Saif Khattak

Work Experience

Aug 2023 - Ongoing
FPGA Compiler Engineer, *Intel*, Toronto ON
- Worked on the FPGA backend for the Intel oneAPI C++ compiler, which compiles high level **SYCL** or **OpenCL** code to an executable where a CPU host can execute kernels on an FPGA
- Designed a feature to generate Avalon-based **RTL** interface for compute kernels that includes registers to control and monitor hardware interrupts, ensuring usability for customers developing device drivers by engaging with FAEs for early feedback
- Created and troubleshooted FPGA-specific loop optimization passes in the **LLVM** framework to improve throughput and area usage for user designs targeting Intel FPGAs as a heterogenous acceleration platform, improving throughput/area performance by 15% on a user benchmark suite
- Debugged complex issues across the hardware-software boundary, including investigating compiled binaries, LLVM IR, OpenCL runtime libraries, Quartus compilation pipelines, Modelsim simulations, and HAL functionality

Sep 2022 - Dec 2022
Software Engineering Intern, *Snowflake*, San Mateo CA
- Developed data privacy features at the SQL query engine level for Snowflake's cloud database platform
- Added rules to an **ANTLR 3** grammar to enable managing data aggregation policies in SQL, enabling customers to share data while maintaining their users' privacy
- Implemented compiler changes in **Java** to parse and generate code for applying policies to a table
- Implemented changes to a custom **FoundationDB** layer to store information about policies

Jan 2022 - Apr 2022
Compiler Engineering Intern, *Groq*, Toronto ON
- Worked on the middle-end of Groq's compiler, which converts ONNX models to an instruction stream for the Tensor Streaming Processor, a custom neural network inference accelerator
- Increased neural network inference throughput by up to 20% by designing algorithms in **C++** to efficiently utilize hardware resources for common tensor operations (e.g. convolutions)
- Created optimization passes in **C++** using the **MLIR** compiler framework to manipulate neural networks described in **ONNX** format
- Created machine learning models in **PyTorch** to run end-to-end compiler tests and measure cycle-accurate performance when run on custom neural network accelerator hardware

Jan 2020 - Aug 2020
Software Engineering Co-Op, *RadComm Systems*, Oakville ON
- Developed software systems for controlling RadComm's radiation detection devices, which are used in recycling and medical facilities for monitoring and safety purposes
- Implemented algorithms in **C#** to analyze radiation patterns using the **ReactiveX** library to handle real-time data emitted by an embedded device, processing energy histograms every 100ms
- Automated the device calibration process using **C#** to allow parallel setup of many devices

Projects

May 2022 - Dec 2022
Bayesian Network Acceleration Compiler, Scala — Chisel — Python — Verilog
- Created a compiler in **Scala** to convert Bayesian network specifications into a **Verilog** module that can answer queries on the network with real-time evidence based on Markov-Chain Monte-Carlo techniques
- Created **Protobuf**-based specifications for model description and elaboration
- Created language for expressing Bayesian networks parsed using an **ANTLR 4** grammar
- Utilized **Chisel** to construct hardware modules dynamically and generate RTL for various backends

Feb 2022 - Mar 2022
CHIP-8 Emulator, C++ — SDL2 — ImGUI
- **C++** interpreter for CHIP-8 instruction set, runs publicly available ROMs
- Includes graphical and audio interface using the **SDL2** library
- Designed live debugger using **ImGUI** to inspect memory dumps and processor state

Education

Sep 2018 - Apr 2023
University of Waterloo, *Computer Engineering B.A.Sc*, Waterloo ON
Cumulative average **92%**

saifkhattak.00@gmail.com
linkedin.com/in/skhattak00/ — github.com/saifk00 — skhattak.dev