Senior Design Project

Fiona O’Connell, Kelly Deal, Danni Krauser

Hardware Acceleration: benefits of migrating compute-intensive software to hardware

- Purpose
  - The purpose of this project is to investigate how migrating a compute-intensive algorithm from a software-only implementation to a hardware-accelerated design can improve performance
- Goals
  - Develop a CPU-only baseline of a compute-intensive algorithm and analyze performance bottlenecks
  - Implement a hardware-accelerated version on an FPGA
  - Measure and compare performance metrics, including execution time, throughput, and energy efficiency, between software and hardware implementations to quantify acceleration benefits and trade-offs

Team Members

Fiona O'Connell	oconnefa@mail.uc.edu	CS
Kelly Deal	dealka@mail.uc.edu	CS
Danni Krauser	krausedc@mail.uc.edu	CS

Project Advisor:

Prof. Yu Zhao	zhao3y3@ucmail.uc.edu

Project Abstract

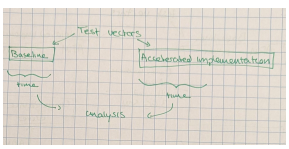
This project will investigate the performance benefits of migrating compute-intensive algorithms from a software-only implementation to a hardware accelerator. We will first develop a benchmark CPU-only version, analyzing the performance and identifying performance bottlenecks. From there, using high-level synthesis (HLS), the algorithm will be converted to VHDL, simulated, and integrated with an ARM processor. The design will then be implemented on an FPGA, with an application developed to handle data transfers between the CPU and hardware accelerator. Finally, we will analyze and visualize the performance metrics to quantify improvements, and discuss the advantages and trade-offs of hardware acceleration.

User Stories

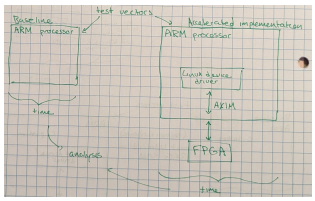
- As a research scientist, I want to accelerate my algorithm’s inference time so that I can process large datasets more quickly.
- As a data analyst, I want to reduce the power consumption of a neural network so that I can deploy my model in an energy-constrained environment.
- As a researcher, I want to measure the performance of a software algorithm on an FPGA so that I can prove the benefits of hardware acceleration.

Design Diagrams

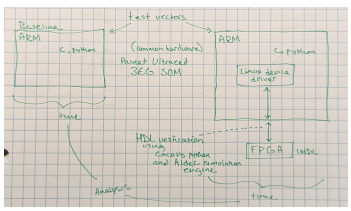
Design D0



Design D1



Design D2



Major Project Constraints

- Economic Constraints
  - All hardware and verification software will be provided on loan by Fiona's co-op
  - We will otherwise be relying on open-source software
- Professional Constraints
  - Quasonix specialist can assist with hardware issues
  - Project will otherwise be completed using our own skills along with supplemental research and advice from professors
- Environmental Constraints
  - Project focuses on improving efficiency of energy intensive algorithms through hardware acceleration
  - Will increase energy efficiency with needing hardware as a tradeoff

Review of Project Progress

Focusing on the software implementations, research and design.

- Currently working on the software baseline.
- AXI bus interface design for data transfer and control signals.
- Developing initial hardware implementations and testing them in simulation.

Expected Accomplishments - End of Term

- Finished software-only baseline.
- Established hardware architecture design.
- Research for moving forward to the hardware accelerated FPGA implementation.
- Defined the AXI bus interface and register map.

Division of Work

Fiona

- Design hardware accelerator architecture
- Design and configure AXI-Lite and AXI interfaces for control and data transfer
- Synthesize, implement, and deploy hardware design on Zynq FPGA
- Develop and optimize C++ host application for accelerator interaction

Kelly

- Research and select software algorithm
- Develop the CPU-only software implementation in C++
- Create test vectors and datasets for validation
- Measure and analyse software results, including baseline performance and bottlenecks

Danni

- Test end-to-end functionality vs. software baseline
- Measure latency and throughput on Zynq board
- Evaluate power consumption of both versions
- Compare performance gains from acceleration
- Visualize results

Expected Demo at Expo

- Software Baseline: Algorithm that runs on C++ CPU implementation.
- Hardware Acceleration: The same task offloaded to the FPGA IP Core.
- Visualizations and metrics to show the quantifiable benefits.