

AI at the Edge Is Here. Is Your Infrastructure Ready?

Edge AI adoption is accelerating fast, but legacy infrastructure can't keep up. Here's what organizations need to know.



Brandon Butler
Senior Research Manager,
Enterprise Networks, IDC

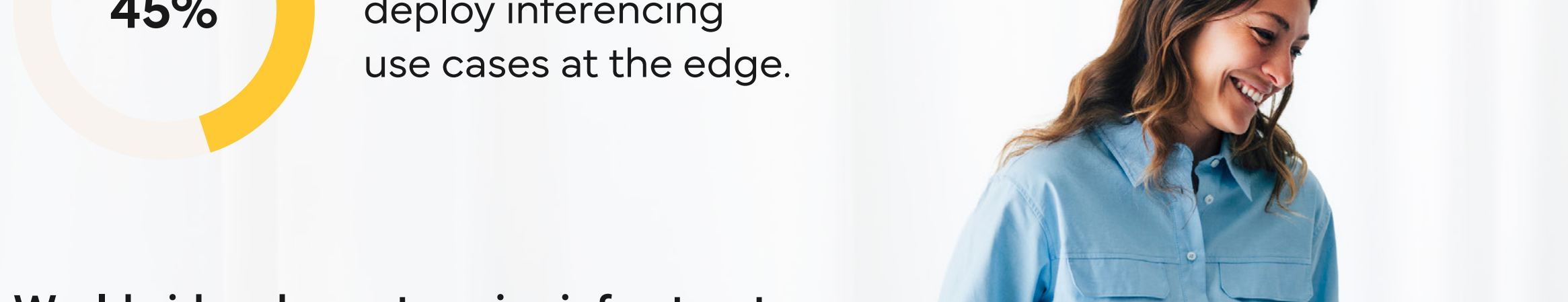


Max Pepper
Senior Research Analyst,
Enterprise and Emerging Workloads,
Worldwide Infrastructure Research, IDC

AI is moving to the edge

Enterprises are rapidly shifting AI workloads closer to where data is generated and decisions must be made.

Organizations have deployed or plan to deploy edge AI:



By 2028:



Worldwide edge enterprise infrastructure spending is projected to grow:

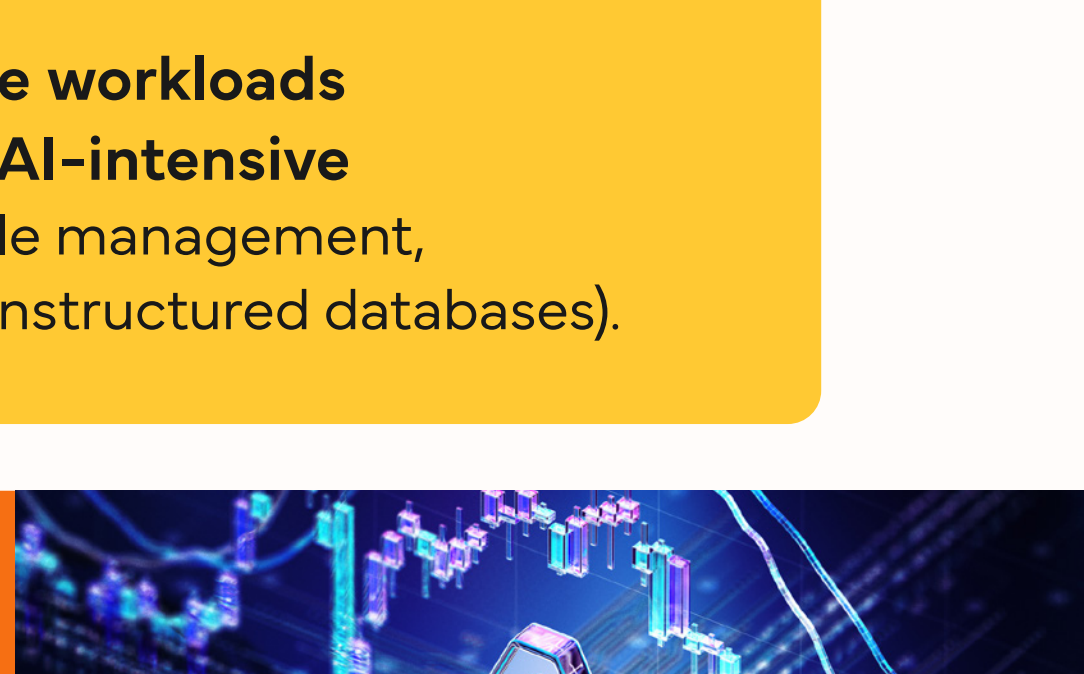


Real-time apps can't wait

Latency, data sovereignty, and bandwidth costs are forcing compute out of the datacenter and onto the edge.

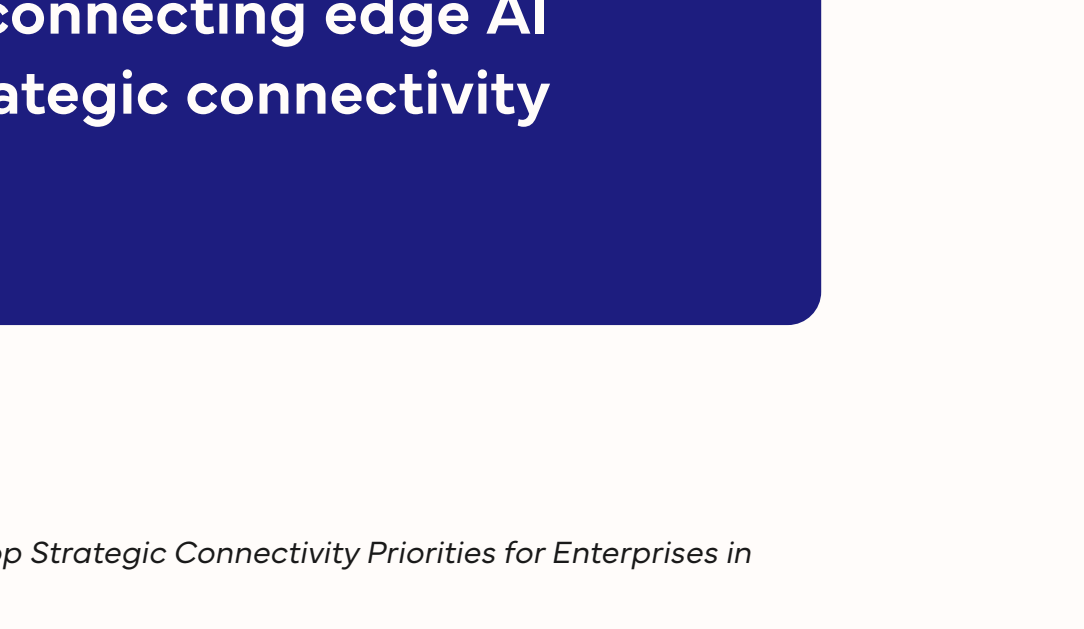
The top edge AI use cases are:

- Fraud detection
- Computer vision quality control
- AI-assisted diagnostics
- Drive-thru optimization



4 of the top 5 enterprise workloads running at the edge are AI-intensive (content delivery, AI lifecycle management, text and media analytics, unstructured databases).

AI lifecycle management is expected to see the greatest growth in edge spending over the next two years.



26% of enterprises cite connecting edge AI workloads as a top strategic connectivity priority for 2025.

Edge deployments are breaking legacy models

Distributed infrastructure at scale exposes critical gaps in management, skills, and security.



Four capabilities that define modern edge

Purpose-built platforms must unify hardware, software, management, and security in one integrated system.

- Zero-touch provisioning:**
Hardware pre-installed with OS and hypervisor shipped directly to remote sites; no onsite IT required
- Centralized SaaS management:**
Single dashboard spanning hardware telemetry, cluster visibility, lifecycle management, and security across all distributed sites
- Hyperconverged software as the foundation for modernization:**
Consolidation of compute, storage, and networking to minimize footprint and simplify IT operations
- Workload flexibility:**
Support for both VM-based and containerized workloads (including Kubernetes-based AI inferencing) on the same platform



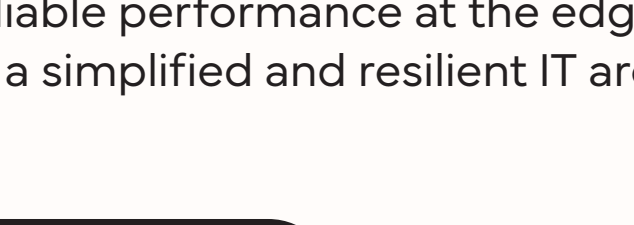
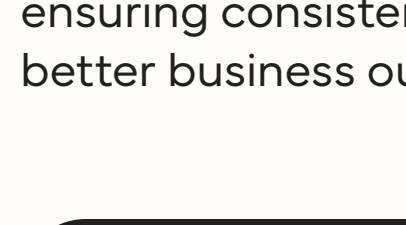
Simplify, secure, and scale the edge

Organizations must act now; the infrastructure choices made today will determine competitive readiness for AI at the edge.

- Co-engineering platforms:**
Co-engineered platforms combining AI-ready hardware, hyperconverged software, and centralized management close the gap between edge ambition and operational reality.
- Automating upgrades:**
Automated, rolling firmware and software upgrades eliminate configuration drift and reduce security exposure across distributed sites.
- Licensing:**
VM-based licensing (per VM, not per CPU core) lowers total cost of ownership for edge deployments.
- Supporting models:**
Proactive support models with synchronized case management and automated RMA reduce resolution times and operational burden at remote sites where onsite intervention is not an option.

Source: IDC's Edge View Survey, 2025; Worldwide Edge Enterprise Workloads Infrastructure Market Forecast, 2026–2029, March 2026

Message from the sponsor



The rapid expansion of edge computing is driven by the need for real-time data processing and localized AI capabilities.

Customers today are increasingly concerned with the complexity of managing distributed infrastructure, ensuring high availability at remote sites, and maintaining consistent security policies across their entire footprint. The Cisco and Nutanix edge solution addresses these challenges by integrating Cisco's AI-ready Unified Edge platform and cloud-based Intersight management with the industry-leading HCI software Nutanix Cloud Platform. This hyperconverged infrastructure (HCI) provides a unified, software-defined environment that simplifies operations from the datacenter to the edge. By combining Cisco's robust networking and compute hardware with Nutanix's flexible software stack, organizations can achieve greater agility and improved cost efficiency. This collaboration empowers businesses to deploy applications rapidly, ensuring consistent, secure, and reliable performance at the edge, ultimately driving better business outcomes through a simplified and resilient IT architecture.

To learn more, read IDC Spotlight