A Rm 202	
5:00 – 7:30PM		Lecture Sec B Rm 202			
					Weekly HW due by 11:55 PM

Detailed Course Description:

ECE 6913 provides a solid foundation for graduate students to understand modern computer system architecture and to apply these insights and principles to design computer systems given the emerging age of domain specific architectures and an unprecedented growth in markets for pervasive, energy-efficient and ultra-low-cost computing.

The course begins with a focus on the Open-source RISC-V ISA given its proliferation over the last 5-7 years across the IoT, mobile and HPC systems and because its compact Base Integer ISA of a mere 47 instructions can support popular software stacks and programming languages. The course continues with IEEE 754 representation of Real numbers for Floating Point arithmetic components, reviewing the RV32FD - Single/Double Precision Floating Point RISC V extensions, RV32C extensions for Compressed instructions and RV32V - Vector extensions of the RISC V ISA

We will look at the basics of pipelining and its implications on the data path, the classic five-stage RISC pipeline in detail, its implementation & control and will examine its performance considering hazards- how these degrade performance and how they are resolved. Branch prediction is introduced as an advanced technique to reduce direct stalls attributable to branches. Out of order instruction execution is introduced as a technique where an instruction executes as soon as its data operands are available – reducing stalls seen in an in-order execution pipeline.

Because high-end processors have multiple cores, the bandwidth requirements on the memory hierarchy are greater than for single cores with the gap between CPU memory demand and on-chip bandwidth continuing to grow with the number of cores. The course details an example memory hierarchy of the Intel quad Core i7 6700 that delivers a total peak data and instruction reference demand bandwidth of 409.6 GiB/s at a clock rate of 4.2GHz - accomplished by multiporting and pipelining the caches using three levels of caches with two private levels per core and a shared L3 with separate instruction and data arrays at the first level.

Architecture innovations over the past few decades may not be a good match for the emerging market domains supporting image or speech recognition, for example, using machine learning methods. The class reviews an example accelerator, a custom CMOS ASIC - the TPU for the data center. Cost-performance comparisons of the TPU with CPUs and GPUs using DNN benchmarks reveal the opportunity of an upcoming renaissance for computer architecture

Online Course Content Schedule: Weekly lecture videos, slide sets and HW assignments, reading assignments to be made available on NYU Brightspace. Homework Assignments are due weekly. Please submit HW assignments as PDFs of Word documents with your identifying information and not on hand-written sheets of paper. Please prefix your HW assignment submission file name with your netID followed by the HW assignment. For example, I would submit HW 4 as a PDF document with filename: *ajb20_HW4.pdf*

Course structure:

Your performance in the course will be assessed via **weekly HW assignments (10% of total grade**) Project (**10% of total grade**): RISCV processor simulator [`Project A`], **2 Midterms (50% of total grade**) and a **final (30% of total grade**). There will also be extra credit HW assignments. Participation in these activities is highly encouraged.

Course Textbooks: [1] Hennessy and Patterson (RISC-V edition) of "Computer Organization & Design, Hardware-Software Interface" [2] Hennessey and Patterson, "Computer Architecture: A Quantitative Approach" [6th Edition], Morgan Kaufmann.

Policy on Academic Honesty:

In pursuing these goals, NYU expects and requires its students to adhere to the highest standards of scholarship, research and academic conduct. Essential to the process of teaching and learning is the periodic assessment of students' academic progress through measures such as papers, examinations, presentations, and other projects. Academic dishonesty compromises the validity of these assessments as well as the relationship of trust within the community. Students who engage in such behavior will be subject to review and the possible imposition of penalties in accordance with the standards, practices, and procedures of NYU and its colleges and schools. Violations may result in failure on a particular assignment, failure in a course, suspension or expulsion from the University, or other penalties.

More details about specific actions that constitute a violation of the NYU policy can be found here. https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/academic-integrity-for-students-at-nyu.html

Moses Center Statement of Disability:

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 www.nyu.edu/csd. The Moses Center is located at 726 Broadway on the 2nd floor.

Course Schedule:

Week	Weekly HW Release date	ECE 6913 Content	Assignment
1	Sept 6	Quantitative Design & Analysis, Physical limits on scaling CMOS – End of Dennard Scaling and Moore's Law. Evolution of System architecture, cost, energy efficiency and performance driven by emerging workloads dominated by AI. Hardware limitations and opportunities from ISAs, DSAs and Wafer Scale/ 3D Heterogenous Integration	HW 1
2	Sept 13	Introduction to the RISC-V, RISC-V Instructions, Instruction formats, memory management. Open-Source RISCV ISA review. Comparisons with older ISAs	HW 2, 3
3	Sept 20		
4	Sept 27	*Introductory Review of Project, Midterm 1 Review Problems*	HW 4
5	Oct 4	**Midterm I** (Tue Oct 1, INET)	-
6	Oct 11	No Class on Tue Oct 15 (Sec B). Class will be held on Thu Oct 17 (Sec A) *Monday Schedule at University on Tue 10/15*	HW 5
7	Oct 18	IEEE 754 representation & Floating-point Arithmetic for Computers. RV32FD – FP registers, FP load/stores/arithmetic, FP moves/converts	HW 6
8	Oct 25	Pipelining: Basic & Intermediate concepts. Classic 5 stage RISC processor pipeline.	HW 7
9	Nov 1	Pipeline Hazards, Pipelining implementation	HW 8
10	Nov 8	*Introduction to Phase II of Project, Midterm 2 Review Problems*	**Phase I report of Project Due**
11	Nov 15	**Midterm II** (Tue, Nov 12th Sec INET)	-
12	Nov 22	Introduction to Memory Hierarchy: Memory technologies, Caches, Cache performance. Comparison of ARM Cortex A-53 and the Intel quad Core i7 6700 Memory hierarchy. Virtual Machines, Virtual Memory, FSM for simple cache controller, Cache coherence	HW 9
13	Nov 29	**Thanksgiving Recess**. *No class on Thu 11/28 (Sec A)* Class will be held on Tue 11/26 (Sec B)	-
14	Dec 6	Introduction to Parallel Processing SISD, MIMD, SIMD, SPMD, and Vector machines, Hardware Multithreading Multicore and Other Shared Memory Multiprocessors. Introduction to the GPU	HW 10
15	Dec 13	*Closing Topics, Review Problems for Final*	**Phase II report of Project Due**
16	Dec 20	**Final: Tue, Dec 17 (Sec INET)**	