



Modernizing Edge Infrastructure for the Era of AI

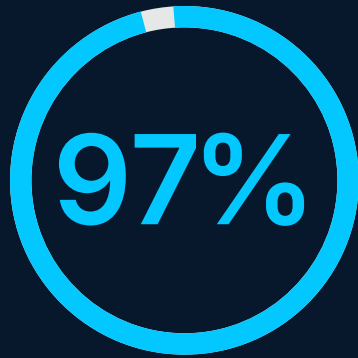
Moving beyond pilot projects to production AI at scale

Table of contents

Rethinking edge investment for agentic AI	3
Plan before you build. Architect for operational consistency and scale	4
Modern edge infrastructure: A full-stack approach	5
The modular advantage	6
Unified management platform and automation	7
Build solutions with ecosystem partners	10
Security: Built in, not bolted on	12
Observability: Turn data into action	13
Unify your infrastructure to build an AI-ready future	14

Rethinking edge investment for agentic AI

Agentic AI at the edge is an infrastructure investment decision, not just a technology trend. Unlike predictive models that surface insights for human review, agentic AI systems can reason, plan, and act in real time at the source. That raises the bar for edge infrastructure and capital planning: organizations need AI-ready capacity faster, less risk of stranded edge spend, and a more defensible path than piecemeal site-by-site buildouts.



of organizations are planning to deploy AI inferencing at the edge within 12 months.



of organizations are moving to integrated AI solutions for edge/on-premises within 12 months.



Unlock the power of AI at the edge

Source: [“Where Should AI Live? The AI Infrastructure Report,”](#) Omdia, August 2026.

Plan before you build. Architect for operational consistency and scale.

The biggest barrier to scaling edge AI isn't the technology—it's the "DIY" trap. Building infrastructure from scratch, site by site, invites the cost surprises, delays, and rework that early adopters now regret. Without a repeatable foundation, every new site adds budget unpredictability, governance friction, and a recommendation that is harder to defend to stakeholders.

94%

of organizations with active AI deployments report regrets with their initial AI infrastructure strategy.

Key insight 1

Plan before you build. The top regret is skipping a rigorous cost analysis.

82%

of organizations agree that their AI infrastructure needs to keep pace with security threats.

Key insight 2

Security needs to be built in, not bolted on, as the edge becomes an attractive target for cybercrimes.

67%

expect their network to hit its limits as an AI bottleneck within 12 months.

Key insight 3

Performance matters—if your infrastructure isn't built for the edge, your AI is essentially "data-starved."

The pivot

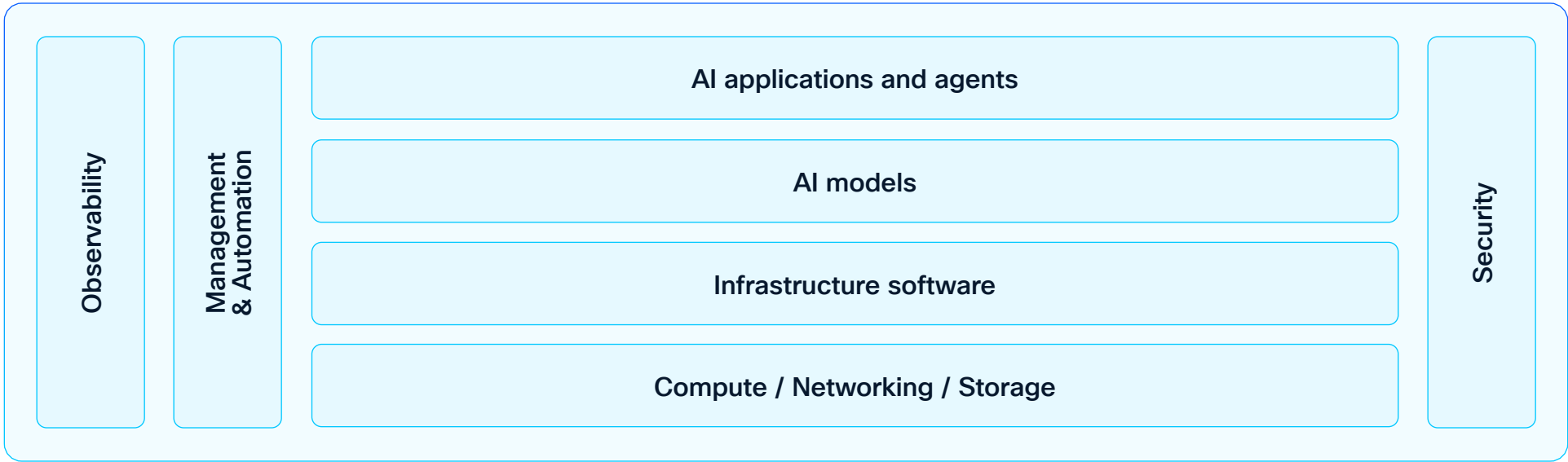
Replace one-off, "snowflake" edge buildouts with an integrated, pre-validated platform that makes costs more predictable, reduces operational risk, and speeds the move from pilot to repeatable deployment.

Source: ["Where Should AI Live? The AI Infrastructure Report,"](#) Omdia, August 2026.

Modern edge AI infrastructure: A full-stack approach

To thrive in the agentic AI era, infrastructure must function as a cohesive system, not isolated parts. Siloed hardware and software cause fragmented performance, inconsistent security, and scaling challenges. A full-stack approach treats the edge as a true data center extension. The ideal solution uses a unified platform that abstracts complexity at every layer. This approach must:

- Integrate the full stack:** Compute, networking, storage, software, and security must work together.
- Ecosystem integration:** Hardware plus infrastructure software and ISVs ensure consistency and automation for AI workloads at the edge.
- Centralize operations:** Enable zero-touch provisioning and fleet-wide visibility.
- Embed security:** Build protection into the architecture, not as an afterthought.



The modular advantage: Decoupling performance from obsolescence

Cisco Unified Edge is built to evolve with you, not against you. Reuse chassis power and cooling infrastructure across multiple generations of compute, networking, and GPU technologies, enhancing longevity, sustainability, and serviceability. This modular architecture lets organizations adopt technology at their pace, simplifying lifecycle management with less risk. It's a unified foundation that matures alongside your AI strategy, keeping your operations agile long after the first deployment.

Cisco Unified Edge

Not a server, a platform

Unified operations from core to edge

AI-ready edge

COMPUTE

NETWORKING

STORAGE

SOFTWARE

Security fused into the network

Cisco Unified Edge Unveiled

[Watch video](#)

Modular architecture: Edge-optimized system integrating compute, network, storage, software, and security for simplified edge deployment and operations.

Run any workload: Personalization for any use case through interchangeable support for compute and network nodes.

Future-ready performance: An integrated 25-gigabit, high-performance network for fast node-to-node communications.

Seamless serviceability: Redundant, hot-swappable modules and easy node replacement simplifies maintenance for remote sites.

Unified management platform and automation

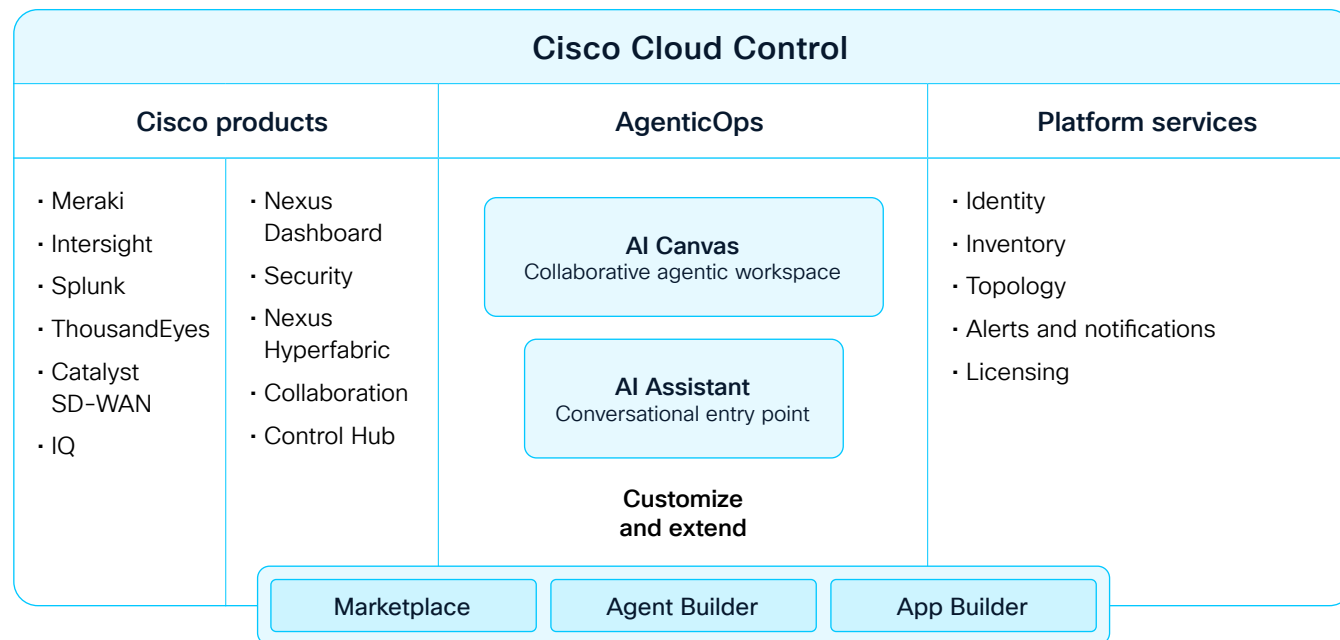
Edge IT relies on specialized tools, yet operational challenges often span multiple domains. [Cisco Cloud Control](#) unifies Cisco's deep networking, security, and compute into a single environment. This integration provides shared management and consistent access, allowing teams to monitor assets, understand connectivity, and resolve issues efficiently without the burden of rebuilding context across platforms.

Cisco Cloud Control helps organizations:

Manage: Access and manage the Cisco estate through one operational environment.

Resolve: Use shared context and AI agents to find and fix issues across products and IT domains.

Extend: Build custom apps and agents, connect supported third-party tools, and add capabilities from Cisco and the broader ecosystem.



See Cisco Cloud Control in action: [Watch video](#)

Simplifying deployment and operations at edge scale

Managing edge infrastructure can be a logistical challenge because unique “snowflake” configurations and manual processes hinder scale and AI readiness. [Cisco Intersight](#) centralizes management into a single, cloud-based control plane, transforming how you operate edge infrastructure across its entire lifecycle:

Day 0 (Provisioning): Zero-touch provisioning leveraging Cisco edge blueprints enables rapid, error-free deployment.

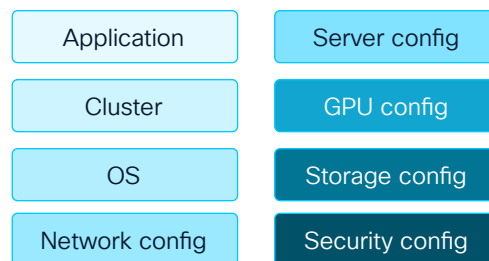
Day 1 (Onboarding): Intuitive mobile app onboarding empowers non-technical staff to deploy infrastructure and eliminates costly truck rolls.

Day 2 (Deployment): Standardized configurations and automated policies ensure consistent, repeatable deployments across every site.

Day N (Lifecycle): End-to-end visibility and observability with real-time analytics enhances reliability and operational insight at AI scale.

[Watch demo](#)

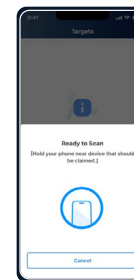
DAY 0 Zero-touch provisioning



DAY 2 Deployment at scale



DAY 1 Easy onboarding with mobile app



DAY N Lifecycle management



Cisco Intersight



splunk>
a CISCO company

Fleet management

Observability

Deployment automation

Your business is ready to innovate and scale, but your IT team is hindered by fragmented, manual deployment processes that limit agility. Stack Automation by Quali, co-developed with Cisco, removes these barriers by orchestrating physical hardware, virtual compute, and ISV applications into production-ready environments within hours. This intelligent platform eliminates manual trial and error, accelerating time to value while ensuring system stability and serviceability. By automating complex configurations, it frees teams to focus on high-value service delivery instead of infrastructure management.

Reduce deployment time:

Automate the “rack-to-app” process, shifting from weeks to hours.

Eliminate configuration drift:

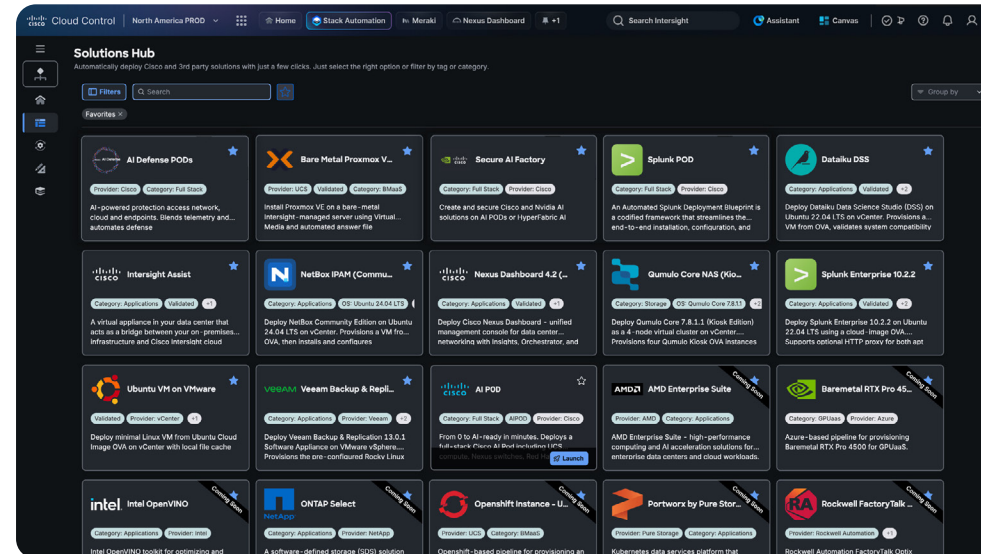
Use Cisco-validated blueprints to ensure secure, consistent deployments.

Drive enterprise flexibility:

Unify infrastructure-as-code assets and agentic intelligence into a governed workflow for scalable customization.

Stack Automation by Quali

Co-developed with Cisco



Available in
Cisco Cloud Control

Plan and fully deploy Cisco
and third-party solutions

[Explore Stack Automation](#)

Technology ecosystem

Build solutions with ecosystem partners

Organizations grapple to replicate data center-grade robustness and consistency in distributed edge environments, leading to fragmented architectures that are difficult to secure, manage, and scale. To capture the full potential of AI, enterprises require more than hardware; they need a foundation that integrates seamlessly with existing software ecosystems, ensuring consistency and automation for complex AI workloads at the edge.

Cisco partners with industry leaders to deliver edge infrastructure as robust as core data centers, providing reliable, scalable, and automated infrastructure software essential for AI workloads. These partnerships help enterprises overcome operational challenges and accelerate AI adoption at the edge.

Partner spotlight

NUTANIX

Nutanix on Cisco Unified Edge offers a hyperconverged platform that supports virtual machines, containers, and AI workloads while simplifying edge deployment and lifecycle management.

[Explore Nutanix](#)



Cisco Unified Edge utilizing Red Hat's advanced compute platform and AI Inference Server, delivers a consistent and reliable foundation for AI workloads across data center and edge environments.

[Explore Red Hat](#)



ISV ecosystem

Solving the fragmented AI software market

Organizations often struggle to deploy AI solutions and achieve fast ROI because of a fragmented software market that complicates vendor selection. This challenge is compounded by a lack of rigorous compatibility testing, which creates uncertainty when scaling at the edge. Furthermore, integration friction from managing disparate systems creates operational bottlenecks that stall deployment and significantly delay time to value for the business.

Cisco is transforming this experience through [Cisco Compatible Solutions for AI](#), a curated ecosystem of independent software vendors (ISVs). We evaluate and test these partners on our infrastructure to ensure deployment confidence and accelerated time to value.

	CROSS-INDUSTRY	RETAIL & SPORTS	MANUFACTURING	HEALTHCARE
ALL APPLICATION				
CISCO COMPATIBLE SOLUTIONS PARTNER	 	 	 	 
APPLICATION DEV / AGENT PLATFORM				
				

Partner spotlight



Cisco and Rockwell Automation integrate Rockwell's manufacturing expertise with Cisco edge infrastructure to transform factory floors into intelligent, self-optimizing environments, enhancing operational outcomes at the edge.

[Watch video](#)



Transforming retail and venue experiences with real-time crowd management and intelligent analytics, WaitTime leverages Cisco edge infrastructure to optimize customer flow and operational efficiency.

[Watch video](#)



Sickbay uses Cisco edge infrastructure to collect and correlate bedside data, delivering real-time, augmented diagnosis systems that empower medical professionals with critical patient insights.

[Watch video](#)

Comprehensive security: Built in, not bolted on

Security at the edge remains a critical vulnerability. AI at the edge places models, applications, and devices closer to both physical and cyber threats. Expanding attack surfaces, bolt-on security tools, and manual patch cycles make these sites attractive targets, while configuration drift across thousands of nodes silently weakens defenses. Cisco multilayered AI security capabilities help protect edge infrastructure, data, and AI models and agents across the enterprise, from devices to data centers.

Security at every layer to protect the edge

Protect with Cisco Security

Built-in

TRUSTED
INFRASTRUCTURE
Physical security



SECURITY
OPERATIONS AND
OBSERVABILITY*

TRUSTED ACCESS*
Segmentation

AI MODEL SECURITY*
Application protection

Cisco Unified Edge

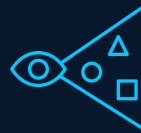
End-to-end observability

Turn machine data into agentic action

Organizations deploying AI workloads at the edge face challenges from fragmented data silos across compute, network, and security layers, causing visibility gaps that lead to reactive troubleshooting, increased downtime, and stalled innovation. Disconnected infrastructure hinders the performance needed for modern AI models.

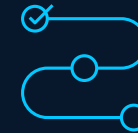
Splunk addresses these issues by transforming raw machine data into a unified, real-time intelligence layer. Integrated with Cisco Cloud Control, Splunk correlates cross-domain data at the source, enabling a shift from reactive monitoring to predictive observability. This allows for system resilience and resource optimization, and supports autonomous agentic operations at scale to meet evolving AI demands.

Cisco Intersight add-on for Splunk available in Cisco Cloud Control



Enhanced infrastructure visibility

Combines Intersight's real-time monitoring with Splunk log analytics and AI-driven insights.



Optimized IT operations

Visualize network traffic paths for faster troubleshooting and resolution.



Proactive monitoring

Use predictive analytics to identify potential issues, enabling proactive remediation.



Security and compliance

Strengthen security posture with Splunk's advanced threat detection and compliance monitoring.

[Discover Splunk](#)

Cisco brings together network, compute, storage, software, security, and observability in one integrated approach creating a cleaner path for organizations to reduce complexity, support multiphase modernization, and move forward with an infrastructure recommendation that is easier to defend.



Deep networking
expertise combined with
compute innovation



Comprehensive SaaS
management for global
fleet operations



Built-in, multilayered
security from device
to data center



Proven track
record in enterprise
infrastructure

To learn more, please visit cisco.com/go/unifiededge

