

FRED W. NIX JR.

Principal Solutions Engineer | Infrastructure & AI
Platforms

Atlanta, GA

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

SUMMARY

Enterprise solutions engineer with 20+ years across infrastructure and presales (EMC, Dell, Pure Storage, Reduxio) and a current practice in private AI. Most recently Senior Solutions Architect at BlueAllly: led technical discovery, architecture, sizing, and the business case for data center, storage, virtualization, network, and security engagements across the Southeast, and carried the private GenAI work alongside that practice. 24 customer engagements in five months, ~\$150K closed in the first 90 days, a seven-figure qualified pipeline, and AI use-case portfolios worth \$21M+ in modeled annual value. Designs GPU nodes, all-flash storage, and virtualization platforms by workload, not by logo, and builds with AI daily on a persistent Claude Code platform of his own design.

PROFESSIONAL EXPERIENCE

Senior Solutions Architect · BlueAllly Technology Solutions, Atlanta, GA

Apr 2026 – Oct 2026

Infrastructure architect for enterprise, healthcare, higher education, state and local government, and nonprofit customers. Owned discovery, architecture, BOM, TCO/ROI modeling, RFP responses, POC design, and executive narrative; carried the AI use-case and private GenAI work alongside the infrastructure practice. Customers anonymized; figures real.

- **VMware exit across three continents** for a global corporate travel company (Atlanta, Frankfurt, Singapore): discovery, hosted POC, licensing strategy, and a 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 that closed as a signed \$95K order in the first 90 days, legacy SAN retirement (34 TB), flash array capacity planning at 91% utilization, backup/DR discovery, five-year TCO; follow-on storage modernization in flight at ~\$210K gross profit.
- **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.
- **Additional infrastructure engagements:** hyperconverged-to-disaggregated refresh (two-socket servers, all-flash NVMe array) for a media ministry; 215 TB, 130-camera video archive design for a state water district; all-flash storage RFP for a major international airport; SIP trunking and SBC eRFP for a state research university with BlueAllly as accountable prime; 48-host virtualization TCO model; hypervisor migration evaluation for an aerospace and defense contractor.
- **Security and identity:** cloud RADIUS and EAP-TLS migration for a multi-state HIPAA healthcare provider (~100 sites, 600+ daily users), retiring on-prem NPS and 100 branch VPN tunnels with certificate lifecycle as a service; IAM engagement at a global agriculture company (150,000 employees, 77 countries, ~2,000 legacy apps) reaching into Microsoft E5 posture, secure service edge, and AI governance.
- **Private GenAI reference architecture:** 2U four-GPU node (4× 96 GB, 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.
- **AI use-case portfolios:** nine use cases for an AZA-accredited research institution (~\$12.1M modeled annual value, 17,800 reclaimed hours, five-year lease above \$2M, appliance sized 109 to 343 TB); six use cases for an HBCU academic medical center (\$2.4M modeled savings); contract intelligence agents for a healthcare payments company (\$290K to \$305K proposal); Copilot Studio vs agent platform evaluation for an office REIT (\$6.7M perceived value); signed AI readiness and Copilot Studio SOWs (~\$55K) at an EPC firm; SOW owner for a \$62.5K OEM-funded AI demand-generation program.
- **~1.5-day demo and prototype turnaround** on a personal toolchain of 150+ repositories and customized Claude Code instances; coached engineers on Claude and MCP; monthly tech talk (~50 attendees).

AI Solutions Architect, Storage & Compute Practice Manager · Business Technology Architects, Alpharetta, GA

Mar 2025 – Apr 2026

- 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; automated 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 vLLM, AWQ/GPTQ quantization, tensor parallelism, and hot-swap switching; OpenAI-compatible API and streaming chat UI for LangChain/LlamaIndex RAG; live GPU metrics via DCGM, Prometheus, Grafana. Partner enablement through hands-on labs and implementation guides.

Partner, VP Sales · EMOTO Electric Motorbikes 2023 – 2025

Entrepreneurial operating role; sales up 225% in year one, expanded the product line, scaled to regional market leader.

Enterprise Sales Engineer · Pure Storage 2021 – 2023

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

Solutions Architect, Converged Systems · Dell Technologies 2019 – 2020

Technical presales for software-defined storage, cloud architecture, and data center consolidation.

Global Presales Director · Reduxio 2015 – 2019

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

Global Presales Operations & Technical Enablement · EMC 2010 – 2015

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

Earlier technical leadership

CTO, solutions consulting, presales, professional services, and systems administration at Syscom, MTI, GT Software, Vitria, Windward, Kilpatrick Stockton, Medaphis.

SELECTED AI ENGINEERING PROJECTS

Agent engineering platform (Larry)

github.com/nixfred

Four-host Claude Code platform with MCP integrations and local-first model routing (12 GB local GPU before any cloud call); persistent agent memory in TypeScript/Bun with SQLite full-text plus vector search, lifecycle hooks, a written constitution of operating rules, and scheduled recovery of missed session extraction.

AI software factory

Reusable AI-assisted delivery workflow (TypeScript/Bun, Python, Cloudflare Workers/Pages, GitOps) that has shipped 10+ production sites and 30+ desktop plugins, and turns a customer conversation into a demo, a sizing, and a deck in about a day and a half.

AI Systems Workbench

tools.nixfred.com

Fourteen browser-based, on-device instruments for designing and evaluating AI applications: prompt lab, token and cost planner, retrieval lab, evaluation workbench, agent designer, least-privilege permission planner, failure investigator, drift monitor, AI stack mapper, and more.

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 and DR, colocation and immersion cooling, VDI, Linux (Arch, Ubuntu), Bash, systemd, Docker, K3s, Tailscale.

Inference & GPU: GPU sizing from 48 GB and 96 GB cards to 384 GB nodes and clusters; private AI appliance stacks; vLLM, Ollama, multi-model serving, AWQ/GPTQ, tensor parallelism, air-gapped AI, DCGM, Prometheus, Grafana.

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

Solutions engineering: technical discovery, solution strategy, use-case portfolios and value modeling, POC design and evaluation, executive presentations, competitive positioning, TCO/ROI/break-even, BOM, RFP and eRFP responses, lease and pass-through structures, OEM partner funding, partner enablement.

Agents & developer platforms: AI agent workflows, Claude Code, MCP, TypeScript/Bun, Python, Cloudflare Workers/Pages, OpenAI-compatible APIs, RAG, LangChain, LlamaIndex, Copilot Studio, Git/GitOps, Linux (Omarchy, Hyprland) plugin author.

EDUCATION & SERVICE

B.S. Business Administration, Information Systems · Auburn University

U.S. Army veteran

RECOGNITION

Pure Storage Global Rookie of the Year, 2021 (250% of quota). Annual quota achieved every year of career.