

Abhishek Muthukumar

muthukua@purdue.edu · (+1) 7657720340 · github.com/abhishekm2610 · linkedin.com/in/abhishekm2610

Summary

Graduate student at Purdue University with a strong background in Kubernetes and **open source** cloud-native technologies. My research focuses on **optimizing GPU utilization** for Kubernetes-based inference of generative AI models by leveraging Large Language Model (LLM) metrics. I am seeking full-time opportunities to apply my expertise in **LLM inference, Kubernetes, storage systems, software engineering, networking, and cloud systems**.

Education

Purdue University, West Lafayette **Aug 2024 - May 2026**
MS in Computer and Information Technology GPA: 4.0/4.0
Relevant Coursework: Cloud, DC, & Carrier Networks, Cloud Infrastructure, HCI, and Applied Statistics

Anna University, India **Aug 2016 - May 2020**
B.Eng in Computer Science and Engineering CGPA: 8.28/10
Relevant Coursework: Programming and Data Structures I & II, Design and Analysis of Algorithms, Artificial Intelligence, Software Engineering

Experience

Veeam **May 2025 - Present**
Kasten Engineering Intern Remote, US

- Designed and implemented core portability features for the **open-source Kopia** backup tool, enabling repositories to be moved seamlessly across storage back-ends and boosting user adoption and deployment flexibility.

Purdue's Rosen Center for Advanced Computing (RCAC) **Aug 2024 - Present**
Graduate Research Scientist West Lafayette, IN

- Building Purdue GenAI Studio with dynamic pod autoscaling in a GPU-enabled **Kubernetes** cluster on the **Anvil Supercomputer**, **optimizing GPU resource utilization** and enhancing system performance.
- Developing a **Kubernetes Operator for real-time resource accounting** and billing, leveraging the Kubernetes Metrics API and time-series storage to track CPU/memory usage per user, enabling cost attribution.

Cisco Systems, Inc. **Aug 2023 - Aug 2024**
Software Engineer Bangalore, India

- Implemented APIs in a **microservice-based application**, developed a queue-based notification service to enhance real-time communication and system alerts.
- Developed applications that reduced manual effort by **70%**, **saving more than 3000 hours a quarter**.

Technical Skills

Languages & Frameworks: Go, C, Python, JavaScript
Technologies: Docker, Kubernetes, Ansible, CI/CD, GitOps, Prometheus, Grafana, OpenShift, Jenkins, RabbitMQ, MongoDB, PostgreSQL, SQLite, GKE, AWS (EC2, S3, EKS), Linux, TCP/IP

Certifications:

- Lean Six Sigma Black Belt** from Purdue University (Dec 2024)
- Cisco Certified Network Associate (CCNA)** (Jul 2020)

Projects

- Developed an automated dashboard to visualize network traffic flows in real-time to identify congestion and mis-configuration issues, used by over 30 customers at Cisco (Sep 2020 - Dec 2020)
- Streamlined service centered network migration automation through Ansible, designed the driver engine and dashboards (Oct 2020 - Jun 2021)
- Developed a Smart Alarm Service for remotely triggering alarms in smartphones over SMS in remote-sensing products (Feb 2019)

Awards & Extracurricular

- Co-authored the peer-reviewed paper "Providing On-Prem GenAI Inference Services to a Campus Community" presented at **PEARC25** at Columbus, Ohio (Jul 2025)
- Represented Purdue's Rosen Center for Advanced Computing at **SC24** at Atlanta, Georgia (Nov 2024)
- Attended the **KubeCon + CloudNativeCon North America 2024** at Salt Lake City, Utah (Nov 2024)
- Board Member of **Cloud Native Association at Purdue (CNAP)** (Aug 2024)
- Awarded at IEEE YESIST12 MakerFair, for **Remote Water Quality Monitoring System**, held at Thailand (Sep 2019)