ONSTAK

Start Executing AI. Stop Experimenting.

AI INFRASTRUCTURE · MENTORED INSTALL

We integrate. We're not an integrator.

We design, deploy, validate, and hand over your AI POD infrastructure — one team, no handoff gap. We build it, and we stay to help you run it.

Built For

Healthcare
Financial Services
Public Sector
Manufacturing
Retail

4	Day 1	Full stack	Your team
Stages	**Validated under real load**	**Compute to GPU apps**	**Trained to operate it**
Design → deploy → test → transfer	Tested before you trust it	One validated environment	Hands-on knowledge transfer

The Problem

Hardware racked. Value stalled.

AI infrastructure arrives, then sits — misconfigured, unvalidated, owned by no one. We build and run as one team, so the stack is production-ready and your people can own it.

What It Does

- → **Validate the design**
  POD sizing, hardware / software interoperability, and architecture review for scale.
- → **Deploy the full stack**
  Compute, unified management, virtualization, containers, and GPU-accelerated AI frameworks.
- → **Test under real load**
  End-to-end validation and a live model deploy to confirm production readiness.
- → **Transfer ownership**
  Hands-on training and runbooks so your team operates and scales it confidently.

AI POD Mentored Install · what's included

ENGAGEMENT — 4 STAGES

01

Solution design & architecture validation

- AI POD sizing tailored to generative-AI workloads
- Hardware & software interoperability validation
- Architecture review for scalability, performance, reliability

02

Deployment & configuration

MANDATORY SERVICES

- Unified Computing System (UCS X-Series / AI POD) — compute & storage scalability
- Nexus Dashboard & Intersight — unified management
- Virtualization layer — VMware vSphere or Nutanix
- Red Hat OpenShift — containerized GPU applications
- NVIDIA AI Enterprise — accelerated AI frameworks

OPTIONAL SERVICES

- External storage integration (NetApp or Pure)
- External connectivity (DCN / SAN / Internet)
- Security — perimeter, host, certificates
- Monitoring & observability — logs, metrics, traces

03

Testing & validation

- End-to-end testing under production-like conditions
- Deploy a model to confirm workload readiness

04

Knowledge transfer

- Hands-on training for IT and operations teams
- Comprehensive guides for managing and scaling the AI POD

ALSO

Deliverables

- Infrastructure setup
- Software stack installation
- Operationalized AI use case
- Knowledge transfer and guidance

Assumptions

- AI PODs racked & stacked by customer / channel partner
- Remote access provided for deployment
- Required hardware & software licenses in place

Duration: **4–10 weeks** · Engineers: **Cisco and NVIDIA-certified.**

1525 McCarthy Blvd, Suite 1230, Milpitas CA 95035 - USA +1 855 811 2345 sales@OnStak.com