

Predicts 2024: Edge Computing Technologies Are Gaining Traction and Maturity

18 October 2023 - ID G00800052 - 18 min read

By Thomas Bittman, Tony Iams, [and 3 more](#)

Edge computing interest and adoption is increasing across vertical industries, and enabling technologies are maturing. Infrastructure and operations leaders need to work with their lines of business on edge computing, and evaluate solutions that accelerate deployments and support extensibility.

Overview

Key Findings

- Technical innovation has drastically lowered the cost to develop and deploy edge systems to collect and analyze data from IT and non-IT assets.
- Containers and cloud-native development practices provide agility benefits that are especially valuable with customer-facing applications for digital business in edge locations, such as retail, manufacturing, utilities, and transportation and logistics.
- The vast majority of customers are approaching edge solutions from an edge-in perspective, rather than cloud-out.
- In the aftermath of the pandemic, retailers are leaning toward store automation and reevaluating the human-machine mix in their stores; this drives early adoption of edge architectures in the store.
- Management and orchestration of large numbers of distributed edge computing nodes, with associated software and zero touch, are critical edge computing requirements.

Recommendations

- Choose machine learning technologies that enable extensibility for future use cases.

- Align the use of containerization at the edge with adoption-related best practices, such as DevOps and platform engineering, to ensure that the agility benefits of containers can be fully realized.
- Harmonize your edge and cloud strategies and architectures, and ensure consistent technologies when possible (but this does not require distributed cloud solutions).
- Kick off a retail store trial conducive to edge computing, to be piloted during the next 12 months by investigating new store Internet of Things workflows and use cases, such as smart check-out and loss prevention.
- Work with business units to understand future edge computing workloads, and to choose an edge management and orchestration solution that supports that kind of extensibility.

Strategic Planning Assumptions

By 2026, at least 50% of edge computing deployments will involve machine learning (ML), compared with 5% in 2022.

By 2028, 80% of custom software running at the physical edge will be deployed in containers, which is a major increase from 10% in 2023.

By 2027, approximately 5% of large enterprises will deploy a hyperscaler distributed cloud solution for edge computing workloads outside data centers.

By 2027, two-thirds of Tier 1 multichannel retailers will have edge computing deployed in their stores.

By 2027, 20% of large enterprises will have deployed an edge management and orchestration (EMO) solution, compared with fewer than 1% in 2023.

Analysis

What You Need to Know

Edge computing is maturing in technology and adoption. According to the 2024 Gartner CIO and Technology Executive Survey, 19% of respondents have already deployed edge computing, and an additional 32% expect to deploy during the next three years (see Figure 1). On the other hand, many enterprises are still unsure what role edge computing will take in their digital transformation – 30% of enterprises responded that they weren't sure about deploying edge computing.

The vast majority of enterprises will find use cases in which edge computing makes sense, as:

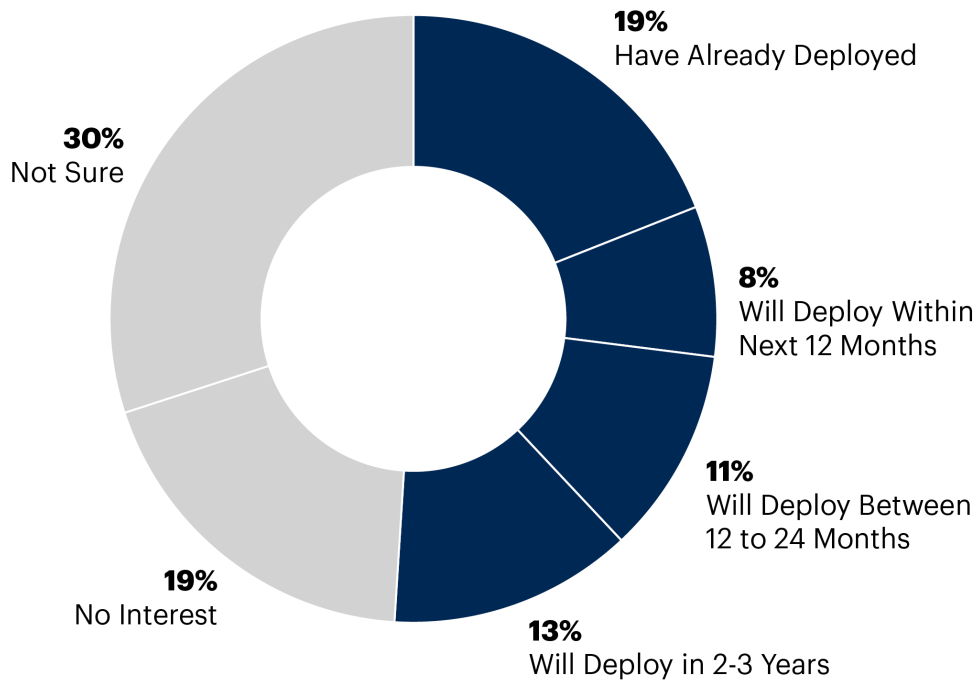
- Digital data and computing become pervasive.
- The role of edge computing in an enterprise's digital transformation becomes apparent.
- Technologies to enable enterprises to deploy edge computing solutions mature.

- Leading-edge enterprises in various industries show results.

Figure 1: State of Deployment for Edge Computing



State of Deployment for Edge Computing Percentage of Respondents



n = 2,443 CIOs and technology executives answering

Q. What are your enterprise's plans for Edge Computing?

Source: 2024 Gartner CIO and Technology Executive Survey

800052_C



This year's edge computing predictions include three areas in which technology and markets are maturing and two projections of edge growth (see Figure 2).

Figure 2. Five Predictions for Edge Computing



Five Predictions for Edge Computing



Source: Gartner
800052_C

Gartner.

- **EMO:** This is a critical platform and management layer for edge computing. EMO is maturing, with a growing set of vendors that have specific areas.
- **The role of hyperscalers:** The cloud complements edge computing, and the role of hyperscalers and the distributed cloud in physical edge deployments is firming up.
- **Edge application platforms:** Edge-native applications have unique requirements, and edge application architectures are becoming clear.
- **Artificial intelligence/machine learning (AI/ML) expansion:** The digital edge is evolving from observability and reporting to richer automation and intelligence, driving growth of AI and ML for the edge.
- **Industry deployments:** Edge computing is being adopted in many vertical industries, and the retail industry promises to be a major edge computing adopter.

Strategic Planning Assumptions

Strategic Planning Assumption: By 2026, at least 50% of edge computing deployments will involve machine learning (ML), compared with 5% in 2022.

Analysis by: Tom Bittman and Eric Goodness

Key Findings:

- Edge use cases are either latency- or data-intensive, or they require a level of autonomy for local decision making or some combination.
- The number of sensors and devices deployed to collect data at the edge from non-IT devices and assets and the software deployed to the edge to generate value from the data are both growing rapidly.

- Devices and software (e.g., AI) at the edge increasingly operate as systems, requiring richer automation and intelligence.

Market Implications:

The role of the Internet of things (IoT) at the edge has evolved from simple observability to reacting to, predicting and avoiding events. From technologies for operationalizing ML (MLOps), to technologies enabling inference models in the smallest edge footprints (e.g., TinyML embedded in IT and non-IT products), ML is becoming a de facto part of many new edge use cases. The hype around AI and generative AI (GenAI) is tremendous. In the 2024 CIO survey, AI/ML was the technology with the most interest and with the most activity during the next three years (34% of CIOs said they had already deployed, growing to 90% by 2026). However, the return on investment (ROI) for avoiding downtime and reducing costs through deep automation is validating the hype, as well as creating clear business cases.

Recommendations:

- Ensure a good business case and ROI. Asset-intensive industries, such as manufacturing, utilities, and transportation and logistics, have shown significant value when augmenting traditional OT systems with AI/ML for advanced anomaly detection and asset management.
- Choose edge-based versus cloud-based solutions for more flexibility, to reduce the amount of data and workload that must move to the cloud, and to enable operational control when disconnected from the cloud.
- Choose ML technologies that can be leveraged for future use cases, and assess the viability and capabilities of vendors for ML at the edge. Although demand is high, many of the vendors in this space are startups, with limited depth for support or customization.
- Identify OT-centric environments where broadening sensorization, data acquisition and event-driven analysis could provide improvements in the performance and availability of non-IT assets.

Related Research:

[Hype Cycle for Artificial Intelligence, 2023](#)

[Innovation Insight: How Enterprises Are Applying Edge Computing](#)

Strategic Planning Assumption: By 2028, 80% of custom software running at the physical edge will be deployed in containers, which is an increase from 10% in 2023.

Analysis by: Tony Iams

Key Findings:

- Containers and cloud-native development practices provide agility benefits that are especially valuable with customer-facing applications for digital business in edge locations, such as retail stores and quick-service restaurants.
- The consistent and predictable runtime environment in containers is useful for running applications in embedded and remote environments, where continuous software updates are required, but debugging and testing are difficult.
- Some engineering challenges remain with delivering containers at the edge, including support for disconnected (i.e., “air gapped”) operation of container infrastructure and ensuring secure supply chains for containerized software. When Kubernetes is used to orchestrate containers at the edge, new tools and operational processes are required to manage large fleets of clusters that can run on compact hardware with minimal local support.

Market Implications:

Containers promise an ideal fit for edge computing, because they provide a consistent environment for applications from development to production. This consistency will be particularly valuable where the production environment is remote or otherwise difficult to access, because container use will increase confidence that applications behave as expected in these environments. The immutability of containers helps to build new applications and to modernize some existing applications with cloud-native patterns. It also enables releases and updates to be delivered to the edge with optimal agility.

These benefits will drive the adoption of containers for custom-developed software with most edge use cases. If their resource requirements and operational overhead can be sufficiently reduced, it may also be possible for Kubernetes-based platforms to orchestrate containers at the edge.

Alternatives that slow the adoption of containers include microVMs, which improve performance and have lower resource requirements than traditional virtual machines (VMs). WebAssembly (Wasm) is a lightweight virtual stack machine and binary code format designed to support secure, high-performance applications. Both are compatible with Kubernetes, and could provide a better fit for running existing and next-generation applications at the edge.

Recommendations:

- Simplify the implementation of containers with integrated stacks versus building and adding containers over another stack. Leverage the container integration and management/orchestration for the entire stack from a single vendor, if possible.
- When starting with an edge-in approach, run containers on a hyperconverged infrastructure (HCI) stack. Containers complement VMs, and they can be operated together. However, it may be simpler and more resource-efficient to run containers without VMs (i.e., on “bare metal” hosts).

- Align the use of containerization at the edge with the adoption-related practices, such as DevOps and platform engineering, to ensure that the agility benefits of containers can be fully realized.

Related Research:

[Solution Path for Cloud-Native Infrastructure With Kubernetes](#)

[Guidance Framework for Deploying Containers and Kubernetes](#)

[Designing and Operating DevOps Workflows to Deploy Containerized Applications With Kubernetes](#)

[Decision Point for Selecting Virtualized Compute: VMs, Containers or Serverless](#)

Strategic Planning Assumption: By 2027, approximately 5% of large enterprises will deploy a hyperscaler distributed cloud solution for edge computing workloads outside data centers.

Analysis by: Tom Bittman

Key Findings:

- Cloud services and cloud-based applications complement edge deployments, especially for enterprise integration, IoT management, streaming analytics and ML.
- Hyperscaler distributed cloud solutions are primarily targeting applications and data that need to remain in enterprise data centers and cannot migrate to hyperscaler data centers.
- Closer to the edge, the diversity of use cases, compute requirements, environmental requirements, form factors and vertical industry requirements reduce the efficacy of hyperscaler distributed cloud solutions, even as back-end cloud services remain valuable.
- The vast majority of customers are approaching edge solutions from an edge-in perspective, not cloud-out.

Market Implications:

Edge computing deployments are growing, and hyperscalers will provide some level of back-end services and integration for the vast majority of them. However, this doesn't mean they need to provide distributed cloud footprints to the edge itself.

Since the inception of cloud services, hyperscalers have addressed region-specific and latency issues through regional data centers, colocation providers and content delivery networks (CDNs). Starting in 2019, hyperscalers began adding 5G core and multiaccess edge computing (MEC) support, and created partnerships with wireless providers — more than three dozen partnerships with Amazon Web Services (AWS), Google and Microsoft alone. The initial distributed cloud offerings

(starting in 2016) were strictly focused on enterprise data centers, and large footprint deployments. For hyperscalers, this is an important onramp for cloud migration. Likewise, early data transfer devices, such as AWS Snowball, focused on physical data transfer to the cloud. The initial IoT offerings from the hyperscalers targeted the management of IoT devices and the processing of streaming data in hyperscaler data centers.

All of these offerings are essentially “cloud-out” – getting cloud services closer to users. However, the vast majority of edge customers are approaching the digital edge from an industry-specific “edge-in” perspective: automating equipment and systems at the edge, improving the shopping experience, etc.

Customer requirements for lower latency, data gravity (processing data where it is created to reduce bandwidth costs), and resilience when disconnected push requirements for on-location processing. Edge computing environments are constrained (in terms of power and environmental), have domain-specific requirements, are less tolerant to disconnection, and have applications designed specifically for the edge.

For hyperscalers, the opportunity is less about the on-location processing than it is about the back-end services used for those deployments. There will be edge computing market niches in which distributed cloud footprints make sense, and high-volume niches that hyperscalers can address with targeted solutions. For the most part, hyperscalers will focus on partnerships and on making their back-end services ideal for edge integration.

Recommendations:

- Harmonize your edge and cloud strategies and architectures, and ensure consistent technologies when possible; however, this does not require distributed cloud solutions.
- Choose distributed cloud offerings that meet your requirements for resilience to disconnection, with appropriate feature sets, solutions and partners for your specific requirements.
- Talk with reference customers that have similar requirements first; they may be hard to find in these early days of distributed cloud for edge.
- Evaluate offerings from vendors focused on edge computing that meet your specific requirements, and integrate well with hyperscale-based applications and services.

Related Research:

[Hyperscalers Stretching to the Digital Edge](#)

[Market Guide for Edge Computing](#)

Comparing On-Premises Public Cloud Appliances: AWS Outposts, Microsoft Azure Stack Hub and Google Distributed Cloud Edge

Magic Quadrant for Cloud Infrastructure and Platform Services

Strategic Planning Assumption: By 2027, two-thirds of Tier 1 multichannel retailers will have edge computing deployed in their stores (see Note 1).

Analysis by: Sandeep Unni

Key Findings:

- Edge computing can efficiently streamline the deluge of data from a number of store IoT solutions. These include smart shelves, smart freezers, item-level radio frequency identification (RFID), smart robots and smart check-out, as well as traditional store systems, such as point of sale (POS) devices.
- Combined with AI/ML, edge computing optimizes the inference and analysis of large datasets generated at the store edge.
- Leveraging edge AI solutions in the store vastly improves in-store execution by accurately tracking inventory quantities at store locations in real time. Inventory accuracy also enables a dramatic reduction in required safety stock, driving more accurate upstream processes, such as demand forecasting and significantly freeing up retailers' working capital.
- The adoption of edge computing enables retailers with innovative real-time insights to increase operational efficiencies, decrease operational costs, and enhance personalized, immersive consumer and associate experiences in store.
- In-store, edge AI solutions proactively support the monitoring and reduction of shrinkage, which was estimated by the National Retail Federation (NRF) to have cost the industry almost \$100 billion in FY21 through theft, shoplifting and returns fraud.

Market Implications:

The retail adoption of edge architectures for store technology solutions continues to show promising early adopter momentum. In the aftermath of the pandemic, retailers are applying automation and undertaking a significant reevaluation of the human-machine mix in stores to compensate for labor and talent shortages and satisfy the need for faster, more-accurate decision making.

Despite the early traction, retailers are moving ahead largely on an individual solution basis, driven by specific store use cases. Moreover, retailers are constrained by lack of existing staffing and domain knowledge at the retail stores and corporate IT. Deployments are being led by retail solution

providers, systems integrators (SIs) or independent software vendors (ISVs) serving the various store workloads.

During the next three to five years, we anticipate the edge computing market in retail to mature from individual store use cases to a more-holistic strategy led at the retail enterprise level, and driven by standardized and extensible architectures. In concert, we expect the vendor ecosystem in the industry to also consolidate capabilities and extend into more total solutions and frameworks spanning management, workload orchestration and security to enable deployments at scale. Retailers need to start planning now with dedicated IT investments in edge resources and talent.

Recommendations:

- Devise an edge computing strategic plan for your retail business by establishing a dedicated edge innovation team in your enterprise architecture function.
- Commission your edge or store architects to undertake an evaluation of store workloads and needs in your current footprint suited for edge computing.
- Kick off a store pilot during the next 12 months by investigating new store IoT workflows, as well as such use cases as smart check-out and loss prevention.

Related Research:

[Hype Cycle for Retail Technologies, 2023](#)

[Infographic: The Retail Store of Tomorrow](#)

[Top Trends Shaping the Retail Store Associate Experience](#)

[Innovation Insight: How Enterprises Are Applying Edge Computing](#)

Strategic Planning Assumption: By 2027, 20% of large enterprises will have deployed an edge management and orchestration solution, compared with fewer than 1% in 2023.

Analysis by: Tom Bittman

Key Findings:

- Managing and orchestrating a large number of distributed edge computing nodes with zero touch is a critical edge computing requirement.
- As edge computing expands, platforms that enable edge extensibility beyond a specific use case will be necessary.

- To enable edge agility, enterprises require the ability to deploy new applications and capabilities rapidly.

Market Implications:

Early edge computing demand is focused on enabling and managing very specific use cases, which are often addressed by tightly integrated solution stacks. Although this solves an immediate need and can often be deployed quickly, it also creates inflexibility and technical debt. As early adopters of edge computing have learned, use cases and applications at the edge multiply. Enterprises are often forced to deploy additional solution stacks, or to replace and consolidate to a more flexible platform.

Ultimately, enterprises will evolve to use EMO platforms that have several or all of these capabilities:

- Fleet monitoring, management and automated operations
- Edge design studios
- Edge application deployment and updates
- Edge orchestration
- Edge software platforms and platform as a service (PaaS)
- Edge security technologies

EMO is an emerging market with a growing number of vendors. However, the challenge is often that customers want to buy solutions, rather than platforms. This has forced EMO vendors to focus on specific verticals, use cases and partners. In addition, vendors that focus on tightly integrated solutions will partner with EMO vendors to expand their market opportunities and reduce development costs.

As IT and OT organizations become stronger partners for edge computing architectures, EMO will become a critical building block that enterprises will use to increase edge extensibility and agility, while reducing management and operations costs.

Recommendations:

- Work with business units to understand future edge computing workloads, and choose an EMO solution that supports that kind of extensibility.
- Ensure that the EMO solution is effective at the volume and footprint scale you envision, in the geographies you require and with differences in connectivity.

- Evaluate the long-term viability of the EMO solution, understand the state of the partner ecosystem and ensure that the EMO solution integrates well with your chosen hyperscaler provider.
- Check references for real-world deployments that are similar to yours, with equivalent vertical industry requirements.

Related Research:

[Hype Cycle for Edge Computing, 2023](#)

[Market Guide for Edge Computing](#)

A Look Back

In response to your requests, we are taking a look back at some key predictions from previous years. We have intentionally selected predictions from opposite ends of the scale – one where we were wholly or largely on target, as well as one we missed.

On Target: 2021 Prediction — By 2025, bandwidth cost will be the primary driver for new edge computing deployments, versus latency in 2021.

Latency remains an important driver for use cases in 2023. However, the growth of data production at the edge, and the growth and evolution of edge AI (mostly inference, but even some training at the edge) has shifted the balance already. In industrial verticals (such as manufacturing or energy), data and AI growth — as well as the need for disconnected operations — have already made bandwidth cost the primary driver of edge use cases. In less asset-intensive verticals (e.g., retail), latency remains more important, but the growth of local data, data sovereignty requirements and disconnected state are causing more data to be processed and filtered locally.

In the end, drivers for edge computing have become more diverse. Latency remains important for many use cases, but those that require data to be processed locally for sovereignty, cost or resilience reasons are growing faster.

Missed: 2020 Prediction — By year-end 2023, 20% of installed edge computing platforms will be delivered and managed by hyperscale cloud providers, compared to less than 1% in 2020.

After a flurry of announcements from hyperscaler cloud providers, there was a tremendous amount of hype that their distributed cloud offerings would dominate edge computing deployments. Our 2020 prediction was intended to deflate the hype. We didn't deflate it enough. As we stated then, "edge requirements can be highly diverse and are difficult to address with general-purpose solutions." We continue to believe that "edge computing will be delivered by ecosystems of partners that vary across use cases." However, this year, we have updated our prediction to show that, although hyperscalers will almost always be a part of an edge computing solution stack, their role closer to the edge will remain more limited.

Evidence

2024 Gartner CIO and Technology Executive Survey. This survey was conducted online from 2 May to 27 June 2023 to help CIOs determine how to distribute digital leadership across the enterprise and to identify technology adoption and functional performance trends. Ninety-seven percent of respondents led an IT function. In total, 2,457 CIOs and technology executives participated, with representation from all geographic locations, revenue bands, and industry sectors (public and private).

Disclaimer: The results of this survey do not represent global findings or the market as a whole; rather, they reflect the sentiments of the respondents and companies surveyed.

Note 1. Tier 1

Gartner defines the Tier 1 market as large, global retailers that conduct business in multiple geographic locations. Tier 1 global retailers operate extensive store real estate as one of many channels and touchpoints, and generate annual retail revenue of at least \$3 billion per year.

Learn how Gartner can help you succeed.

Become a Client ↗

© 2024 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by [Gartner's Usage Policy](#). Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "[Guiding Principles on Independence and Objectivity](#)." Gartner research may not be used as input into or for the training or development of generative artificial intelligence, machine learning, algorithms, software, or related technologies.

[About](#) [Careers](#) [Newsroom](#) [Policies](#) [Site Index](#) [IT Glossary](#) [Gartner Blog Network](#) [Contact](#) [Send Feedback](#)

Gartner[®]

© 2024 Gartner, Inc. and/or its Affiliates. All Rights Reserved.