

Alejandro Rojas

Software Engineer | Distributed Systems | Data Platforms | AI Infrastructure

Denver, Colorado rojas.alejandro303@gmail.com www.rojasalejandro.com
linkedin.com/in/arojas-swe github.com/Alejandro-Rojas

SUMMARY

Software engineer with 5+ years writing and operating production software for distributed data platforms — a multi-cluster Trino / Starburst estate across AWS and on-prem, spanning source-level Java / Python engineering, release engineering, and production reliability. Professional agentic AI / security engineering, and hands-on local LLM inference infrastructure as a technical project.

EXPERIENCE

Comcast, Software Engineer II

Current

Production software for a distributed data platform · multi-cluster Trino / Starburst estate across AWS and on-prem

- Engineer and operate a multi-cluster production Trino / Starburst estate across AWS and on-prem, providing federated data access across a heterogeneous data ecosystem.
- Java, Python, SQL, and shell across the Trino codebase — proxy / integration components, query performance, code-level platform troubleshooting, and client-side spooling.
- Cut MTTR from roughly 3 hours to roughly 1 hour through production troubleshooting, fleet observability, incident response, and runbook / automation that removes repeat failure classes.
- Ship weekly zero-downtime Trino / Starburst deployments via GitHub Actions CI/CD, using pre-release query replay and validation to catch regressions before rollout.
- Implemented Snowflake on-behalf-of authentication through Trino for identity preservation, integrated with Privacera / OAuth data-layer access control.
- Supported a heterogeneous data ecosystem — hands-on Databricks / Unity Catalog integration and troubleshooting, plus Snowflake, Spark, Tableau, Python connectors, SSIS, and Iceberg.
- Kept AWS and on-prem infrastructure declarative with Terraform + Ansible, plus Kubernetes / Helm and S3-related migration work.
- Contributed to a professional agentic security / compliance platform: repository / context ingestion, multi-step agent workflows, structured findings, deterministic guardrail / policy evaluation, risk / impact analysis, and remediation guidance.

TECHNICAL PROJECT

Personal AI Infrastructure

Designed and operate heterogeneous local LLM inference infrastructure across multiple NVIDIA GPU systems, serving LLMs behind OpenAI-compatible APIs with SGLang, vLLM, and llama.cpp. Hands-on work in quantization, long-context / KV-cache tuning, speculative decoding, and containerized serving.

SKILLS

Languages: Java, Python, SQL, Shell / Bash

Distributed / Data: Trino, Starburst, Snowflake, Databricks, Unity Catalog, Iceberg, Spark

Cloud / Platform: AWS, Terraform, Ansible, Kubernetes, Helm, GitHub Actions, Docker

Observability / Security: Prometheus, Grafana, OAuth, Privacera, Snyk, CrowdStrike

AI Infrastructure: SGLang, vLLM, llama.cpp, LLM inference, quantization, KV-cache / context tuning, speculative decoding

EDUCATION

Drexel University — M.S., Software Engineering

Graduate Certificates: Software Architecture · AI & Machine Learning · Software Security