SPOTLIGHT

Sponsored by: Lenovo

Using modernized edge resources and AI, digitally savvy retailers are finding ways to deliver better customer experiences and overcome disruption prevalent in the industry.

Innovation Begins at the Edge

February 2023

Written by: Jennifer Cooke, Research Director, Edge Strategies

How AI at the Edge Supports Innovation

Digital technologies are being infused into all aspects of operations and engagement. Real-time decisions and operations depend upon modernized compute platforms and the ability to extract insights from data. Often, these platforms need to be near where data is created, where people work and play, and where "things" operate. Infrastructure to support digital-first business and operations is increasingly deployed outside of a core datacenter.

IDC forecasts that by 2026, 20% of servers running AI will be located outside of core datacenters — at the edge. Edge computing enables organizations to acquire, analyze, and act upon data near where it is created. Access to modernized edge resources is a fundamental requirement for organizations shifting to digital-first business.

This shift to digital first has accelerated across all industries as organizations have had to engage with customers and partners in new ways and streamline operations. The retail industry has been particularly impacted by major market forces such as supply chain disruptions, labor shortages, and changing customer expectations. Those retailers seeking to adapt and thrive amid these market changes are leaning on technology to create new ways to engage with customers. Forward-thinking retailers are investing in innovative new ways to improve the customer experience. This innovation is built on edge infrastructure, often fueled by AI and ML.

The way consumers shop has been radically changed. Customers seek seamless digital and physical experiences. They are influenced digitally and require personalized experiences across online and in-person engagements. Convenience and speed are high priorities when delivering a great customer experience. Retailers are forced to embrace digital channels or risk losing customers.

AT A GLANCE

KEY STATS

According to IDC research:

- **66%** of organizations are or plan to run AI workloads at the edge. Many will be investing in more powerful infrastructure, technologies to enable remote monitoring, and hardware acceleration to run AI and ML workloads.
- Early adopters of AI report **30–35%** improvement in business outcomes with the rollout of AI solutions.
- **57%** of organizations plan to use computer vision and AI solutions for fraud prevention, traffic intelligence, and queue management.
- **44%** of organizations believe they will require investments in high-performance infrastructure to accomplish their goals.

KEY TAKEAWAY

Digitally savvy retailers are leveraging AI workloads at the edge to radically revamp the customer experience, providing greater convenience, speed, and personalized products.

The infusion of digital technologies into the consumer experience will be invisible. Access to high-performance infrastructure, fast and secure connectivity, and AI platforms to generate real-time insights will create greater convenience and personalization for customers. In a myriad of ways, modernized edge and cloud resources are improving the quality of products, streamlining supply chains, and enhancing operations and logistics.

Retail: Why Edge and AI Is Strategic and Necessary

Creating more personalized experiences and improving the omni-channel shopping experience requires the ability to gather massive amounts of data and apply AI to glean insights from that data. Modernized, high-performance infrastructure is required to support the demanding new workloads at the edge — located in close proximity to where data is created and gathered.

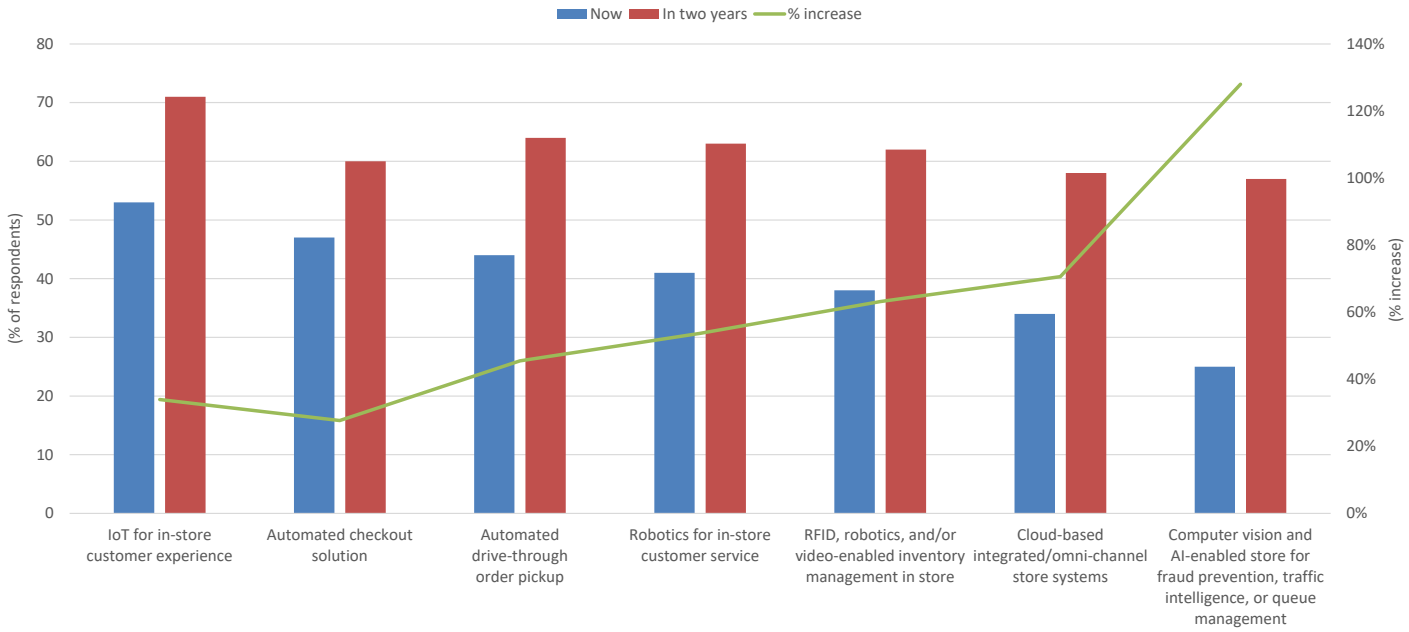
Retailers are planning a more than 2x increase in investment in computer vision and AI-enabled fraud prevention, traffic intelligence, and queue management technologies (see Figure 1). By 2024, two-thirds of retailers plan to use robotics for in-store customer service, RFID, or video-enabled inventory management systems. These new workloads are built on an underpinning of modernized infrastructure extracting insights from massive data sets. Using existing data sources such as video streams, retailers are making stores and warehouses "smarter" and more efficient and improving the safety of workers and customers. Modernized edge resources that integrate seamlessly with core and cloud resources deliver the following benefits:

- **Improve performance and reduce latency.** Data-intensive workloads can be run near where data is created, enabling near-real-time decision making. For operational workloads, onsite compute that can operate without connectivity is often required.
- **Support data sovereignty and security.** 37% of organizations worldwide are using edge infrastructure to remain compliant and keep data residency requirements. One of the top reasons to use edge resources is to improve data security and control. With a strong focus on data and physical security, edge solutions are designed and purpose-built for deployment outside of the four walls of a datacenter. Advanced security technologies and processes, along with considerations for the physical environment, are table stakes for edge infrastructure.
- **Reduce costs.** Edge infrastructure enables data to be processed onsite and avoid high data transmission costs. In an ever-vigilant business environment that tracks infrastructure spending against business value received, edge investments often reduce spending on connectivity services.

The retail industry is under pressure to transform now, or risk being surpassed or replaced by retailers that can deliver a better, more personalized experience and operate more efficiently. The industry is experiencing large-scale store closures, difficulties finding and retaining employees, and reduced profit margins. To overcome these challenges, many retailers are planning to spend more than 2x what they are today in AI technologies to manage queues and store traffic and reduce fraud. In two years, 71% plan to invest in IoT solutions to improve the in-store experience (see Figure 1).

FIGURE 1: ***Retailers Planning Significant Increase in Technology Use to Improve Store Operations***

Q Regarding your company's strategy for store operations, are you using the following technologies now? In the next two years?



n = 504

Source: IDC's Global Retail Operating Models Survey, August 2022

Edge Innovation Examples

Edge computing and AI is an important foundation for the retail industry as it seeks to unify the customer experience online (via phone or PC) or in the store. Mastering omni-channel commerce requires access to advanced core, edge, and cloud resources to build a resilient business that enables customers to experience products online in a similar way to an in-store experience. Customers seek convenience and speed. They will gravitate to retailers that can provide a frictionless experience throughout the entire process, from product education and selection to delivery and ongoing service. To deliver this experience, retailers need access to cloud and edge resources that improve and drive insights into everything from product demand to transportation and logistics. AI is used to understand weather patterns and where and when customer demand may spike, for example.

In addition, digital twin technology has delivered measurable benefits in the manufacturing industry to improve products, safety, and operational efficiency for several years. Today, this technology is used in retail to improve the customer experience and deal with labor shortages. American retailer Kroger, which operates supermarkets and department stores throughout the United States, is using digital twin technology to reduce checkout times and improve operational efficiency in stores. By simulating the checkout process and testing out various scenarios prior to implementing them, Kroger was able to reduce average checkout times from 80 to 53 seconds. This technology often requires accelerated graphics capabilities and higher-performance compute infrastructure. Kroger uses an AI platform and an AI software stack to aggregate and analyze very large video input data sets in real time.

The following are some in-store retail examples of retail innovations:

- **Smart shelves.** These use RFID technology and weight sensors to manage inventory in stores, helping retailers understand products that are low or out of stock and prevent theft. Digital displays show prices and promotions in real time, often tailored to the shopper in close proximity. These capabilities impact retail operations in a number of ways — they personalize the shopping experience, reduce the labor required to reprice items with paper labels and stickers, and streamline supply chain operations by managing inventory levels.
- **Digital signage.** Retail signage can allow for the digitization of shelves. A solution that includes software to connect devices and equipment can in turn connect stores with machine learning and big data functions to drive more cross selling and create audience analytics. The result is the ability to deliver more immersive and interactive experiences.
- **Computer vision.** This helps retailers improve the in-store experience in many ways — heat maps to understand store foot traffic flows, image recognition to understand buyer sentiment, virtual mirrors, in-store personalized advertisements, and inventory management and loss prevention. A computer vision solution can streamline the ability to gain actionable insights from video data with inferencing and analytics.

All these innovations require modernized, secure edge resources that can leverage data from the cloud and process large amounts of image-based data onsite.

Special Considerations for Edge Infrastructure to Support AI

Edge needs to operate seamlessly with core and cloud resources. Modernized edge platforms need to support hybrid, heterogenous environments on a broad scale. These platforms must be able to gain insights from often diverse data sources that reside in multiple regions to enable the following capabilities:

- **Coordination and consistency.** Assembling the infrastructure, platforms, applications, and expertise to gain value from AI at the edge requires a great deal of coordination. One of the challenges organizations face is establishing a comprehensive AI strategy that spans heterogenous, hybrid environments — from edge to core to cloud datacenter. Modernized platforms that support these diverse environments pave the way to greater innovation.
- **Streamlined management.** The scale and geographic reach of edge deployments increases the need for systems that can be automatically or remotely updated. Open, cloudlike systems that can be updated without the need for manual intervention will be more secure and resilient. For example, consider what's involved in updating 100 servers in one datacenter versus 100 servers in all different locations. The logistics involved make it difficult, if not impossible, to support advanced workloads in remote locations without a high degree of automation and standardization.
- **Edge as a good neighbor.** Especially in retail and smart spaces, infrastructure will be placed where people shop, work, and reside. Edge compute needs to be a good neighbor and not be intrusive or disruptive to the surrounding environment. Noise levels are an important consideration for edge infrastructure.
- **Security and resilience.** Consider purpose-built infrastructure to ensure that the equipment is protected from the environment and unauthorized access as well as operates quietly and unobtrusively. This can be achieved by encrypting all communications, ensuring confidentiality and the integrity of data and authentication. Edge AI systems are starting to use some of the most advanced cryptography techniques available, including homomorphic encryption and attribute-based encryption. IDC's research points to one in five organizations experiencing edge

equipment damage due to temperature, humidity, dust, or vibration in the deployment location; 18% report equipment being tampered with (either intentionally or unintentionally) by unauthorized users. For many edge deployments, ruggedized servers or secure enclosures are necessary.

Lenovo's Edge Solutions

Lenovo's edge portfolio is multifaced and broad — consistent with its customers' diverse needs for distributed IT solutions. The company has been supplying servers and storage to remote and back offices (in addition to corporate datacenters) for years. Digital-first business has accelerated customers' need to integrate data and workloads across multiple cloud, core, and edge sources. Along the edge journey, Lenovo aims to be a trusted innovation partner, providing technology-developing solutions to meet industry-specific needs for edge computing.

Lenovo aims to be a trusted innovation partner, providing technology-developing solutions to meet industry-specific needs for edge computing.

In 2022, Lenovo introduced a new line of small form factor, ruggedized systems for edge deployments, called ThinkEdge. Lenovo made a clear distinction from its core ThinkSystem servers with new features that address the unique challenges of compute deployed outside of a datacenter, yet maintain the computing superpowers of Lenovo servers including NVIDIA GPUs. Locking bezels, a wider range of operating temperatures, shock and vibration resistance, and low noise-emission are a few of the features built into Lenovo's ThinkEdge line. These features address the physical demands of edge IT service.

Lenovo's ThinkAgile hyperconverged infrastructure (HCI) is designed for rapidly deploying hybrid multicloud and edge solutions. Available as a service with Lenovo TruScale, ThinkAgile streamlines and simplifies edge use cases with preconfigured infrastructure that can be managed remotely.

Lenovo is perceived as a trusted provider that can help drive innovation and support customers along the digital-first journey. ThinkEdge systems with NVIDIA GPUs are purpose built for AI workloads at the edge. Partnerships and collaboration with NVIDIA and other industry solution providers build on Lenovo's ability to innovate at the edge.

Challenges

One of the most frequently cited customer pain points is integrating legacy platforms with new solutions. Edge innovation is happening rapidly — but a company's strategy for extending to edge locations should be done in lockstep with the company's core and cloud strategies to ensure data security and compliance.

Another key customer challenge is managing infrastructure across many platforms and in many different locations. Lenovo's ability to help organizations embrace automation of their edge infrastructure is key to long-term success. By educating customers on a solid, long-term edge strategy and helping customers develop and execute the frictionless flow of data and insights from edge to core datacenter, Lenovo has the opportunity to expand its role as a trusted provider.

Conclusion

The retail industry is undergoing an especially disruptive transition. The industry is experiencing large-scale store closures, difficulties finding and retaining employees, and reduced profit margins. Edge resources are the critical foundation the retail industry needs to deliver exceptional customer experiences, improve operational efficiency, overcome staff shortages, and shift to omni-channel commerce. The pressure is on to evolve or risk the consequences of a rapidly changing market. Modernized edge resources are an integral part of a retailer's digital-first strategy. Lenovo's firsthand experience in driving digital innovation at the edge provides the company with a significant opportunity for success.

About the Analyst

Jennifer Cooke, *Research Director, Edge Strategies*

Jennifer Cooke is a research director within IDC's worldwide cloud and edge infrastructure practice, where she leads IDC's Edge Strategies research. Jennifer's research provides insights into the ecosystem of physical infrastructure, software, and services that support secure and resilient operations at the edge. With a background in datacenter research and a 25+ year career as a technology analyst, she has a keen interest in the evolving role of technology in supporting efficient operations and innovation.

MESSAGE FROM THE SPONSOR

More About Lenovo and NVIDIA

In partnership with NVIDIA, Lenovo is developing world-changing technologies to create a more efficient, connected, and digital society. By designing, engineering, and building the world's most complete portfolio of innovative, AI-ready edge devices and infrastructure, Lenovo and NVIDIA are leading an Intelligent Transformation – to create better experiences and opportunities for millions of customers worldwide.

Accelerating AI relies on GPUs, and NVIDIA delivers enterprise GPU acceleration everywhere it's needed – to datacenters, desktops, laptops, and the world's fastest supercomputers. As companies are increasingly data-driven, the demand for AI technology grows. From speech recognition to recommender systems and supply chain management, AI technology provides enterprise teams with the power, tools, and algorithms to work effectively.

Lenovo Edge Computing empowers users to solve real-world challenges with robust infrastructure solutions that generate faster insights. Complex organizational and business decisions can be made quickly, and with a higher level of confidence with ThinkEdge servers that are rugged and secure with physical tamper-proofing, data encryption, and the ability to withstand conditions of all kinds. So no matter what you need, we'll find the right solution for you.

Lenovo and NVIDIA bring innovative solutions and intelligent infrastructures to solve the most significant challenges facing modern day retailers.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
140 Kendrick Street
Building B
Needham, MA 02494, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com