

KELVIN NGO

PRODUCTION OPERATIONS ENGINEER | HYPERSCALE AI INFRASTRUCTURE

kelvinngok@gmail.com | linkedin.com/in/kelvinngo

Production operations engineer with 9+ years of experience operating large-scale data center infrastructure and deploying Meta's first hyperscale AI training data centers. Combines hands-on Linux, GPU platform, hardware, firmware, network, and storage troubleshooting with global incident command, fleet analytics, automation, and cross-functional technical leadership.

CORE EXPERTISE

Hyperscale AI Data Centers | GPU Training & Inference | Linux Systems | Fleet Reliability | Incident Command & RCA | Python & Bash Automation | OCP Hardware | Firmware & Out-of-Band Management | Telemetry & Grafana | Technical Leadership

PROFESSIONAL EXPERIENCE

Production Operations Engineer

May 2022 - Present

Meta

- Led deployment and production stabilization for Meta's first hyperscale AI training data centers, supporting GPU training and inference infrastructure across 17+ global campuses.
- Owned platform health for NVIDIA, AMD, and MTIA environments, driving safe scale-up, high availability, and rapid recovery for mission-critical AI workloads.
- Served as Global Incident Commander for high-severity, multi-region events; coordinated technical response, stakeholder communications, risk decisions, and service restoration.
- Drove root cause analysis across Linux, networking, storage, firmware, and OCP server hardware, translating local failures into fleet-wide corrective actions and long-term design changes.
- Built Python and Bash automation and diagnostic tooling to identify failure patterns, reduce manual troubleshooting, and scale operational response across thousands of servers.
- Analyzed fleet telemetry, failure rates, utilization, and SLA performance to detect systemic risk and prioritize reliability and serviceability improvements.
- Partnered with AI, hardware, firmware, architecture, and vendor teams to move new platforms from prototype and site introduction into sustained mass production.
- Authored global runbooks, standardized operating processes, and mentored engineers in a 24/7 production environment.

PROFESSIONAL EXPERIENCE - CONTINUED

Senior Data Center Engineer athenahealth

May 2018 - May 2022

- Planned data hall designs and deployment strategies across three large data centers.
- Led the Dallas data center shutdown and Richardson migration while protecting operational continuity.
- Maintained the firmware catalog and deployed fleet updates across approximately 10,000 servers.
- Configured and troubleshot Linux servers, hardware, storage, networking, and performance issues.
- Managed vendors, structured cabling, operating budgets, burst labor, and onsite delivery.
- Built and trained the onsite team and documented repeatable operations in Confluence SOPs.

NOC Technician II TD Ameritrade

Jan 2017 - May 2018

- Monitored trading services for 11M+ client accounts holding more than \$1T in assets.
- Resolved hardware, software, application, and automated-job incidents in production.
- Coordinated outage response across internal support groups and third-party service providers.
- Provided first-level diagnosis for automated job applications and business-critical workflows.
- Communicated impact, escalation status, and restoration progress during service events.

TECHNICAL SKILLS

Systems & Automation: Linux (Red Hat, Ubuntu, openSUSE), Python, Bash, Ansible, Grafana, Nagios, JIRA, Confluence

Compute & Hardware: OCP, NVIDIA/AMD GPU, MTIA, Dell PowerEdge, HPE, Supermicro, IPMI/RACADM, OS imaging, firmware lifecycle

Network & Storage: Cisco switching/routing, structured fiber, Pure Storage, NetApp, RAID, DNS, DHCP, HTTP, VMware

EDUCATION

Bachelor of Applied Arts and Science, Information Technology | Tarleton State University, 2017