Jie Liu

471F Frank H. T. Rhodes Hall, Ithaca, NY 14850
jl3952@cornell.edu • +1 (607) 216-7193

EDUCATION

Cornell University, Ithaca, NY, USA

- Ph.D. student in Computer Systems Lab, ECE — Aug 2019 – Present
  - Advisor: Prof. Zhiru Zhang
  - Courses: Computer Architecture (A), Complex Digital ASIC Design (A), Advanced Compilers (A)

Tsinghua University, Beijing, China

- B.Eng in Microelectronic Science and Engineering — Sep 2015 – Jul 2019
  - Overall GPA: 3.62/4.0; Rank: 5/24
  - GRE: Verbal 157, Quantity 170, AW 4.0
  - Relevant Courses: Digital Integrated Circuit Design (A-), Very Large Scale Integrated Circuit CAD (A), Computer Networks (A-), Database (A), Quantum Mechanics and Statistical Mechanics (A), Microelectronics Fabrication Technology (A), Communication Systems and Circuits (A).

RESEARCH EXPERIENCE

Backend Optimization for HeteroCL
With Prof. Zhiru Zhang, ECE, Cornell University — Sep 2020 – Present

- Background:
  - *HeteroCL is a Python-based domain-specific language compiler infrastructure, with the philosophy of decoupling algorithm specification from hardware customization. HeteroCL enables programmers to explore design space in a productive and systematic way.*
- Implemented LLVM unroll codegen for HeteroCL.
- Add Catapult HLS backend for HeteroCL.
  - Supported basic arithmetic operations, loops, arrays, bit slicing and optimizations including loop pipelining, loop unrolling and streaming interface. Each feature comes with a unit test.
  - Implemented a count-min data sketch algorithm with HeteroCL which runs successfully with the new Catapult HLS backend.

A Sparse-Sparse Matrix Multiplication Accelerator Based on Row-Wise Product
With Prof. Zhiru Zhang, ECE, Cornell University — Nov 2019 – Sep 2020

- Background:
  - *Sparse-sparse matrix multiplication (SpGEMM) is a compute kernel widely used in numerous applications. Our approach is based on row-wise product which offers better trade-off on data reuse and storage requirement.*
  - *We designed a hardware friendly sparse format which better utilizes limited memory bandwidth.*
  - *We proposed a specialized accelerator architecture to implement the row-wise product. Our gem5 simulation shows 1.8× speedup over the previous state-of-the-art SpGEMM accelerator OuterSPACE .*
- Implemented the accelerator with a memory system using PyMTL.
  - Including a) PE with sub-modules of Multiplier, Merger and Drainer b) data loaders that support the sparse format and c) scratchpad memories to implement the priority queues.
  - Run simulation and pushed into ASIC tools.
- Implemented the accelerator using Mentor's Catapult HLS tool.

Efficient CNN Dataflow Mapping on CGRAs (B.Eng Project)
Undergraduate, with Prof. Shouyi Yin, Tsinghua University — Oct 2018 – Jun 2019

- Background:
  - *CGRAs are situated between FPGAs and custom ASICs in the spectrum of programmability. In recent years, various kinds of CGRAs have been proposed in both academia and industry, but a comprehensive mapping space exploration has not yet been well studied.*
  - *The high-level picture is to implement a large-scale CGRA based on a manually designed ISA with the aim of accelerating compute-intensive applications. Efficiently mapping high-level programs to optimized configuration information for each PE is the key problem in the project.*
  - *I focused on the mapping space exploration of a regular nested loop in the CNN application, which will provide inspirations for back-end optimizations.*
- Built estimation models for latency performance and power consumption.
  - Leveraged the polyhedral model to simplify analysis of nested loops.
- Compared dataflow patterns with different combinations of spatial reuse and temporal reuse.
  - Calculated latency and energy based on simulation results of meta operations.

Systolic Array Design Implementation for Matrix Decomposition Algorithms [Github]
Undergraduate Intern, with Prof. Jason Cong, UCLA — Jul 2018 – Oct 2018

- Background:
  - *Automatic systolic array generation has been an interesting research topic. It attempts to produce highly optimized programs while greatly reduces the design efforts.*
  - *The goal of this project is to build a polyhedral-based automatic systolic array High-Level Synthesis (HLS) design generation framework targeting FPGA platforms.*
  - *I focused on implementing and optimizing best-performance-candidate systolic array designs for matrix decomposition applications, which will be compared with the automatically generated results.*
- Designed 1D and 2D highly-parameterized systolic array architectures based on polyhedral analysis.
  - Manually mapped Cholesky, LU and QR decomposition algorithms to multiple 1D and 2D systolic array architectures.
  - Implemented highly-parameterized systolic arrays in HLS and evaluated using Xilinx Vivado HLS tools.
  - Optimized the HLS designs for better performance, achieved good throughput compared with HLS linear algebra library on FPGAs and LAPACK on CPUs.
  - Leveraged the Synchronous Dataflow Graph (SDFG) model to make a coarse prediction of the latency performance for dataflow systolic array applications in Vivado HLS tools.
- Presented as a poster in FPGA'19.

IMU-based System for Real-Time Pelvis Plane Measurement
Undergraduate, with Prof. Hong Chen, Tsinghua University Oct 2017 – Apr 2018

- Background:
  - *Total Hip Replacement (THR) surgery is one key solution to deal with hip joint diseases. However, in modern THR the attitude of the prosthesis is mainly determined by human experience, resulting in high failure rate in the surgery.*
  - *In this project, we analyzed clinical experimental data and built an inertial measurement unit (IMU) based real-time aiding system to help measure the implant angles of the acetabular prosthesis with high accuracy and stability.*
- Built an attitude measurement system with IMU sensor modules.
  - Put sensors at specific locations; Use the movement of one sensor to determine the initial angles of the pelvis plane; Use the other two sensors to track the real-time attitude of the pelvis
  - Sensor data were obtained from a host program on a PC.
- Optimized the attitude estimation algorithm with better stability.
  - Acquired the attitude angles based on quaternion transformations. Algorithm was implemented using Matlab.

PUBLICATIONS

CONFERENCES

[1] N. Srivastava, H. Jin, J. Liu, D. Albonesi, and Z. Zhang, **MatRaptor: A Sparse-Sparse Matrix Multiplication Accelerator Based on Row-Wise Product**, 2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), pp. 766-780. IEEE, 2020.

JOURNALS

[1] H. Chen, Z. Cao, S. Su, J. Liu and Z. Wang, **Measurement System for Attitude of Anterior Pelvic Plane and Implantation of Prothesis in THR Surgery**, *IEEE Transactions on Instrumentation and Measurement*,vol. 67, no. 8, pp. 1913-1921, Aug 2018.

POSTER PRESENTATIONS

CONFERENCES

[1] Jie Liu, Hong Chen, Zhihua Wang, **IMU-based Real-Time Acetabular Prosthesis Implant Angles Measurement in Total Hip Replacement Surgeries**, *IEEE Biomedical Circuits and Systems Conference (BioCAS 2018)*, October 17-19, 2018, Cleveland, Ohio, USA.

[2] Jie Liu, Jason Cong, **Dataflow Systolic Array Implementations of Matrix Decomposition using High Level Synthesis**, *27th ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA 2019)*, February 24-26, 2019, Monterey Bay Seaside, California, USA.

COURSE PROJECTS

Multicore Processor Fall 2019
ECE 4750: Computer Architecture, Cornell University

- Implemented a 4-core system using SystemVerilog.
  - Including pipeline processors, blocking i/dcaches and networks.
  - Evaluated using PyMTL simulation.

AWARDS

Excellent SRT (Student Research Training) Project Award, Tsinghua University Dec 2019

SKILLS

Programming Languages: C/C++, Python, Matlab, Verilog / PyMTL, LaTeX, HSpice, SQL.
Framework: LLVM
CAD Tools: Mentor Catapult HLS, Cadence Stratus HLS