CISCO Connect

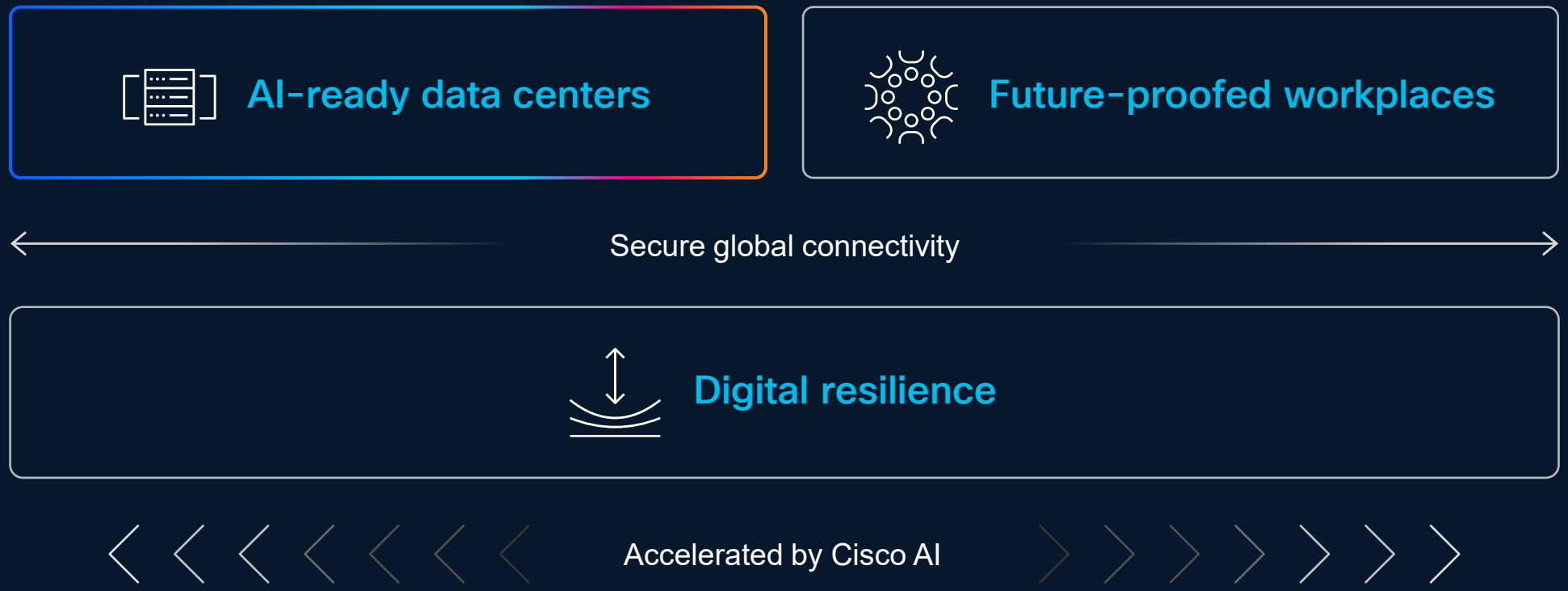
Full Stack AI-Ready Data Centers

Built for AI, ready for what's next



Charlie DeMarco
Obserservability & AgenticOps SE Director, @social

Cisco powers how people and technology work together across the physical and digital worlds



© 2026 Cisco and/or its affiliates. All rights reserved.

AI demand is spreading across every industry

Manufacturing

Predictive maintenance

Quality control

Demand forecasting



Public sector

Smart cities

Security and safety

Services improvement





Retail

Personalization

Inventory optimization

Sales forecasting



Financial services

Fraud detection

Risk assessment

Trading



Healthcare

Diagnosis

Drive-thru optimization

Patient support

Education

Learning & teaching experiences

Smart & secure facilities



© 2026 Cisco and/or its affiliates. All rights reserved.

AI adoption is high.
Scaling to production is still hard.

88%
Enterprises use AI today

<40%
Have scaled beyond pilots

70–80%
AI projects still fail





What breaks when AI becomes a production data center workload?

Infrastructure
GPU, storage, power, cooling, and fabric

Platform & data path
Kubernetes, pipelines, placement, dependencies

Security
Trust zones, policy, risk and governance

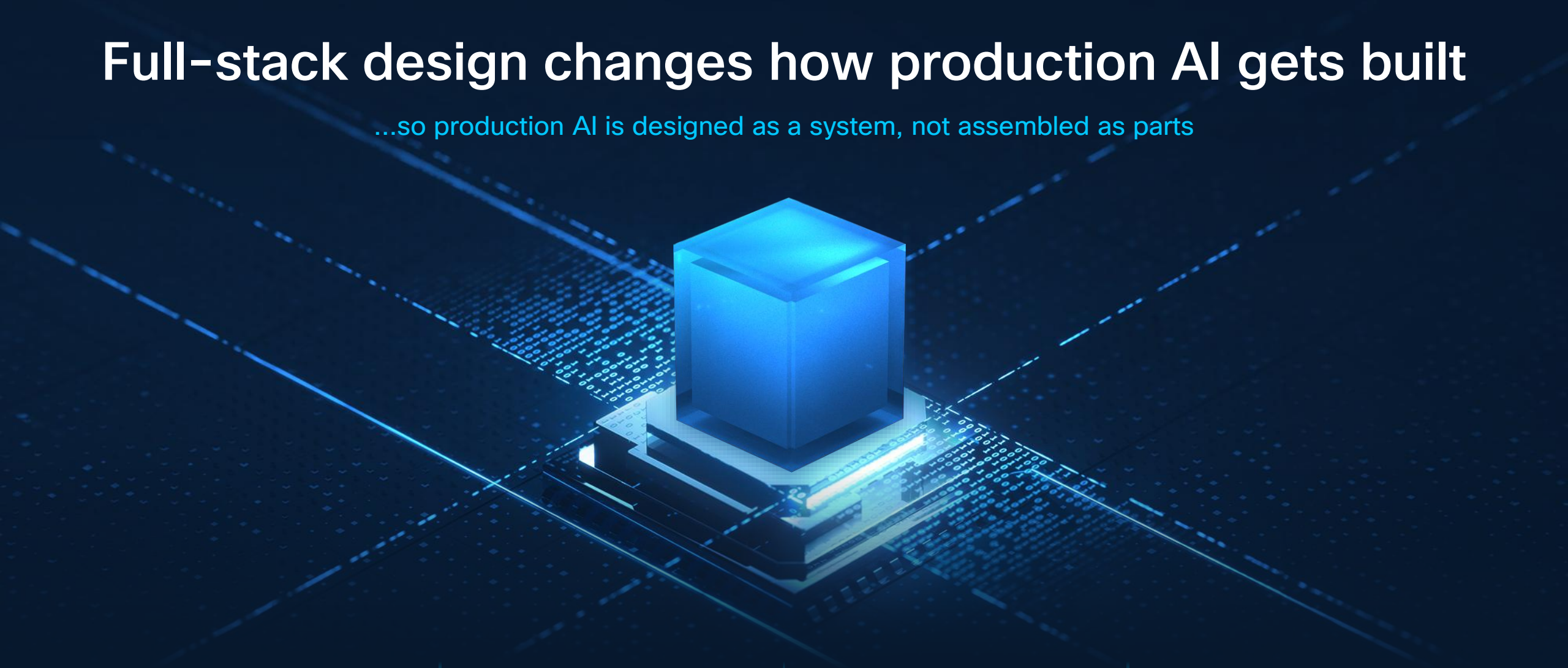
Operations
Visibility, ownership, lifecycle

© 2026 Cisco and/or its affiliates. All rights reserved.

Sources: Deloitte State of AI 2026, McKinsey, NVIDIA State of AI 2026, Cloudera/HBR 2026

Full-stack design changes how production AI gets built

...so production AI is designed as a system, not assembled as parts



Validated architecture

Capacity, fabric, security, operations

Faster integration

Fewer handoffs, fewer surprises

Day-one Operations

Visibility, lifecycle, change control

Trust & accountability

Policy, ownership, governance

© 2026 Cisco and/or its affiliates. All rights reserved.

cisco

Cisco AI-Ready Data Centers

Full-stack infrastructure, engineered as one system

Distributed security fused into every layer

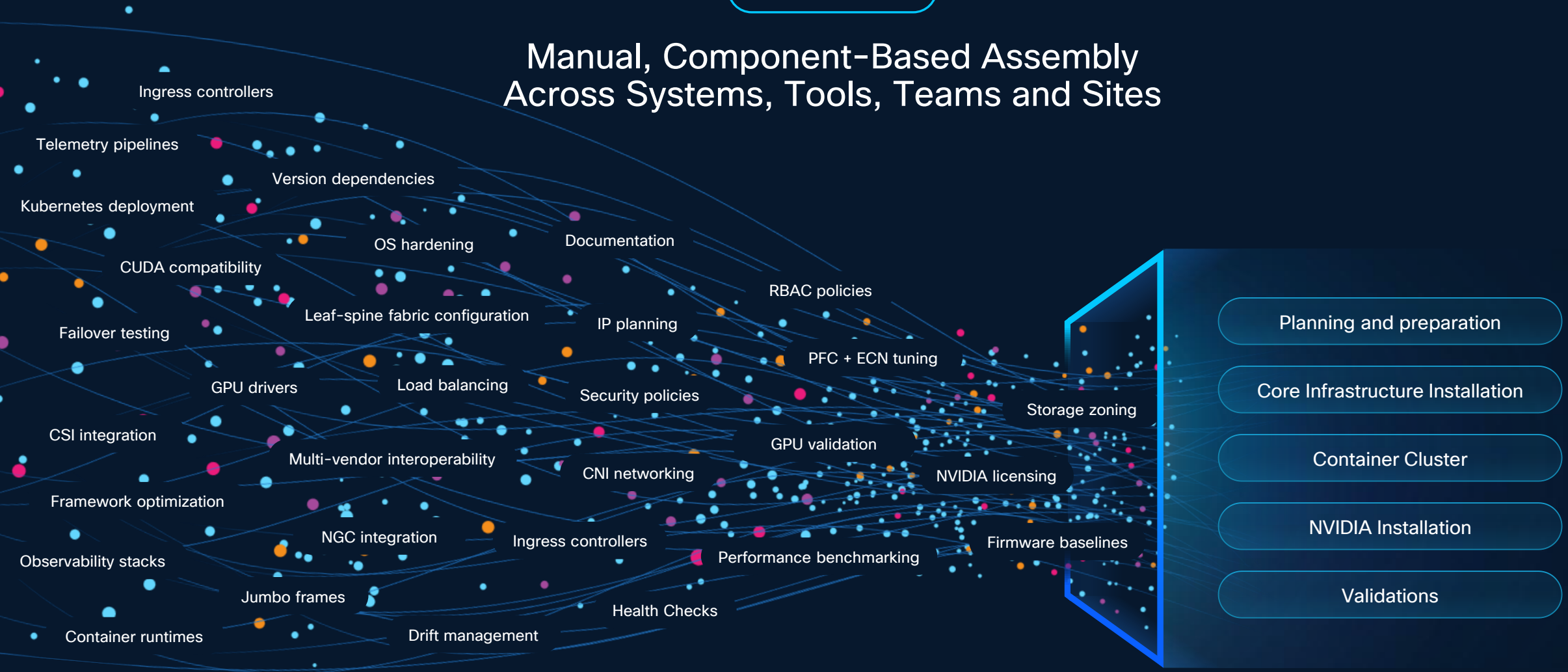
AgenticOps aligning speed with precision

© 2026 Cisco and/or its affiliates. All rights reserved.

cisco

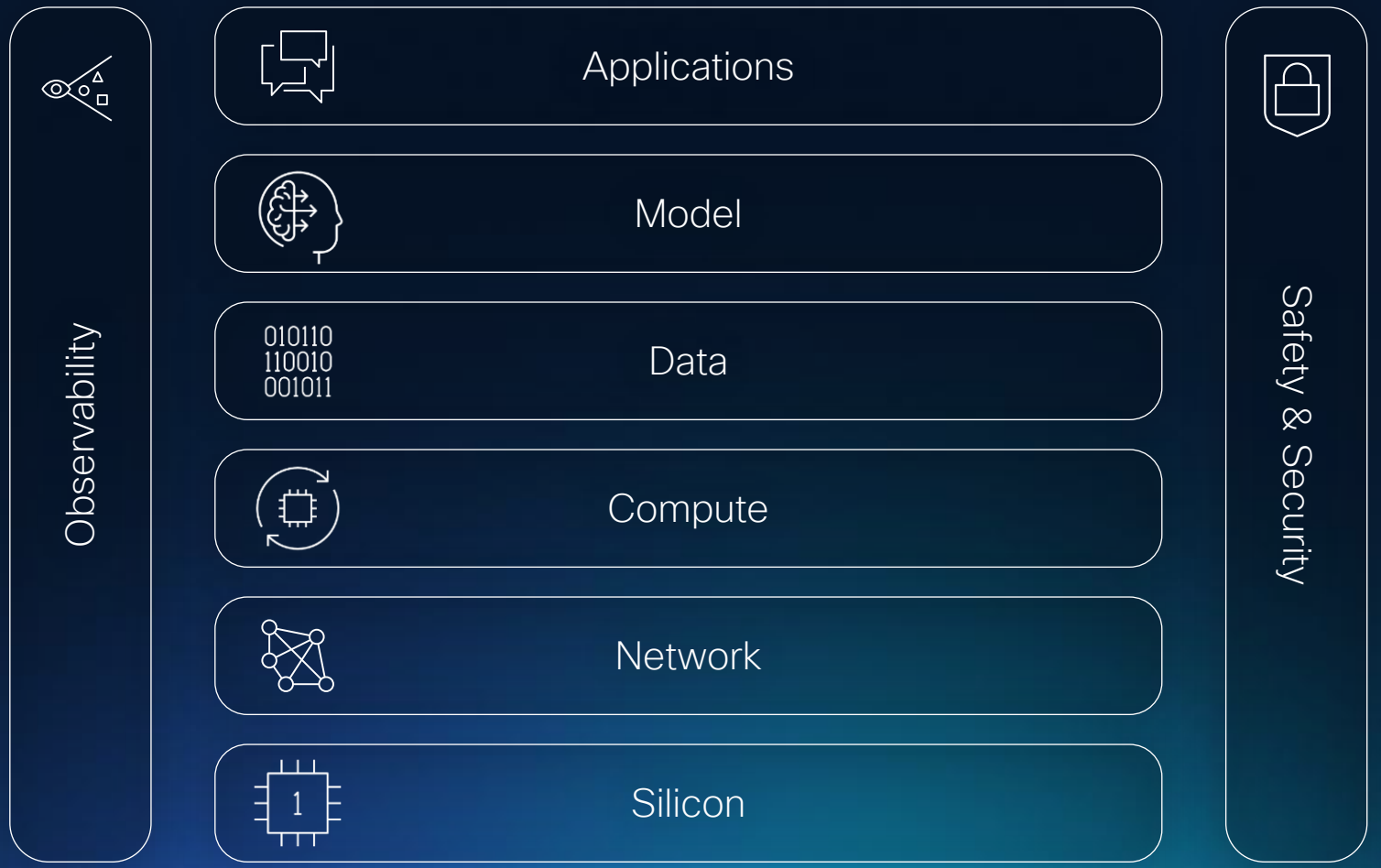
BEFORE

Manual, Component-Based Assembly Across Systems, Tools, Teams and Sites



25-40	300-700+	40-100+	High	WEEKS
Tool configurations	Infrastructure actions	Handoffs & approvals	Operational Risk	Deployment Time

Cisco is the **critical infrastructure for the AI era**

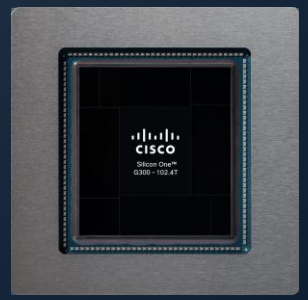


A full-stack approach to AI

© 2026 Cisco and/or its affiliates. All rights reserved.

Recent AI networking innovations

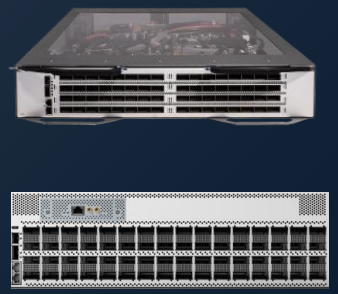
SILICON



Cisco Silicon One G300

102.4 Tbps
Network Switching
Silicon

SYSTEMS



Advanced Scale-Out Systems

Cisco N9000 +
Cisco 8000 Systems

Cisco N9100 Systems
with NVIDIA Spectrum-X
Ethernet

OPTICS



High-Quality Optics

Cisco OSFP 1.6T
Pluggable Optics

Cisco OSFP 800G Linear
Pluggable Optics

SOFTWARE



Enhanced Software Features

AI Job Observability
Native Splunk Integration
AgenticOps Capabilities

OPERATING MODEL

Cisco
Nexus
One

Nexus One

Unified operations across on-premise and cloud-based deployments with flexible OS support

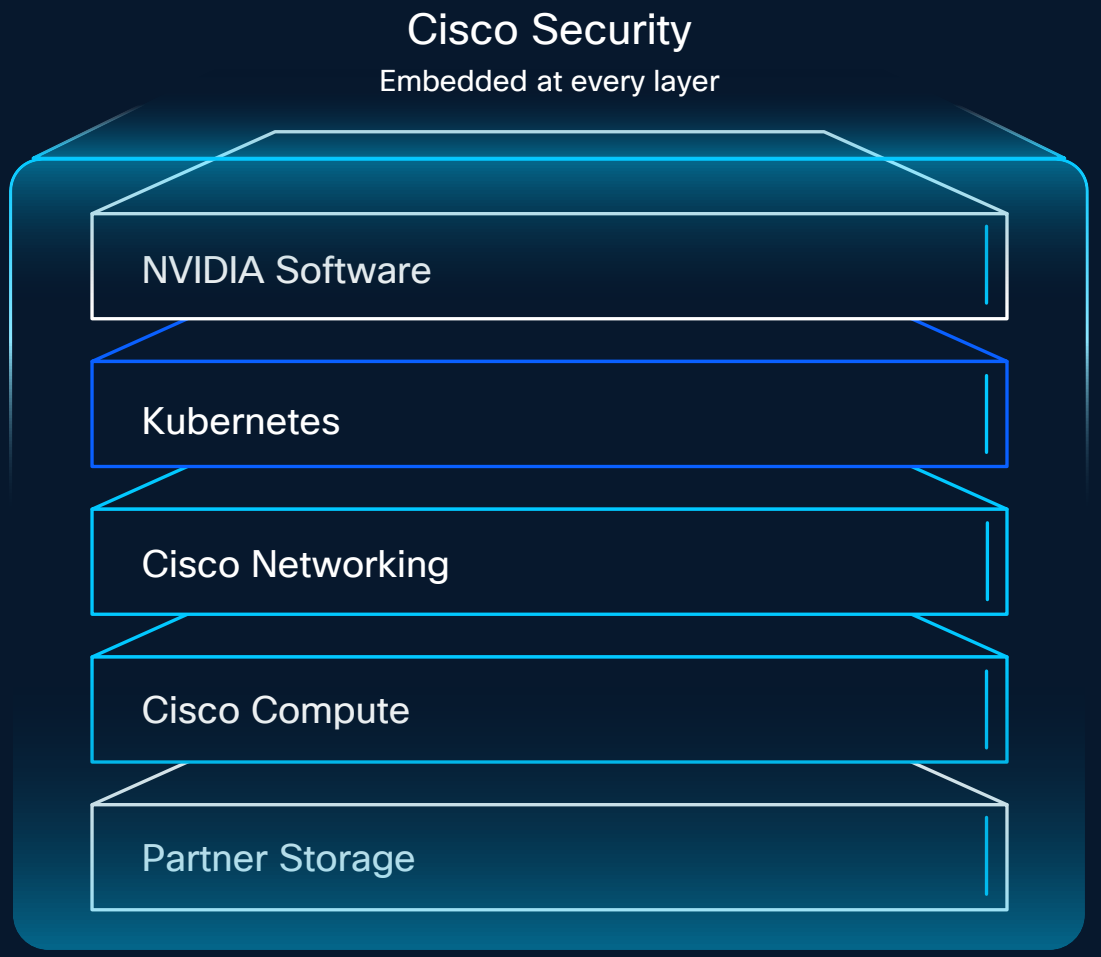
SCALE-OUT | HYPERSCALE | ENTERPRISE | NEOCLOUD | SOVEREIGN CLOUD | SERVICE PROVIDERS

© 2026 Cisco and/or its affiliates. All rights reserved.

Cisco Secure AI Factory with NVIDIA

A reference design with validated architectures to **accelerate AI adoption**

- Security-first architecture
- High-performance networking, compute, storage, and AI software
- Flexible deployment options: Vertically integrated or modular system



© 2025 Cisco and/or its affiliates. All rights reserved.

Unified approach to enterprise and AI workloads

Broad spectrum of Cisco AI servers to meet any requirement

TRAINING

OPTIMIZATION & INFERENCING

INFERENCING



Dense GPU Servers



HGX Servers w/
(8) B300 GPUs



RTX Pro 6000 & 4500 GPU
Servers w/ up to 8 GPUs

Modular Servers



RTX Pro 6000 and
4500 GPU Servers

Unified Edge



RTX Pro 6000 & 4500,
L40S, L4 GPU Platform

Cisco Intersight – Simplified operations and fleet management

CORE — EDGE

© 2026 Cisco and/or its affiliates. All rights reserved.

cisco

Cisco IT deploys AI-ready data center in weeks

See the full story

Challenges

Critical business need with short build timeline

Existing data center facilities weren't built to power modern AI

Extensive technical requirements including ethernet-based lossless fabric

Operations for new environment needed to be automated

Solutions

Cisco 8100 Series embedded with Silicon One

Cisco N9000 Series running NX-OS

Cisco Nexus Dashboard

SONiC running on Cisco 8000 series switches

Results

Speed of deployment: AI Cluster was designed and deployed in 3 months, 80% faster than alternative rebuild

Enhanced AI capabilities: AI Cluster supports more than 25 use cases with more added each week

Designed to support future needs: Maximized agility to power any workload, traditional or AI, at massive scale, no matter where the data lives

> Not only were we able to support the needs of the business today, but we're designing how our data centers need to evolve for the future.
>
> The modularity and flexibility of Cisco's infrastructure gives us confidence that we can keep scaling with the business."

Jon Woolwine
Director, Technical Systems Engineering, Cisco IT

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Cisco AI Network Fabric keeps the AI loop moving

Front-end services feed the AI loop. Back-end fabric keeps accelerators moving.

Cisco fabric building blocks

Build the fabric

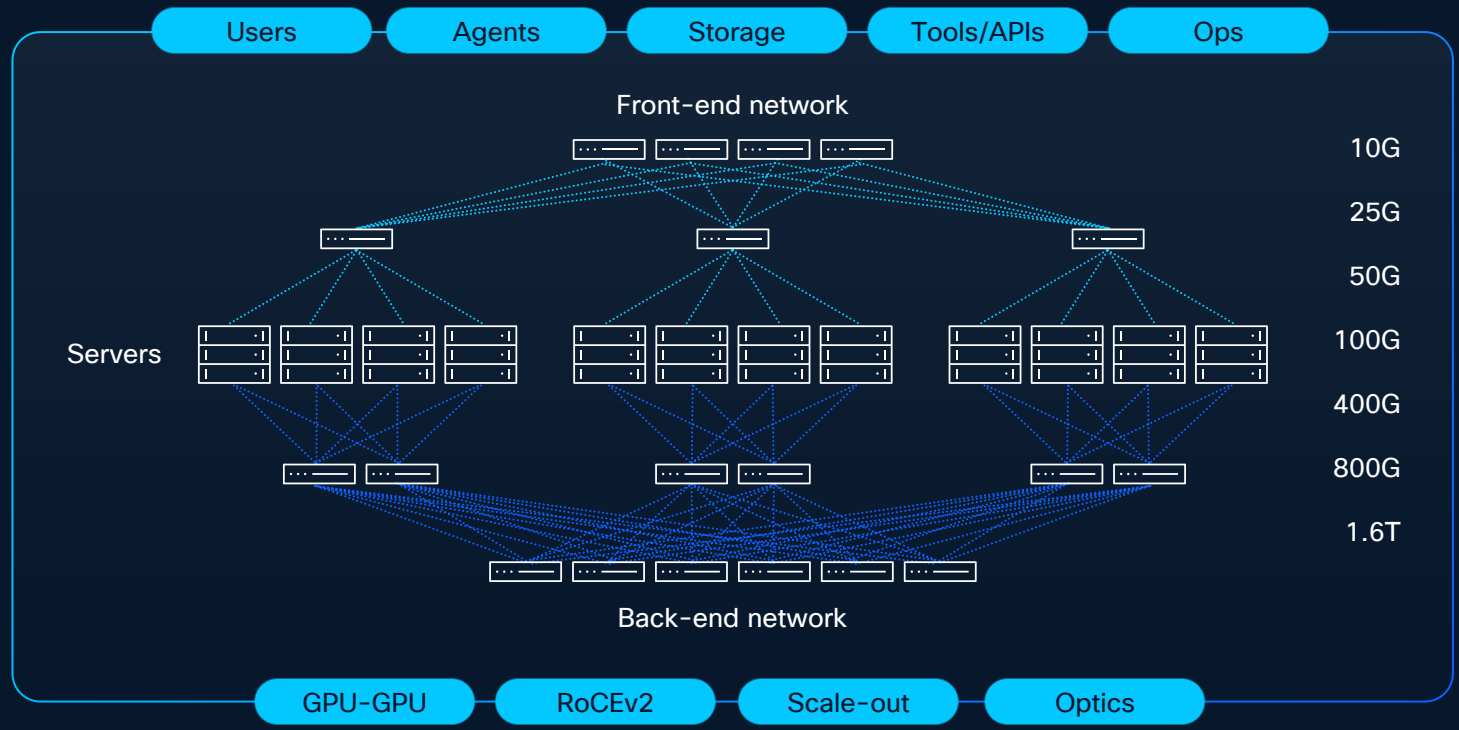
- Cisco N9000
- Cisco Silicon One | NVIDIA Spectrum-X Silicon
- Cisco Optics
- NX-OS | ACI | SONiC

Tune the AI traffic

- RoCEv2 lossless Ethernet
- Intelligent Packet Flow
- Intelligent buffering | deep and shared

Operate it

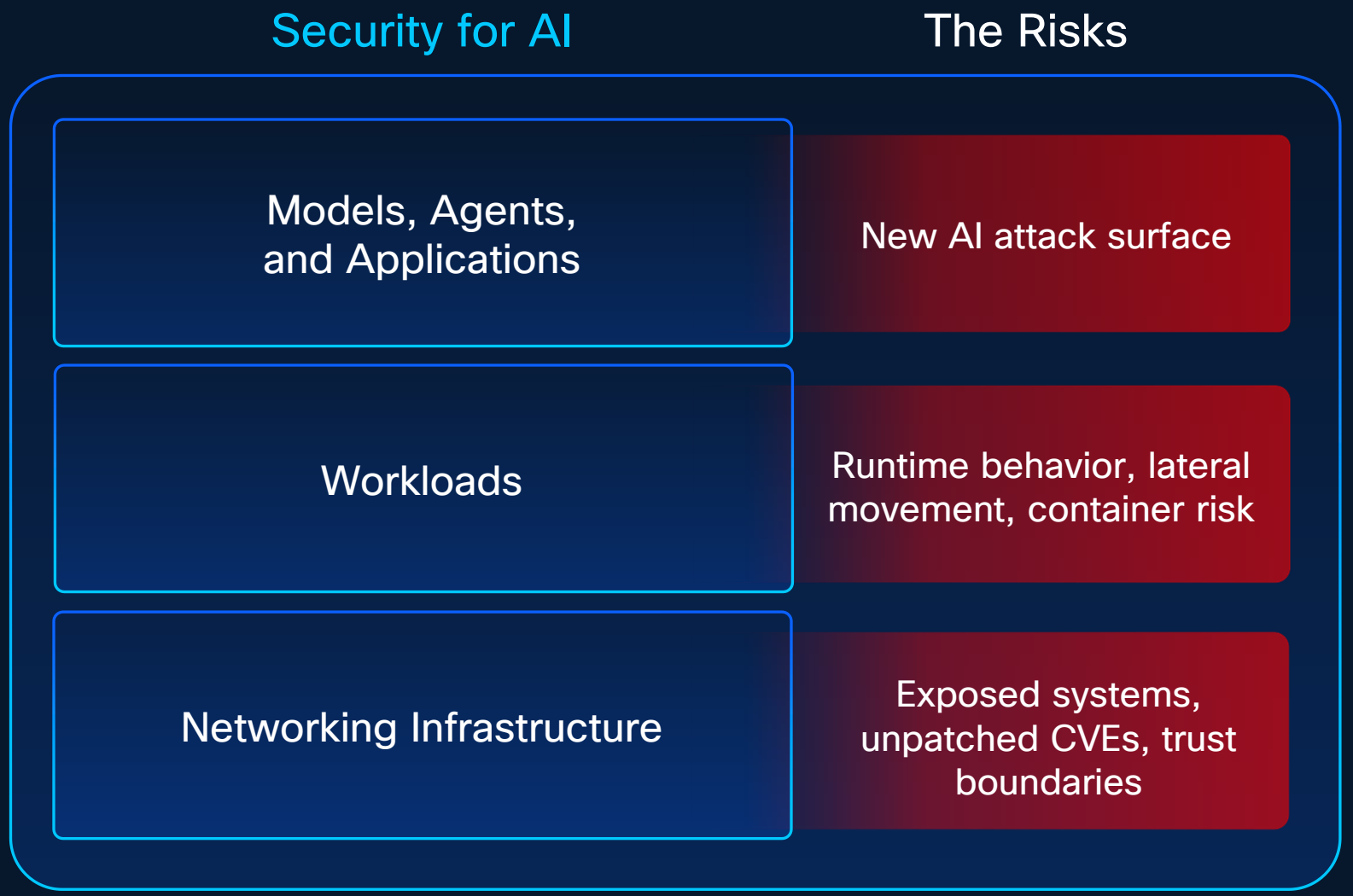
- Cisco Cloud Control
- Cisco Nexus One
- Automation | telemetry | congestion insight



Lossless | High throughput | Low latency | Low jitter | Telemetry-ready

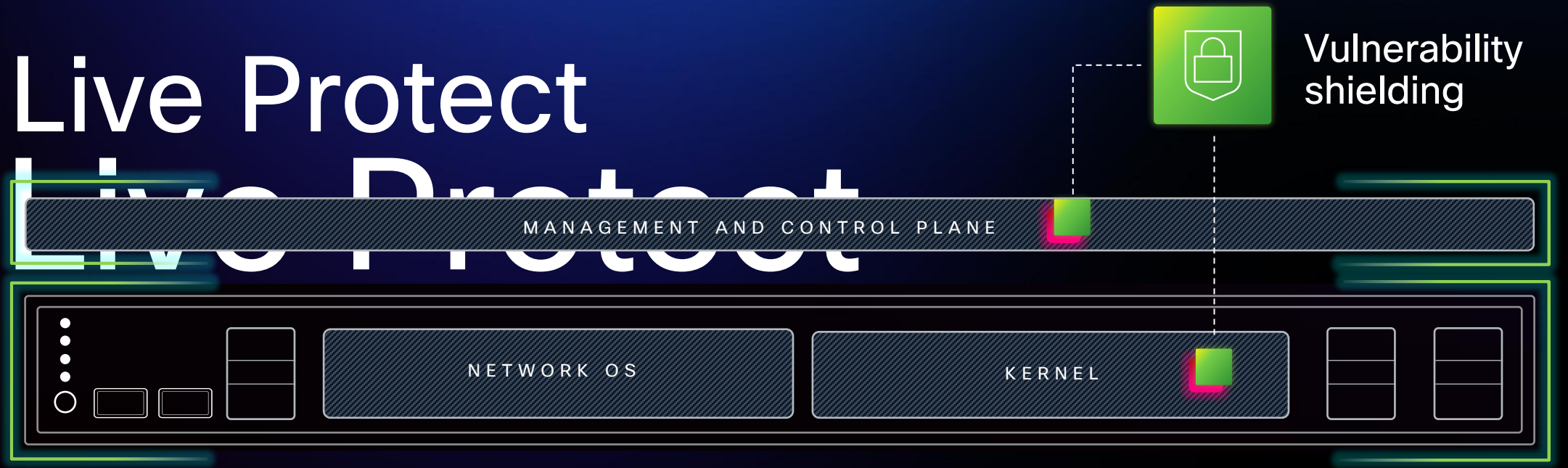
© 2026 Cisco and/or its affiliates. All rights reserved.

Secure the AI era: new stack, faster attackers



© 2026 Cisco and/or its affiliates. All rights reserved.

Live Protect



EXTENDING ACROSS CISCO NETWORKING AND SECURITY

Apply policy to live system	Protect now, patch in planned window	No downtime or switch reboot	Continuous re-validation

Threat containment at massive scale built into your network



Hardware-accelerated segmentation on every port

© 2026 Cisco and/or its affiliates. All rights reserved.

Cisco Hybrid Mesh Firewall

Distributed security at the speed of change



Cisco Firewalls

Physical | Virtual | Cloud | FWaaS



3rd Party Firewalls



NVIDIA Bluefield DPU Firewall on AI Servers



Smart Switches

eBPF



Workload Agents

Define policy once, enforce everywhere

© 2026 Cisco and/or its affiliates. All rights reserved.

cisco

Cisco AI Defense

Securing Custom Agents and AI Apps



Discover



Detect



Protect

NEW

Adaptive red teaming and guardrails

NEW

Protecting the full agent supply chain

NEW

Ecosystem to work with any cloud, any GPU

© 2026 Cisco and/or its affiliates. All rights reserved.

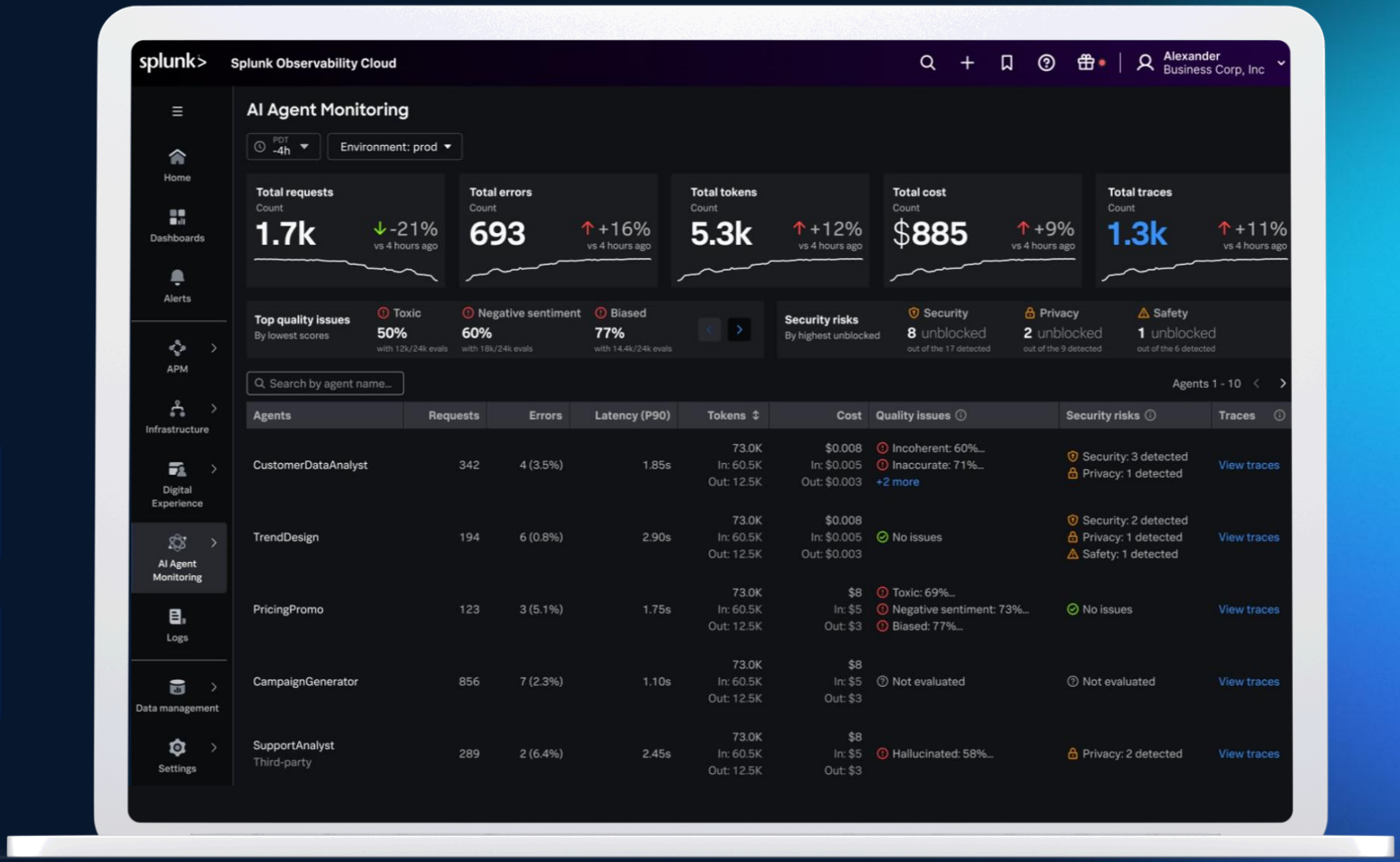
cisco

Splunk Observability enables unified visibility

Monitoring the performance, quality, security, & cost

AI agent monitoring

AI infrastructure monitoring



© 2026 Cisco and/or its affiliates. All rights reserved.

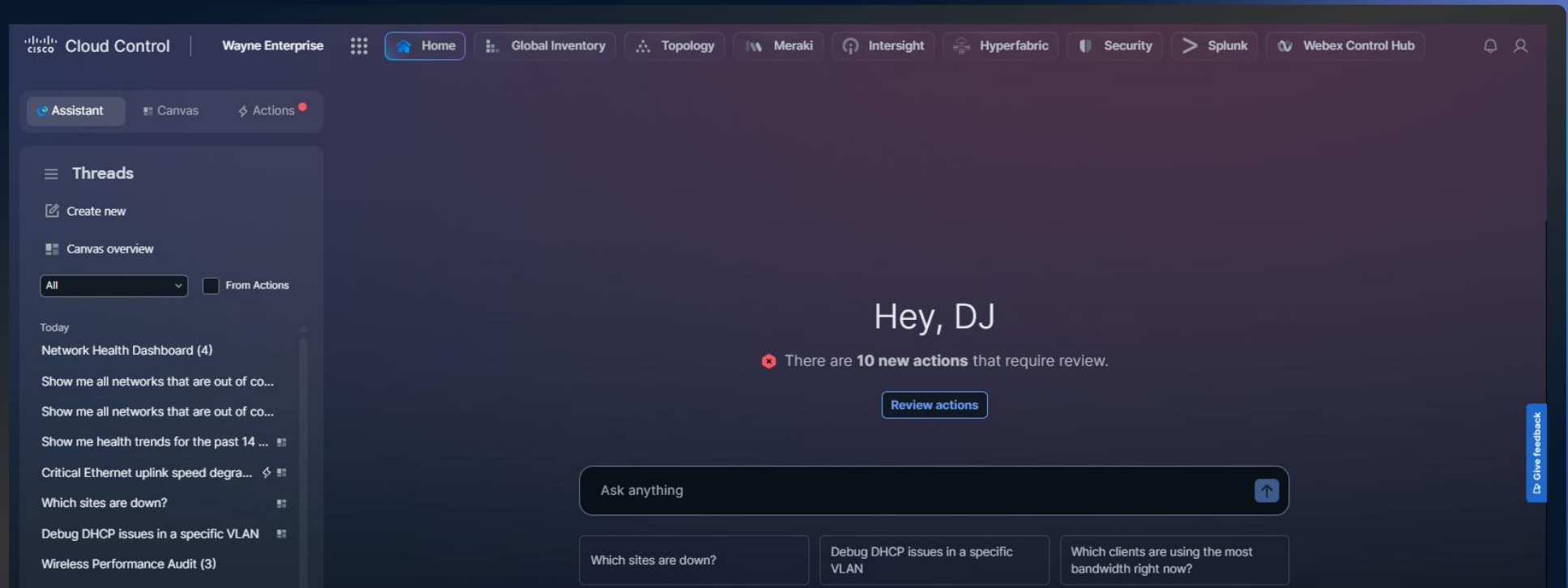


© 2026 Cisco and/or its affiliates. All rights reserved.

Cisco Cloud Control

One platform. Every domain. The foundation everything else is built on.

ThousandEyes · Security Cloud Control · Splunk · Nexus Dashboard · Nexus Hyperfabric · Catalyst SD-WAN · Intersight · & more





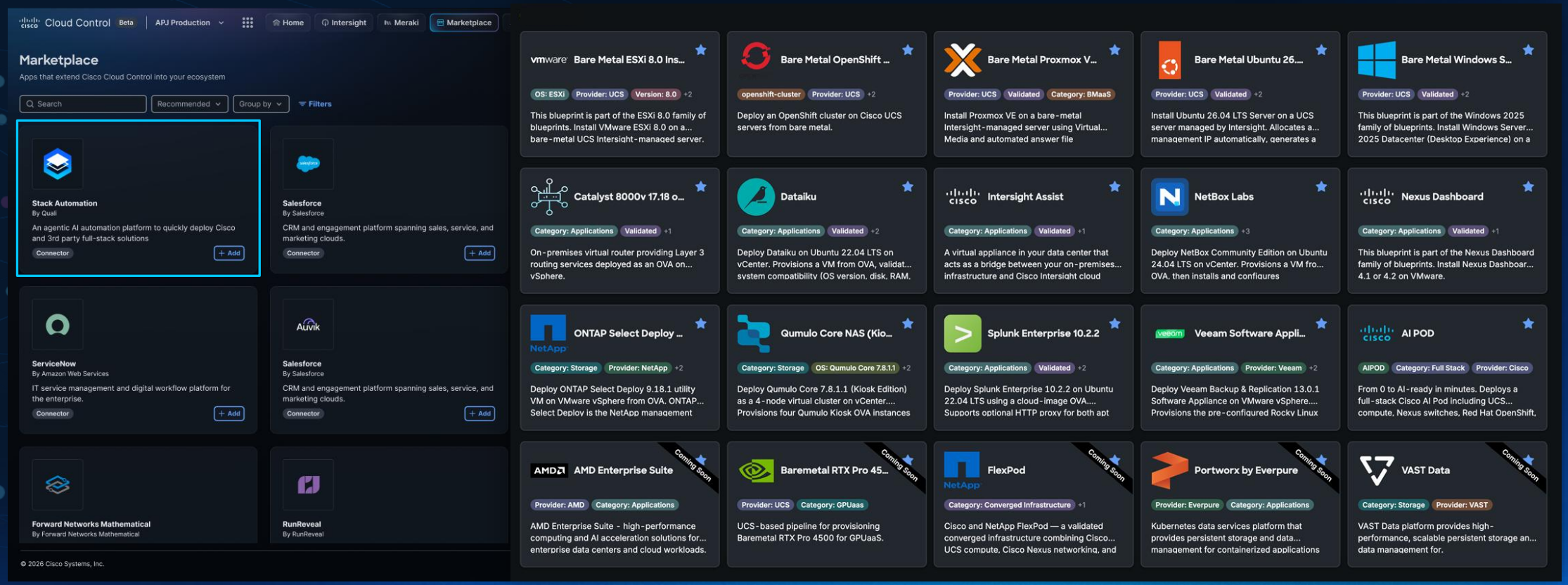


© 2026 Cisco and/or its affiliates. All rights reserved.

NOW

Stack Automation by Quali

Click-To-Deploy Solutions Hub, Cisco Cloud Control Marketplace-Ready



Blueprints embedded with policies, best practices, validated designs, and agentic intelligence

© 2026 Cisco and/or its affiliates. All rights reserved.

Cisco delivers the critical components to modernize data centers



Full-stack infrastructure engineered as one system

- Cisco 6000
- Cisco 8000
- Nexus 9000
- Cisco MDS
- Smart Switch
- Silicon One
- Cisco Optics
- NVIDIA/Intel/AMD
- UCS-X
- UCS AI Servers
- Unified Edge
- AI PODs
- Hyperconverged Infrastructure
- Converged Infrastructure
- Virtualization
- Kubernetes
- Frameworks and automation

Distributed security fused into every layer

- AI Defense
- Hypershield
- Isovalent Enterprise
- Live Protect
- Secure Firewall
- Secure Workload
- Multicloud Defense
- Network segmentation
- Identify services engine
- Security Cloud Control

Agentic operations align speed with precision

- Cisco Cloud Control
- Intersight
- Nexus One
- ThousandEyes
- Splunk Observability Cloud
- Splunk Enterprise
- NVIDIA AI Enterprise & NIMs

* enhanced with CVDs

© 2026 Cisco and/or its affiliates. All rights reserved.

Next steps

Pick one AI use case to pressure-test

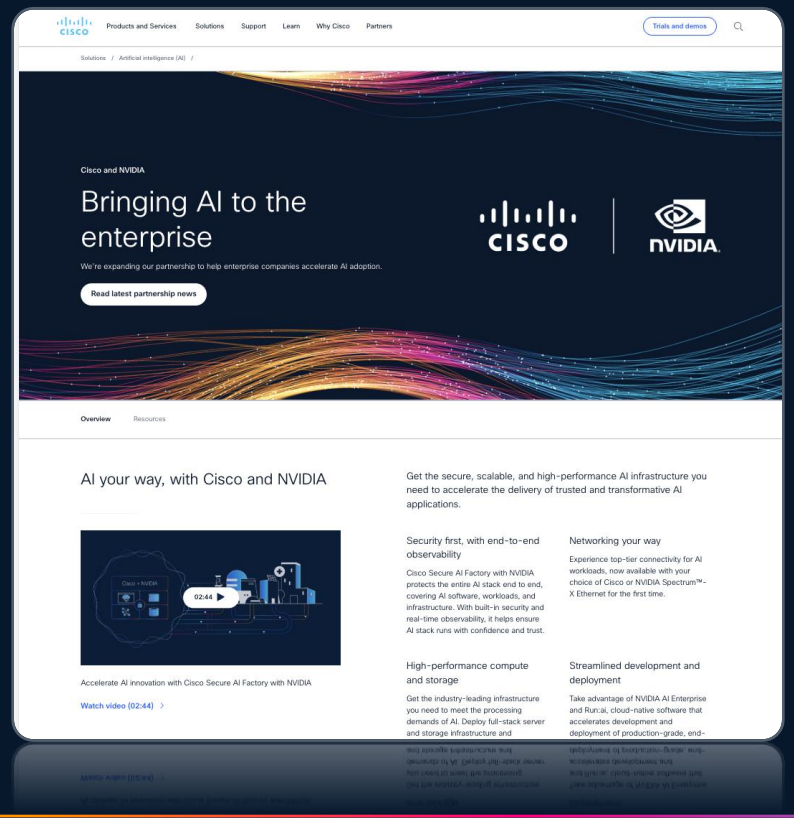
- Begin the next technical discussion with your Cisco team
- Map AI goals to solutions
- AI POD / SAIF architecture, fabric visibility, AI security, or day-2 ops

Start the next technical discussion and achieve your AI goals.

© 2026 Cisco and/or its affiliates. All rights reserved.

cisco

Resources



[SecureAIFactory.com]

[FAQ]

Recent Cisco Validated Designs

- [FlexPod with Cisco AI POD: Infrastructure for AI Training and Fine-Tuning Deployment Guide]
- [Cisco AI POD for Enterprise Training and Fine-Tuning with Everpure Deployment Guide]
- [Cisco AI Defense on AI PODs]
- [Cisco AI POD for Enterprise Training and Fine-Tuning Design Guide]
- [FlashStack for Enterprise RAG Pipeline with NVIDIA NIM]
- [FlexPod for Accelerated RAG Pipeline with NVIDIA NIM]

NVIDIA Reference Architecture Compliant Cisco RAs

- [Cisco Enterprise RA with Cisco Nexus 9000 switches, compliant with NVIDIA ERA]
- [Cisco Nexus Hyperfabric Enterprise Reference Architecture (RA), compliant with NVIDIA Enterprise RA]
- [Cisco Cloud RA with Cisco Nexus 9000 switches, complaint with NVIDIA NCP RA]

© 2026 Cisco and/or its affiliates. All rights reserved.

CISCO Connect

Thank you



Cisco Unified Computing System

Unified approach to Enterprise and AI workloads

Validated solutions for AI with compute, network, storage, and software



GPU Accelerated

GPU Optimized

Unified Edge

© 2026 Cisco and/or its affiliates. All rights reserved.


CISCO