

Vijay Gupta

github.com/vijaygupta18 | linkedin.com/in/vijaygupta18 | vijayrauniyar1818@gmail.com | +91-8565017450 | Web Portfolio

Experience

Juspay – NammaYatri || *Software Development Engineer* (May 2023 – Present)

- Designed and architected the **end-to-end flow** for a **multimodal transport feature**, enabling real-time GPS tracking across **3 cities** with **10K+ DAU**
- Architected and scaled **Redis/Valkey KV infrastructure** to **5M+ daily transactions** and **5K+ events/sec** via autoscaling and table-level sharding; executed **zero-downtime Redis→Valkey migration** with zstd compression and zone-aware routing, cutting memory by **50%** and costs by **40%**
- Drove **\$126K annual savings** and **20% AWS cost reduction** through EC2 right-sizing, AZ-aware routing, and storage efficiency improvements
- Owned **production reliability and on-call**, driving **75% reduction in 5xx error rates** through incident response, RCA, and alert tuning; maintained **99.9% availability**
- Built **Vishwakarma** — autonomous SRE agent ingesting alerts from **AlertManager** and **CloudWatch**, running **16 parallel investigations** across Prometheus, Kubernetes, Elasticsearch, and databases; delivers PDF RCA reports to Slack, cutting investigation time by **70%**
- Improved **P95 end-to-end API latency** by **45%** by profiling critical paths, optimizing I/O and caching layers, and eliminating high-latency downstream dependencies
- Built **high-performance in-memory GTFS service** with GraphQL-based startup preloading, reducing hot-path read latency by **60%** across **5K+ req/sec** workloads
- Implemented **multi-cloud routing infrastructure** with cloud-aware Redis and Kafka connections and client-aware routing, enabling **zero-downtime cross-cloud deployments**
- Designed and owned **core backend systems** for a real-time mobility platform serving **200K+ rides/day**, driving architecture decisions across **latency, reliability, and cost**
- Mentored **3+ engineers**, led design reviews and interview loops, directly contributing to **40% team growth** while maintaining hiring bar

Vahan | *Software Development Engineer - I* (June 2022 – April 2023)

- Redesigned microservice architecture for **AI-driven WhatsApp Bot** supporting **100K+ job seekers**; improved customer satisfaction through **45%** faster responses and **35%** higher engagement
- Implemented concurrent chat processing with **RabbitMQ**; reduced telecalling costs by **32%** and handled **10K+** conversations
- Automated application** status tracking via data warehouse integration; reduced manual monitoring effort by **70%**

Technical Skills / Achievements

Languages: Haskell, Python, C++, JavaScript, TypeScript, SQL, Java

Backend/ Distributed systems: Node.js, REST APIs, Kafka, RabbitMQ, Redis (Valkey), PostgreSQL, ClickHouse

Monitoring/CI/CD Tools: Grafana, Prometheus, VictoriaMetrics, Jenkins

Cloud Technologies: AWS, Kubernetes, Docker, CI/CD, Lambda, Autoscaling, Observability

Developer Tools / Frontend: Git, GitHub, VS Code, HTML5, CSS, Bootstrap, React.js

Computer Science Foundations: Data Structures, Algorithms, OOP, Competitive Programming

Solved **400+** problems on Data Structures and Algorithms on **Leetcode and Codechef**

Projects

Vishwakarma — Autonomous SRE Investigation Agent | *Python, FastAPI, LLM, Kubernetes, Prometheus, Elasticsearch*

- Built an autonomous agentic SRE platform** that ingests alerts from AlertManager and CloudWatch, runs **multi-step investigations** with up to **16 parallel tool executions** across Prometheus, Kubernetes, Elasticsearch, and databases, and delivers structured **PDF RCA reports to Slack**; reduced incident investigation time by **70%**

Multi-Cloud DB Manager | *Node.js, React, TypeScript, PostgreSQL, Redis, Docker*

- Built a unified multi-cloud database management platform** to query **PostgreSQL and Redis** across AWS, GCP, and Azure simultaneously — features parallel query execution, side-by-side result comparison, **3-tier RBAC**, Monaco SQL editor, and password-protected destructive operations

Location Tracking Healthcheck System | *Haskell, Kafka, Redis*

- Developed real-time healthcheck pipeline** using **Kafka** location streams to detect stalled drivers and built an **FCM-based alerting service** to trigger notifications when location updates stop, improving GPS accuracy and reliability for **200K+ rides/day**

Education

Kamla Nehru Institute of Technology

Bachelor of Technology in Information Technology — 8.5 CGPA (Top 5% of class)

Sultanpur, India

Aug. 2018 – Jun 2022