CHUNYANG (FRANK) ZHU

+1 (236) 999-9661 ⋄ Burnaby, BC, Canada ⋄ chunyang.zhu1990@gmail.com

SUMMARY

Senior SRE with deep experience in AI platform reliability, Kubernetes infrastructure, and cloud automation. Led global operations for Microsoft AI Translator, driving 99.9% availability, secure model releases, multi-region Kubernetes operations, and SLO-driven incident reduction.

EDUCATION

M.S. Cybersecurity, New York Institute of Technology - Vancouver 2023 – 2024
M.S. Computer Science, Washington University in St. Louis 2014 – 2016
B.S. Computer Science, University at Buffalo, SUNY 2010 – 2014

SKILLS

Cloud & Infra: Azure, Kubernetes (AKS), Docker, OpenShift
Programming: Python, Java, Go
Observability: Prometheus, Grafana, Splunk, Datadog
Databases & Msg: MongoDB, MySQL, Couchbase, Kafka
DevOps: Azure DevOps, CI/CD Pipelines, VMware, Terraform

EXPERIENCE

Microsoft Feb 2025 – Present
Site Reliability Engineer Lead (Contract via CSI Interfusion) *Remote*

- Operated multi-region Linux/Kubernetes environments for global AI inference services, ensuring 99.9% availability while performing node and VM-level troubleshooting beyond Kubernetes workloads.
- Managed lifecycle operations for Linux VMs underpinning AKS clusters, including provisioning, patching, scaling, node replacement, and resolving OS/network/storage issues impacting service health.
- Designed and optimized production-grade AKS clusters with security hardening, autoscaling improvements, and enhanced diagnostics, reducing MTTR by 50%.
- Built automated CI/CD pipelines for microservices and AI models, enabling safe canary deployments, traceable rollouts, and weekly global releases.
- Led large-scale Linux patch automation and CVE remediation across distributed VM fleets, meeting Cyber EO compliance and standardizing secure release governance.
- Developed observability using Prometheus/Grafana, implementing SLO-based alerting, performance monitoring, and integrating Exporter metrics into modern monitoring workflows.
- Automated operational workflows using Python, Go, and Bash, improving deployment, monitoring, and troubleshooting efficiency.
- Coordinated a distributed 6-engineer SRE team, establishing DRI/on-call workflows, escalation processes, and operational best practices across global services.

Branch Metrics 2020 – 2021
Solution Engineer *Remote*

- Built and deployed mobile attribution integrations for enterprise clients, improving onboarding efficiency and reducing integration time by 40%.
- Designed end-to-end SaaS data flows using REST APIs, webhooks, and event-based pipelines to support high-scale marketing analytics.

NVIDIA 2017 – 2018
System Software Engineer *Shanghai, China*

- Developed and maintained microservices for the NVIDIA Gaming Platform (50M+ users), enabling scalable game distribution and user services.
- Rebuilt CI/CD pipelines with automated testing and structured logging, reducing deployment time and improving release reliability.
- Implemented service-level monitoring and debugging tools to enhance production visibility and accelerate issue resolution.

Walmart Inc. 2016 – 2017
Programmer Analyst *Bentonville, AR, USA*

- Built Azure-based Java microservices supporting robotics-driven inventory automation, improving system efficiency and operational throughput.
- Implemented event-driven workflows and messaging using Azure services to support large-scale warehouse operations.
- Collaborated with robotics and infrastructure teams to ensure reliability, service scalability, and smooth CI/CD delivery.