

# Arafath Mohammed

+1-9018604938 | Memphis, TN | arafathamer.mohammed@gmail.com | [LinkedIn](#) | [Portfolio & Git hub](#)

## PROFESSIONAL SUMMARY

---

Site Reliability and DevOps Engineer with 6 years of experience. In my current role, I bridge the gap between raw hardware power and usable research platforms by treating infrastructure as code. Built and supported HPC environments used for research and large-scale data processing across on-prem and remote infrastructure. Worked on automating day-to-day operational tasks, system configuration, monitoring, and deployment processes to reduce manual effort and improve reliability. Managed and optimized GPU-based systems used for AI and machine learning workloads, along with research storage, backup, and disaster recovery services, ensuring performance and compliance. Hands-on experience deploying CUDA-enabled GPU workloads, containerized AI runtimes (Docker, Singularity), and distributed training frameworks (PyTorch Distributed, DeepSpeed) on Kubernetes. Designing observability stacks (Prometheus, Grafana, OpenTelemetry), building CI/CD pipelines, and automating infrastructure provisioning with Ansible and Terraform across enterprise and research-grade environments.

## TECHNICAL SKILLS

---

**HPC & Compute:** Linux (RHEL, Ubuntu, CentOS), HPC Workload Schedulers (Slurm, IBM LSF), Distributed Job Workflows, Multi-GPU/Node Training, Accelerator-Aware Scheduling, Cluster Configuration & Validation

**AI/ML & GPU:** NVIDIA CUDA, cuDNN, NCCL, DCGM, PyTorch, TensorFlow, Jupyter, DeepSpeed, Horovod, Ray, vLLM, Hugging Face Transformers, MLflow, Experiment Tracking, Checkpoint Management, MLOps Workflows

**DevOps & Automation:** Ansible, Terraform, Git, GitHub Actions, Jenkins, CI/CD Pipelines, Infrastructure-as-Code, Bash, Python, Shell Scripting, Self-Service Automation, Systemd

**Containers & Orchestration:** Docker, Singularity, Kubernetes (EKS), Helm, Containerized AI Runtimes

**Observability & Monitoring:** Prometheus, Grafana, Loki, Tempo, OpenTelemetry, NVIDIA DCGM Exporters, CloudWatch, PagerDuty, Alerting, Dashboards, KPI Tracking, Capacity Planning

**Storage & Data:** High-Performance Storage (NFS, Lustre-adjacent), SAN/NAS, Backup & Archival Services, Data Lifecycle Management, Data Cataloguing, Tiered Storage, Disaster Recovery

**Cloud & Security:** AWS (EC2, S3, SageMaker, EKS, IAM), GCP (Compute Engine, GKE), Azure, Vulnerability Assessment, Patch Coordination, Access Control, HIPAA-Compliant Environments, Incident Response

**Generative AI Productivity:** Prompt Engineering, AI-Assisted Scripting & Troubleshooting, LLM-Powered Documentation, Log Analysis Automation, Workflow Generation, GenAI Output Validation & Security Awareness

## PROFESSIONAL EXPERIENCE

---

### AI-ML Platform Engineer

Nov 2023 – Present

FedEx (via Ameya Infotech) | Memphis, TN

- Operated and maintained GPU-accelerated compute infrastructure supporting AI/ML training and inference workloads across a hybrid on-premises and cloud environment, ensuring 99.9% platform availability for data-intensive research and production pipelines.
- Deployed NVIDIA Tesla T4 GPU environments with CUDA, cuDNN, and NCCL stacks, serving containerized AI models (PyTorch, TensorFlow) via FastAPI with API key authentication and systemd service management.
- Built comprehensive GPU observability using NVIDIA DCGM Exporters, Prometheus, and Grafana dashboards, tracking GPU utilization, memory, thermal metrics, and scheduler activity across multi-node clusters, reducing incident detection time by 60%.
- Designed and implemented CI/CD pipelines using Jenkins, GitHub Actions, and Ansible for automated infrastructure provisioning, configuration management, platform testing, and zero-downtime deployments across compute and storage tiers.
- Automated recurring HPC operational tasks through Python and Bash scripts — including node health checks, storage allocation workflows, job troubleshooting diagnostics, and self-service provisioning mechanisms — reducing manual effort by 45%.
- Managed containerized AI runtimes using Docker and Singularity, ensuring reproducible compute environments with pinned dependencies, validated CUDA compatibility, and consistent GPU driver versions across training and inference nodes.
- Implemented performance and utilization monitoring for compute nodes, GPU resources, storage systems, and backup operations building dashboards, health checks, and alerting mechanisms via Prometheus and Grafana.
- Reduced average job queue wait times by 30% by tuning Slurm scheduler policies adjusting fair-share weights, parameters, & partition priorities across multi-node GPU clusters running distributed PyTorch training workloads.

- Supported data lifecycle management across active storage, backup systems, and archival tiers implementing retention policies, access controls, traceability, and tier-to-tier data movement workflows for compliant and research data.
- Coordinated incident response and root cause analysis for infrastructure disruptions, performing vulnerability assessments, patch coordination, and backup validation to maintain secure and resilient HPC services.
- Applied generative AI tools (Claude, ChatGPT) to accelerate scripting, automation development & log analysis.
- Collaborated with research teams, storage engineers, and security teams to translate operational requirements into scalable infrastructure solutions, maintaining detailed technical documentation and runbooks.

### **Site Reliability Engineer**

*Jun 2023 – Nov 2023*

*Grant Thornton | Remote*

- Achieved 99.9% uptime across 12 production services on Kubernetes clusters with HPA handling 3x traffic spikes; reduced incident resolution time by 40% via structured root cause analysis framework and escalation.
- Built modular Terraform + Ansible infrastructure-as-code pipelines for automated provisioning of compute, storage, and networking resources, cutting environment setup from 3 days to 4 hours, used by 20+ engineers.
- Implemented observability stack using Prometheus, Grafana, and Loki for centralized metrics, log aggregation, dashboards, and alerting tracking scheduler activity, resource utilization, and service health KPIs across distributed infrastructure.
- Containerized multi-tier applications using Docker and Kubernetes, enabling 50% faster deployments, 30% infrastructure cost reduction, and zero environment drift between staging and production through reproducible container images.
- Rolled out automated patch cycles and configuration hardening across 200+ Linux hosts using Ansible playbooks triggered through Jenkins, covered SOX, PCI-DSS, and internal audit controls with zero remediation items two quarters running.
- Supported data management workflows including backup validation, restore readiness testing, and disaster recovery drills for critical production services and regulated data environments.

### **ERC Associate**

*Feb 2021 – Aug 2021*

*Amazon | Hyderabad IN*

- Improved the efficiency and user experience of internal HR tools by offering feedback on system usability.
- Identified gaps in application performance, collaborated with the product and development teams to suggest actionable improvements.

### **Site Reliability Engineer**

*Jan 2019 – Feb 2021*

*Web Initiate | Remote*

- Developed Python backend services deployed on AWS EC2 and Elastic Beanstalk, building REST APIs and automation scripts for enterprise applications supporting high-throughput data processing workflows.
- Authored Terraform infrastructure-as-code for AWS environments VPCs, subnets, security groups, NAT gateways, and compute instances establishing repeatable provisioning patterns for development, staging, and production tiers.
- Built SQS-based asynchronous message processing pipelines and DynamoDB data models for high-throughput workloads, optimizing read/write capacity with TTL-based data lifecycle policies.
- Created monitoring dashboards using CloudWatch and Grafana with alerting thresholds for API health, database performance, storage utilization, and message queue depth establishing early observability practices.
- Implemented distributed tracing and code quality workflows using Sleuth, Zipkin, and SonarQube, improving troubleshooting efficiency and service reliability across microservice architectures.

### **Cloud Engineer**

*Jun 2018 – Dec 2018*

*Zobibox | Remote*

- Wrote Python and Bash scripts for infrastructure automation integrating with Ansible and Terraform for resource deployment and plugging into CI/CD pipelines for repeatable provisioning workflows.
- Supported Linux system administration, shell scripting, and containerization using Docker and ECS, assisting in migrating legacy on-premises applications to AWS cloud infrastructure.
- Managed MySQL and DynamoDB databases, built Spring Boot microservices integrated with AWS services (SQS, RDS, Lambda), and contributed to infrastructure documentation and operational runbooks.

## **EDUCATION**

### **Masters in Computer Science**

*May 2023*

*Christian Brother University, Memphis TN*

**CERTIFICATIONS**

---

- AWS Solutions Architect - Associate
- AWS Cloud Practitioner