

Aafaq Rana

Senior DevOps / SRE • AI & Agentic DevOps Engineer

+92 304 4869282 • m.aafaq.rana@gmail.com • linkedin.com/in/aafaq-rana-369176156 • aafaq-rana.space

PROFESSIONAL SUMMARY

Senior DevOps / SRE and AI engineer with 6+ years building and operating cloud platforms across AWS, Azure, GCP, and OCI, currently leading the DevOps team at Netsol Technologies. I run Kubernetes from managed EKS and AKS to fully air-gapped K3s on customer bare metal, serve LLMs on GPU node groups with vLLM and DCGM-based alerting, and owned the infrastructure control set through a SOC 2 audit. I also build and review AI agent evaluation work: Harbor-format benchmark tasks with deterministic verifiers, Terraform and AWS tasks graded against live cloud state, and pod leadership against a 14-dimension rubric. Python developer (FastAPI, Flask, Django), Terraform first, and a daily user of Claude Code and Cursor.

CORE COMPETENCIES

AI Engineering & Agentic DevOps: LangChain, LangGraph, LlamaIndex, Kubernetes MCP Server, AI Agent Development, Agentic Workflows, OpenAI API, Ollama, Autonomous Incident Response, Multi-Agent Orchestration

GPU & LLM Infrastructure: GPU Node Groups on EKS, NVIDIA Device Plugin, vLLM Model Serving, DCGM Metrics, KEDA, Karpenter

LLM Evaluation & ML: RLHF, pass@k, HumanEval, Preference Ranking, Reward Modeling, Prompt Engineering, Benchmark Task Authoring

Programming & Scripting: Python (FastAPI, Flask, Django), Go, TypeScript, Bash, SQL, C/C++

Cloud Platforms: AWS, Microsoft Azure, Google Cloud Platform (GCP), Oracle Cloud Infrastructure (OCI), DigitalOcean, Cloudflare (DNS, WAF)

Infrastructure as Code & Automation: Terraform, Pulumi, AWS CDK, CDKTF, CloudFormation, Ansible

CI/CD & Platform Engineering: Azure DevOps (Repos & Pipelines), GitHub Actions, Jenkins, SVN to Git migration, Docker, Kubernetes, Helm, ArgoCD, GitOps, Google Cloud Build

Kubernetes: Amazon EKS, Azure AKS, Google GKE, Oracle OKE, K3s, Air-Gapped Deployments, Harbor, Longhorn, Operators, RBAC, NetworkPolicy

Security & DevSecOps: IAM, KMS, Secrets Management, SOC 2 Compliance, OPA, Checkov, tfsec, Snyk, Sysdig, GitGuardian, Policy-as-Code, OWASP, Vanta

Monitoring & Observability: Prometheus, Grafana, Loki, Tempo, EFK/ELK Stack, OpenTelemetry, CloudWatch, Azure Monitor, Sentry, Chaos Engineering

Data & Messaging: AWS RDS and Aurora (MySQL, MariaDB, PostgreSQL), PostgreSQL pooling, replicas, failover and restore drills, Kafka, Lambda, SQS, SNS, Step Functions, EventBridge, Snowflake Cortex

PROFESSIONAL EXPERIENCE

Netsol Technologies

Dec 2021 – Present

Senior DevOps Engineer (DevOps Team Lead)

- **Leadership & DevOps Program Management:** Lead the DevOps team and manage multiple concurrent projects across cross-functional engineering groups, owning delivery planning, technical direction, stakeholder coordination, and the mentoring of DevOps and SRE engineers.
- **Legacy Modernization (Program Supervisor):** Supervised end-to-end modernization of legacy products, migrating source control from SVN to Azure DevOps Repos (git-svn) and re-platforming CI/CD from Jenkins to standardized Azure DevOps YAML pipelines, cutting build-and-ship time by 50% and improving delivery throughput by 30%.
- **GPU & LLM Infrastructure:** Built and operate GPU workloads on Amazon EKS with dedicated GPU node groups, the NVIDIA device plugin, and taints and tolerations for workload isolation, serving models with vLLM and streaming DCGM metrics into Prometheus so alerting fires on real GPU utilization rather than CPU proxies.
- **Agentic Kubernetes Operations (MCP-Based Failure Prevention):** Engineered a Kubernetes AI agent on the Kubernetes MCP Server to autonomously monitor cluster health, detect anomalies, and predict failures before production impact, driving proactive LangGraph-based remediation and reducing unplanned outages.
- **AI-Driven Documentation & DORA Uplift:** Applied agentic AI (LangChain, LlamaIndex) to automatically document legacy and internal codebases into a centralized RAG knowledge base on ChromaDB, accelerating onboarding and incident response to reduce MTTR and improve DORA metrics.
- **SOC 2 Compliance & Infrastructure Security:** Owned the infrastructure control set through a SOC 2 audit, covering IAM and KMS key policy, secrets management, network segmentation, and evidence collection, and embedded policy-as-code validation (OPA, Checkov, tfsec, Vanta) into CI/CD.
- **On-Premises Kubernetes (Air-Gapped):** Built and manage on-premises K3s clusters on customer bare metal, including fully air-gapped installs with mirrored images, charts, and operators, plus Harbor, Longhorn, PostgreSQL operators, RBAC, ingress, and cluster hardening.
- **Cloud & Hybrid Infrastructure Provisioning:** Design scalable cloud-native and hybrid infrastructure with Terraform and AWS CDK across AWS, Azure, GCP, and OCI, with modular code, per-environment state, reviewed plans in CI, secure networking, IAM, and encryption.

- **CI/CD, GitOps & DevSecOps:** Engineer multi-cloud CI/CD via Azure DevOps, GitHub Actions, and Google Cloud Build with ArgoCD GitOps across EKS, AKS, GKE, OKE, and on-premises K3s, integrating automated testing, Snyk scanning, and Sysdig monitoring aligned to OWASP best practices.
- **Autoscaling & Compute Cost:** Run Karpenter node pools with Spot capacity and consolidation, and KEDA for scaling on external signals, on Amazon EKS, so compute tracks real demand instead of static over-provisioning.
- **Observability & Chaos Engineering:** Built centralized observability with the Grafana LGTM stack (Loki, Tempo, Grafana) and Prometheus, plus Azure Monitor and Sysdig, for proactive alerting, incident investigation, resilience testing, and chaos engineering.

Turing

Jun 2025 – Jul 2026

Pod Lead, AI Agent Evaluation (Contract)

- **Pod Lead, Agent Benchmark Programs:** Pod lead on two agentic benchmark programs (OBI and Company Bench): review trainers' tasks against the client quality standard and a 14-dimension rubric, set pod standards, run pod-lead onboarding, and maintain the runbook that takes a task from claim to manual review.
- **Harbor Task Authoring:** Build Harbor-format agent tasks end to end, from fixture, instruction and golden answer to deterministic verifier and harness, gated on an oracle scoring exactly 1.0 and a multi-model pipeline (GLM, Claude Opus, GPT) whose difficulty bands reject tasks every model solves.
- **Evaluation Tooling & Failure Analysis:** Write Python tooling for Harbor checks, handoff verification and batch runs on remote Linux nodes, and read failure shapes in trajectories, for example treating a bimodal pass/fail pattern as instruction ambiguity rather than difficulty.

Vera Platforms

Jul 2026 – Sep 2026

Infrastructure Benchmark Engineer, IaC (Contract)

- **Terraform & AWS Benchmark Tasks:** Authored Terraform and AWS infrastructure tasks for AI agent training, built to fail frontier models at pass@4: setup provisions real AWS resources and seeds applied Terraform state into the agent's sandbox, the answer key stays hidden, and teardown cleans up.
- **Live-State Verification:** Wrote boto3 graders that check the live cloud end state rather than the agent's report, signing real API calls as different identities and retrying so eventual consistency is never misread as failure.
- **Trajectory Analysis & Fairness Review:** Analysed agent trajectories and failure causes across task revisions, caught a harness setting that made a task look easy, and wrote fairness rebuttals defending each task's solvability to client reviewers.

Devsinc

Jul 2020 – Dec 2021

Associate Software Engineer

- **CI/CD Pipelines:** Deployed Python FastAPI applications to EC2 via Azure DevOps pipelines with secure, automated release processes.
- **Infrastructure Management:** Used Terraform for infrastructure-as-code deployments across AWS and Azure, enabling consistent and reliable environments.
- **AWS Deployments:** Deployed frontend and backend applications on AWS with high availability and redundancy across multiple environments.
- **Application Development:** Built a web-based document processing application in Ruby on Rails, improving processing speed by 20%.

KEY PROJECTS

- **GPU Inference Platform on Amazon EKS:** Dedicated GPU node groups serving LLMs with vLLM, NVIDIA device plugin with taints and tolerations for scheduling isolation, DCGM metrics in Prometheus with Grafana alerting on real GPU utilization.
- **Agentic Kubernetes Failure-Prevention Agent:** Kubernetes MCP Server and LangGraph agent that monitors cluster health, predicts failures before production impact, and drives remediation with kubectl-native tooling.
- **Infrastructure Benchmarks for AI Agents:** Terraform and AWS tasks with hidden answer keys, live-resource graders and teardown, designed so a correct agent can win and a shortcut cannot.
- **Air-Gapped Kubernetes for Regulated Customers:** K3s on customer bare metal with no internet access, mirroring every image, chart, and operator including PostgreSQL operators, packaged as a repeatable install.

EDUCATION

B.Sc. Computer Science

2017 – 2021

School of Electrical Engineering and Computer Sciences (SEECs), NUST

CERTIFICATIONS

- AWS Certified Solutions Architect – Associate
- Microsoft Azure Fundamentals (AZ-900)