

HEMANT SHARMA

Bengaluru, India | +91 9873194941 | hemantsharma1498@gmail.com | linkedin.com/in/hemantsharma14 | github.com/hemantsharma1498

Experience

Clientell — Senior Software Engineer

October 2025 – Present

Python, FastAPI, PostgreSQL, Redis, GCP, Langfuse, OpenTelemetry

Bengaluru, Karnataka

- Own the **agent execution path end to end**, from tool-call latency to stream delivery to incident response. Wrote the design doc for the Salesforce org indexing platform, took it through review with the founding engineers, and built it as an **orchestrator-worker system on Pub/Sub**: ~28k work items per run, workers autoscaling to 100 Cloud Run instances, **fan-out bounded by customer API quota**. Tool-call latency dropped from ~4s to ~800ms.
- Took over a p99 regression (~1.4s against a ~400ms baseline at flat RPS) that had escalated past initial triage with no hot path in the profile, and root-caused it to **SQLAlchemy sessions held across an async boundary**, so pool checkouts outlived request scope. Fixed session lifecycle, sized the pool against measured service time, and added a pool-saturation alert so the same class of bug surfaces before customers see it. **p99 has held at baseline since**.
- Made SSE streams survive bursty upstream emission (5 to 15 second OpenAI token bursts) with a **bounded queue**, **fixed-batch draining**, **throttled delivery**, and retry on transient faults. Forced-kill failures went from ~4/day to none in the last ~500 streams. Wrote the runbook and handed operational ownership to two other engineers.
- Built **observability from zero across all 9 services** and set the SLOs the team now runs against. Median time-to-detection went from **over a day to under 10 minutes**; no incident has reached us via a customer since.
- **Mentor two engineers** on backend design; run the weekly design review that gates anything touching the agent execution path.

Fae — Lead Backend Engineer

January 2025 – October 2025

FastAPI, Go, PostgreSQL, Docker, AWS ECS, GitHub Actions

Bengaluru, Karnataka

- Rewrote a legacy NestJS codebase into a two-service FastAPI monolith with dependency injection across API, service, and data layers, dropping refactors worth 8-10 story points the monolith forced per sprint.
- Closed permission gaps for ground-ops on the internal platform, implementing **ABAC authorization** with an audit log mapping every operation to the acting identity, enabling monthly access and performance audits for ground personnel.
- Moved slow external payment calls off the critical transaction path and tightened transaction boundaries, halving payment execution latency from ~5s to ~2.5s and eliminating timeout failures.

Career Break — Admitted to Masters program, NYU (declined)

November 2023 – December 2024

Go, Distributed Systems, Operating Systems, Networking

- Built a **Redis-protocol-compatible** in-memory server supporting GET/SET, HGET/HSET, and ZADD, with automatic key expiration and durability to SQLite3.

Gida Technologies — Software Engineer

October 2022 – November 2023

Python, Node.js, Google BigQuery, PostgreSQL

Bengaluru, Karnataka

- Built a daily ETL pipeline ingesting ~600k **health insurance records** into BigQuery with source-level segregation and PGP encryption, meeting Ayushman Bharat Digital Mission (ABDM) regulatory requirements.
- Automated recurring analyst reporting (Python jobs orchestrated from Node.js), removing ~20 hours/week of manual effort.

Projects

Consistent Hashing Library | Go

- Implemented a consistent hash ring with 20 virtual replicas per node and binary search over sorted hashes for O(log N) lookups; pluggable CRC32/FNV-1a hashing, thread-safe reads and writes, and atomic full-ring replacement.

Distributed MapReduce (MIT 6.5840) | Go, RPC

- Built a coordinator-worker MapReduce system with exactly-once task semantics and crash recovery, using heartbeat monitoring and idempotent operations to handle worker failures under concurrent execution.

Technical Skills

Languages: Python, Go, TypeScript, SQL

Backend & Systems: FastAPI, REST APIs, gRPC, WebSockets, Distributed Systems, Concurrency

Infrastructure: PostgreSQL, Redis, BigQuery, OpenTelemetry, Docker, GitHub Actions, AWS, GCP, Linux

Education

Manipal Institute of Technology

July 2017 – June 2021

Bachelor of Technology

Manipal, Karnataka