

Om Kelkar

Software Engineer based in New York, NY

✉ Available upon request | ✉ om.u.kelkar@gmail.com | 🌐 www.omkelkar.com | 📱 okel17 | 🌐 okelkar

Experience

Meta Platforms, AI Platform (Meta Superintelligence Labs)

September 2023 - present

SENIOR SOFTWARE ENGINEER

New York, NY

- Delivered production-ready Kubernetes cluster-autoscaler designed to manage O(100K) GPU hosts across 30 on-prem and cloud regions, thereby improving compute fungibility and reducing fragmentation across isolated research and product engineering teams
- Shipped gang-scheduling and distributed inference model serving capabilities for Llama4 and Muse Spark launches, involved from early design to driving the rollout.
- Designed and built LLM model evaluation orchestration system, supporting 9K weekly evals with 2 9's reliability, safely leveraging off-peak production compute. Navigated project from ideation to broad internal adoption.
- Led complex migrations of all Facebook app GPU predictions (1 billion QPM) to the next generation inference platform. Built operational tooling, developed customer-specific integrations, conducted training sessions, and worked across 8 engineering teams to ensure a reliable transition. Deep core infrastructure debugging and MLOps experience, working on heterogeneous compute, auto-scaling, and regional fail-over.
- Quality champion - developed team SLOs, improved security posture, and hardened system reliability with improved observability, simulation tooling, integration testing, and chaos testing.

Meta Platforms, Distributed Tracing Platform

November 2021 - August 2023

SOFTWARE ENGINEER

New York, NY

- Extended Java context propagation library support in service of company-wide cost attribution project. Primary driver unblocking attribution of 45MW of compute, working across multiple partners teams in the Messaging org. Significant debugging/learning in other teams' code.
- Built load-shedding knobs for prioritizing higher criticality tracing data for re-processing during disaster events.
- Upgraded sink format of service tracing data set from JSON to thrift binary to decrease cost from 1.15GB/s to 809MB/s at peak load.
- Designed solution to offload expensive customer-integrated UDFs off the hot-path and into async jobs.

Microsoft Corporation, Personal Insights

August 2018 - October 2021

SOFTWARE ENGINEER

Redmond, WA

- Coded a dynamic landing page for Teams Insights web app, packaging existing insights into one-click solutions to increase user engagement.
- Delivered critical update to the licensing API, decoupling and reducing underlying dependencies based on caller identity. Prevented downtime of flaky dependency affecting every upstream equally, improving overall app availability by 40%.
- Integrated insights backend with Outlook Web UI hooks to display push notifications, adding 15 million monthly clicks.

Education

Carnegie Mellon University

Pittsburgh, PA

B.S. COMPUTER ENGINEERING, ADDITIONAL MAJOR IN BIOMEDICAL ENGINEERING, UNIVERSITY HONORS 3.71/4.0

Aug 2014 - May 2018

Projects and Teaching

Supplemental Instruction / EXCEL Program

Pittsburgh, PA

STUDENT SUPERVISOR, STUDENT TEACHER

August 2016 - May 2018

- Managed 30 student teachers by providing ongoing enrichment, hosted teaching seminars on the seven learning styles, hands-on problem solving, and reinforcing first principles.
- Taught 15 students each semester for Signals & Systems and Digital Systems coursework.

Lock-free Priority Queue

Pittsburgh, PA

DEVELOPER

May 2017

- Implemented a skiplist-based priority queue in C++, leveraging a lock-free paradigm to support concurrency, resulting in 2x speedup over traditional fine-grained locking approaches in insert heavy loads.
- Wrote a dirty pointers scheme to handle conflicting insert and deletion operations to preserve accuracy.
- Developed a test harness in C++ and an insert/delete trace generator in Python to assess performance.

Skills

Computer Languages: C++, Python, C-sharp, C, SQL, SML, TypeScript, CSS, HTML

Technologies: Kubernetes, Pulumi, Terraform, Docker, Thrift, ReactJS, Git, Bash