

IRFAN SHARIF

irfansharif.io — irfan@irfansharif.io — in/irfansharifm — github.com/irfansharif

SUMMARY

- Experience with infrastructure, distributed systems, ML, storage and search engines
- Go, C++, C, Java, Python, Rust, Haskell, Ruby; NumPy, SciPy, Pandas, TensorFlow, Keras

EDUCATION

University of Waterloo

Sept '14 – Apr '19 (*expected*)

Honours Bachelor of Applied Science in Computer Engineering

GPA 3.9

EXPERIENCE

LinkedIn

Sunnyvale, CA

Software Engineering Intern (Machine Learning Infrastructure)

Sept – Dec '18

- Worked cross-team on model serving, versioning, runtime environments and deployment systems
- Published designs on model state machines/deployment orchestration across thousands of hosts
- Introduced dynamic request buffering mechanism to mitigate bursty, high-throughput traffic load

Amazon

Palo Alto, CA

Software Engineering Intern (Search Relevance)

Jan – Apr '18

- Led multiple design reviews for proposals outlining measures to improve barcode search relevance
- Authored query reformulation modules to improve system recall and click-through rate by 2.49%
- Leveraged sequence-to-sequence learning to generate alternative search suggestions for users
- Incorporated OCR text around barcodes into ranking models to reduce defects/variance in results

Cockroach Labs

New York, NY

Backend Engineering Intern (Storage/Performance)

May – Sept '17

- Designed a dedicated persistent state storage engine for Raft, increased throughput by 14.6%
- Implemented flow-control mechanisms for writes, significantly reducing system tail latencies
- Patched etcd/raft's PreVote extension from the thesis paper, formally verifying using TLA+
- Forked grpc/grpc-go, increased throughput by 12.8% batching syscalls/reducing GC pressure

Cockroach Labs

New York, NY

Backend Engineering Intern (Distributed SQL)

Aug – Dec '16

- Constructed core components of new distributed query execution engine, including distributed aggregations, sort, row-deduplications, and performant n -way parallel SQL joins
- Improved JOIN performance by an order of magnitude, the subject of an engineering article

Shopify

Ottawa, ON

Production Engineering Intern (Infrastructure)

Jan – Apr '16

- Built an auto-scaling build system with dynamic, load-balanced workload redistribution
- Optimized the overall scheduling system to handle 1,000+ builds per day saving 60,000 USD/mo

RESEARCH

Computer Aided Reasoning Lab (*SAT/SMT solvers*)

Oct – Dec '16

- Researched search space pruning, clustering and parallelization strategies with Prof. Vijay Ganesh

PROJECTS

- **Kura** – An end-to-end encrypted, synchronized, distributed file system for capstone project
- **CFilter** – Cuckoo Filter implementation with 730+ stars and 35+ forks on GitHub

Open-source contributions: **coreos/etcd**, **grpc/grpc-go**, **uber/go-torch**, **cerebrum**, **mqueue**