

SHARAT CHANDRA MANCHI SARAPU

SOFTWARE DEVELOPMENT ENGINEER II | BACKEND, DISTRIBUTED SYSTEMS AND AI

Hyderabad, India | +91 9059726556 | mssharchandra@gmail.com
linkedin.com/in/mssharchandra | github.com/mssharchandra | mssharchandra.com

SUMMARY

AI-native builder and Software Development Engineer II with 3+ years delivering **distributed fintech, developer platforms, and agentic systems**. Built payment infrastructure processing **INR 30B+**, Go services at 1K TPS, and **sub-100 ms real-time analytics**; provided **technical leadership for 12 engineers** across system design, reliability, scalability, and end-to-end delivery.

EXPERIENCE

ASBL | Hyderabad, India

Dec 2025 - Present

Software Development Engineer II (Founding Engineer)

ASBL moved product engineering in-house from Inncircles; continued ownership of the same product suite.

- Integrated **HDFC Bank APIs** into idempotent payment and bank-reconciliation workflows using NestJS, MongoDB, and AWS, processing **INR 30B+ (approximately USD 360M)** in payment volume.
- Launched ASBL's enterprise **financial platform (Financial OS)** for invoicing, statement reconciliation, budgeting, loan tracking, and prepayments; captured **95% of organization-wide transactions** and prevented up to **INR 100M per month** in revenue leakage.
- Engineered a shared **Go communications platform** using **Redis Streams** and sustaining **1K TPS** and **500K-recipient campaign bursts** with goroutines, channels, Kubernetes horizontal pod autoscaling, and multi-channel vendor failover.
- Designed a high-scale **OLAP pipeline** using Amazon Kinesis Data Streams, Kinesis Data Firehose, Apache Iceberg, and S3 for event ingestion and analytics.
- Optimized real-time analytics with **MongoDB aggregations, materialized views, and change streams**, cutting dashboard response time from **2 seconds to under 100 ms** with near-real-time updates.
- Developed a **FastAPI AI gateway** with LLM and MCP tool calling workflows for helpdesk and visitor operations, reducing support tickets by **40%** and response time by **70%**.
- Created an internal **developer productivity platform on Linear** using AWS Lambda and Bitbucket webhooks to automate CI/CD, ephemeral environments, velocity tracking, and releases, cutting lead time by **96% (48 hours to 2 hours)**.
- Automated cash-flow and loan forecasting with scenario simulators, eliminating **8 hours of weekly manual modeling** and supporting budget and prepayment decisions.

Inncircles | Hyderabad, India

May 2023 - Dec 2025

Software Development Engineer II (Tech Lead)

- Architected event-driven **Java and Spring Boot distributed services** with Kafka and gRPC for invoicing, cash flow, and workflows, handling **100K+ requests per day** while reducing **p99 latency by 30%**.
- Led a monolith-to-microservices migration across Node.js and Spring Boot services with Docker and AWS ECS, enabling fault isolation, horizontal scaling, and **60% faster deployments**.
- Earned **2 promotions in under 2 years**; scaled engineering from **1 to 26** and led **12 engineers** across hiring, mentoring, cross-functional delivery, code reviews, system design reviews, and backend standards.
- Mentored **40 interns** and standardized onboarding tracks, reducing time to first contribution by **75%**.

TECHNICAL SKILLS

Languages: Go, Java, TypeScript, JavaScript, Python, SQL

Backend and Systems: Spring Boot, Node.js, NestJS, FastAPI, REST APIs, gRPC, Microservices, Distributed Systems, System Design

Data and Streaming: MongoDB, Aggregation Pipelines, Change Streams, Materialized Views, PostgreSQL, Redis, Kafka, Redis Streams, Amazon Kinesis Data Streams, Kinesis Data Firehose, Apache Iceberg, S3, OLAP

Cloud and Delivery: AWS, Lambda, ECS, Docker, Kubernetes, CI/CD, Bitbucket, Horizontal Scaling, Fault Tolerance, Idempotency, Observability

AI Engineering: LLM Integration, Real-Time Voice Agents, Streaming ASR/TTS, MCP, Tool Calling, RAG, Guardrails, Evaluation

Foundations: Data Structures and Algorithms, Object-Oriented Programming, Design Patterns, Concurrency, asyncio, API Design, Automated Testing

PROJECTS

Duet Real-Time Voice Agent | github.com/mssharchandra/duet

- Prototyped an open-source, controlled-duplex voice agent in Python with streaming ASR, Gemini reasoning, TTS, and idempotent tool calls; measured **192 ms barge-in cancellation** in an end-to-end synthetic test.
- Implemented deterministic consent, grounding, claim safety, and action guardrails with evaluation and observability; passed **132/136 checks (97.1%)** across **17 adversarial scenarios** and maintained **149 automated tests**.

Hermes Brain Multi-Agent Learning System | github.com/mssharchandra/hermes-brain

- Built a Git-backed agent memory and learning harness across **5 AI clients** with graph-based memory retrieval through MCP, source provenance, schema validation, human approval gates, and regression tests.
- Implemented agentic engineering controls for source-grounded research, claim-to-evidence mapping, deterministic lifecycle gates, reviewed publishing, recall scheduling, and regression evaluation.

ACHIEVEMENTS

- Placed **2nd among 400 experienced builders** at Google for Developers PromptWars 2026, an AI product-building competition.
- Won **Smart India Hackathon 2022** and authored **3 Scopus-indexed research papers** on applied AI and distributed applications.

EDUCATION

B.E. Computer Science, Vasavi College of Engineering | CGPA 9.08/10 | Aug 2019 - Jun 2023