Sarthak Kumar

+91-8279436744 — kumar.v.sarthak@gmail.com — linkedin.com/in/sarthakchaudhary — github.com/sarthakvk — India

Career Highlights

- Scaled Meta Reality Labs spatial audio data pipeline by 1000x (10K → 10M scenes/run), enabling production-scale ML training data generation
- Reduced distributed ATS sync time by 90% at impress.ai
- Built a transactional email pipeline at impress.ai sustaining 100K+ emails/day on AWS SES/SQS/SNS with automated bounce and spam handling
- Owned end-to-end infrastructure build at Highbreed from scratch — ECS/ECR orchestration, automated deployments, and ELK-based observability and search

Experience

Meta Mar. 2024 – Present

Software Engineer (Contractor)

Built spatial audio simulation infrastructure at Meta Reality Labs — a pipeline that simulates sound propagation across millions of acoustic scene configurations to generate synthetic training data for audio ML models.

- **Scaled scene simulation pipeline from prototype to production: 10K → 10M scenes/run**
  - **Memory-bounded scene generation:** Eliminated scene count ceiling with async incremental batch uploads to blob storage with backpressure, fully decoupling generation from upload
  - **Decoupled result offloading:** Eliminated worker memory accumulation by streaming simulation results to Hive in batches via a per-worker async uploader, keeping workers continuously unblocked
  - **Distributed write coordination:** Eliminated partition explosion at scale by routing all workers to write distinct files under a shared partition directory, with a coordinator performing a single atomic commit across all workers via two-phase commit
  - **Post-commit compaction:** Coordinator runs an `INSERT OVERWRITE` via Presto after the final commit to consolidate all worker files into optimal partition sizes for read performance
- **DAG-based scene config engine:** Designed and built a generator dependency graph engine enabling controlled parameter variation across large-scale spatial audio simulation runs — users define interdependent generators in YAML and the engine resolves them into diverse scene configurations at scale; implemented cycle detection, multi-parent resolution, and Kahn's algorithm-based topological traversal
- **21× scene upload throughput:** Identified remote blob storage I/O as the serial bottleneck; parallelized scene uploads with a thread pool — 21× throughput gain (1.16 → 24.44 scenes/sec, 95% latency reduction)
- **Cross-team label propagation:** Built a system preserving time-accurate event metadata through all signal transformations during signal simulation; directly adopted by the Sound Understanding team for Audio Event Detection model training

Highbreed Development India Jul. 2022 – Aug. 2023

Principal Software Engineer

Freelancing marketplace platform connecting businesses with vetted professionals.

- Returned to take full technical ownership; designed and deployed the entire AWS infrastructure from scratch — ECS/ECR container orchestration with automated service definitions, independent service releases, and horizontal scalability
- Set up ELK stack serving dual purpose: centralized observability and full-text search; built a periodic sync layer between Elasticsearch and the relational DB to keep search indexes eventually consistent and fast
- Mentored 3 engineers through structured code reviews and engineering best practices

impress.ai Jul. 2021 – Jun. 2022

Software Engineer

AI-powered end-to-end recruitment automation platform — candidate screening, assessment, and engagement via conversational AI.

- Re-engineered a distributed data sync with client-hosted ATS systems; identified synchronous nested query loops as the bottleneck and parallelized every processing stage — 90% reduction in sync time
- Built a high-throughput transactional email system using AWS SES, SQS, and SNS sustaining 100K+ emails/day with automated bounce and spam handling

Highbreed Development India May 2020 – Jun. 2021

Software Engineer

Freelancing marketplace platform connecting businesses with vetted professionals.

- Built Django-based backend services and containerized the application stack with Docker; established CI/CD pipelines for automated testing and deployment

Technical Skills

Languages	**Python** (primary), Go, C++, Rust, SQL, JavaScript / TypeScript
Systems	Distributed systems, DAG-based systems, stream processing, concurrent & parallel programming, producer-consumer pipelines, event-driven architectures, memory-constrained design
Data	Relational databases, Hive, Apache Parquet, DuckDB, columnar storage, time-series data
ML & Audio	Audio signal processing; foundational ML (neural networks, backpropagation, random forests, gradient descent)
Tools & Technologies	Kafka, Apache Airflow, FastAPI, Django, Docker, Kubernetes, WebSocket, NumPy, Pandas, PyArrow, Redis, Git, Linux, AWS, GCP, Azure

Projects

gokey | *Go* github.com/sarthakvk/gokey

- Fault-tolerant distributed key-value store in Go using HashiCorp Raft — implemented the FSM interface (`Apply`, `Snapshot`, `Restore`), cluster membership management, and an HTTP API layer for multi-node deployments with leader election and log replication

algo-trade | *Python, AWS S3/EC2* github.com/sarthakvk/algo-trade

- Streams full-mode tick data for all NSE instruments via 3 parallel Broker WebSocket connections; producer-consumer pipeline with bounded backpressure queue writes ZSTD-compressed Parquet files in Hive-style date partitions, queryable via DuckDB
- Fully autonomous daily lifecycle — starts pre-market, collects data, uploads to S3 with SHA-256 verification, shuts down post-close; reverse-engineered TOTP authentication flow for fully headless unattended execution

Real-time bidirectional translation | *Python* github.com/sarthakvk/rtt

- FastAPI WebSocket backend streams chunked audio through Azure Cognitive Services STT → translation → TTS pipeline with async concurrency for low-latency multi-device playback

Education

AKTU Noida, India

Bachelor of Technology, Computer Science 2017 – 2021