

SAI HARSHA VEMULA

Tuscaloosa, Alabama

☎ 816-614-4801 ✉ vemulasaiharsha@gmail.com 🌐 [sai-harsha-vemula](#) 📁 [Portfolio](#) ✍ [Blog](#)

DevOps Engineer with 5+ years of experience building and operating production cloud infrastructure on AWS. Currently run the cloud platform behind a national-scale hydrologic forecasting system, with deep expertise in Terraform, Kubernetes, GitHub Actions, CI/CD cost optimization, and cloud security automation. AWS Community Builder (Serverless).

Experience

Alabama Water Institute / CIROH, University of Alabama

July 2025 – Present

DevOps Engineer

Tuscaloosa, Alabama

- Own production AWS infrastructure for the NextGen Research DataStream (NRDS), a cloud-based numerical prediction system for the NextGen National Water Model, operating **1,213 EventBridge schedules, 24 Lambda functions, 26 AMIs, 27 IAM roles, and 20 S3 buckets** that power daily nationwide hydrologic forecast runs.
- Migrated CI/CD builds from GitHub-hosted x86 runners to AWS CodeBuild on ARM/Graviton, reducing CI/CD compute costs by **~95%** while enabling native multi-architecture (amd64/arm64) Docker image builds.
- Consolidated **24 fragmented GitHub Actions workflows into 1** reusable, parameterized pipeline, cutting maintenance overhead and standardizing build, scan, and deploy stages across all repositories.
- Detected, contained, and remediated a live crypto-mining security incident in **under 2 hours**, performing forensic analysis of compromised resources and rotating all exposed credentials with zero forecast downtime.
- Hardened the platform post-incident with Trivy container image scanning wired to Slack alerts, weekly Gitleaks secret scanning across all repositories, and Checkov IaC policy checks enforced in CI.
- Modernized Terraform across the organization: restructured scattered configurations into a monorepo of reusable modules, replaced long-lived AWS keys with GitHub OIDC federation, and enabled remote state locking with automated drift detection.
- Built the NRDS observability stack: CloudWatch dashboards tracking Lambda health and forecast pipeline status, plus SSM session logging to S3 for audit and compliance trails.
- Developed an automated S3 output-validation pipeline with an HTML status dashboard that verifies forecast completeness across all Vector Processing Units (VPUs) after every national run, replacing hours of manual verification.
- Drove AWS cost optimization on a **~\$2,800/month** account through S3 Intelligent-Tiering archive tiers and VPC Gateway Endpoints, eliminating avoidable data-transfer and storage spend.
- Engineered an ARM-based self-hosted GitHub Actions runner fleet with Lambda-driven start/stop scheduling and automated disk cleanup, eliminating idle compute costs outside working hours.
- Introduced AI-assisted code review (CodeRabbit) for pull requests across the organization, accelerating review cycles and raising code-quality consistency.
- Automated bidirectional sync of GitHub issues and PRs to the org-wide project board via GitHub Actions, and built utilities for NetCDF file extraction from S3, improving cross-team visibility and researcher self-service.

Vitrana, Inc.

July 2021 – June 2023

Software Engineer – DevOps

Bangalore, Karnataka, India

- Built and refined Kubernetes manifests to streamline automated deployment, scaling, and management of containerized healthcare applications, boosting deployment speed by **20%** and cutting system downtime by **30%**.
- Created and managed Helm charts for Kubernetes deployments, ensuring consistent, versioned, and scalable application rollouts across multiple environments.
- Implemented Argo CD for automated, GitOps-based continuous delivery, eliminating manual release steps and improving deployment consistency across Kubernetes clusters.
- Designed and managed multiple AWS VPCs, configuring subnets, CIDR allocations, and Network ACLs to enhance network segmentation, security, and performance across cloud environments.
- Deployed robust monitoring and alerting with Prometheus and Grafana, slashing incident response times by **40%** and materially improving service availability.
- Applied Infrastructure as Code with Terraform to automate cloud provisioning, halving environment setup time and eradicating configuration drift.
- Containerized microservices with Docker and Docker Compose for consistent, repeatable deployments across all environments, reducing deployment issues by **35%**.
- Designed an efficient cache management strategy using Redis and AWS ElastiCache, improving application performance and reducing database load through optimized data retrieval.
- Established security best practices by integrating Trivy and SonarQube into CI/CD workflows, enhancing code quality and mitigating vulnerabilities before release.
- Configured and managed AWS WAF to protect web applications against common threats, strengthening security posture and reducing exploitation risk.
- Adopted AWS Secrets Manager for secure storage and automated rotation of credentials, ensuring compliance with security best practices and minimizing exposure.
- Managed AWS RDS for scalable, secure, and reliable database operations, implementing automated failover and disaster-recovery processes that cut database downtime by **20%**.

Vitrana, Inc.

August 2020 – July 2021

Associate Software Engineer

Bangalore, Karnataka, India

- Integrated and managed Azure Active Directory for secure identity and access management, implementing single sign-on (SSO) and multi-factor authentication (MFA) for a client-facing project.
- Automated critical system monitoring and deployment workflows using Bash and Python, minimizing manual intervention by **40%** and improving deployment consistency.
- Developed AWS Lambda scripts to automate EBS snapshot backups and retention, remove outdated ECR images, schedule log cleanup, and dynamically scale EC2 instances based on system metrics, improving cloud efficiency and cost management.
- Contributed to the migration of legacy applications to cloud-native architectures, improving scalability and simplifying operational complexity for a **15%** reduction in maintenance costs.
- Designed and enforced IAM roles, policies, and Kubernetes RBAC, ensuring secure access control and maintaining compliance across the organization.
- Operated and optimized **300+ Docker containers** in production environments, implementing best practices that elevated application stability and reduced deployment errors by **25%**.
- Automated configuration management with Ansible, decreasing manual configuration errors by **35%** and streamlining deployment of microservices in non-Kubernetes environments.
- Designed a real-time analytics dashboard using AWS QuickSight with Amazon Redshift as the data warehouse and AWS Glue for ETL, enabling seamless access to S3-stored data.
- Deployed and configured the ELK stack (Elasticsearch, Logstash, Kibana) to centralize logging and monitoring, accelerating issue detection and resolution by **30%**.
- Implemented automated failover strategies and disaster-recovery runbooks, strengthening system reliability across production services.
- Supported release management processes by coordinating multi-environment deployments and maintaining release schedules, improving release accuracy and reducing rollback incidents by **30%**.

Education

University of Missouri-Kansas City

August 2023 – December 2024

Master of Science in Computer Science

Kansas City, Missouri

Vellore Institute of Technology

June 2016 – June 2021

M.Tech (Integrated) in Software Engineering

Vellore, Tamil Nadu, India

Technical Skills

Languages: Python, Bash/Shell, Java, SQL, HCL, JavaScript

AWS: EC2, Lambda, EventBridge, S3, EKS, ECS, CodeBuild, RDS, VPC, IAM, CloudWatch, SSM, SNS, Route 53, WAF, ElastiCache

IaC & CI/CD: Terraform, GitHub Actions (OIDC, self-hosted runners), Argo CD, Jenkins, GitLab CI, Ansible, CloudFormation

Containers & Orchestration: Docker (multi-arch builds), Kubernetes, Helm

Security & Compliance: Trivy, Gitleaks, Checkov, SonarQube, HashiCorp Vault, AWS Secrets Manager

Monitoring & Logging: CloudWatch, Prometheus, Grafana, ELK Stack, Datadog

Databases & Caching: PostgreSQL, MySQL, Redis, Elasticsearch, AWS RDS, Redshift

Azure: Azure AD (Entra ID), Virtual Machines, AKS, Azure Functions

Selected Projects & Presentations

NRDS Forecast Validation Dashboard | *Python, AWS Lambda, S3, GitHub Actions*

- Designed an automated pipeline that scans S3 forecast outputs across all VPUs after each daily national run, flags missing or malformed NetCDF outputs, and publishes an HTML health dashboard, turning hours of manual verification into a zero-touch check.

Self-Hosted ARM CI/CD Runner Platform | *AWS EC2 Graviton, Lambda, GitHub Actions*

- Built an ARM-based self-hosted GitHub Actions runner fleet with Lambda-driven start/stop scheduling and automated disk cleanup, enabling native arm64 builds and eliminating idle compute spend outside working hours.

Serverless Web Application | *AWS Lambda, API Gateway, DynamoDB, S3*

- Engineered a scalable serverless web application leveraging AWS Lambda and DynamoDB for low-latency CRUD operations via API Gateway, with S3 static hosting for seamless frontend-backend interaction and minimal operational overhead.

Research Poster: “The NextGen Research Datastream (NRDS)” | *Cloud-Based Numerical Prediction*

- Presented NRDS cyberinfrastructure research at **AGU 2025** (New Orleans) and **CIROH DevCon 2026** (Salt Lake City); accepted for poster presentation at **ALWRC 2026** on accelerating research-to-operations in hydrology.

Certifications & Recognition

AWS Community Builder – Serverless Category, 2026 Cohort

AWS Certified Solutions Architect – Associate | **HashiCorp Certified: Terraform Associate**

Certified Kubernetes Administrator (CKA) – KodeKloud Course | **Microsoft Certified: Azure Fundamentals**