ghostdevhv
himanshu-verma

Himanshu Verma

Bengaluru, India

himanshu@ghostdevhv.me
ghostdevhv.me

Work Experience

Oct 2023 – Feb 2023 | **Senior Software Engineer** | **Onehouse, India**

Agentic AI

- Designed and built **Spark Agent** end-to-end as the sole engineer (Scala/Java driver plugin + supervised Rust subprocess + React UI), an in-driver AI assistant for live Spark debugging publicly launched at quanton.dev.
- Architected the two-process system over versioned UDS-IPC with multi-provider LLM orchestration across **34 driver + Spark History Server tools**; validated at **−0.11% TPC-DS wall-clock impact** and **p99 listener cost 233 μs** via JMH microbenchmarks.
- Hardened the agent's knowledge-base IP: **AES-256-GCM** encryption of 82 documents with per-file canary sigils, CI-gated leakage tests, citation-path validation, and per-tenant prompt watermarks.
- Led a new AI team to build and ship the **RCA agent** (architecture, LangGraph orchestration, MCP tool integration, access-control), now handling first-level investigation on **100% of production Sev1 alerts at 60–70% accuracy** via a runbook-gated deterministic flow with bounded non-deterministic fallback.
- Engineered a three-layer token-cost reduction (response trimming, tool-history compaction, prompt caching) cutting per-investigation cost by **∼30%**; built secure multi-cloud dataplane connectivity across AWS and GCP.

Platform & Infrastructure

- Built the infra foundation as the sole engineer, migrating all customer environments to Terraform-based automation, and shipped CI/CD pipelines reducing infra deployment time by **99% (2 days → 1 hour)** with self-hosted GitHub runners cutting ops costs by **40%**.
- Established a company-wide development environment using Terraform Cloud and GitHub with consistent resource tracking, TTL-based cleanup policies, and full auditability, resulting in a **30% reduction in cloud costs**.
- Designed Java services for infrastructure orchestration (Terraform module updates) and customer onboarding (in-house Redis-based workflow engine automating Kubernetes cluster creation), with extensions for private clusters, Calico networking, and custom AMIs.
- Led the GCP control plane rollout expanding the platform from AWS-only to multi-cloud – defining secure infra patterns, hardening API request flows, coordinating external penetration testing, building a disaster recovery framework, and achieving **SOC 2 compliance**.
- Led a multi-phase initiative to autoscale EBS volumes based on Spark workloads (RFC, architecture, and rollout); evaluated NVMe-based alternatives, resulting in a **45% reduction in EC2 + EBS costs** for a large customer.
- Designed a Java-based Spark router and workload-driven autoscaling algorithm to dynamically scale Spark applications based on query and resource demand.
- Lead infra and on-call standups, unblocking engineers and coordinating across teams; drove operational improvements that reduced weekly alert volume from **∼250 to ∼60** across Sev1–Sev3 incidents.
- Mentored infra engineers through code reviews, incident troubleshooting, and reliability practices, supporting the team as it scaled from **1 to 6 engineers**.

Oct 2022 – Sept 2021 | **Software Engineer II** | **GitHub, India**

- Built the captcha framework for npmjs.com (JavaScript SDK for Octocaptcha integrated across multiple user-facing flows) and a GDPR-compliant data export pipeline on AWS (SQS, Lambda, EC2) with scoped IAM permissions.
- Hardened platform security: JWT secret rotation without traffic impact, MFA enforcement for package maintainers, and resolution of accessibility audit findings across npm surfaces.
- Optimized Azure CI by replacing default pipeline agents with custom VMSS-based agents (**+50%** pipeline performance), and built Geneva-metrics + ICM monitoring for new instance creation.

Jan 2021 – June 2018 | **Senior Member of Technical Staff** | **Adobe Inc., India**

- Developed Acrobat "Liquid Mode", an adaptive PDF reading experience using content classification and extraction techniques.
- Tuned ML model parameters using a Genetic Algorithm (C++), improving precision for good and bad pages by **1% and 5%**, and automated checkpoint selection in Python, reducing manual effort by **66%**.
- Owned deployment pipelines for team components, including code reviews, model registration, and end-to-end verification.
- Led development for Adobe Connect mobile apps, implementing SSO, upgrading Adobe AIR SDK, and shipping iOS v2.7.0.
- Migrated file sharing and web link features from Flash to HTML using React and Java with full accessibility support.

Technical Skills

Programming	**Proficient:** Java, Python, C++, C **Competent:** Scala, Rust, JavaScript
AI / Agents	LangGraph, LangChain, MCP, Langfuse
Frameworks / Libraries	Apache Spark, React, Node.js, Express, FastAPI, TensorFlow, Keras, scikit-learn, pandas
Data / Storage	PostgreSQL, MySQL, MongoDB, Firebase, OpenSearch, Prometheus
Infrastructure / Platform	Terraform, Kubernetes, Docker, AWS, GCP, Azure, Linux, Git, CI/CD

Education

Aug 2013 - Apr 2018 | Bachelors + Masters in **Computer Science and Engineering** - Indian Institute of Technology, Kharagpur
Teaching assistant for Compilers (CS31003) and Data Structures (CS19001)