

FRED W. NIX JR.

Principal Solutions Engineer | AI Applications & Developer Platforms

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

SUMMARY

Enterprise solutions engineer with a foundation in Linux systems and infrastructure and a current practice building AI agents and developer platforms. Connects technical discovery and hands-on prototypes to executive business cases: at BlueAlly, 24 customer engagements in five months, AI use-case portfolios worth \$21M+ in modeled annual value, a private GenAI reference architecture sized to ~14M peak monthly tokens, and a ~1.5-day demo-to-prototype turnaround on a personal Claude Code platform. Twenty-plus years of presales across EMC, Dell, Pure Storage, and Reduxio underneath it.

PROFESSIONAL EXPERIENCE

Senior Solutions Architect · BlueAlly Technology Solutions, Atlanta, GA

Apr 2026 – Oct 2026

Supported enterprise, healthcare, higher education, and public-sector customers across the Southeast. Architected solutions spanning GPU compute, NVMe storage, VMware and private cloud, networking, security, and AI. Customers anonymized; figures real.

- **Owned technical discovery**, architecture, workload sizing, BOMs, TCO/ROI, RFP responses, and POC design; presented recommendations to executives. Built tailored demos and prototypes in ~1.5 days using Claude Code and reusable AI workflows. ~\$150K closed in the first 90 days and a seven-figure qualified pipeline.
- **Nine use-case AI portfolio** for an AZA-accredited research institution: ~\$12.1M modeled annual value, 17,800 reclaimed hours; translated operational requirements into infrastructure sizing (109 to 343 TB of vector and KV retention) and a five-year lease business case built on private AI and data sovereignty.
- **Private GenAI reference architecture** on a 2U four-GPU node (4× 96 GB, 384 GB per node) spanning appliance stack, orchestration, and governance for ~14M peak monthly tokens; structured a fixed-price 60 to 90 day POC to support phased investment decisions.
- **AI discovery workshops and platform evaluations**: six use cases at an HBCU academic medical center (\$2.4M modeled savings); contract intelligence agents and vendor SOC security for a healthcare payments company (platform vs custom build evaluation, \$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. Coached engineers on Claude and MCP; monthly tech talk (~50 attendees).
- **VMware exit across three continents** for a global travel company: discovery, hosted POC, licensing strategy, three-site migration plan; NVMe/TCP stretched storage design avoiding ~\$500K in Fibre Channel fabric.
- **Infrastructure and security engagements**: 190+ property hotel chain modernization (signed \$95K VMware order, ~\$210K GP storage follow-on in flight); 18 MW immersion-cooled colocation BOM; cloud RADIUS and EAP-TLS migration across ~100 HIPAA healthcare sites; IAM at a 150,000-employee global agriculture company; 215 TB video archive; airport storage RFP; state university SIP/SBC eRFP.

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

Mar 2025 – Apr 2026

- Built an air-gapped AI inference platform for enterprise deployment on GPU clusters, all-flash storage, and 100GbE; automated one-command deployment from bare metal to running inference with 100% success across partner rollouts.
- Implemented multi-model serving for LLaMA, Mistral, and custom fine-tunes using vLLM, AWQ/GPTQ quantization, tensor parallelism, and hot-swap model switching; exposed an OpenAI-compatible API and streaming chat UI supporting LangChain/LlamaIndex RAG; integrated live GPU metrics with DCGM, Prometheus, and Grafana.
- Delivered partner enablement through deployment labs and implementation guides for channel partners.

Partner, VP Sales · EMOTO Electric Motorbikes

2023 – 2025

Entrepreneurial operating role; sales up 225% in year one, expanded the product line.

Global Presales Director · Reduxio

2015 – 2019

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

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 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

Built a four-host Claude Code platform with MCP integrations and local/cloud model routing (12 GB local GPU before any cloud call). Developed 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

Created a reusable AI-assisted software delivery workflow using TypeScript/Bun, Python, Cloudflare Workers/Pages, and GitOps; shipped 10+ production sites and 30+ desktop plugins. The same tooling accelerates customer demos and prototypes.

AI Systems Workbench: application evaluation & agent governance

tools.nixfred.com

Fourteen browser-based tools for designing and evaluating AI applications, all processed on device: compare outputs against explicit rubrics, flag critical failures and coverage gaps, investigate behavior drift, decompose workflows, design agents with least-privilege permissions, budget latency, and map the full request path.

TECHNICAL SKILLS & SOLUTIONS ENGINEERING

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.

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

Solutions engineering: technical discovery, solution strategy, POC evaluations, executive presentations, competitive positioning, value engineering, TCO/ROI, RFPs, OEM partner funding, partner enablement.

Linux & infrastructure: Linux (Arch, Ubuntu), Bash, systemd, SSH, Docker, K3s, Tailscale, backup/recovery, infrastructure automation, VMware/VCF, Proxmox, hybrid and private cloud, all-flash NVMe, 100GbE, networking, IAM, zero trust, 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.