Vamshi Saggurthi

P: +1 2016871827 | vamshisaggurthi@outlook.com | LinkedIn | Github | vamshisaggurthi.com

Software Engineer – Distributed Systems & Reliability

Software engineer specializing in building and validating large-scale distributed systems with a focus on correctness, availability, latency SLOs, and production testing infrastructure. Experienced in real-time streaming systems, observability, and synthetic workload validation at multi-million events/hour scale.

EDUCATION

Rutgers University, New Brunswick , NJ
Master's in Computer Science **May 2023**
Coursework: Machine Learning,Design and Analysis of Algorithms, Distributed Systems,Data Visualization

Osmania University , Hyderabad, India
Bachelor's in Information Technology **Jun 2017**
Coursework: Data Structures and Algorithms, Operating Systems,Computer Networks,Database Systems,

---

Publications

TR-2-PATH and MYC as a marker of Enza resistance Nature Communications impact score **17.69 [link]** 2023
Real time PII detection in streaming systems at scale - Sigmod Conference [link] 2024

Skills

Languages: Python, Java, R, JavaScript
Libraries/Frameworks: Airflow,Kafka,pySpark,ReactJS, Redux, Firebase, TensorFlow, Sci-kit Learn, PyTorch, Pandas, Dash **Databases**: MySQL, Postgresql, MongoDB
MISC: Git, Docker, Kubernetes, RESTful, Jenkins CI/CD, AWS

WORK EXPERIENCE

Amazon Inc | Software Engineer II **June 2025– Present**

- Led modernization of a large monolithic React application into a modular proxy-based architecture, launching to **25+** regions with production validation and AppSec approval within 8 weeks.
- Built system health validation and observability tooling for AWS Outposts (on-prem distributed systems), enabling early detection of failures and reducing operational risk.
- Developed telemetry pipelines to monitor system correctness, availability, and performance across geographically distributed deployments.

Striim Inc | Senior Software Engineer **July 2023–June 2025**

- Designed and implemented real-time PII detection in distributed streaming systems processing **3M+ events/hour** on a single GPU.
- Built synthetic workloads and stress tests to validate correctness and latency under peak throughput.
- Defined throughput and latency SLOs and implemented automated validation pipelines for streaming services.
- Introduced gRPC-based microservices into a monolithic on-prem platform, enabling interaction testing across distributed components.
- Designed and implemented real-time envelope encryption using cloud KMS, supporting **100k+ events/sec** with correctness guarantees.
- Built production observability infrastructure using OpenTelemetry, replacing legacy logging-only debugging with metrics and traces.
- Developed multi-agent Copilot systems to automate pipeline creation and diagnose customer issues at scale.

Research Assistant | Big Data — Jan 2021 –May 2023

- Designed and operated Docker-based Airflow pipelines over HPC (Slurm) to ingest and validate large research datasets.
- Built fault-tolerant trigger-based operators for distributed workloads.
- Developed scalable graph-mining algorithms improving computational performance by 4×.

Striim Inc | SDE Intern — July 2022 – Aug 2022

● **Real Time streaming Analytics:** Reduced 37% bugs , by building end to end analytics application which includes data scraping, ETL, and web interface for visualization, which formed as the baseline data for prioritizing development.
● **Library Diagnosis:** Developed library diagnosis module for analyzing jar dependencies installed on customer deployment**.**
● Built multiple ETL pipelines tailored for customers to tackle specific tasks demonstrating the real time analytics of the data

Dassault Systemes Inc | Site Reliability Engineer — Aug 2020 – Jul 2021

● Developed & maintained cloud microservices for COVID research and clinical trials supporting Moderna and other institutes
● **Custom Observability Tooling**: Designed and Developed custom tooling for observability, monitoring and metrics reporting for advanced analytics.
● Built Predictive alert systems which reduced scaling failures by 15%
● **MicroService Orchestration:** Transformed kubernetes based custom tooling for local microservices orchestration boosting developer productivity by 40%.
● Resolved scalability bottlenecks for backend services in AWS which improved service performance by 10%

Hexagon Inc |Senior Software Development Engineer — Aug 2017 – Aug 2020

● **Smart Form Generator AI:** Developed complex cloud based web applications for mechanical engineers' design workflows.
● Optimized data flow in the application by building custom middleware using React-Redux ,boosting load time by 30%.
● Lead "edge computing feature" improving performance by 40% in mobile browsers and devices with limited compute power.
● Engineered and implemented proof of concept for Computer Vision based text extraction from instrument specification images.

Academic Projects

Criminalytics[link] — **Jan-May-2021**

- Built ETL pipelines ingesting 15 years of crime data (7M+ records).
- Designed interactive dashboards and optimized query paths to return analytics under 1.5 seconds.

Maze Solver [link] — **Sep-Dec 2020**

● Designed multiple AI agents using Repeated A*, Inference, and Bayesian Networks for optimally searching a hidden target within a maze using Python and NumPy
● Optimized these agents to find the targets in (101) *(101) dimension mazes under 20 milliseconds
● Built a CNN with Dense layers using PyTorch to imitate these agents obtaining accuracy of 92% in solving the mazes.

Image -to -Image Translation (GAN's) [link] — **Jan-May 2021**

- Explored Supervised I2I using Pix2Pix GAN to translate Street View Images to Aerial View Images and vice-versa
- Implemented CycleGAN framework for the task of translating Real Pizza to Synthetic and Live Pizza Image Domains
- Researched the drawbacks of CycleGAN framework & proposed an enhanced CycleGAN by 10% with reduced artifacts