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

- Redesigned 20+ compiler stages with a collective blast radius of 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 zero 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 aggregation 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 5000 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 throughput by 15% across standard **OpenCL** benchmarks by using scalar evolution analysis to minimize the bitwidth of induction variables.
- 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

- Designed an **MLIR**-based compiler optimization in **C++** to eliminate extraneous padding in tensor operations, improving throughput by up to 20% for key **ONNX** models on custom ML hardware.
- Built end-to-end validation pipeline in **PyTorch** to track compiler QoR metrics for benchmark models, ensuring effectiveness of implemented optimizations.

Projects

May 2022 - Dec 2022 **Bayesian Network Inference Accelerator**, Scala — Chisel — Python — Verilog

- Designed a compiler in **Scala** to convert Bayesian network specifications into a **Verilog** module that performs real-time probabilistic inference.
- Defined **Protobuf**-based specifications for model description and elaboration.
- Implemented domain-specific 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