Nikhil Pothuru

Pleasanton, CA • pothurunikhil@gmail.com • (209) 814-9664 • linkedin.com/in/nikhil-pothuru

SKILLS

Languages: Java, Python, SQL, C++, JavaScript
Frameworks & Libraries: Spring Boot, Spring Cloud, gRPC
Infrastructure: MySQL, Memcached, Redis, ActiveMQ, Kafka, AWS (EC2, S3), Docker
Tools: Jira, Git, JUnit, Jenkins, Checkmarx, Elasticsearch, Logstash, Kibana (ELK)
Core Expertise: REST APIs, Parallel Processing, Batch Processing, Scalability, Fault Tolerance

EXPERIENCE

Veeva Systems Pleasanton, CA
Software Engineer *2020* - Present

Platform Development
- Built CDB's ingestion platform with Java & MySQL to unify Veeva and external clinical data for large-scale aggregation
- Launched CDB from early phase to production, scaling to support 500+ studies on a multi-tenant AWS architecture
- Scaled the platform through multithreading & pessimistic advisory locking to handle multi-terabyte dataset ingestion
- Integrated with Veeva's change data capture (CDC) API to improve data ingestion latency via near real-time data updates
- Developed async ETL pipelines with ActiveMQ and MySQL to support scalable, high-volume third-party data ingestion

Scaling
- Reduced MySQL transaction bloat & replication lag by normalizing transformation data output from 10 GB to 100 MB
- Implemented RESTful API pagination & MySQL query tuning to improve ingestion audit retrieval latency by 80%
- Used async jobs & CRON-based validation to architect migration of 500M+ ingestion audit records with 0 downtime
- Introduced public REST APIs for vendors to prepare their data for ingestion, reducing integration time by 40%
- Architected new async ingestion pipelines enabling customers to integrate vendor metadata for 200M+ clinical records

Technical Leadership
- Served as security POC across 3+ teams, using Checkmarx to monitor vulnerabilities and lead remediation efforts
- Defined technical requirements & ensured alignment with stakeholders regarding functional and non-functional specs
- Owned feature development as tech lead, creating architecture docs and coordinating Agile delivery through Jira
- Coordinated end-to-end testing with QA, boosting coverage & cutting new production defects by 30% over 1 year
- Led org-wide tech talks on system design & performance tuning, standardizing practices across 3 engineering teams
- Mentored junior engineers regarding performance optimization, increasing team capacity to deliver scalable features

PROJECTS

Distributed File System
- Designed a client-server file system supporting fetch, store, list, & stat operations using gRPC and Protocol Buffers
- Implemented whole-file caching & CRC-based consistency checks to maintain data integrity for concurrent writes
- Added failure recovery with gRPC timeouts and conflict resolution to ensure reliability during network failures

EDUCATION

Georgia Institute of Technology, M.S. Computer Science Atlanta, GA
Relevant Coursework: Distributed Computing, Graduate OS, Computer Networks

University of California, San Diego, B.S. Math - Computer Science La Jolla, CA
Relevant Coursework: Algorithms, Advanced Data Structures, Theory of Computation