

Sai Tarrun Pitta

+1 (657) 751-9260 | saitarrunpitta@gmail.com | linkedin.com/in/saitarrunpitta | github.com/saitarrun

Summary

Software Engineer with 3+ years of experience building scalable backend services and web applications. Proficient in Python, Java, and JavaScript with strong foundations in Data Structures, Algorithms, and Object-Oriented Design. Experienced in AWS cloud-native development, microservices, and performance optimization, delivering reliable low-latency systems in Agile teams.

Experience

Software Engineer Intern

Remote

Germania Farm Mutual Insurance Association

May 2025 – Dec 2025

- Designed and implemented Claim Intel, an AI-driven claims risk analytics platform with Angular and React dashboards, reducing average claim triage time by 35% and saving ~5 minutes per claim by surfacing actionable risk signals.
- Scaled the platform to support up to 50 adjusters processing up to 5,000 claims/week by integrating RESTful services and real-time updates into responsive UI workflows, improving throughput and triage consistency during peak volume.
- Improved dashboard performance by 30% by building reusable UI components, optimizing client-side rendering, and hardening data-fetch pipelines using TypeScript, HTML5, and CSS3, reducing time-to-insight for claim reviewers.
- Reduced regression cycles by 40% by implementing automated UI regression suites and CI validation gates using Selenium and JUnit, lowering defect escape rate by up to 25% and reducing hotfixes by up to 3 per release in an Agile Scrum environment.

Software Engineer

Hyderabad, India

Accenture

Jan 2022 – May 2024

- Engineered secure backend services for high-frequency payments and wallet operations using Node.js REST and gRPC microservices, sustaining up to 1,000 transactions per minute and processing up to 2M transactions/day.
- Achieved 99.9% availability by deploying containerized services on AWS EKS, implementing Kafka-based event-driven processing and Redis caching to improve resiliency, reducing incident pages by up to 25% and decreasing MTTR by 30%.
- Reduced end-to-end transaction latency by 20% through batching, cache strategy, and data-structure efficiency, plus observability-led tuning with Prometheus and Grafana, reducing payment timeouts by up to 40% and improving transaction success rate by up to 50 bps.
- Reduced AWS infrastructure costs by 35% per month by right-sizing EKS workloads and optimizing cache utilization while maintaining SLA adherence and peak-load performance.

Software Engineer

Hyderabad, India

Cognizant Technology Solutions

Feb 2021 – Jan 2022

- Developed Spring Boot REST APIs for core workflows with authentication and CRUD services backed by MySQL on AWS, reducing response time by 30% and improving p95 latency to ~150 ms through query optimization and efficient in-memory data handling.
- Strengthened release engineering by implementing CI/CD pipelines and containerized AWS deployments, reducing deployment duration from ~60 minutes to ~10 minutes and minimizing environment drift and manual deployment failures.
- Supported up to 300 internal users across up to 25 workflows by stabilizing service behavior across environments and devices, improving operational reliability for day-to-day execution.
- Improved security and reliability by enforcing role-based UI access controls, validating and sanitizing inputs, and implementing secure API patterns, reducing access-related defects by up to 50% and cutting production UI incidents by up to 25%.

Technical Skills

Languages: Python, Java, JavaScript, SQL, C++

Frameworks: React, Angular, Node.js (Express), Django, FastAPI, Spring Boot

Cloud/DevOps: AWS (EKS, ECS Fargate, RDS, ElastiCache, CloudWatch), Azure, Docker, Kubernetes, Terraform, Helm, Jenkins, GitHub Actions

Data/Messaging: PostgreSQL, MySQL, MongoDB, Redis, Kafka

Testing/Observability: Pytest, JUnit, Selenium, Postman, k6, Prometheus, Grafana

Tools/APIs: Git, REST, GraphQL

Education

Master of Science in Computer Science

Fullerton, CA

California State University, Fullerton

Aug 2024 – May 2026

Graduate program (MS CS); internship completed during program.

Projects

Patient Outreach and Appointment Reminder Platform

Tech Stack: React, TypeScript, Node.js (Express), PostgreSQL, Redis, BullMQ, Docker, Terraform, AWS ECS (Fargate), Prometheus

- Built a multi-tenant patient outreach platform on AWS using React and Node.js; reduced no-shows by up to 25% through automated, configurable appointment reminders and quiet-hour compliance.
- Architected a durable scheduling engine using BullMQ and Redis with custom deduplication and idempotency logic; achieved 99.9% reliability for notification delivery and p95 send-time <60s.
- Implemented tenant-isolated RBAC middleware and Infrastructure-as-Code with Terraform to provision least-privilege IAM roles, ECS (Fargate) services, and RDS databases; integrated Prometheus for RED metrics (rate, errors, duration) observability.

Zero-Downtime Monolith-to-Microservices Migration

Tech Stack: Python (Django, FastAPI), Node.js (Express), Docker, Kubernetes (AWS EKS), AWS (RDS PostgreSQL, MSK Kafka, ElastiCache Redis), Terraform, Helm, GitHub Actions, k6

- Led the migration of a legacy monolith to microservices on AWS EKS using the Strangler pattern and versioned REST APIs, enabling zero-downtime cutover while maintaining backward compatibility.
- Implemented Kafka-based dual-write with backfill and reconciliation checks to maintain data consistency during migration, sustaining 99.9% availability through canary rollouts.
- Built automated Terraform + Helm provisioning and GitHub Actions CI/CD for progressive deployments, achieving sub-5-minute release cycles and validating stability under load with k6 (20 concurrent users, p95 latency <100 ms).