



EE 599: ACCELERATED COMPUTING USING FIELD PROGRAMMABLE GATE ARRAYS (FPGAs)

TUE 5-620PM, LAB/DISCUSSION THU 5-620PM

SPRING 2021

NO. OF UNITS: 2

INSTRUCTOR: VIKTOR K. PRASANNA

Field Programmable Gate Arrays have become a key computing platform to accelerate applications at data center, cloud and at the “edge” including IoT. This course will review the technology and software tools from application acceleration perspective and discuss (application-specific) architectural, software and algorithmic innovations to realize the potential of this technology to optimize latency, throughput and energy efficiency in variety of emerging application workloads.

Prerequisite: (EE 451 or EE 457)) or consent of the instructor. This course does NOT require background in hardware or FPGA.

Text: Course will be based on recent research publications, survey articles and lecture notes by instructor.

Course Grade: based on lab assignments (programming homeworks) (25%), project proposal (25%), presentation (20%), and final report (30%).

Project: The focus of the course is in designing accelerators using FPGAs. The project will be focused on specific application areas of interest to the students to identify a problem that needs acceleration, design an application specific architecture, develop scalable parallel algorithm and map it onto a target FPGA device. The project will consist of literature survey, problem definition, solution idea, hardware design and use of software tools to map the design to a FPGA. It will consist of proposal preparation, discussions with the instructor and the TA, present details of the design and implement it and report the resulting acceleration.

Sample project: Parallelizing LSTM models on FPGAs with coherent memory. Identifying opportunities for parallelism, surveying state of the art techniques for kernels and primitives, performance modeling and estimating projected performance. Implementation in VHDL or Verilog or high level synthesis tools, place and route results. Summarizing latency and throughput performance and energy dissipation.

Course Outline:

1. FPGA basics, architectural characteristics
2. FPGA abstractions and computational models
3. Accelerating Dense Algebra
4. Accelerating Networking (SDN, NFV)
5. Graph Analytics at the edge
6. Accelerating ML Kernels
7. Accelerating ML Kernels
8. Accelerating AI at the Edge
9. Accelerating secure and privacy preserving computations
10. FPGAs in the Cloud
11. FPGAs for edge computing
12. Programming abstractions and automatic hardware synthesis from software