

VASU SHARMA

☎ +91-9654374405 | ✉ in.tech.vasu.sharma@gmail.com | [in](#) [Linkedin](#) | [G](#) [Github](#) | [M](#) [Medium](#)

SUMMARY

Software Engineer with 5 years architecting and deploying production systems end-to-end, spanning high-throughput backend microservices, document ingestion pipelines, agentic AI systems, and cloud infrastructure on AWS. Owns ambiguous, deeply technical problems end-to-end and ships solutions that run reliably in production and constrained deployment environments.

TECHNICAL SKILLS

Languages: Python, Go, JavaScript, C/C++, Java, Ruby
OS & Systems: Unix/Linux, AIX, macOS, Agile/Scrum
Backend: Spring Boot, JPA, Hibernate, FastAPI, RoR
APIs & Messaging: REST, gRPC, Kafka, RabbitMQ

Databases: PostgreSQL, MySQL, SQLAlchemy, Redis
Cloud/Infra: AWS, ECS, SQS, Docker, Kubernetes, CI/CD
Frontend: TypeScript, React.js, Redux, HTML/CSS
AI/ML: PyTorch, Pandas, NumPy, TensorFlow, RAG

EXPERIENCE

Senior Software Engineer

IBM, India Software Labs

Nov. 2024 — Present

Bengaluru, India

- Designed and built an **AI**-powered regression-triage system with a document ingestion and retrieval-augmented generation (**RAG**) pipeline, processing unstructured test logs, embedding them into a vector store, and auto-classifying failures by root cause across thousands of test cases, cutting manual triage time by around **40%**.
- Ideated, prototyped, and deployed **BuildMonitor**, an **AI**-powered merge-conflict resolver. Led technical discovery with engineering teams, built proof-of-concept demos, drove org-wide adoption, and mentored engineers on integration, cutting branch reconciliation time by around **20%**.
- Engineered **MemorySanitizer** for a new **OS** target (**AIX**), designing shadow-memory data structures, a custom heap allocator, and platform-specific interceptors. Scoped complex requirements into phased deliverables under tight timelines.
- Contributed to **LLVM open-source**, collaborating asynchronously with distributed maintainers across review cycles.

Software Engineer - II

Phyllo

June 2024 – Nov. 2024

New Delhi, India

- Engineered an **open-source** managed service for API integrations in **Python** and **Go**, reducing third-party integration time by **60%** and achieving **99.9%** uptime with **Kubernetes** orchestration.
- Designed extensible API infrastructure with **JWT** based authorization (**RS256**), **OAuth 2.0** flows, and role-based access controls, enforcing auditability and data governance standards while maintaining **50ms** latency.
- Built data orchestration pipelines using **Kafka** for async event processing and **Redis** caching, handling **1M+** concurrent **API** requests with **90%** test coverage using **pytest**.

Software Engineer II — Software Engineer I

Scaler by Interviewbit

Oct. 2021 — June 2024

Bengaluru, India

- Designed and built the **Class & Contest Creation Platform** with **monolithic backend** and **micro-frontend architecture**, reducing system error rates from **7.5%** to **0.1%** and saving **40%** of manual operational hours for content teams.
- Developed features like **Batch-change** tool, **Cohort-switch** tool, an admin panel for Ops and other UI touchpoints that lead to a significant **15.0%** increase in user retention.
- Developed the **Course Pause/Resume** feature end-to-end, building an accounting ledger with double-entry bookkeeping and idempotent API patterns reducing support tickets by around **50%** and improved revenue generation retention by **15%**.
- Led **RCA** for production breakages, drove query optimizations on critical paths, and established **CI/CD** pipeline standards with **Prometheus** based monitoring across services.
- Awarded **The People Saviour** (2023) and **Needle Mover of the Month** (2022) accolades for impactful contributions.

PROJECTS

AI based Job Search Agent ([link](#))

AI agent searching jobs, scoring fits, updating resume and generating alerts.

Python, SQLite, Agentic LLM, ReportLab

Cross Platform Coverage Gap Scanner ([link](#))

C/C++ scanner finding coverage gaps and using LLMs to triage findings.

Python, Static Analysis, LLVM LIT, llvm-cov

Deterministic Motion Graphics Renderer ([link](#))

Deterministic pipeline ensuring byte-identical motion graphics via a REST API.

TypeScript, React, Redux, Remotion

EDUCATION

Bachelor of Technology, Information Technology

Guru Gobind Singh Indraprastha University, Delhi, India

2017 - 2021

Percentage: **88.0**