

FRED W. NIX JR.

Infrastructure Solutions Architect | AI Infrastructure & Presales Engineer

Atlanta, GA
frednix@gmail.com · (404) 953-8305
linkedin.com/in/frednix · github.com/nixfred ·
resume.nixfred.com

SUMMARY

Twenty plus years of enterprise infrastructure and presales engineering (EMC, Dell, Pure Storage, Reduxio), now applied to private AI. Senior Solutions Architect at BlueAlly leading technical discovery, architecture, sizing, and business case for data center, storage, virtualization, network, and security across the Southeast, with private GenAI carried alongside the infrastructure practice. In 2026: 24 customer engagements spanning VMware exits, storage refreshes, data center builds, identity, and AI use-case portfolios worth \$21M+ in modeled annual value. Designs GPU nodes, all flash storage, and virtualization platforms by workload, sized to the customer's numbers. Builds with AI daily through a persistent Claude Code infrastructure of his own design.

EXPERIENCE

Senior Solutions Architect · BlueAlly Technology Solutions, Atlanta, GA

Apr 2026 – present

Infrastructure architect for enterprise, healthcare, higher education, state and local government, and nonprofit customers: data center, storage, virtualization, network, and security. Own discovery, architecture, BOM, TCO/ROI modeling, RFP responses, POC design, and executive narrative; carry the AI use-case and private GenAI work alongside the infrastructure practice. Customers below are anonymized; figures are real.

Infrastructure and data center

- **VMware exit across three continents** for a global corporate travel company (Atlanta, Frankfurt, Singapore): discovery, hosted POC, licensing strategy, parallel three-site migration plan against staggered Broadcom renewals; stretched storage over NVMe/TCP on all flash arrays avoiding ~\$500K of Fibre Channel fabric.
- **Infrastructure modernization** for a 190+ property extended-stay hotel chain: 480-core VMware entitlement review, legacy SAN retirement (34 TB), flash array capacity planning at 91% utilization, backup/DR discovery, five-year TCO for leadership.
- **Multi-vendor data center BOM** (servers, storage, firewalls) for an immersion-cooled colocation operator: 250,000 sq ft, three data halls, 18 MW with a 24.5 MW site cap, GPU hosting relationship behind it.
- **Hyperconverged to disaggregated refresh** (five node cluster to two socket servers, all flash NVMe array, private cloud automation) for an Atlanta media ministry, remaining on VMware by design.
- **215 TB video surveillance archive** storage discovery for a Florida state water management district: 130 cameras, four sites, commercial VMS, tiering and protocol design.
- **TCO and RFP work:** 48-host virtualization break-even model (materials manufacturer); hypervisor migration evaluation (aerospace and defense contractor); all flash storage RFP (major international airport); SIP trunking and SBC eRFP for a State of Georgia research university with BlueAlly as accountable prime integrator.

Security and identity

- **Cloud RADIUS and EAP-TLS migration** for a multi-state HIPAA healthcare provider (~100 sites, 600+ daily users): retire on-prem NPS and 100 branch VPN tunnels, certificate lifecycle delivered as a service on the existing cloud managed wireless estate.
- **Identity and access management** at a global agriculture and food company (150,000 employees, 77 countries, ~2,000 legacy apps): Microsoft E5 posture, secure service edge, physical/logical access convergence, AI governance.
- **Contract compliance save:** identified a managed-services decommission that would have dropped monitored assets below contracted minimums at a higher-education account; escalated and preserved revenue and standing.

Private AI and GenAI

- **On-prem GenAI reference architecture:** 2U four GPU node, 4× 96 GB GPUs (384 GB per node), private AI appliance stack, orchestration and governance layers; sized to ~14M peak monthly tokens; packaged as a fixed-price 60 to 90 day POC to decouple from CIO budget gating.
- **Nine use-case AI portfolio** for an AZA-accredited, USDA Class R research institution: ~\$12.1M modeled annual value, 17,800 reclaimed hours (8.5 FTE); appliance sized from 109 TB to 343 TB of KV retention; five-year lease above \$2M; data sovereignty as the wedge against cloud-only competitors.
- **Multi-domain AI campaign** at an HBCU academic medical center (clinical documentation, research ops, student success, reporting): six prioritized use cases, \$2.4M modeled annual savings, ~1,300 reclaimed hours.

- **Contract intelligence agents and vendor SOC security** for a healthcare payments technology company: platform vs custom build evaluation, pass-through delivery model, \$290K to \$305K budgetary proposal.
- **AI workshops and opportunity assessments:** office REIT (\$6.7M perceived value, Copilot Studio vs agent platform on Fabric), commercial HVAC platform (1,200 employees, 33 sites), county government, state financial regulator, retirement services firm, EPC engineering firm (AI readiness and Copilot Studio SOWs, VDI to 350 users, HA/DR), youth services nonprofit.
- **OEM partner funding:** own the SOW for a \$62.5K OEM demand-generation program (customer discovery workshop plus AI certification training for the BlueAlly sales and SA team). Coach engineers on Claude and MCP; present at the monthly tech talk (~50 attendees).
- **~1.5-day demo and prototype turnaround** using a personal toolchain of 150+ repositories and customized Claude Code instances with persistent memory and local model routing.

Storage & Compute Practice Manager, AI Pod Architect · Business Technology

Mar 2025 – Apr 2026

Architects, Alpharetta, GA

- Built a production-grade, air-gapped AI inference platform for channel partners on modular blade compute, 48 GB GPU clusters, all flash storage, and 100GbE; one-command deployment from bare metal to running inference with a 100% success rate across partner rollouts.
- Multi-model serving (LLaMA, Mistral, custom fine-tunes) with hot-swap switching, vLLM, AWQ/GPTQ quantization, tensor parallelism; OpenAI-compatible API and ChatGPT-style UI with streaming and live GPU metrics (DCGM, Prometheus, Grafana) for LangChain/LlamaIndex RAG.
- Trained channel partners on AI infrastructure deployment with hands-on labs and documentation.

Partner, VP Sales · EMOTO Electric Motorbikes

2023 – 2025

- Sales up 225% in year one; expanded the product line and scaled to regional market leader.

Enterprise Sales Engineer · Pure Storage

Global Presales Director · Reduxio

2015 – 2019

- Fortune 100 accounts across the Southeast. 250% of quota in 2021, Global Rookie of the Year. 100% uptime through a hurricane disaster recovery event.

- First American employee; led US market entry. 150%+ of quota annually, 90%+ conversion across 50+ POC projects.

Converged Systems Presales vArchitect

Global Presales Operations, vSpecialists PM · EMC Corporation

- Operations for 5,500 presales engineers worldwide; trained 1,000+ new engineers; grew vSpecialists from 30 to 150+.

· Dell Technologies

- Software-defined storage and cloud presales; major data center consolidation initiatives.

Earlier senior roles

Syscom, MTI, GT Software, Vitria, Windward, Kilpatrick Stockton, Medaphis: CTO, solutions consulting, presales, professional services, IT administration.

SKILLS

Infrastructure: Two socket and four GPU rack servers, modular blade compute, all flash NVMe arrays, scale out NAS, VMware vSphere/VCF, hyperconverged and disaggregated designs, private cloud automation, Proxmox, NVMe/TCP, 100GbE, SD-WAN, backup & DR, colocation & immersion cooling, VDI.

AI & GPU: GPU sizing from 48 GB and 96 GB cards to 384 GB nodes and multi node clusters; private AI appliance stacks; vLLM; AWQ/GPTQ quantization; tensor parallelism; RAG; OpenAI-compatible APIs; LangChain, LlamaIndex; Microsoft Copilot Studio; MCP servers; Ollama local inference.

Security: zero trust access, EAP-TLS and cloud RADIUS, air-gapped deployment, HIPAA and PCI scoped design, IAM, AI governance.

Presales & business: discovery workshops, use-case portfolios and value modeling, TCO/ROI/break-even, BOM, RFP and eRFP responses, POC design, lease and pass-through commercial structures, OEM partner funding programs, executive briefings.

Development & automation: TypeScript and Bun, Python, Claude Code and Personal AI Infrastructure, git-based knowledge systems, Cloudflare Pages and Workers, Linux (Omarchy, Hyprland) plugin author, Prometheus and Grafana, GitOps.

AI-augmented practice: Since January 2026, runs "Larry," a Personal AI Infrastructure on Claude Code across four hosts with version-controlled memory, local-GPU-first model routing, and a written constitution of operating rules; 10+ production sites and 30+ desktop plugins shipped.

Education: B.S. Business Administration, Information Systems, Auburn University. **Recognition:** Pure Storage Global Rookie of the Year, 2021 (250% of quota); annual quota achieved every year of career.