New York University Tandon School of Engineering

Electrical and Computer Engineering

Course Outline **ECE 6913 Section B** [Computing Systems Architecture], Fall 2020

Instructor: Azeez Bhavnagarwala

Instructor Contact : ajb20@nyu.edu

Instructor Office hours: Tue: 11 AM - 1:30 PM or by appointment

Course Assistants: Karan Parikh **kap9580@nyu.edu**, Sahil Chitnis **ssc9983@nyu.edu**, Kewal Jani **kj2062@nyu.edu**, Zhiming Fan **zf2035@nyu.edu**, Haotian Zheng **hz2687@nyu.edu**, Shan Hao **sh6206@nyu.edu**

Course Assistant Off. Hours:

Room 808, Monday, Tuesday, Wednesday, Thursday & Friday 9 AM - 11:00 AM

Instruction: In-person Lectures (Section B) Thursdays 11 AM – 1:30 PM. INET, INET2: videos of lectures posted on Brightspace

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 see course instructor*.

Detailed Course Description:

ECE 6913	MON	TUE	WED	THU	FRI
9:00 -11:00	Office Hours CA1 Rm 808	Office Hours CA2 Rm 808	Office Hours CA3 Rm 808	Office Hours CA4 Rm 808	Office Hours CA5 Rm 808
11:00 – 1:30		Office Hours Instructor Rm 817		Lecture 370 Jay, Rm 202 Instructor	
2:00 – 4:00					
4:00 – 5:50					
6:00 – 9:00					Weekly HW due by 11:55 PM

ECE 6913 aims to provide 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 ubiquitous, 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 years across 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 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 dealt with. 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 Classes. 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 HW **weekly assignments (20% of total grade**) that includes a project (HW 5 - 5% of grade) and Publication Reviews (HW 10 - 5% of grade), **2 Quizzes (50% of total grade**) and a **final (25% of total grade**). Class Participation (5% of grade). There will also be (extra credit) pop-quizzes and (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

Course Schedule:

Week	Date	ECE 6913 Content	Assignments
1	Sept 2	Quantitative Design & Analysis, Physical limits on scaling CMOS – End of Dennard Scaling and Moore's Law. Evolution of System architecture and cost with Open Source ISA, DSAs and Wafer Scale Engines	HW 1
2	Sept 9	Introduction to the RISC-V, RISC-V Instructions, Instruction formats, memory management	HW 2
3	Sept 16	Floating point Arithmetic for Computers – IEEE 754 representation,	HW 3
4	Sept 23	Pipelining: Basic & Intermediate concepts. Classic 5 stage RISC processor pipeline.	HW 4
5	Sept 30	Pipeline Hazards, Pipelining implementation.	HW 5 (RISCV project)
6	Oct 7	**Quiz 1**	
7	Oct 14	Open-Source RISCV ISA review. Comparisons with older ISAs. RV32FD – FP registers, FP load/stores/arithmetic, FP moves/converts, RV32C - Compressed Instructions, RV32V -Vector Extensions,	HW 6
8	Oct 21	Introduction to Memory Hierarchy: Memory technologies, Caches, Cache performance.	HW 7
9	Oct 28	Virtual Machines, Virtual Memory, FSM for simple cache controller, Cache coherence	HW 8
10	Nov 4	Review of ARM Cortex A-53 and the Intel quad Core i7 6700 Memory hierarchy	HW 9 Quiz 2 Review
11	Nov 11	**Quiz 2**	
12	Nov 18	Introduction to Parallel Processing SISD, MIMD, SIMD, SPMD, and Vector machines, Hardware Multithreading	HW 10 Publication Review
13	Nov 25	**Thanksgiving (no class)**	
14	Dec 2	Multicore and Other Shared Memory Multiprocessors	
15	Dec 9	Review for Final	
16	Dec 16	Final	

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.