

SRIRAM RAO

@ reach@sriramrao.com

+1 (949) 560-3250

San Jose, CA

linkedin.com/in/sriram-rao

github.com/sriram-rao

EXPERIENCE

Dremio

Software Engineer - PhD Intern

Jun 2022 - Sep 2022

Remote, CA

Industry

- Devised a proof-of-concept (POC) to progressively improve query response in data lakes. Familiarized with **Calcite** and **Iceberg**.
- Improved row-count estimation in **query planning** via accurate statistics observed in previous executions. (From LEO by Markl V., VLDB 2001).

Microsoft

Software Engineer 2

Jun 2016 - Sep 2020

Bengaluru, India

- Redesigned a workflow manager used for Extract-Transform-Load (ETL) in 100+ workflows, reducing deployment time from 1h to < 5s.
- Piloted a **Spark** Streaming POC pipeline to compute the statistical significance for A/B tests 3x faster than existing batched methods.
- Refactored API cache configuration system using **Aspect-Oriented** Programming. Decreased config. code 5x and codebase size by 300 lines.
- Contributed to teammates' success with detail-oriented guidance on 40+ design reviews, 100+ **code reviews** and live issues while on call.

Microsoft

Summer Intern

May 2015 - Jun 2015

Bengaluru, India

- Analyzed insert & response times of 3 data stores under stress loads. Stores: Azure Data Explorer (**Kusto**), **MongoDB**, **column-store SQL**.
- Concluded that Kusto suited the use-case of log analysis (response < 5s) and column-store SQL for aggregation-based queries (response < 1s).
- Enabled migration from **OLAP** cubes to column-store SQL. Cut ETL time from 10 days to 1 hour, with response time < 1s (vs. instant in cubes).

University of California, Irvine

Graduate Student Researcher

Sep 2020 - Sep 2024

Irvine, CA

Academia

- Designed database plugin to balance latency and server costs by **allocating** compute resources between current and **decision-tree**-forecast loads.
- Developed database framework to call **data generators** implicitly during analysis. Integrated simulator into **PostgreSQL** for no-setup follow-up.
- Created a **pipeline execution** system for workflows defined as directed acyclic graphs of tasks. (On GitHub as sample.)

University of California, Irvine

Teaching Assistant

Sep 2020 - Dec 2024

Irvine, CA

- Facilitated instruction across 200+ students in advanced courses: Data Management Principles, Beyond SQL, **Database & Web App** Projects.
- Rated 4/5 in anonymous feedback from students in courses, with appreciation for database expertise and **straightforward explanation**.
- Led 3 weekly problem-solving sessions (30 students each) and 2 office hours to improve understanding, coding, and debugging.
- Collaborated** with professors and TAs to create lecture slides, 30 quiz questions, 7 code assignments, and 20 exam questions, per course.

SKILLS

Languages

*: Novice

Python, C#, Java, C++, C, Ruby, Lisp, Prolog, SQL
UI/UX: HTML, CSS, JavaScript, SwiftUI*
Automation: Bash, Powershell, Lua*

Technologies

Databases: Big Data, NoSQL, MongoDB, OLAP, PostgreSQL, Column-stores.
Compute Platforms: Spark, ETL, DAG, Query Engine, Apache Calcite, Iceberg.
Backend: .NET, Spring, Flask, REST, Microservices, AOP, Architecture, Caching, Distributed Systems.
Infra/DevOps: Docker, Azure, CI/CD, Custom IaC.

EDUCATION

University of California, Irvine

MS in Computer Science (Data Management)

Sep 2020 - Mar 2025

Irvine, CA

University of California, Irvine

PhD in Data Management - ABD

Sep 2020 - Sep 2024

Irvine, CA

- All But Dissertation. Advised by Prof. Sharad Mehrotra on workload-aware pre-computation.

National Institute of Technology, Karnataka

B. Tech. in Computer Engineering

Jul 2012 - Mar 2016

Surathkal, India

PUBLICATIONS

- S. Rao**, M. Boissier, and S. Mehrotra, "Genie generator-driven iterative data exploration," Integrating data generators, like simulators or benchmark data producers, into databases; poster presented at alumni meet (Paused).
- S. Rao**, M. Boissier, and S. Mehrotra, "Janus autonomous resource allocation and pre-compute for future workloads," (Paused).
- S. Dinesh, **S. Rao**, and K. Chandrasekaran, "Traceback a forensic tool for distributed systems," in *Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics ICACNI 2015, Volume 2*, Springer, 2016, pp. 17-27.