Publication date:
27 May 2020
Author:
Julian Bright

Telcos and Edge Computing: Opportunity, Threat, or Distraction?

In partnership with:

Brought to you by Informa Tech

Contents

Summary

In brief

The market for edge computing is experiencing rapid growth and has naturally captured the attention of communications service providers (CSPs). Although still in its infancy, edge computing will almost certainly prove the most significant new network architecture of the coming decade.

By providing the capability to process, manage, and analyze digital content in near-real time, close to where applications are accessed and devices located, edge computing has the potential to unlock new applications and service opportunities. Often closely tied to 5G, the edge is not in fact dependent on any single access technology and may not prove to be the monopoly of any single class of provider. As commercial deployments start to be rolled out, it is still not clear if CSPs are among those who stand to benefit, and if they are, how much it will cost to participate at the edge.

Omdia view

Edge is broadly viewed as an opportunity for CSPs, but if not approached in the right way it could be a distraction for some players and even a threat to others.

Scaling down the collection, storage, and processing of data for deployment at the network edge ought not to present a major technical challenge in and of itself. Indeed, many vendors agree that this is not where edge computing faces its major hurdle.

It is rather the practical, logistical, and commercial obstacles of edge computing that need to be overcome. Most importantly, critical questions such as who will pay for the infrastructure (potentially comprising many thousands of edge nodes in a single network) and who stands to derive the greatest benefit from such an investment are yet to be answered.

CSPs believe they see a clear opportunity in the emerging edge computing market. But other potential players, including public cloud providers, over-the-top (OTT) content players, and software-as-a-service (SaaS) vendors, are equally attracted by the prospect of delivering solutions at the edge.

Key messages

- Edge computing represents a far-reaching change to existing network architectures and is attracting interest from a broad ecosystem of potential participants. CSPs are well placed to become central players in the edge computing market, but their role is not clearly defined.

- CSP networks are already evolving towards a virtualized, cloud-based model that can be the basis of future edge computing deployments. The European

Telecommunications Standards Institute (ETSI) multi-access edge computing (MEC) architecture is a blueprint for service provider network evolution that can deliver on the full promise of edge computing in the context of 5G.

- The edge computing and edge services plans of several CSPs including AT&T, BT, CenturyLink, Telefonica, and Verizon are already well advanced.
- However, the edge is complex and as yet not fully defined. Most likely it will be multilayered and with a mix of hardware and software, varying edge device and equipment types, and diverse edge-installation form factors and enclosures. What form these take will be largely use-case driven.
- The scale and speed of growth at the edge will demand efficient and agile approaches to retrofitting existing edge-of-network sites for new edge applications as well as deploying new edge infrastructure.
- As such, the market for edge node equipment such as micro and prefabricated data centers and enclosures is already outgrowing other sectors of the cloud data center market, reflecting growing enthusiasm for the edge. However, given the size and repeatability of the market, some degree of standardization will be necessary.

CSPs should be targeting edge services

Why edge computing?

Edge computing can be applied wherever bandwidth, low latency, or the localized management of large volumes of data are critical to the delivery of high-quality services. It works by moving the storage, handling, and processing of data out to edge nodes, close to where the data is generated by the application, device, or end user and away from the centralized cloud or core network. These edge nodes can take the form of micro data centers, network nodes, devices, or sensors.

Thus, in an application such as a gaming service, users will no longer rely on connecting directly to the core data center processing the game, which brings higher latency the further away from the data center they are. Instead, by connecting locally to the edge network, all users will experience the same, ultra-low level of latency.

Risks and rewards

Many CSPs are already transforming their networks by moving towards a software-based, cloud architecture. The combination of network function virtualization (NFV) and telco cloud is providing a platform for edge computing, given additional support by the move from today's virtual machine (VM) model to a cloud-native, container-based reference architecture.

Telcos are also pushing their new network infrastructure and services into regional data centers and local metro networks located closer to customers. The shift to 5G is driving more of this transformation and lets providers inject new network features in different locations across their footprint.

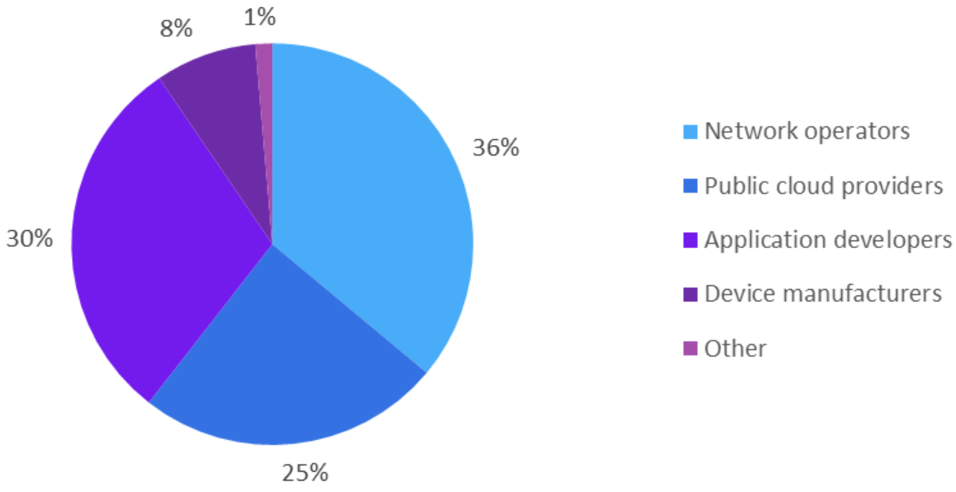
As 5G technology and networks evolve, edge computing can provide a high-performance, on-demand, and cost-effective platform capable of supporting a growing number of use cases. However, when it comes to new edge services solutions, CSPs will need to establish their role in the value chain by determining the right business model and partner(s) and also where new revenues will come from. There will be many choices and opportunities, which CSPs will have to vet and manage. Some questions that have not yet been resolved include the following:

- How will CSPs monetize the better performance, better costs, and lower latency of edge computing, and what is the customer's willingness to pay?
- How can CSPs prioritize and define the best business model and partners?
- Will edge services deliver new net revenues, or are CSPs diluting revenues from their existing services?
- Who will own the primary customer relationship?
- If a CSP opens up services such as co-location to key partners, what happens when other third parties want the same access?
- How do CSPs manage complex service requests from many different providers, covering many locations in their networks?

Fail to find answers to these questions and, even if they embrace the edge opportunity, CSPs run the risk of losing out on new revenue streams and becoming mere connectivity providers. Over commit to their investments in edge computing before the market is ready or demand for services is established, however, and they risk jeopardizing more secure revenue opportunities from existing services and developments such as 5G.

Given the balance of risks and rewards involved, it is unsurprising that when surveyed, many CSPs are cautious about the extent of their role in the creation of new edge services (see Figure 1).

Figure 1: Which ecosystem segment will be most important in the creation of new revenue services from the edge?



Source: Omdia survey of 147 operators

Edge overview and market update

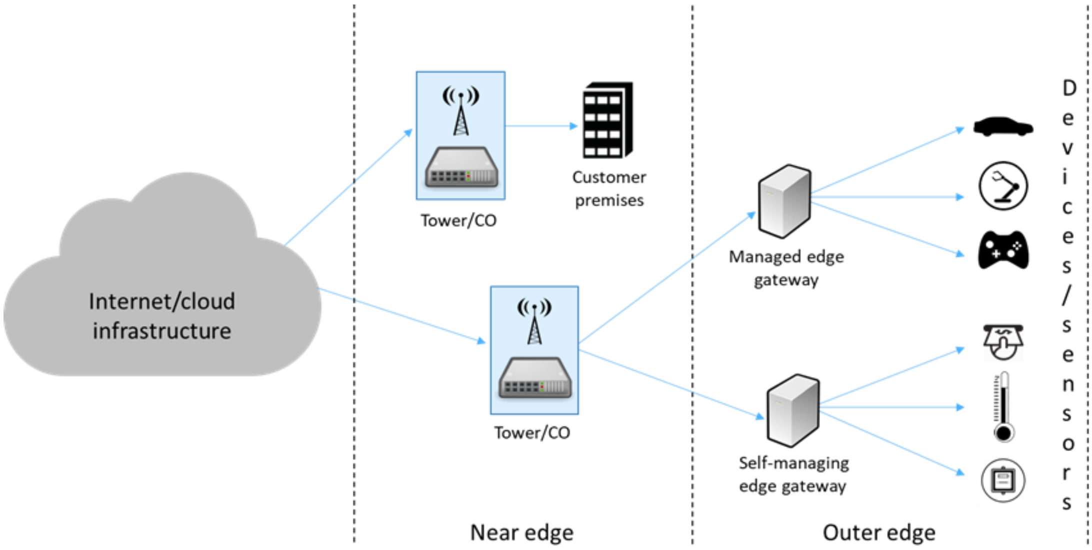
Defining the edge cloud

Physical definitions of the edge are complex and only just beginning to be formulated. In its broadest sense, the edge can be expected to take many forms and to comprise multiple layers. Today, many consider the edge to be a device with compute and storage in a remote location, like a mobile phone mast or retail outlet, but as technologies such as application-specific integrated circuits (ASICs) evolve and put more intelligence into devices, the edge can be expected to encompass a greater diversity of devices and endpoints, and to move out closer to the consumer or sensor.

Devices, infrastructure, and connectivity will represent the base layer of an edge deployment, with orchestration and operational management among the additional layers, and with application development and service creation driving the ecosystem.

O expects that the edge will initially come in two main types: the near edge and the outer edge. The near edge will be based on traditional servers, storage, or hyper-converged infrastructure (HCI) devices and be remotely managed, although hardware incidents will require a physical presence. The outer edge will feature gateway devices, either fully managed or developed as completely immutable and self-managing entities and connected through mobile (4G and 5G) networks (see Figure 2).

Figure 2: Initial edge architecture



Source: Omdia

It is likely that both approaches for the outer edge (fully managed software-based solutions and self-managed hardware-based solutions) will proceed as use cases demand.

As edge networks evolve, it is possible that a "middle" edge might also start to appear for situations where compute, storage, and networking are required in a remote location, but the form factor is much smaller. Sitting between the fully managed near edge and the self-managed swappable outer edge, the middle edge would most probably combine remote management with an immutable operating environment.

The impact of these evolving edge scenarios is that organizations will need to change their management processes to significantly expand their physical device management capabilities.

Edge computing use cases

The network requirements for edge computing in CSP networks are defined by standards body ETSI, as part of the MEC reference architecture. This is designed to allow applications and services that can benefit from proximity to the customer and from receiving local radio-network contextual information to be hosted in a multivendor, mobile-edge computing environment.

Numerous potential use cases are being identified that will benefit from MEC. They include the control and monitoring of industrial machinery and processes (Industry 4.0), connected and autonomous vehicles, augmented and virtual reality (AR/VR), high-quality video and gaming, and a range of services for enterprise and industry verticals.

These use cases for MEC share certain key characteristics. Data center infrastructure provider Vertiv summarizes them in terms of four key archetypes:

- **Data intensive** includes use cases where data volume, cost, or bandwidth issues make it impractical to transfer over the network directly to the cloud or from the cloud to point of use. Examples include smart cities, smart factories, smart homes/buildings, high-definition content distribution, high-performance computing, restricted connectivity, virtual reality, and oil and gas digitization.

- **Human-latency sensitive** use cases include those where anything but the speedy delivery of data would negatively impact a user's technology experience – potentially reducing a retailer's sales and profitability, for example. Use cases include smart retail, augmented reality, website optimization, and natural language processing.
- **Machine-to-machine latency sensitive** use cases also feature speed as a defining characteristic because machines are able to process data much faster than humans. Examples include the arbitrage market, smart grid, smart security, real-time analytics, low-latency content distribution, and defense force simulation.
- **Life-critical** use cases encompass those that directly impact human health and safety and where speed and reliability are vital. They include smart transportation, digital health, connected/autonomous cars, autonomous robots, and drones.

Progress to date

Numerous trials and proofs of concept for MEC are being carried out by CSPs across a range of use case scenarios. Examples include

- the control of automated guided vehicles (AGVs) in factory environments
- MEC-enabled augmented reality experience for visitors to public venues and events
- parking management at major shopping malls
- a 5G-ready digital pedestrian road safety trial
- automated quality inspection and monitoring of manufactured consumer products
- the interpretation and analysis of video data sourced from surveillance cameras.

Who will own the edge, and who will pay?

How will the market evolve?

Because MEC represents the convergence of two worlds – connectivity and computing – at the network edge, no single company or class of player can be expected to deliver the full set of edge resources or capabilities. Edge networks need to handle a combination of IT and networking technologies, including real-time data processing and management, ultra-low latency connectivity, and localized content caching, in a highly cost-effective and efficient manner. Edge-network nodes or micro data centers equipped with their own compute, processing, storage, and management capabilities may well be required to service one or many applications on a localized basis.

CSPs are readying their networks for the edge opportunity

With the commercial roll out of 5G, CSPs are developing the capability to deliver unprecedented network speeds, greater capacity, and massively reduced latency.

Network evolution and the transition to 5G are together providing them with the necessary basis to support the emerging MEC architecture.

As they transition their networks towards greater virtualization and cloud-native network platforms, CSPs are moving more applications into the cloud where they can start to benefit from a more agile service creation environment than was previously possible. Network virtualization lays the foundations for the MEC-based architectures that are the next logical step in the process of building support for new applications and use cases, based on 5G, which require features such as high-speed connectivity and ultra-low latency.

Among those beginning to deploy edge cloud infrastructure in readiness for the introduction of commercial services are operators in Korea, Japan, and China, where preparations are particularly well advanced. Both SKT and KT are deploying regional edge data centers with plans to target both the business-to-business (B2B) and business-to-consumer (B2C) markets. In China, all three leading operators have either begun large-scale, precommercial edge cloud network construction or are engaged in pilot projects using MEC.

The edge computing and edge services plans of several other CSPs are already well advanced:

- AT&T believes that a logical step in its program of network transformation and virtualization will be for edge computing nodes to be located within its own network. The operator has recently announced work on network edge compute (NEC) in cooperation with Microsoft. This will allow customers to use Microsoft Azure cloud services closer to AT&T's network edge locations focused on its 5G network.
- BT is increasing the number of its metro exchanges tenfold in anticipation of providing an expanded edge computing service offering for retail, enterprise, and wholesale customers.
- CenturyLink announced in August 2019 that it had more than 100 locations ready to be utilized for edge compute services and is planning to invest several hundred million dollars in this new program. It sees many use cases that are transport agnostic, not tied to 5G services.
- Telefonica has developed plans for edge services as part of its network transformation and service deployment. Content delivery network (CDN), storage, video, and gaming applications are under review and being tested already. Telefonica sees other opportunities including future Internet of Things (IoT) and 5G applications.
- Verizon announced its Intelligent Edge Network (iEN) some years ago, before its new initiative in edge services. Today it is focused on developing edge services that leverage its 5G network and is working with a range of partners and vendors. In December 2019 Verizon announced a partnership with Amazon Web Services (AWS) to participate in its Wavelength program and place AWS cloud resources closer to its customers.

In reality, many CSPs acknowledge that creating the edge cloud will represent a broad ecosystem and require support from a number of players. Not least, the engagement with application developers will be crucial in driving the development of new applications for the edge and must include features such as service APIs in order to enable the exposure of underlying network information and capabilities.

Other potential players are targeting the edge

The CSPs' claim to be best placed to build and run the edge cloud is not unchallenged, however, and a number of potential stakeholders are already in contention for the space. As well as cooperating with CSPs, providers of IT products and services may want to build their own edge clouds either to sell to CSPs for their networks or to sell directly to end users. Other ecosystem partners such as system integrators or tower companies may also consider that they have the skills or the physical assets required to build the edge cloud (see Figure 3).

Figure 3: The growing edge ecosystem

Source: Omdia

Public cloud providers such as AWS, Google, and Microsoft are also targeting the market for edge by developing scaled-down versions of their hyperscale cloud infrastructures in order to meet the demands of smaller workloads and the more diverse and often physically constrained environments of the edge cloud. These solutions will ensure consistency and continuity in terms of operating environment, APIs, hardware, and functionality, while being located closer to the end user.

AWS Outposts, launched in 2018, is a hybrid, scaled-down version of AWS' hyperscale cloud solution, which supports the cloud provider's environment in third-party data centers and other co-location spaces, providing fully managed, on-premises compute, storage, and processing functionality. Microsoft's Azure Stack is also part of this trend to deploy cloud resources closer to the customer. It provides a portfolio of products that extends Azure cloud services from the data center to edge locations and remote offices. After the cloud provider's edge deals with AT&T and Verizon, more such partnerships can be expected to follow.

SaaS providers have also started to use edge architecture for services, allowing customers to have hybrid arrangements with a mix of cloud-based, regionally based, and on-premises deployments for their customers. Many are promoting the use of hybrid deployments of certain SaaS enterprise resource planning (ERP) applications on premises and in private data centers or for their own cloud-based services.

Large network equipment providers are helping CSPs develop and support edge services. Ericsson has created its Gravity division, which is primarily focused on helping service providers develop edge services to support an ecosystem of partners, letting telcos leverage their assets and capabilities.

Startup firms such as EdgeMicro, EdgeInfra, and Vapor IO are developing metro network infrastructure and equipment to support edge services deployment across US metros.

Competition or cooperation?

No single class of provider appears ideally placed to provide a complete edge solution, but by electing to use their existing networks of cell towers, central offices, and aggregation points as edge data centers, mobile and wireline operators can build a platform for the development of edge services while saving massively on costs. The only proviso is that these assets should provide a suitable environment in terms of location, space, physical environment, power, and scope for future expansion.

Cloud and SaaS providers have a strong commercial imperative to deploy edge computing in order to extend their services out to existing enterprise customers, industry verticals, and even telcos. Although acquiring the real estate required to build a basic edge infrastructure will be costly, and they will need the 5G connectivity that CSPs can provide, their plans are already well advanced.

Other infrastructure players such as tower companies and data center providers targeting the edge can help to build the physical infrastructure and work with multiple edge service providers to facilitate their service offerings, yet they cannot build the edge ecosystem without further collaboration.

Ultimately, success will depend on determining the right business model, understanding where new revenues will come from, and building an edge network according to these criteria. This may well only be achieved through establishing partnerships.

Building the edge cloud

Demand for edge data centers is growing

As the number of MEC deployments increases and CSPs look to improve response times and performance through their networks, so the market for smaller, more compact data centers that can be placed closer to end users is growing. These edge data centers have the capacity to cache bandwidth-intensive content, aggregate data from local connected devices, and enable new latency-sensitive applications.

Because they are deployed across multiple cities, these edge data centers need to be of a design that is repeatable, scalable, and easily serviceable, or even capable of functioning without the need for frequent servicing. There is as yet no industry-wide standard definition for these edge data centers, but many take the form of prefabricated modular data centers (PMDCs) that can be built and deployed to a remote location in as little time as a few months, or even be prefabricated and ready to go, creating a turnaround time of less than a month in some cases (see Table 1).

Table 1: Drivers of prefabricated modular data centers

Speed of deployment
Outsourced data center design
Mobility
Offsite integration and manufacturing
Save space and are suitable for harsh environments
Repeatable and scalable
Single point of control for all systems
Mobility
Prefabricated modular data centers are depreciable assets

Source: Omdia

All-in-one modules that have the IT, power, and cooling infrastructure within one module, are the preferred solution among telcos, content providers, and media for remote locations including mobile applications and "edge of network" applications that need a small data center presence close to end users. These all-in-one modules are relatively evenly split between 1–5 racks and 6 –10 racks of IT load, but as edge-of-network data center demand grows, shipments of modules with 1 –5 racks are forecast to grow the fastest.

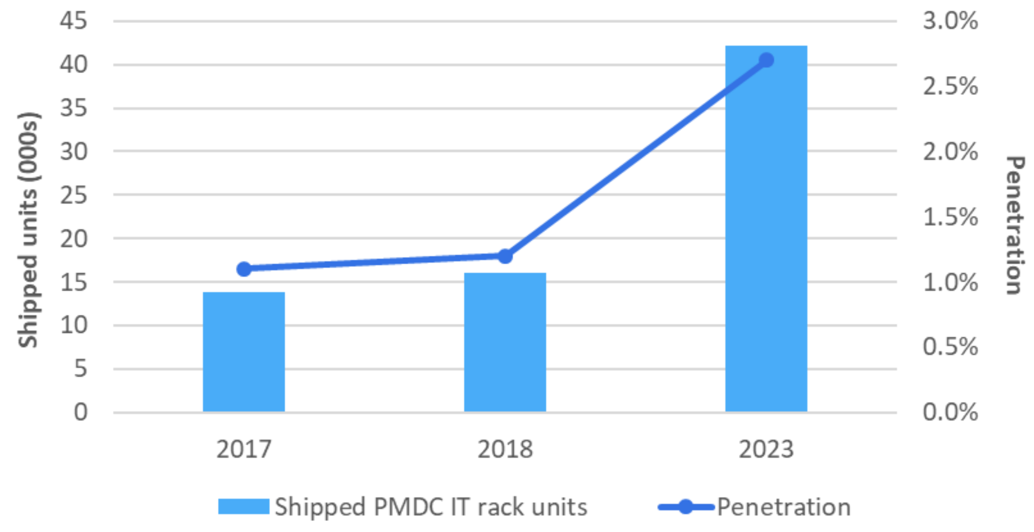
PMDCs can address both the need for rapid build out of new data center capacity and the harsh environmental conditions and remote regions with limited infrastructure where data center design expertise may be in short supply. Edge data centers do not need to be innovative or cutting edge in design; a basic unit with moderate storage and IT load capacity will be sufficient for most use cases. However, innovative approaches are emerging, such as the use of direct liquid cooling and of on-site generation from fuel cells, solar, and gas turbines.

As the network edge develops and companies deploy edge data centers in increasing numbers, standardization will play a critical role. Edge data centers and PMDCs will need to be uniformly managed and interoperable. Many companies are currently in the proof-of-concept phase when it comes to PMDC edge deployments, and as these companies' business models evolve and they figure out which concepts work, standardization of edge data center deployments will become a significant driver for the PMDC market.

The edge data center market is growing faster than previously forecast

The PMDC market is forecast to grow from $1.2bn in 2018 to $4.3bn in 2023, with shipments already growing faster than had been forecast just one year ago (see Figure 4). This growth is driven largely by cloud service providers continuing to build out their data center footprints as well as by telcos and the growth of edge computing and by government, industrial, healthcare, education, and retail end users.

Figure 4: Shipped PMDC IT rack units



Source: Omdia

In Asia, China is accelerating to triple-digit growth rates, and the more developed countries in the region are also expected to show rapid growth. North America leads in the adoption of PMDCs, where a key trend with a longer-term impact on the market is the exploration of proof-of-concept and adoption of PMDCs associated with edge-located data centers.

All-in-one module growth in EMEA is forecast to accelerate from 2021 because of edge computing, although this edge-related growth is forecast to be less than in North America due to the currently more distributed data center landscape in Europe.

Middle East and Africa will see high demand for PMDCs throughout the forecast period, driven by the burgeoning demand for data center capacity from co-location, government, and CSPs, as well as the large international cloud providers such as AWS and Microsoft.

Conclusion

What should operators' strategy be for the edge?

- Operators' existing assets, such as central offices, aggregation points, or base station sites, can potentially work as a starting point for edge computing deployments. However, they should bear in mind a number of factors when developing a strategy for the edge network, as computing and storage capacity is likely to be more constrained the closer to the network edge resources are placed.

- Consideration must be given to how these sites will be used in the longer term as 5G is rolled out. Network hardware, whether for enterprise or network applications, still needs to be capable of scaling in terms of processing, memory, and storage, and of withstanding varying environmental conditions. Virtualized core network components and associated services may also need to be hosted later, on the same platforms.

- Applications that require the caching or processing of large amounts of data at the edge may be better sited at a central office or at metro aggregation points, rather than at the base station or on customer premises. As edge networks extend outwards to encompass more devices and endpoints, traditional compute and storage infrastructures will give way to a greater diversity of devices, and endpoints will likely be subject to greater autonomy, or even be self-managing entities.

- If they can successfully manage these practical challenges, identify the right business models, and be prepared to work with strategic partners, CSPs can emerge as owners of the value chain for edge services and the primary drivers of the development of the edge network. In this way, edge computing will be an opportunity rather than a threat or distraction.

Appendix

Author

Julian Bright, Senior Analyst, Network Infrastructure and Software
Julian.bright@omdia.com

Get in touch

www.omdia.com
askananalyst@omdia.com

Omdia consulting

Omdia is a market-leading data, research, and consulting business focused on helping digital service providers, technology companies, and enterprise decision-makers thrive in the connected digital economy. Through our global base of analysts, we offer expert analysis and strategic insight across the IT, telecoms, and media industries.

We create business advantage for our customers by providing actionable insight to support business planning, product development, and go-to-market initiatives.

Our unique combination of authoritative data, market analysis, and vertical industry expertise is designed to empower decision-making, helping our clients profit from new technologies and capitalize on evolving business models.

Omdia is part of Informa Tech, a B2B information services business serving the technology, media, and telecoms sector. The Informa group is listed on the London Stock Exchange.

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help your company identify future trends and opportunities.

Copyright notice and disclaimer