Cisco Unified Edge: Driving the Future of Intelligent Retail

How AI at the edge is reshaping retail

Retailers are leveraging AI at the edge to fundamentally transform their businesses across various facets, moving beyond traditional analytics to real-time, actionable insights and automated processes. This transformation is driven by the need for immediate data processing, enhanced security, and personalized customer interactions directly at the point of sale or service.

One transformative application is personalizing customer experiences. AI at the edge enables retailers to gather and process real-time customer analytics data directly in the store. This data, which may include browsing patterns, product interactions, and insights like customer preferences, helps enable highly personalized recommendations, relevant promotions, and dynamic pricing to enhance the customer experience. By processing this data locally, retailers can deliver tailored experiences without the delays associated with sending all data back to a central cloud, ensuring that the customer interaction is seamless and relevant.

Another critical area is enhanced security and loss prevention. By deploying AI-powered video analytics at the edge, retailers can monitor store environments in real time. This allows for the immediate detection of unusual behavior, such as potential theft or suspicious activities, and proactively alerts in-store security personnel. This capability not only helps prevent financial losses but also enhances the safety and security of both customers and staff. Low-latency processing at the edge ensures that these alerts are generated instantaneously, enabling rapid response.

Furthermore, retailers are using AI at the edge for optimized store operations and inventory management. AI models can monitor thousands of data points across multiple stores, analyzing everything from foot traffic patterns to shelf stock levels. This enables retailers to optimize product placement, refine promotional strategies, and manage inventory more efficiently. For instance, AI can identify sudden spikes in demand for certain products or detect inefficiencies in stock replenishment, allowing for real-time adjustments. This leads to improved sales, reduced waste, and a more streamlined operational flow, ultimately boosting profitability.



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

Trends and challenges

Retail's AI edge infrastructure imperative

The retail industry stands at the precipice of a profound transformation, driven by the rapid advancements in Artificial Intelligence (AI). Retailers are eager to leverage AI for personalized customer experiences, operational optimization, and enhanced security and fraud detection by tuning models with proprietary data. This path to value creation, however, is not without its challenges, requiring retailers to address the dual demands of modernizing edge infrastructure for traditional workloads while simultaneously supporting new AI workloads.

However, the current state of retail infrastructure often presents significant hurdles to AI adoption. Traditional legacy systems were simply not built to handle the scale, speed, and intelligence required by real-time AI inference and agentic workflows across thousands of store locations. For instance, as millions of AI agents become embedded in business applications, a single user query can trigger a cascade of internal inferences and API calls. This can generate anywhere from 5 to 25 times more network traffic than a basic chatbot, overwhelming conventional edge networks. Organizations are already anticipating that a substantial portion of their significant GPU investments could be lost due to network bottlenecks.

Compounding these technical demands are the practical challenges of products that don't interoperate, deployments that vary by site, and a shortage of skilled resources, which create friction at every stage. This inadequacy manifests as rigid, disconnected systems that can throttle performance, increase operational complexity in managing and deploying distributed retail edge deployments, and increased security risks from expanded attack surfaces.

Security at the retail 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 retail sites attractive targets, while configuration drift across thousands of nodes silently weakens defences.

Perhaps the most significant and often overlooked challenge in retail AI lies in the decentralized nature of AI services. While much of the initial model training typically occurs in centralized data centers, the shift of test-time inference to the edge is making it the new frontier for enterprise AI. A significant amount of AI workloads are projected to be inferencing workloads, and a substantial portion of these will be executed at the edge directly within retail stores, distribution centers, or regional hubs. This necessitates a new vision for an "AI-ready enterprise" that extends seamlessly from the core data center all the way to the edge. The edge will become the cornerstone of scaled delivery of AI services in retail, where low-latency, high-throughput, secure, and cost-efficient operations are paramount for real-time decision-making and customer interaction. Retailers need a full system rethink, not just incremental upgrades, to truly unlock the potential of AI at the edge.

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

How it works

Cisco Unified Edge: the AI platform the retail edge demands

Building an AI-ready infrastructure, especially for the demanding retail edge, requires more than just high-performance compute. It necessitates a full system rethink that can seamlessly integrate into everyday business operations and drive them forward.

Cisco Unified Edge provides a comprehensive, AI-ready enterprise infrastructure designed to meet these demands. It delivers a unified, integrated platform that converges compute, networking, storage, security, observability, and centralized cloud management into a single, modular chassis built specifically for retail edge environments. This approach transforms traditional server deployment by building AI-ready edge infrastructure, ensuring unparalleled operational agility and stopping the “rip and replace” cycle.

Key capabilities:

- Future-ready performance: Cisco Unified Edge offers modular, high-performance compute and networking nodes with GPU acceleration and Intel® Xeon® 6 Processors, optimized for demanding AI workloads such as real-time video analytics, customer personalization, and inventory optimization. Its modular design supports interchangeable compute and network nodes, allowing for easy serviceability and personalization to any edge workload without compromise. This ensures investment protection across a 10-year+ lifespan by accommodating multiple generations of technology, effectively handling the massive data streams and increased network traffic generated by agentic AI workflows.
- Seamless, scalable operations: achieve breakthrough operational simplicity at scale with a software-defined system and unified, fleet-wide operations. Centralized cloud management through Cisco Intersight® provides end-to-end visibility and AI-driven insights from edge to core. It enables zero-touch provisioning, automated day-0 to day-N operations, and pre-validated blueprints for rapid, consistent deployments and updates across hundreds or thousands of retail locations. Cisco Intersight fleet management capabilities democratizes retail edge management, ensuring operational consistency and efficiency from a single pane of glass, and simplifies onboarding and upgrades for any scale of deployment. Plus, end-to-end observability with real-time analytics, including integration with Splunk®, provides deep insights across edge infrastructure, enabling rapid error detection and correction. This enhances reliability, reduces downtime, and lowers operational costs by providing a seamless operating model from core to edge.



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

- Security fused into the platform: security is paramount for Cisco Unified Edge. It integrates multi-layered, zero-trust security embedded directly into the hardware and software, protecting retail workloads, data, and infrastructure. This includes integrated physical and digital anti-tampering features with consistent security profiles and policies to prevent configuration drift across edge systems at retail sites. Additional network segmentation, AI-model, and container security capabilities safeguard against evolving physical and cyber threats across distributed edge locations, enabling confident innovation at scale.

Powering retail with leading ISVs

Cisco is collaborating with several leading Independent Software Vendors (ISVs) to enable the seamless integration of specialized software with Cisco Unified Edge systems, ensuring full-stack solutions are certified, scalable, and easy to deploy for specific industry use cases. For example, in retail, collaboration with ISVs facilitates personalized customer experiences across physical and digital touchpoints, real-time analytics, customer engagement, and efficient operations.

WaitTime is an AI-powered ISV that uses computer vision to analyze crowd flow, wait times, and occupancy in real time. In retail, it helps optimize staffing, reduce checkout queues, and enhance customer experience by turning camera feeds into actionable insights. Retailers use it to improve store layouts, reduce abandonment, and respond to peak traffic instantly.

Quividi delivers an AI- and computer-vision-powered audience measurement platform for Digital-Out-of-Home (DOOH) signage and retail media. Using standard cameras, its software anonymously detects and counts pedestrians, vehicles, and viewers, then generates real-time metrics — impressions, dwell time, demographic segments, and attention analytics across screens and store displays. Quividi's insights enable advertisers, media owners, and retailers to verify who viewed their content, optimize campaigns dynamically, and better understand shopper engagement along the path to purchase.

PreciTaste is an AI-powered kitchen automation platform that transforms how restaurants and food-service businesses plan, prep, and cook. Using computer vision and demand-forecasting algorithms, PreciTaste predicts how much food will be needed, then guides kitchen staff in real time on what and when to prepare, reducing food waste, lowering labor and food costs, and improving order accuracy.

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

Transforming retail with AI

Loss and fraud detection

- AI-enhanced video analytics
- Facial recognition systems
- Real-time fraud detection

Drive-through optimization

- AI-powered order prediction
- Automated voice recognition
- Real-time traffic management

Digital merchandizing

- AI-powered customer analytics
- Dynamic digital signage
- Personalized mobile engagement

Learn more

Learn more at cisco.com/go/unifiededge.

The Cisco Advantage

The future of retail edge infrastructure

The retail industry is at a pivotal moment, the shift of AI to the edge is creating unprecedented opportunities for real-time insights, personalized customer experiences, enhanced security, and optimized operations. However, this shift also brings significant challenges related to legacy infrastructure limitations, network bottlenecks, security vulnerabilities, and the complexities of managing distributed environments.

Cisco Unified Edge, with our technology partners, delivers a full-stack platform that converges compute, networking, security, storage, software and management to address the challenges presented by AI at the retail edge. Cisco Unified Edge delivers modular, high-performance compute for AI workloads, centralized cloud-managed operations at scale, and built-in zero-trust security. Its flexible design supports evolving technologies, simplifies deployment and fleet management across locations, ensures real-time visibility and reliability, and safeguards data and workloads. With Cisco Unified Edge, retailers can unlock the full potential of AI, driving innovation, enhancing customer satisfaction, and achieving meaningful operational efficiencies.

Cisco Capital

Financing to Help You Achieve Your Objectives

Cisco Capital® can help you acquire the technology you need to achieve your objectives and stay competitive. We can help you reduce CapEx. Accelerate your growth. Optimize your investment dollars and ROI. Cisco Capital financing gives you flexibility in acquiring hardware, software, services, and complementary third-party equipment. And there's just one predictable payment. Cisco Capital is available in more than 100 countries. Learn more.

© 2026 Cisco and/or its affiliates. All rights reserved. Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

C22-5700601-01 09/26