

Sumit Kumar

Noida, India | thedeceptio@gmail.com | +91 9650376719 | github.com/thedeceptio

Summary

Senior Software Engineer specializing in ClickHouse and large-scale data systems, with 11+ years of experience designing and optimizing backend and data infrastructure. Hands-on expertise in ClickHouse (25TB+ clusters, 100M+ rows/day ingestion, 20B+ row tables), Kafka event pipelines, and hybrid hot/cold storage policies (S3/ECS/local disk). Proven success in schema evolution, query/index optimization, and cost-efficient infra migrations, alongside strong backend development in Python, Node.js, and MySQL.

Experience

JobHai.com | A Naukri.com Venture

Tech Lead | Apr 2023 – Present

- Designed and operated ClickHouse clusters with 25TB+ data, powering analytics, fraud detection, and ML pipelines.
- Built Kafka → ClickHouse ingestion pipelines processing 100M+ rows/day, equivalent to multi-TB/day streaming workloads.
- Implemented schema evolution and partition strategies (weekly partitions, schema migrations) to handle 20B+ rows in single tables.
- Optimized MergeTree table design, indexing, compaction, and TTL-based hot/cold storage policies, improving query performance and cutting storage by 30GB/day.
- Migrated cold data from AWS S3 → Dell ECS → disk-only infra, achieving █2 Cr annual savings while maintaining query performance.
- Monitored and tuned part merges, dictionary memory usage, and suspicious_broken_parts metrics, ensuring high availability and faster recovery.
- Scaled MySQL master-slave architecture to handle 60M API calls/day, reducing master QPS by 30% through read offloading.
- Reduced recommendation pipeline cron runtime from 12h → 1h, enabling 10M+ targeted notifications daily.

Lead Engineer | Sep 2019 – Mar 2023

- Designed backend architecture for candidate/job recommendations, processing ~5M daily notifications, boosting DAU by 30%.
- Built real-time job suggestion APIs (~1K RPS, 200ms latency) using Redis caching + optimized queries.
- Developed large-scale clickstream ingestion into ClickHouse for behavioral analytics.

TrulyMadly

Backend Developer & Data Analyst | Mar 2014 – Aug 2019

- Developed Recommendation Engine, improving engagement by 60% and reducing cold-start starvation.
- Implemented monetized features that drove the company to break even.
- Integrated AWS services to scale backend infrastructure.

Education

B.Tech, Computer Science | Roorkee Engineering & Management Technology Institute | 2008 – 2012

Skills

ClickHouse (expert) • Kafka • MySQL • Apache Iceberg (familiar) • Trino/Presto (familiar) • Python • Node.js • Airflow • Spark • Hadoop • Redis • RabbitMQ • AWS S3/ECS • Kubernetes (exposure)

Honors & Awards

- Gold Award for Excellence, InfoEdge (2023) – for building the vernacular JobHai app.
- Employee of the Year, TrulyMadly (2018, 2019).