

AVINASH SHARMA

Senior Full Stack Engineer

+91-8877209962 | avi.rohit13@gmail.com | [LinkedIn](#) | [GitHub](#) | avinashsharma.co

CAREER SUMMARY

Full stack engineer with 7+ years of experience building React/TypeScript applications and the cloud infrastructure that powers them. Deep hands-on work across the stack — from micro-frontend architecture and Core Web Vitals optimization to Python FastAPI services, real-time WebSocket pipelines, and Kubernetes deployments on GCP. Recent focus on Contact Center AI platforms: live transcript streaming over Pub/Sub and Redis, and agentic AI systems built on LLM tool calling. Comfortable owning a feature from UI component through to the load balancer in front of it.

PROFESSIONAL EXPERIENCE

EXL SERVICE

Sr. Developer | Bengaluru, India — *October 2024 - Present* **Front End Developer** | Bengaluru, India — *May 2022 - September 2024*

Real-time AI Platform (GCP)

- Built a Python FastAPI/WebSocket service on **GKE** streaming live call transcripts to a Salesforce Lightning Web Component agent-assist UI, consuming **Redis pub/sub** with sub-second end-to-end latency
- Migrated production services from **Cloud Run** to **GKE** behind an Internal Application Load Balancer, configuring NEG-backed ingress, `gce-internal` ingress class, and BackendConfig timeouts to sustain long-lived WebSocket connections
- Designed a **Pub/Sub event interpreter** normalizing Dialogflow CX/CCAI conversation events into Redis-backed transcript buffers with TTL-driven lifecycle management and a GCS NDJSON archival pipeline for post-call storage
- Developed supporting **Cloud Run microservices** for per-contact metadata (Redis hash store with agent ownership model) and SIP recording status relay, with bearer-token auth and VPC connectivity
- Diagnosed and resolved edge-layer issues blocking WebSocket upgrades across an F5 Distributed Cloud → ILB → GKE chain, coordinating with network teams on protocol and idle-timeout configuration
- Built an internal monitoring dashboard (FastAPI + Redis scanner + live auto-refresh UI) giving the team real-time visibility into pipeline health across every stage

Cloud Infrastructure & Infrastructure as Code

- Provisioned and managed AWS infrastructure using **Terraform**, authoring HCL configurations for EC2 instances, security groups, and IAM users/roles, and running

the full `(init → plan → apply → destroy)` lifecycle against live environments

- Defined **security group** rules and attached them to instances and users, applying least-privilege ingress/egress policies rather than manual console configuration
- Used Terraform state to manage resource lifecycle repeatably — tearing down and recreating environments on demand to keep dev infrastructure reproducible and cost-controlled
- Stood up and configured **Amazon Connect** as a contact center backbone, handling instance setup, telephony and routing configuration, and integration with downstream application services
- Applied the same IaC discipline across GCP, managing GKE ingress, BackendConfig, and Cloud Build pipelines as version-controlled configuration

Frontend Architecture

- Reduced **CTI call-alert latency by 83%** ($\approx 30s \rightarrow 4-6s$), delivering near-instant real-time alerts on incoming calls and materially improving agent response time
- Architected scalable React applications with **Next.js and Redux Toolkit**, implementing modular micro-frontend structure with custom hooks, HOCs, and context-based state management
- Cut **bundle size by 40%** and improved Core Web Vitals through lazy loading, code splitting, and React Suspense
- Implemented secure authentication flows using **AWS Cognito, Okta SSO, and Microsoft OpenID**, integrating internal microservices via **AppSync GraphQL** and REST endpoints
- Built interactive dashboards and LLM workflow visualizations using **React-Flow and Recharts**, enabling stakeholders to analyze complex conversational datasets
- Established engineering standards: Vite build tooling, ESLint/Prettier, Husky pre-commit hooks, and Sentry error tracking with custom error boundaries
- Developed end-to-end test automation with **Playwright**, reducing regression escapes across releases

NOKIA NETWORKS (Client Location) — ICS

Front-end Developer | Bengaluru, India — *April 2020 - May 2022*

- Developed ReactJS web applications translating UX designs into responsive, production-ready interfaces
- Built **Python FastAPI** services for uploading and retrieving artifacts from Docker containers
- Integrated RESTful APIs for large tabular datasets with server-side data handling
- Worked within Gerrit-based code review and JIRA defect management workflows

INFINITE COMPUTER SOLUTIONS (ICS)

Associate Software Engineer | Bengaluru, India — *May 2019 - April 2020*

- Built an end-to-end internal tool for work-package management using ReactJS, Java, and Spring Boot
 - Led a team of four as UI Developer; owned version control and merge strategy for the repository
 - Deployed applications to DEV environments (JBoss); **Awarded Employee of the Month (February)**
-

TECHNICAL SKILLS

Frontend: ReactJS, Next.js, TypeScript/JavaScript, Redux Toolkit, React Hooks, Context API, HTML5, CSS3, Tailwind, React Native (Expo)

Backend: Python (FastAPI, uvicorn), Node.js, NestJS, REST, GraphQL, WebSockets, Server-Sent Events

Cloud — GCP: GKE, Cloud Run, Pub/Sub, Cloud Build, Memorystore (Redis), Dialogflow CX / CCAI, Cloud Storage, Internal Load Balancing, VPC Connectors, IAM

Cloud — AWS: Amazon Connect, Cognito, AppSync, S3, EC2, IAM, Security Groups, Amplify, CodeCommit

Infrastructure & DevOps: Terraform (HCL, provisioning, state, lifecycle management), Kubernetes (manifests, ingress, NEG, BackendConfig), Docker, Cloud Build CI/CD, GitHub Actions, Nginx, Linux, PM2, Cloudflare Tunnel

AI / Agentic: LLM tool calling, Model Context Protocol (MCP), Google Agent Development Kit (ADK), LiteLLM, Ollama / local inference, prompt engineering, RAG fundamentals

Databases: PostgreSQL, MongoDB, MySQL, Redis, Firestore, Supabase

Testing & Quality: Playwright, Jest, React Testing Library, ESLint, Prettier, Husky, TDD

Data Visualization: React-Flow, Recharts, D3.js

PROJECTS

- **MCP Infrastructure Server** — MCP server exposing VPS operations (PM2 process control, git, disk and memory metrics) as AI-callable tools, with a Telegram bot interface. Deployed via GitHub Actions CI/CD.
 - **Agentic AI Travel Planner** — Multi-tool agent built on Google ADK with LiteLLM model routing, covering tool calling, memory, and observability patterns.
 - **Self-Hosted Homelab** — Ubuntu server running Jellyfin, Home Assistant, Nextcloud, and Uptime Kuma; exposed securely via Cloudflare Tunnel and Tailscale with hardware-accelerated transcoding.
 - **S3 / R2 Bucket Explorer** — Client-side file explorer for AWS S3 with multi-account profile support.
-

EDUCATION

Post Graduate Diploma, Full Stack Development IIIT Bangalore

Bachelor of Technology, Computer Science Institute of Technical Education & Research,
Bhubaneswar | CGPA: 7.6 | 2015 - 2019