Saif Khattak

Overview

2+ years of experience optimizing highly parallel workloads across databases, FPGAs, and ML accelerators.

Work Experience

June 2024 - Ongoing
SQL Compiler Engineer, *Snowflake*, San Mateo CA
Compiler Platform Team
- Rearchitected 20+ compiler stages that are collectively run by 17M+ daily queries as part of a migration to a new expression framework.
- Engineered large-scale workload validation strategies to roll out migrated compiler stages, ensuring minimal regressions across 6B+ daily customer queries.
- Resolved 40-50 customer escalations per week during support rotations, triaging and resolving issues with query performance degradation, correctness bugs, and compilation failures across the full stack.

Data Governance Team
- Led design of new SQL syntax enabling multiple policies on a table, reducing boilerplate and preventing errors in applying privacy constraints.
- Designed algorithms to enforce Join Policy semantics on the parse tree of a SQL query, enabling data sharing for over 3000 tables while preserving sensitive information.
- Designed optimizations on query execution plans to increase query flexibility while maintaining privacy guarantees provided by aggregation policies, which protect over 80,000 tables in production.

Aug 2023 - June 2024
FPGA Compiler Engineer, *Intel*, Toronto ON
- Created an FPGA-specific **LLVM** optimization pass in **C++** that improved performance by 15% across standard **OpenCL** benchmarks by using scalar evolution analysis to narrow induction variables, reducing bandwidth and area requirements on the device.
- Debugged complex issues across the hardware-software boundary, from **SYCL** front-end through **LLVM** IR transformations to RTL code generation

Jan 2022 - Apr 2022
ML Compiler Engineer (Co-op), *Groq*, Toronto ON
- Optimized memory layout for tensor operations by eliminating extraneous padding, improving throughput by up to 20% in certain **ONNX** models on custom ML accelerator hardware.
- Built end-to-end validation pipeline in **PyTorch** to track compiler QoR metrics for important models, ensuring effectiveness of implemented optimizations

Projects

May 2022 - Dec 2022
Bayesian Network Inference Accelerator, 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

Education

Sep 2018 - Apr 2023
University of Waterloo, *Computer Engineering B.A.Sc*, Waterloo ON
Graduated with distinction.

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