KUNAL KEJRIWAL

Backend Software Engineer | Python · Java · Distributed Systems · Cloud-Native
kunal.resolute@gmail.com · github.com/kunal-kejriwal · linkedin.com/in/kunal-kejriwal

SUMMARY

Architected async pipelines processing **20M+ records/day** and built enterprise search infrastructure on GCP at Deloitte, within a 40-person engineering team serving 150+ global clients. Cut API latency by **35%** via N+1 query elimination and O(n²)→O(n) algorithmic refactoring. Independently shipped **APIEngine** (500+ users, 250K+ API calls/user/month) and **MeshEngine** — a distributed mesh network simulation platform with self-healing routing.

TECHNICAL SKILLS

Languages:	Python · Java · JavaScript · SQL
Backend:	Django · FastAPI · Django REST Framework · Spring Boot · Node.js
Systems & Messaging:	Celery · Redis (Pub/Sub + caching) · Kafka · GCP Pub/Sub · asyncio · WebSockets
Cloud & DevOps:	GCP (Cloud Run, Cloud SQL, Vertex AI, Memorystore) · Docker · GitHub Actions · Cloud Build · CI/CD
Databases:	PostgreSQL · MySQL · SQLAlchemy (async)
Practices:	Microservices · REST API Design · System Design · Event-Driven Architecture · Performance Engineering · pytest

EXPERIENCE

Associate Engineer — Deloitte USI

Gurugram, India | Oct 2023 – Present

Backend engineer on ChangeScout — an enterprise OCM platform serving 150+ global clients (40-person eng/PM/QA team).

- **Async Pipeline Ownership:** Architected high-throughput async data pipelines processing 20M+ records/day (load-tested to ~30M), orchestrating Celery workers with Redis-backed task queues under strict API rate limits. Selected Cloud Run over GKE for stateless workloads to minimize operational overhead while retaining auto-scaling.
- **Latency Optimization:** Cut API response time by 35% by eliminating N+1 query patterns (replaced per-record SELECT loops with batched prefetch joins), refactoring O(n²) list-intersection logic to hash-set lookups in O(n), and compressing response payloads across critical endpoints.
- **Enterprise Search (GCP):** Built and owned an end-to-end search feature — integrated GCP Search with a Java/Spring Boot service layer, indexed 50K+ documents stored in Cloud Storage, and served ranked results via a low-latency REST API (p95 < 200ms). Managed Cloud SQL and Vertex AI components in the pipeline.
- **Observability & Logging:** Designed a centralized error-logging system for all REST API failures — structured logs with request context, stack traces, and severity tagging. Built role-based access (admin-only) with indexed lookup by timestamp and endpoint for fast incident triage.
- **CI/CD & Cloud Ops:** Architected deployment pipelines (GitHub Actions + GCP Cloud Build) enabling zero-downtime containerized releases on Cloud Run with automated rollback and release packaging.

Technical Writer — Unite.AI (part-time)

Remote | May 2024 – Present

- **AI/ML Research:** Authored 80+ in-depth technical articles on cutting-edge AI/ML research (LLMs, multimodal models, diffusion models, efficient inference) for a global developer audience.

Project Intern (Backend) — DeepVision

Remote | Jan 2023 – Jun 2023

- **ML Inference APIs:** Built production REST APIs in Python serving computer vision model inference on image/video input; implemented structured error handling, input validation, and latency monitoring.

PROJECTS

APIEngine — Dynamic API Generation & Testing Platform

*Python · Django · DRF · FastAPI · Redis · Celery · Docker · GCP | **Live** · theapiengine.com · **500+ developers (organic) · 250K+ calls/user/month***

- **Problem:** Enables developers to instantly provision globally accessible REST endpoints per record — eliminating the need to stand up backend infrastructure for testing and prototyping.
- **Architecture:** Built a custom URL dispatcher and request multiplexer with per-user namespace isolation, configurable rate limits, and extensible schema support.
- **Distributed Execution:** Orchestrated background jobs via Celery + Redis with retry logic and dead-letter queues for failed tasks; containerized with Docker and deployed on GCP.

MeshEngine — Distributed Mesh Network Simulation Platform

Python · FastAPI · async SQLAlchemy · Redis Pub/Sub · PostgreSQL · Docker · WebSockets · GCP (Cloud Run, Cloud SQL, Memorystore)

- **Problem:** Simulates self-healing drone mesh networks for disaster-recovery and defence scenarios — routes messages across multi-hop topologies and automatically reroutes around failed nodes in real time.
- **Architecture:** Designed a control-plane/worker-plane split: FastAPI control plane handles topology and Dijkstra-based routing [O((V+E) log V)], while independent async workers subscribe to Redis Pub/Sub events for fan-out execution. Chose Redis Pub/Sub over Kafka for lower operational overhead given PostgreSQL-backed message persistence.
- **Observability:** Structured logging (structlog), per-message latency breakdown (hop-by-hop), real-time WebSocket event stream, and GCP Cloud Monitoring uptime checks. 25 automated tests (unit + topology).

EDUCATION

B.Tech, Electronics & Communication Engineering | National Institute of Technology (NIT) Delhi | 2023

Publications: 80+ articles on AI/ML research · unite.ai/author/kunalkejriwal