Shanqing LIN

INJ 239 (Bâtiment INJ), EPFL, 1015 Lausanne

xusine@outlook.com xusine

Education

School of Computer and Communication Sciences, EPFL, Lausanne Sep. 2020 - Present

Ph.D. (expected) in Computer Science

Supervisor: Professor Babak Falsafi

Thesis Direction : Functional warming acceleration for samping simulation

Research Interest :

- Microarchitecture simulation
- Performance and dynamic program analysis
- Reconfigurable hardware (CGRA)

School of Microelectronics, Fudan University, Shanghai Sep. 2016 - Jun. 2020

B.E. in Microelectronics Science and Engineering

Supervisor: Professor Jun Han

Research interest : DNN accelerators

GPA : 3.70 / 4.00

Employment

Internship at NVIDIA Corporation Mar. 2020 - Aug. 2020

Supervisor: Wei Xu, Tegra Architecture Group, NVIDIA Corporation (Shanghai)

Contribution:

- Run DNN traces to profile the new generation GPU on Tegra SoC
- Implemented basic operation fusion on TVM to enable the profiling of internal CUDA kernel

Skills

Programming:

- Proficient languages: C/C++, Rust, Python, and Scala (mainly Chisel)
- Beginner languages: Zig, JavaScript, and Haskell
- Operation system: Linux (Archlinux addicted) and Windows

EDA Tools:

- RTL: Verilator and Synopsys VCS
- FPGA: Xilinx Vivado, with experience in Amazon Cloud FPGA Platform Development
- ASIC: Basic knowledge in Synopsys Design Compiler and IC Compiler

Text tools: Markdown, HTML/CSS, Typst, and LaTeX

Selected Research Achievements

CloudSuite Jul. 2021 - Present

CloudSuite is a set of monolith cloud service workloads for data-center research. We aim at providing guidelines for cloud-native hardware design.

Contribution:

- Compiled and tested workloads on ARM platform and RISC-V platform
- Validated workloads by measuring and reasoning the throughout and 99% latency
- Breaked down performance difference among representative x86 and ARM platforms

DevteroFlex Sep. 2020 - Aug. 2022

DevteroFlex is an emulation accelerator extending QEMU with an interleaved pipeline on FPGA. It is designed to complement software instrumentation framework by enabling zero-performance-overhead dynamic binary instrumentation using FPGA hardware resource.

Contribution:

- Designed and implemented the FPGA memory hierarchy (Cache, TLB, and MMU)
- Conceived and implemented the coherence protocol between QEMU and FPGA
- Built RTL verification platform based on Verilator
- Seted up a prototype on AWS FPGA and successfully run programs (gcc and JVM)

Sireye (Final Undergraduate Project) Sep. 2019 - Jun. 2020

Sireye is a sparse CNN accelerator implementing vectored row-stationary dataflow. It is a derivative of exploring the role of ISA played in engineering effort and clock frequency. The key insight behind is that complex hardware state machines are hard to program and debug while usually laying on the critical path for ML accelerator. By replacing them with dynamically generated instructions from high-level software, a rapid development process and clock frequency can be obtained.

Supervisor: Professor Jun Han, State Key Laboratory of ASIC and System, Fudan University

Contribution:

- Designed and implemented the instruction set and the hardware from scratch
- Proposed the runtime based on the JIT compiler
- Evaluated the hardware with standard ASIC flow (no tape-out)
- Compared with the reference implementation (Eyeriss), we achieve 400MHz (over 200MHz) at the cost of 20% IPC reduction

Teaching Experience

TA for Advanced Multiprocessor Architecture @ EPFL	2023 Fall
TA for Advanced Computer Architecture @ EPFL	2021/2022/2023 Spring
TA for Computer Architecture I @ EPFL	2021/2022 Fall
TA for Computer Architecture (INFO130038.03) @ Fudan University	2019 Spring