



LEADER. BUILDER. TECHNOLOGIST.

Technical Leader and Platform Engineer specializing in GPU-accelerated media AI, edge computing, and real-time video analytics. Currently serving as a technical lead and architect for a consultancy delivering production-grade AI platforms for Fortune 100 clients and professional sports leagues.

Proven track record designing multi-tenant, bare-metal and container based platforms, enabling partner ecosystems, and shipping production-ready solutions leveraging NVIDIA NIMs, Blueprints, and Holoscan for Media. Known for bridging hardware, software, and stakeholders to turn emerging AI capabilities into scalable products that solve real-world media and broadcast challenges.

EDUCATION

American University - Graphic Design
Community College of Allegheny County

SKILLS

Core Expertise

- GPU-Accelerated Media AI and Edge Platforms
- Distributed and Multi-Tenant Systems Architecture
- Real-Time Video Analytics and Low-Latency Pipelines
- Technical Leadership, Architecture Strategy and Partner Enablement
- Customer-Facing Solution Architecture (Enterprise & Media)

NVIDIA & AI Stack

- NVIDIA NIMs, Blueprints, Holoscan for Media
- AI Inference at Scale (LLMs, VLMs)
- GPU Performance Optimization
- Object Detection, Segmentation & VLM Pipelines (YOLO, SAM, RF-DETR)

Platform & Infrastructure

- Kubernetes (OpenShift, EKS, Bare Metal)
- Helm, Docker, Containerization
- GPU-Accelerated On-Prem + Hybrid Systems
- Prometheus, Grafana, Kafka
- Ansible, Terraform

Product & Execution

- Technical Product Leadership and Roadmapping
- Cross-Company Collaboration and Joint Development
- Rapid Prototyping and Proof-of-Concept Delivery

Code Languages

Python | React (Next.js) | Tailwind CSS | REST API
MySQL | PostgreSQL | SQLite | PNMP

Technical Lead / Architect / Staff Engineer monks.com

May 2023 - Present

- Technical Lead and Platform Architect for a Fortune 100 customer, designing a multi-tenant edge AI media platform that bridges bare-metal GPU infrastructure and AI inference services to deliver LLM and VLM workloads at the edge. Enabled customers to self-serve and launch media applications within the NVIDIA Holoscan for Media ecosystem, accelerating adoption of GPU-accelerated edge workflows.
- Drove cross-company technical alignment across NVIDIA partners and customers, leading collaborative architecture and design sessions to standardize on NVIDIA NIMs, reference frameworks, and Helm-based deployment patterns. Advanced the containerization and GPU enablement of traditional broadcast software, improving portability, lifecycle management, and performance on NVIDIA GPUs.
- Solution Architect for real-time AI broadcast pipelines, designing systems that dynamically adjust live video bitrate based on subject detection and scene understanding, deployed on a three-node Kubernetes cluster with 8x NVIDIA L40 GPUs, demonstrating scalable, low-latency AI inference for live media.
- Technical Lead and Sole Engineer for an on-premises AI activation supporting unscripted television production, delivering a GPU-accelerated logging and analytics platform for post-production teams. Reduced key scene discovery from hours to minutes using RTX PRO 6000 Blackwell Max-Q, enabling near-real-time editorial workflows.
- Architected and deployed AI video analytics solutions using NVIDIA NIMs and Blueprints, including VSS-driven sports analytics pipelines for football strategy analysis, play outcome detection, and automated statistics generation.
- Platform Architect for Monks Live.Vision, leading proof-of-concept development, reference architectures, and technical strategy for live sports and media use cases, translating customer needs into scalable, GPU-accelerated platform designs aligned with NVIDIA's media and AI roadmap.

Director of Product Development fullmeasure.io

Feb 2020 - May 2023

- Managed distributed team of 20+ engineers, oversaw a platform serving millions of users at hundreds of colleges ensuring platform uptime and release management.
- Applied AWS cloud solutions to optimize cloud usage, reducing monthly spend by over 50% leading to over \$300,00 in annual savings.
- Designed and implemented a comprehensive CI/CD pipeline for a monorepo, overseeing the deployment of 28 distinct services using docker & kubernetes.
- Provided robust infrastructure support and maintenance for six environments that facilitated the sending of 1 million SMS messages monthly.

Technical Product Manager fullmeasure.io

May 2017 - Feb 2020

- Managed over 500 concurrent integrations, including the maintenance of critical code responsible for processing 10 million ETL transactions per day
- Utilized A/B testing and market research to better the process of client integrations, improved implementation timeline from 3 months to 2 weeks.
- Collaborated with creative director using figma to refine new feature sets, led to creating a front-end component library in stencil.js to improve development time, taking an average 3 point story to a 1.
- Lead of transitioning customer base from legacy product to new platform, supported by 28 different micro-services, involved many stakeholders and proper time management on a client by client basis to ensure live sms communications weren't interrupted.

Product Owner & QA Engineer fullmeasure.io

May 2016 - 2017

- Owner of mobile application development, utilized the ionic mobile framework, successfully building and maintaining hundreds of android and iOS applications.
- Utilized rspec tests to implement a rigorous and comprehensive testing strategy, ensuring clean and thoroughly tested code for each feature.
- Lead of qa engineering team of four - responsible for full coverage testing on all delivered feature sets in the stage environment.

Graphic Design & UI/UX Buzzmark Inc.

May 2014 - 2016

- Designed and edited logos, created promotional web and print materials, and provided comprehensive support for marketing initiatives.
- Led the redesign of a website to utilize WordPress, resulting in streamlined internal management and enhanced UI/UX.
- Developed an internal CRM system to manage daily manual updates required for in-app content, improved manual task times from a few hours to minutes.