

Building the Resilient Multi-Access Edge

1. Introduction

Technology companies, service providers, and enterprises are all excited about the new business opportunities enabled by extending cloud computing to the edge of the network. By pushing cloud networking, compute, storage, and distributed applications closer to the edge, new applications are emerging that could result in a wide range of potential new services.

Futuriom research indicates that service providers are interested in deploying edge services to generate new business and revenue streams, while enterprises are eager to consume new edge applications that can improve customer experience and automation. Emerging applications at the edge include industrial Internet of Things (IIoT), business analytics, connected vehicles, video surveillance, augmented reality, and online gaming – just to name a few. These applications would leverage multi-access edge compute (MEC) and connectivity to the cloud, delivering low-latency processing and more efficient use of cloud bandwidth. They can be bundled with communications in an as-a-service model, representing large revenue growth opportunities.

It's important to note that virtualization is essential technology for 5G networks. In the 5G core, virtualization enables network slicing to support multiple virtual networks running over a single physical network infrastructure. 5G virtualization provides a new level of flexibility, for example allowing communications service providers (CSPs) to allocate a network slice to specific kinds of devices. At the edge of the network, 5G virtualization enables CSPs to deploy a wide array of sliced networked services from a single physical network, thereby diversifying, expanding, and increasing their revenue streams in a highly cost-effective way.

In order to deliver these services, a new type of edge infrastructure is emerging. It will require a unique blend of packet-based virtualized networking, compute, storage, and cloud connectivity. The total new market for edge infrastructure will be in the billions of dollars.

Many edge applications can be supported with 4G connectivity as well as wireline networks. A range of new use cases, however, will require 5G wireless in order to maximize connection density, boost bandwidth, minimize latency, or guarantee quality of service. These 5G use cases fall into the areas of ultra-reliable low-latency communications (URLLC) and massive machine-type communications (mMTC). As 5G deployments reach mass-market scale in 2021, the networking infrastructure will be in place to introduce these new use cases, especially in areas of high-demand density, thereby opening up new business opportunities for companies throughout the MEC ecosystem.

Edge compute, storage, and networking technology will also play a large role in enterprise deployments of edge applications – for example, to support analytics in warehouse or retail locations, regardless of the connectivity – as well as in regional colocation facilities or points of presence (PoP). This means that edge applications and MEC can be connected through a variety of networking infrastructures, including 4G/5G wireless and wireline.

Edge cloud deployments are progressing through three overlapping phases. First, edge-hosted content delivery networks (CDNs) were introduced as highly distributed server clusters that help minimize delays in loading web page content by reducing the physical distance between the server and the user. This helps users view the same high-quality content without slow loading times. Currently, CSPs worldwide are introducing the 5G telco cloud, which enables the delivery of value-added services to both consumers and enterprises, leveraging the unique capabilities of 5G that are discussed in more detail below. Finally, “Industry 4.0” brings on-premises applications and services to smart factories, enabling real-time data collection and analysis in support of highly efficient, automated process control and quality inspection functions.

In this independent market brief, we explain key market drivers and business opportunities behind edge services and MEC, outlining some notable recent deployments by service providers. We identify specific use cases that can be enabled at the edge, highlighting some of their unique business models.

2. Key Cloud Edge Demand Drivers

The basic concept of the edge is straightforward: Local compute and storage are needed at the edge of the network, close to the physical location where data is both created and consumed by humans and machines, enabling that data to be processed and analyzed locally rather than in a central data center or in the cloud. This architecture minimizes the latency associated with decision-making in real-time applications, reduces connectivity costs by sending only the information that matters to the cloud, improves security by keeping sensitive data close to its source, and enables autonomous operation in the event that connectivity is lost or intermittent.

As an extension of cloud computing, MEC enables real-time data collection and analysis, while complementary applications running in the cloud itself provide centralized analytics and interpretation. Combining MEC with cloud compute enables businesses to maximize their overall operational efficiency while deploying new types of services and applications that depend on real-time responsiveness to data from both consumers and devices.

But that also brings challenges: By pushing compute and storage to the edge infrastructure, new networking capabilities are needed. The network needs to be programmed to meet application demand – providing critical quality of service (QoS) and security for connected cloud applications. A MEC node can be located at a radio access network (RAN) cell site, a factory floor, a central office (CO), a mine, an oil or gas rig, a metro aggregation site, a roadside micro-data center, a vehicle, a consumer device, or anywhere that constitutes the network edge for the relevant application or service. This will require an enormous expansion of edge data centers and nodes to accommodate these services. According to PricewaterhouseCoopers (PwC), the global market for edge data centers is expected to nearly triple to \$13.5 billion in 2024 from \$4 billion in 2017, thanks to the potential for these smaller, locally located data centers to reduce latency, overcome intermittent connections, and store and compute data close to the end user.

Companies in a wide range of vertical markets have started to adopt MEC as a key technology driving the transformation of their businesses, as described below.

Industrial IoT (IIoT)

In an industrial environment, MEC reduces connectivity costs by processing data from sensors locally and sending only relevant information, instead of raw data streams, to the cloud.

Within smart factories, MEC, along with 5G connectivity, can minimize end-to-end latency, which comprises the total time for data to travel from a sensor to the data center, the information processing at the data center, and the response time back to the factory. With the most time-critical modules located close to the sensors, the total latency is dramatically reduced thanks to 5G, making the manufacturing process more agile and more responsive to changes. For maximum efficiency and operational cost savings, the modules necessary for real-time decision-making can be hosted on-premises, while those that are less time-critical run in the cloud.

Finally, security represents a key concern for enterprises operating in the industrial segment. A cloud edge architecture ensures that the majority of real-time, process-dependent, and machine-specific information never leaves the factory, minimizing the threat of it being intercepted or hacked.

Video Surveillance

Over the past decade and enabled by the proliferation of low-cost, remotely managed, high-resolution networked IP cameras, video surveillance technology has become prevalent across a diverse set of applications. Beyond surveillance of people in transportation hubs, city streets, retail environments, events, and other public places, video surveillance is also widely used for commercial and residential security, as well as for monitoring processes in applications such as manufacturing, agriculture, energy, and logistics.

Processing video data from cameras remotely in the cloud requires a high-bandwidth network connection to transmit a constant, high-resolution stream for remote analysis; or, if bandwidth is limited, forces only low-resolution, compressed video to be transmitted, which reduces the accuracy of the analysis.

MEC addresses these problems by enabling high-resolution video data to be processed either within the camera itself or at a local edge server, rather than remotely in the cloud. High-end IP cameras have enough processing power to run algorithms such as facial recognition, leveraging analytics based on artificial intelligence (AI) and/or deep learning technologies. With this approach, only selected events and/or short video sequences that are flagged as important — such as a specific individual, vehicle, or defective component — need to be transmitted to the cloud. This significantly reduces the required backhaul bandwidth while maintaining high-quality analytics with maximum accuracy.

MEC for video also results in significantly reduced latency, which is important for any time-critical detection scenario. Locally processed facial recognition, for example, can recognize an unwanted or suspicious person and immediately send an alarm to the appropriate recipient. Processing video data at the edge also helps to alleviate privacy concerns, especially in surveillance applications that are subject to regulatory constraints or in commercial applications where, for example, manufacturing data constitutes valuable intellectual property.

Virtual Reality (VR)/Augmented Reality (AR)

Applications for virtual reality (VR) and augmented reality (AR) extend far beyond consumer use cases such as Pokémon Go and its successors. Manufacturing, for example, is one segment where AR has the potential to significantly improve productivity through cost reductions along with increased operational efficiency.

AR applications typically analyze the output from a device's camera to supplement a user's experience when visiting a point of interest. The application needs to be aware of the user's position and the direction they are looking in — with this information provided via the camera view, positioning techniques, or both. Following analysis, the application is then able to offer

additional information in real time to the user. As soon as the user moves, that information needs to be refreshed.

For AR experiences, a key advantage of MEC along with 5G connectivity is the ability to reduce dizziness associated with high latency and slow frame refresh rates. This dizziness leads to an experience that is frustrating, potentially nausea-inducing, and ultimately disorienting. Moving compute power into edge servers located close to the network edge allows an AR application to eliminate the need for high-processing bandwidth on goggles that therefore become expensive, power-hungry, and too heavy for comfortable use over an extended period. One of the biggest benefits of MEC-hosted AR is that latency is predictable, resulting in a consistent experience for users instead of the constantly changing delays that result from cloud-hosted implementations. MEC can also help to address privacy concerns around AR, since user data is processed in a local edge data center rather than being sent to the cloud over the public network.

VR and AR are poised to deliver truly immersive media experiences for sports fans while at the game. Sports such as baseball, cricket, football, and soccer have already held successful trials in “smart stadiums,” allowing spectators to stream video from unique, custom camera angles, including from drones and spider-cams, all connected by 5G. “Virtual cameras” present views from within the field of play, giving spectators the opportunity to experience the action from the perspective of the players themselves. All these use cases require MEC in order to guarantee the responsiveness that spectators expect while eliminating the need to backhaul prohibitive amounts of data to the cloud.

Cloud Gaming and E-Sports

Although the size of the mobile gaming market was estimated at over \$18 billion in 2018, growth of downloaded games and in-app purchases has flattened over the past few years. This is partly due to gamers’ expectations of seamless and sophisticated gaming scenarios that demand a truly immersive, real-time experience requiring large amounts of storage and high-processing power.

Massively multiplayer games (MMPGs) typically involve players controlling their avatars, with any movement of an avatar needing to be communicated as quickly as possible to all players who have that avatar in their field of view. Latency has a major impact on the overall user experience, to the point where perceptible delays can render a game effectively unplayable. A video game must appear to respond instantaneously to keystrokes and controller movements, implying that any commands issued must complete a round trip over the network and be processed fast enough by the data center for the player to feel like the game is responding in real time. For the best multiplayer experience, the latency must be consistent across all players, otherwise those with the lowest latency have the opportunity to react faster than their competitors.

MEC improves the experience of cloud gaming by significantly reducing latency and providing the necessary storage and processing power in edge data centers. With processing centers for a game running at the edge of the network — for example, in each metro area — the ultra-low latency results in reduced lag time. This enables a more interactive and fully immersive experience than if the game is hosted in a remote cloud data center.

Smart Retail

MEC enables retailers to use the vast amounts of data connected by sensors and cameras to deliver highly responsive and personalized experiences for customers. One example of this trend is Amazon's Go store concept, which uses a wide array of advanced technologies, including sensors, deep learning algorithms, and edge compute to deliver a fully automated retail experience. In order to respond quickly to customers, technologies such as AI, AR, and machine learning (ML) need to be located closer to the edge of the network to deliver real-time results.

AR solutions that show customers how a product can be used, what reviews there are on a particular item, or even what other products would complement an item they're considering can all engage customers in the brick-and-mortar sales experience.

By leveraging MEC, retailers are also better positioning themselves to compete with the likes of Amazon and other online retailers. While physical shopping will never live up to the convenience of online shopping, it can offer other benefits, and it's critical for retailers to preserve a unique and fulfilling buying experience in physical stores by leveraging cutting-edge technologies.

By improving store security and surveillance and understanding consumer trends while providing engaging in-store experiences that drive sales and encourage brand loyalty, the low latency associated with MEC enables brick-and-mortar retailers to transform their customers' shopping experience.

Application Development Environments

Developers create edge applications using modern principles, such as continuous integration and continuous deployment (CI/CD). New vendors are emerging to serve this new class of developers with solutions that simplify the cost-effective deployment of edge applications. Indeed, an important early category of "killer apps" for MEC is the development and provisioning solutions that make the new edge experiences possible.

3. Notable Edge and MEC Deployments

Over the past couple of years, CSPs worldwide have deployed MEC to address a variety of use cases, including the following:

- **AT&T** is deploying a MEC network at IBM's lab in Yorktown Heights, New York, to develop and showcase new capabilities for enterprises seeking a private, on-premises 5G+ millimeter wave (mmWave) service. This will allow IBM and eventually other enterprises to run workloads on-premises with low latency, increased privacy, and improved security.
- **AT&T** also deployed its AT&T Network Edge (ANE) platform to enable Game Cloud Network to deliver its new multiplayer, interactive "Tap and Field" mobile gaming experience, providing players with near-real-time interaction.
- **China Unicom** used MEC to transform the experience of fans during concerts at the Mercedes-Benz Arena in Shanghai, delivering personalized live video at low latency to thousands of users in real time. Fans were able to stream HD video from four unique camera channels using a mobile application on their device, with a negligible latency of just 0.5 seconds.
- **China Mobile** discussed plans to deploy edge cloud compute on 100 test nodes, open 100 application programming interfaces (APIs) for edge cloud computing capabilities, and introduce 100 cooperative partners to promote commercial applications, all in 2019. It explained that MEC is necessary to meet the requirements of the three main 5G use cases: URLLC, mMTC, and enhanced mobile broadband.
- **Rakuten** launched its 5G mobile service in Japan, leveraging virtualization and cloud-native technologies throughout the network, including an edge-hosted virtual RAN solution from Altiostar.
- **Deutsche Telekom** is working with lighting manufacturer Osram to deliver factory and warehouse automation solutions.
- **Korea Telecom** has delivered MEC-based machine vision solutions for smart factories in partnership with Cognex.
- **Orange** is collaborating with Dell on a MEC initiative to (1) define use cases and business models, (2) validate infrastructure accelerators such as field-programmable gate arrays (FPGAs), graphics processing units (GPUs), and smart network interface cards (smart NICs) for edge workloads, and (3) develop AI/ML-enabled software for remote automation of a heterogeneous edge.
- **NTT** partnered with Clouidian on a MEC-based solution to interpret and analyze surveillance camera footage, leveraging a compact, high-speed AI data processing device equipped with a camera and LTE/WiFi connectivity.

- **Lumen** (formally CenturyLink) is investing in MEC capacity to deliver edge cloud services in sectors such as retail, hospitality, food services, logistics, transportation, and manufacturing where the businesses are spread over a wide geography. While one customer had been considering buying servers to host a business application in each of its 2,000 sites, Lumen deployed it for them on 100 MEC nodes, ensuring the low round-trip latency the application requires.
- **SK Telecom** delivered its MEC-based Detective Inspection solution to South Korean auto parts manufacturer Myunghwa Auto Parts for factory automation and process inspection applications.
- **Telefónica** launched its MEC-based cloud gaming platform in Spain, replacing gaming consoles with MEC nodes from which the video output is sent to the user’s display.
- **Telstra** has deployed MEC-based IoT services for agricultural use cases in Toowoomba, Australia.
- **Verizon** teamed up with Amazon Web Services (AWS) to deliver MEC-based cloud gaming services in partnership with video game publisher Bethesda Softworks and the National Football League (NFL).
- **Vodafone** is preparing its networks for distributed MEC services in industrial use cases such as video analytics, real-time asset inspection, AR for remote field teams and drone control, as well as consumer-oriented applications like AI-powered media editing and immersive experiences in retail, events, and tourism.

4. Edge Opportunities in 5G

While edge compute capabilities can certainly be delivered without 5G, there is no doubt that 5G will accelerate edge capabilities and services. The business value of 5G connectivity for edge compute use cases derives from the improved operational efficiencies that it brings to edge applications as well as the new types of services that it enables.

	4G/LTE	5G
Peak Data Rate	1 Gbps	20 Gbps
Latency	10 ms	<1 ms
Peak Connection Density	70K devices/km ²	1M devices/km ²
Network Slicing	No	Supported in standalone (SA) configuration

The capabilities of 5G will enable MEC use cases that were either not achievable or not cost-effective using 4G technology. Some examples include the following:

- **Highly immersive AR applications** will leverage the deterministic, ultra-low latency of 5G to produce new levels of user experience in industrial, consumer, and telemedicine use cases.
- **Connected vehicles** will leverage the ultra-low latency and high bandwidth of 5G for both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. These applications will also take advantage of MEC to avoid the latency associated with round trips to the cloud, enabling massive amounts of data to be processed with AI and analytics in near-real time. An example use case involves embedding sensors in roads to flag hazardous road conditions and send that data to connected vehicles.
- **Manufacturing quality inspection** will leverage the high bandwidth and low latency of 5G to ensure near-real time detection of process deviations or manufacturing flaws, enabling production lines to be stopped or control parameters to be adjusted quickly enough to minimize wastage.
- **Drones** used for applications such as surveying, package delivery, and video surveillance will leverage highly reliable, low-latency 5G connections for control, while also enabling object recognition for navigation to be performed on the ground, resulting in reduced costs and extended battery life.
- **Mission-critical media streaming** will proliferate due to the capability of processing multiple video streams at the edge thanks to 5G network slicing, accelerating the deployment of use cases such as AI-driven security, facial recognition, safety, and compliance in the workplace, as well as autonomous traffic flow control based on computer vision.
- **Cloud gaming** will benefit from the combination of 5G and MEC, which increases path diversity for “core hosted” game servers and results in better network connections for gamers. Highly latency-sensitive, multi-player games perform better over mobile networks, so 5G together with MEC will spur the development of new games optimized for mobile devices.

Beyond the kind of individual applications mentioned above, much of the value of 5G technology for MEC applications is derived from 5G’s characteristics as a platform that removes complexity for the developers of applications for IoT and other edge services. And in many scenarios, MEC is the key enabler for 5G networks, rather than the other way around. AT&T, for example, has stated that many of its mmWave 5G customers have used on-premises MEC nodes to deploy a private 5G network for their businesses.

5. Business Challenges for Edge Deployments

The distributed edge will require an integrated approach that can connect MEC with packet networking, optical, and RAN technologies that can handle massive scale. Some of the crucial new technologies that will play a role include Open RAN, distributed routing, and segment routing (SR); edge orchestration and automation software; integration with optical networks; and edge-data management and operations. SR in particular will be important in the use of next-generation IP technologies to support classes of service and prioritization of traffic flows across multiple domains, including optical, IP, and the RAN. The concept of adaptive IP uses SR and other modern IP technologies to eliminate a node-level protocol and provide optimized traffic engineering, troubleshooting, and service orchestration across a multi-vendor, end-to-end network.

At first glance, edge-compute applications could offer many compelling benefits for enterprises and/or consumers. However, they present important business challenges that must be acknowledged and addressed in order for deployments and operations to become financially viable.

Ecosystem Complexity

There are very few companies with the capability to develop, deploy, and operate a complete end-to-end edge storage and compute service. And even if these companies did exist, the industry push toward open architectures and away from proprietary, single-vendor solutions means that the majority of use cases will involve some combination of application software vendors, platform software vendors, hardware vendors, edge data center operators, CSPs, systems Integrators (SIs), installers, cloud service providers, regulators, and even municipal utilities. With the advent of 5G, it will be increasingly important to implement multi-vendor orchestration capabilities that can deliver end-to-end services.

Sustaining a successful outcome for the end user of the edge service implies an end-to-end revenue model that ensures each company in the ecosystem can meet its business objectives for the project. If only one company in the chain fails to deliver, support, or operate their product or service, the entire use case may fail or at least miss its business goals.

Multi-vendor partnerships have already been formed around end-to-end service chains. For example, at Mobile World Congress in 2019, Affirmed Networks, Akamai, Ciena, Equinix, Tiledmedia, and Wind River provided a live demo of the automated delivery of mobile video services across a distributed carrier cloud from the edge to the core.

Edge Application Development

Edge services will flourish only when high-quality software applications become available for commercial deployment that address high-value enterprise and consumer business opportunities. While a vast community of companies has spent many years developing and optimizing applications for the cloud environment, this is not the case for the edge.

Standard APIs and programming primitives will be key to enabling MEC services to run on any infrastructure platform worldwide, allowing developers to focus on the application logic rather than specifics of deployment locations and individual infrastructure.

Integration and Operation

From the perspective of the end user that is ultimately paying the bills, who is the “single throat to choke” that’s responsible for the deployment and operation of the edge service?

Very few companies understand the hosting of a 5G RAN as well as the integration, deployment, and operation of MEC-hosted cloud software. In particular, the large SIs with decades of experience in installing and operating equipment in massive industrial environments such as factories, oil rigs, and mines typically do not have the expertise in telecom working or cloud applications. There are interesting opportunities for alliances and partnerships between different types of integrators with complementary expertise, which can ensure a viable business that involves integrating, supporting, and operating 5G MEC services in specific vertical markets.

Cloud-Native Architectures

The evolution of virtualization technology from virtual machines (VMs) delivering “simple virtualization” to containers and container network functions (CNFs) enabling cloud-native architectures produces a more efficient infrastructure that can host more functions per node, while doing so in a more performant manner than legacy architectures, with positive attributes such as lower latency. Newer applications, such as the 5G virtual RAN (vRAN), are being developed directly within cloud-native architectures and will accelerate the adoption of this approach. Service providers are increasingly faced with the complex operational and technological challenges of a physically distributed, cloud-native vRAN infrastructure.

CSP Virtualization Expertise

While many CSPs have made strides toward becoming software-oriented organizations through their initiatives in software-defined networking (SDN) and network functions virtualization (NFV), few would claim to have completed that transformation. Indeed, most CSPs struggle to

recruit talented software engineers in the face of competition from both the webscale giants and from scrappy startups flush with stock options.

MEC applications enabled by 5G represent significant opportunities for CSPs to expand their business into new vertical markets such as smart cities, smart factories, and cloud gaming. Unless they achieve the right level of software expertise, however, they risk losing these commercial opportunities to public cloud providers like Amazon, Equinix, Google, Microsoft, and others.

High-Performance Networking

While many potential edge applications such as autonomous vehicles will only be possible when 5G is broadly available, edge compute applications can employ any connectivity available – including Ethernet, IP, and optical fiber. Because of the diverse networking needs of edge services, edge nodes need to be adaptable to a wide variety of packet-based networking technologies. The network will have to include intelligence and software-based management that can route and segment applications based on performance needs.

This could become a chicken-and-egg problem: Without an URLLC network, there is no killer app, while without a killer app no one will pay for a URLLC network and deploy edge compute. But the main challenge is that without high-performance, virtualized networking technology, these edge apps will not deliver the necessary performance.

6. Conclusion: Key Differentiators for the Edge

The emergence of the new edge cloud architecture is expected to deliver many new capabilities for both enterprises and service providers, making it a key strategic growth platform for the communications technology industry. The business impact will be in the trillions of dollars.

The arrival of 5G will serve as a catalyst to accelerate network architectures built on technologies such as virtualization and microservices. This new edge cloud will be the first to be built on an entirely packet-based, software-defined architecture – or adaptive IP. The emergence of the edge is also expected to transform the public cloud by shifting the center of gravity away from public cloud, providing opportunities for cloud providers, communications service providers (CSPs), and enterprises alike. Service providers have a new opportunity to monetize networks.

The complexity and the performance demands of edge applications will be significant, however. These applications will require high-performance, software-based orchestration and management of networking services using adaptive IP. This will be necessary to deliver end-to-end performance of applications over multi-vendor networks, using modern virtualization and

IP networking techniques. This integrated IP-based networking approach will provide high-performance routing, virtualization, and connectivity over multi-domain RAN, IP, and optical networks.

In the past six months, Futuriom has gathered end-user survey feedback about development of the cloud from a diverse range of CSPs, cloud operators, and large enterprises. Some of the key technology differentiators for delivering the high-performance cloud edge are detailed below.

Edge compute will be a key driver of CSP automation. Service providers will not be able to keep pace with cloud-scale expansion of bandwidth and connections without using scale-out cloud automation techniques. Ninety-five percent of respondents to the Futuriom 2020 CSP Automation survey said they think that data-driven network automation will be crucial to successfully launching 5G and edge networks.

Edge compute drives virtualization. The diverse range of edge applications and the data demand of MEC mean that the cloud edge will require high connection density, low latency, and high bandwidth. In order to support a wide range of demanding edge applications, the edge infrastructure also needs to support high-performance network slicing and network virtualization.

The cloud edge has high-performance routing requirements. In order to support edge services such as 5G, MEC, and URLLC apps on virtualized platforms, the cloud edge will have to run on high-performance packet-based networks. Futuriom believes this will require IP that can operate at line rate and enable integration with the cloud RAN (C-RAN) environment. Emerging technologies such as segment routing multiprotocol label switching (SR-MPLS) will likely be employed to meet the needs of high-performance apps. SR capabilities will enable applications and services to be assigned different levels of performance that can be managed using end-to-end traffic flows and traffic engineering, all delivered in a multi-vendor, heterogeneous, and multi-domain environment.

Cloud-native mindset is required. Service providers will need to make the leap to a cloud-native mindset, which involves building infrastructure based on a programmable, scale-out mindset. The move to cloud-native platforms for edge compute can improve automation, auto-scaling, and service velocity while at the same time reducing both capital expense (capex) and operational expense.

The operators that can implement an agile, high-performance edge infrastructure will be best positioned to deliver high-value, edge-focused use cases for both consumer and enterprise segments.