

Sudhakar Chundu

P: +1 513-666-0099

E: chundubabu@gmail.com

December 16, 2025

Dear Hiring Team,

I am excited to apply for the role focused on GPU clusters and Kubernetes operations. With over 20 years of experience in AI/ML infrastructure and DevOps, I have a proven track record of building and managing large-scale, high-performance systems.

At Trackonomy Systems, I architected serverless GPU infrastructure, reducing inference costs by 65% while maintaining 99.97% uptime for 5M+ daily predictions. I also deployed multi-platform inference pipelines and built real-time observability stacks using Prometheus and Grafana, aligning closely with your focus on GPU telemetry and monitoring. My hands-on experience with Kubernetes, GPU scheduling, and Infrastructure-as-Code tools like Terraform and Helm has consistently delivered reliable, scalable solutions.

Your emphasis on optimizing high-performance networking and leading on-call rotations resonates with my work at AWS, where I ensured 99.95% availability for ML workloads and implemented GitOps-based releases across 15+ regions with zero downtime.

I am eager to bring my expertise in containerized GPU stacks, CI/CD pipelines, and DevOps practices to support your mission of advancing generative models. Let's discuss how I can contribute to your team's success.

Thank you for considering my application. I look forward to the opportunity to connect.

Best regards,
Sudhakar Chundu