Xavier Routh

xrouth2@illinois.edu | github.com/xavierrouth | https://xavierrouth.com

Education

University Of Illinois, Urbana-Champaign **May 2026**
M.S. *Computer Science* - 4.0 GPA
Advised by Professor **Vikram Adve**

B.S. *Computer Engineering* - 3.86 GPA **May 2025**
James Scholar Honors Program, HKN ECE Honors Society

Coursework: Advanced Algorithms, Advanced Parallel Programming (CUDA), Advanced Compilers, Advanced Computer Architecture, ML Compilers, Performance Engineering, Advanced VLSI Design

Experience

LLVM Group, with Professor **Vikram Adve**, **May 2023 - Present**
HPVM-HDC - ISCA 2025

- A heterogeneous programming language and compiler built for hyperdimensional computing (HDC).
- Developed custom bufferization and memory management pass for HDC intrinsic operations using LLVM.
- Wrote a CUDA runtime library to support HDC operations for NVIDIA GPU targets.
- Wrote compiler backends for two programmable HDC accelerators.
- Ongoing work on auto-tuning framework for the HDC programming domain and compiler.

Hercules

- Built a compiler targeting heterogeneous hardware based on a sea of nodes intermediate representation.
- Implemented loop transformation passes such as, tiling, fusion, fission, interchange.
- Developed compiler backend for custom coarse-grained reconfigurable accelerator.

Annapurna Labs (AWS), Neuron Compiler Intern **June - August 2025**

- Implemented a dependency analysis for ML compiler IR for use in transformations such as software pipelining.
- Two-stage design, first using classical dependence checking, then using the polyhedral model and ISL.
- Improved kernel developer iteration time by 7x by detecting dependence violations early.

NVIDIA, CuDNN Compiler Intern **June - August 2024**

- Developed python tooling to lower deep learning computation graphs from frontend graph API to MLIR.
- Worked with CMake and Pybind to enable faster testing and iteration of DL compiler flow through usage of python instead of C++, for writing both compilation and execution tests.

Projects

Grape CGRA

- Researched, designed and taped out a custom course-grained reconfigurable array.
- Data-flow based design using an 8x8 grid of highly configurable functional units, with fast reprogramming speed.
- Responsible for development of a custom compiler to lower computation graphs to our architecture.

LC-3 Compiler

- Developed a C compiler that targets LC-3 assembly. Implemented in C and Rust.
- Includes custom memory allocation strategies to avoid usage of dynamic memory and improve cache efficiency.
- Provided as an official course resource for ECE 220, Introduction to Computer Systems & Programming.

RISC-V Pipelined Out-of-Order Processor

- Implemented and verified an Out-of-Order processor for the RISC-V rv32im ISA using SystemVerilog.
- Used a custom designed explicit register renaming based microarchitecture.
- Parameterizable I-Cache, D-Cache, and L2 shared cache with pseudo LRU replacement policy.

CUDA Convolutional Neural Network

- Implemented the forward propagation stage of the LeNet-5 CNN architecture using CUDA and Nsight Compute.
- Used optimization techniques including tiling, kernel fusion, and streaming to improve memory bandwidth.
- Placed in the top 15 of ECE 408 final competition submissions, out of around 200 students.

Skills

Languages and Libraries: C, C++, CUDA, LLVM, NKI, MLIR, SystemVerilog, x86, RiscV, Rust, Python