Edge compute

Coming to a place near you

Sponsored by

AECC
AUTOMOTIVE EDGE COMPUTING CONSORTIUM

GSMA Intelligence

The GSMA is a global organisation unifying the mobile ecosystem to discover, develop and deliver innovation foundational to positive business environments and societal change. Our vision is to unlock the full power of connectivity so that people, industry, and society thrive. Representing mobile operators and organisations across the mobile ecosystem and adjacent industries, the GSMA delivers for its members across three broad pillars: Connectivity for Good, Industry Services and Solutions, and Outreach. This activity includes advancing policy, tackling today's biggest societal challenges, underpinning the technology and interoperability that make mobile work, and providing the world's largest platform to convene the mobile ecosystem at the MWC and M360 series of events.

We invite you to find out more at gsma.com

Follow the GSMA: @GSMA

GSMA Intelligence is the definitive source of global mobile operator data, analysis and forecasts, and publisher of authoritative industry reports and research. Our data covers every operator group, network and MVNO in every country worldwide — from Afghanistan to Zimbabwe. It is the most accurate and complete set of industry metrics available, comprising tens of millions of individual data points, updated daily.

GSMA Intelligence is relied on by leading operators, vendors, regulators, financial institutions, and third-party industry players, to support strategic decision-making and long-term investment planning. The data is used as an industry reference point and is frequently cited by the media and by the industry itself.

Our team of analysts and experts produce regular thought-leading research reports across a range of industry topics.

gsmaintelligence.com

info@gsmaintelligence.com

This report was authored by GSMA Intelligence with support from AECC.

The Automotive Edge Computing Consortium (AECC) is an association of cross-industry, global leaders working to explore the rapidly evolving and important data and communications needs involved in instrumenting billions of vehicles worldwide. The AECC's goal is to find more efficient ways to support the high-volume data and intelligent services needed for distributed computing and network architecture and infrastructure. The AECC's members are key players in the automotive, high-speed mobile network, edge computing, wireless technology, distributed computing, and artificial intelligence markets. For more information about the AECC and its membership benefits, please visit aecc.org.

Published February 2024

Authors
Silvia Presello, Lead Analyst
Tim Hatt, Head of Research and Consulting
Shiv Prashant Putcha, Director, Consulting

Contents

Executive summary

Surveying the edge as it comes into its own

Interest in edge computing gathered momentum in 2023, in line with the broader enterprise digitisation movement taking place in multiple industries. Several factors are behind the momentum, including growing IoT deployments, compute capacity and processing upgrades, 5G network expansion, and the exponential growth of data traffic. Improving economic conditions (even though GDP growth remains sluggish) and falling inflation have also helped business confidence and, in turn, infrastructure investment rise to a level not seen since before the pandemic. In short, edge is coming into its own after several false dawns.

GSMA Intelligence conducted a survey in June and July 2023 to evaluate investment sentiment, challenges, costs and opportunities for edge computing. The survey involved 400 executives from companies across the value chain, including mobile operators, network equipment vendors, cloud service providers, IoT service providers/systems integrators and car manufacturers. The survey was complemented with analysis of proprietary data and company interviews.

New industries enter the frame

The automotive industry has historically been a major driver for edge investments, and there remains positive sentiment for edge in connected vehicle service areas involving security and data storage, for example. Playing to this demand, connected vehicle services are expected to unlock substantial revenue opportunities for car manufacturers. Revenue gains realised from deploying edge computing in connected vehicles will rise considerably over the next three years, with 15% of car manufacturers expecting to generate a 16–20% uplift on 2022 revenues.

However, other industries are now entering the frame. Consumer electronics is viewed as having the highest sales potential (32% of survey respondents rate it as the No.1 sector), followed by media & TV (17%) and retail (10%). The sentiment speaks to use cases in these industries that require compute power closer to the end user, as well as new (or changed) monetisation models anchored in data analytics. Examples include VR broadcasting of sports matches, on-site training of oil & gas engineers, or car manufacturers monetising data assets (such as traffic information and driving patterns) with insurance companies and fleet managers.

The diversification into new industries reflects a recognition of how localised computing infrastructure can enable new product areas. However, it also reflects the latent but fast-moving experimentation with artificial intelligence (AI) large language models (LLMs) embedded within computing that are bringing functionality to software not previously possible. This is a major new point of competitive differentiation in edge computing, as with cloud and 5G more broadly. It is likely to be the focus of both hard capital investment and research and development (R&D) in 2024/2025.

Creative investment models

On the investment front, telecoms operators continue to be seen as the main buyers of edge infrastructure, with around 50% of surveyed companies rating operators the top group. IoT providers and systems integrators are next (43%), followed by equipment vendors (31%). This reflects the RAN and localised datacentre outlays inherent with edge. Indeed, around two thirds of operators surveyed expect themselves to carry the main financial responsibility.

However, the structure of edge investments is changing to incorporate cost and revenue sharing, which means that the lead contractor may increasingly be an equipment vendor (e.g. Nokia or Ericsson), supported by an operator. This co-opetition reflects the nature of edge deployments. It could also help alleviate financial pressures for operators, who are funding 5G and fibre while in a low-growth environment.

IT integration remains a barrier

The integration of edge computing into existing IT enterprise networks is a key challenge. Enterprises should view edge computing as a component of their broader enterprise digital strategy, which includes other technologies such as AI/ML, cloud computing, private networks and slicing. Fully completing the transformation may take three to five years, but suppliers can mitigate the integration barrier by breaking things into stages.

1

The evolution of edge

1.1 A call for uniformity

Edge computing refers to the movement of computing resources, such as those for storage and networking, from a centralised public/private cloud to the edge of the network. This means bringing resources closer to users and devices, where data is generated. Edge computing is not a substitute for cloud computing but a supplement and expansion; it is particularly suitable for real-time analysis of data and intelligent processing.

Despite rapid development, edge computing is still in its infancy. The edge computing community has yet to agree on standardised definitions, network architecture and protocols. A consensus on edge standards among the community is crucial to enable edge to seamlessly work with other technologies. There are various architectures at the edge, with the boundaries between them blurry:

- **Fog computing** – A term created by Cisco in 2014, this refers to the decentralisation of computing infrastructure by extending the cloud through strategically placing nodes between the cloud and edge devices. It helps filter important information from the large amount of data collected from devices.
- **Multi-access edge computing (MEC)** – A concept defined by the European Telecommunications Standards Institute (ETSI), MEC offers application developers and content providers cloud computing capabilities and an IT service environment at the edge of the mobile network. MEC aims to unite the telco and cloud worlds, providing IT and cloud computing capabilities within the RAN near mobile subscribers. The environment provides ultra-low latency and high bandwidth, as well as real-time access to radio network information that can be leveraged by applications. Operators can open the radio network edge to third-party partners.
- **Cloudlets** – These are mobility-enhanced, small-scale data centres placed near edge devices so they can offload processes. A cloudlet is an architectural element that extends cloud computing infrastructure. It represents the middle of a three-tier hierarchy comprising mobile device, cloudlet and cloud. Cloudlets are designed to improve resource-intensive and interactive mobile apps in particular, through the extra availability of low-latency computing resources.
- **Micro data centres** – These are small-scale data centres that are an extension of hyperscale cloud data centres. Micro data centres include all the compute, storage, networking, power and cooling required for a given workload.

For this analysis, the above emerging technologies are collectively referred to as ‘edge computing’, with a few exceptions where necessary.

1.2 Why edge matters

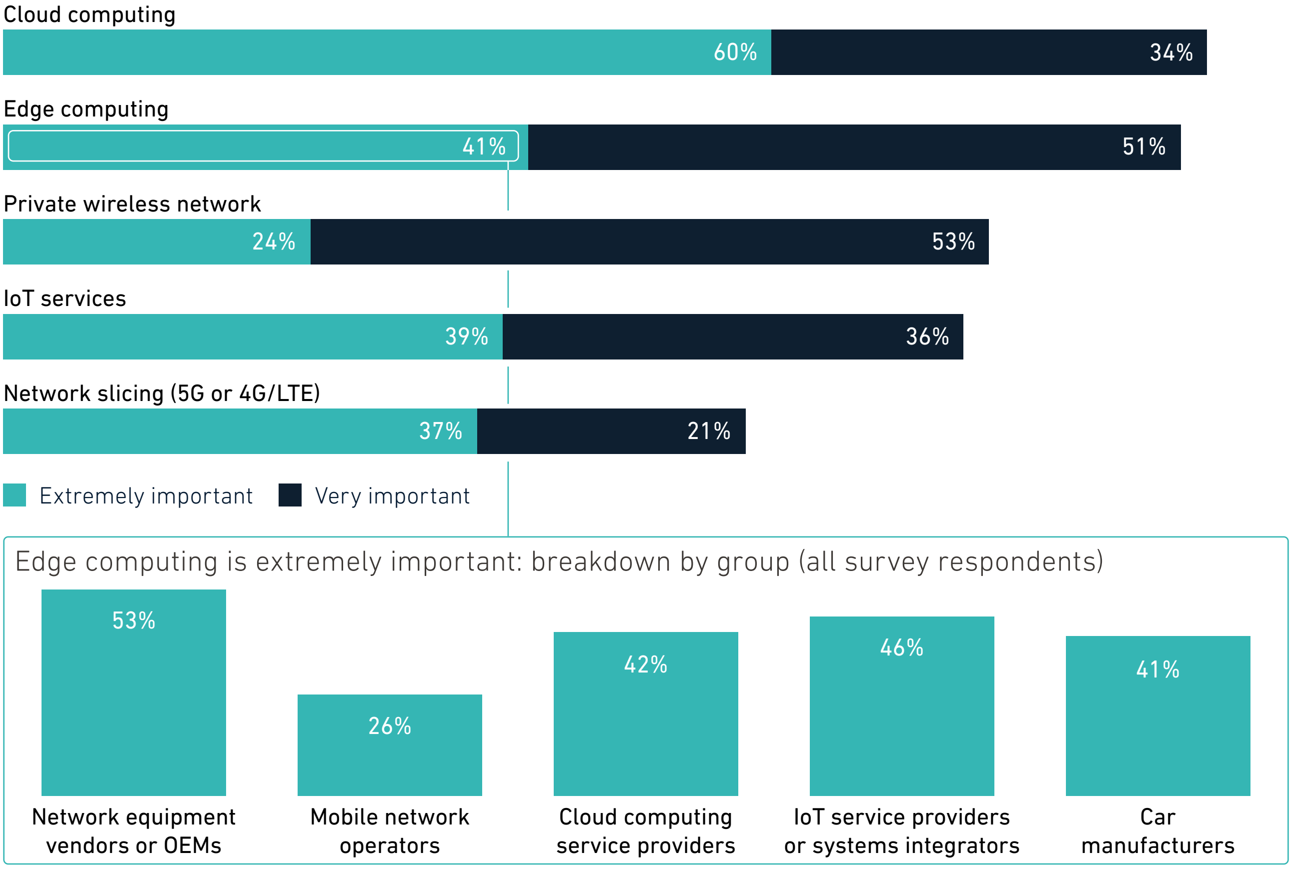
The edge computing concept has been around for a while and has evolved within the 3GPP and IEE standards. With the latest advances in wireless technologies, edge is now in the spotlight. Edge computing enables artificial intelligence (AI) large language models (LLMs) embedded within computing to bring functionality to software not previously possible. It provides IoT with improved speeds, reduced latency and real-time response, as data can be processed locally without being sent to a centralised public or private cloud.

For all the groups surveyed by GSMA Intelligence, edge computing is becoming increasingly important to business-to-business (B2B) sales success. As edge use cases mature, edge players are eyeing the opportunity to capture more value in the B2B segment.

Figure 1

Edge computing is seen as extremely/very important by 92% of survey respondents

How important are each of the following product areas for your B2B sales success? (All survey respondents)



Source: GSMA Intelligence

According to the GSMA Intelligence survey, 64% of all respondents have deployed edge in one country, while a further 28% have done so across multiple countries. These deployments will vary in size. Some will have standard edge servers on-premise; some will link with other technologies. For example, integrating private wireless networks with edge computing enables businesses to support applications requiring low latency and high bandwidth (e.g. autonomous guided vehicles in a car-manufacturing factory). Integrating network slicing with edge computing enables support for edge computing use cases with different latency requirements. This can be achieved through a cloud-native microservices architecture that leverages software-defined networking (SDN).

5G is key to edge

While various technologies can meet the need for IoT connectivity with high bandwidth and low latency (e.g. in robotics for manufacturing and V2X applications), edge computing can process and analyse the data, and provide a fast, data-driven response. As such, the technologies can be used in combination for optimal output.

With 5G standalone (5G SA), a portion of the network can be allocated according to the specific needs of particular edge use cases:

- enhanced mobile broadband (eMBB) is dedicated to use cases requiring high throughput and low latency
- enhanced machine-type communication (eMTC) is dedicated to use cases optimised for small data transmission, extreme power-saving mechanisms, stationary devices, low throughput and high geographic density
- ultra-reliable, low-latency communications (URLLC) enables mission-critical applications for use cases that require a highly reliable control plane, high-performance user plane, very low latency, high mobility and low throughput. URLLC is particularly suitable for enabling V2X applications.

The 3GPP recently added time-sensitive networking (TSN) to the list, introducing the concept of 5G RedCap (reduced capability) or NR-Light. 5G RedCap addresses use cases that cannot be addressed by eMBB, URLLC or eMTC. It enables user devices to benefit from the scale of 5G deployments but leveraging fewer 5G capabilities, for an optimum balance of features versus cost and power consumption. It is particularly useful for wireless industrial sensors, video surveillance and smart wearable technology. For instance, until now, devices such as smartwatches have connected to the internet via a user's smartphone. However, new products will use RedCap to link directly to the cloud.

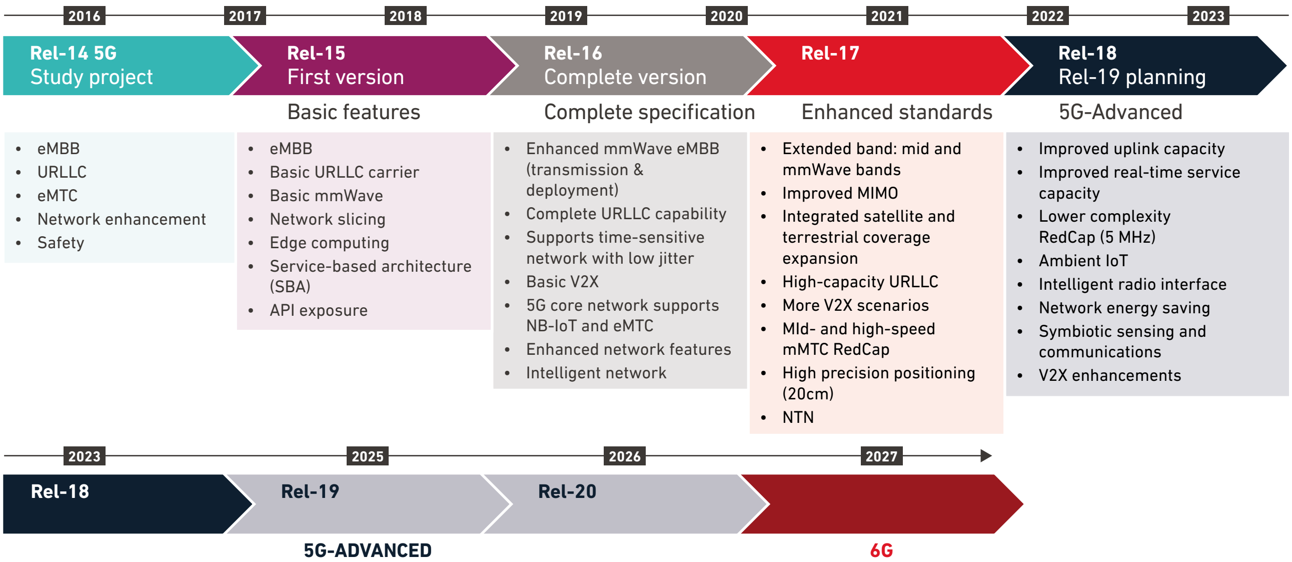
Edge is key to 5G

The transition from 5G to 5G-Advanced networks will require the integration of several key technologies; edge computing will be a crucial part of this mix. It will enable AI/ML capabilities to be pushed from the central public/private cloud to the edge of networks. This will be achieved through edge network architectures using edge nodes and gateways, which will provide a continuum of cloud computing through to a point close to the end-user. APIs will also be a crucial component of this transition, serving as building blocks for digital services that leverage the network. They can enable the creation of new revenue streams and monetisation.

Furthermore, edge computing in conjunction with advancements in satellite communications will play a key role in enabling a fully converged 5G-Advanced architecture that integrates ground, sea, air and space networks.

Figure 2

3GPP standards and 5G evolution



Source: GSMA Intelligence

1.3 Edge drivers

Tech advancements and enterprises are spurring demand for edge computing. Enterprises are undergoing digital transformation, driven by the need for operational efficiency, cost cutting and productivity gains. The growing volume of data produced by connected devices has led to a surge in demand for solutions that enable reduced latency, data processing and analytics, and real-time response or decision-making. Enterprises are looking for solutions that unlock the potential offered by data generated from connected devices, and edge computing taps into this.

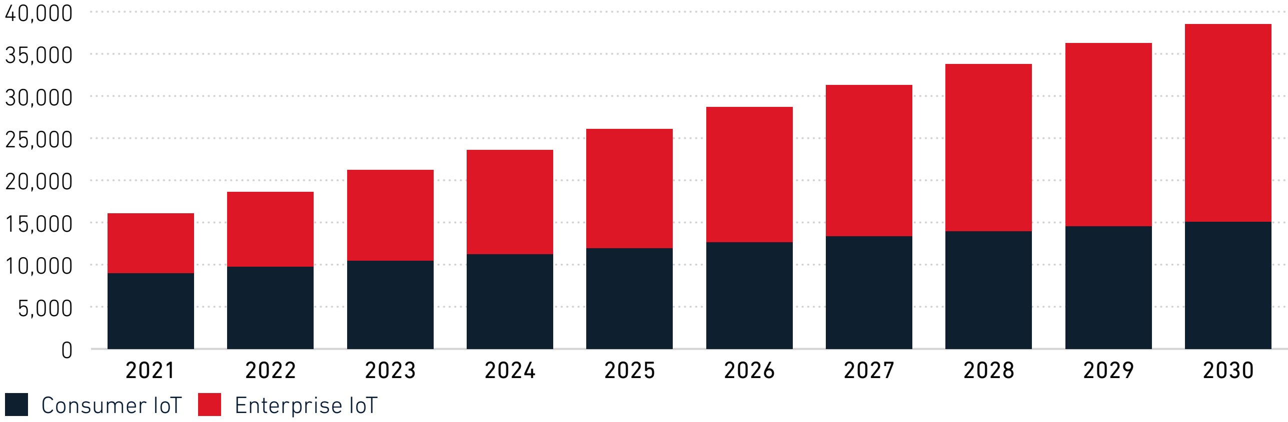
By 2030, GSMA Intelligence forecasts there will be more than 38 billion IoT connections, with over half of them enterprise connections. Around 15% will be licensed cellular IoT. Growth in licensed IoT connectivity technologies will enable edge solutions, particularly MEC – the form of edge computing that employs cellular networks as the primary means of connectivity.

Edge computing enables new capabilities for end-user devices and IoT, across different industries including automotive, transport, retail and manufacturing. The next step for IoT is to incorporate 5G and edge computing into deployments. Together, they can provide the high speed, low latency and massive capacity needed for real-time applications such as autonomous driving.

Figure 3

Enterprises will drive growth in IoT connections

Global IoT consumer and enterprise connections (million)

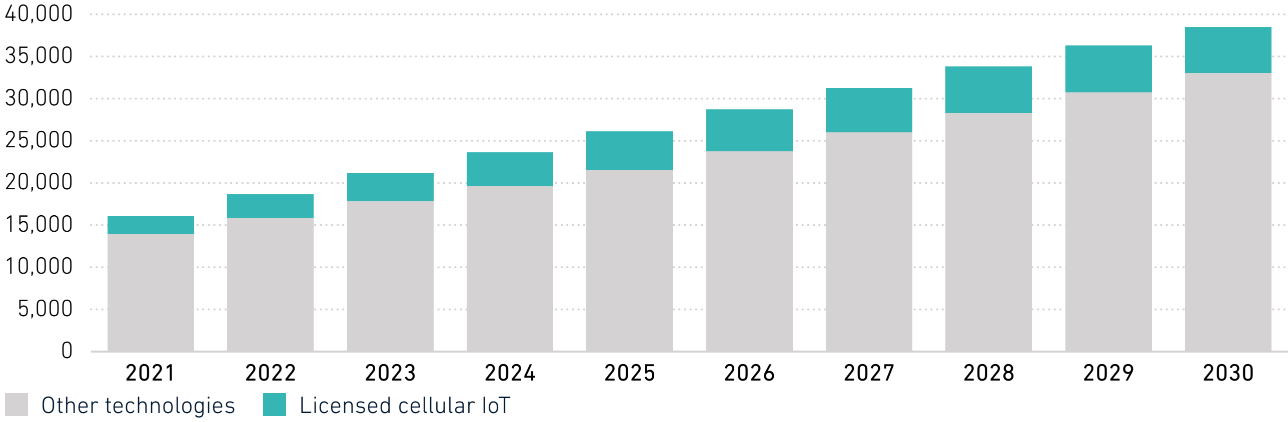


Source: GSMA Intelligence

Figure 4

Growth in licensed cellular IoT connections will enable MEC use cases

Global IoT connections (million)



Source: GSMA Intelligence

1.4 Automotive and beyond

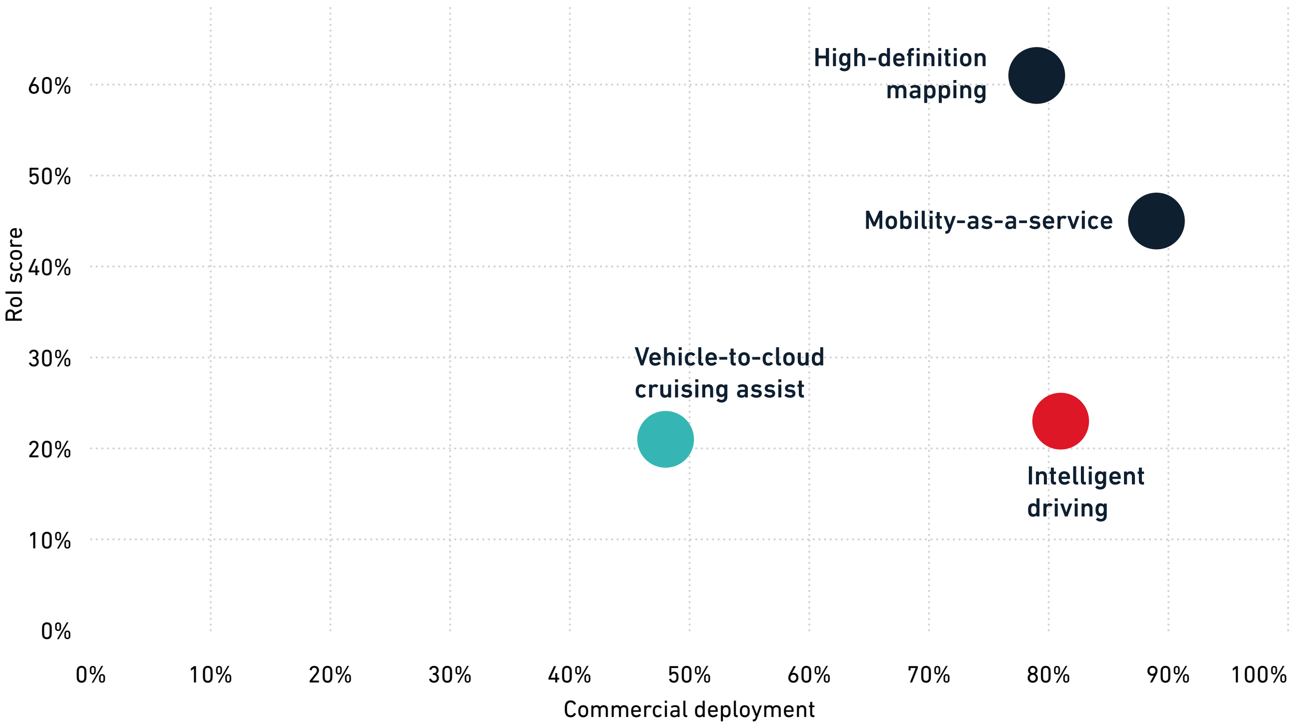
The automotive industry has been leading the way in edge computing applications. With continued advances in IT infrastructure, transportation is evolving into a highly data-driven system that generates real-time data in massive quantities. Using 5G in combination with non-terrestrial networks (NTNs), AI and edge computing nodes on roads and vehicles, cellular vehicle-to-everything (C-V2X) technology is expected to see significant growth. Developed within the 3GPP, C-V2X enables vehicle-to-vehicle and vehicle-to-network communication.

Edge computing (MEC) is key to C-V2X applications as it enables deployment closer to vehicles and pedestrians, reducing latency. Car manufacturers are exploring use of edge computing as it can lead to opportunities such as enabling real-time HD mapping with location-specific data; intelligent driving; advanced driver assistance system (ADAS); and the development of user-based applications (e.g. mobility-as-a-service). According to the car manufacturers surveyed, mobility-as-a-service and HD mapping have the highest potential for return on investment.

Figure 5

Growing demand for high-definition mapping and mobility-as-a-service boost return-on-investment potential

Thinking about connected vehicle services, which of the following services has the highest RoI potential? (Car manufacturer respondents)



Demand for HD mapping and mobility-as-a-service is boosting RoI potential for connected vehicle services among car manufacturers, with services widely deployed.

Vehicle to cloud has not yet progressed into the same RoI quadrant but is slowing advancing. Some 44% of respondents are at the trial stage, while 8% have plans to deploy it but are not yet at trials.

Safety concerns, regulatory frameworks, different national regulations, and liability in the advent of accidents are some of the challenges associated with autonomous vehicles, potentially affecting returns.

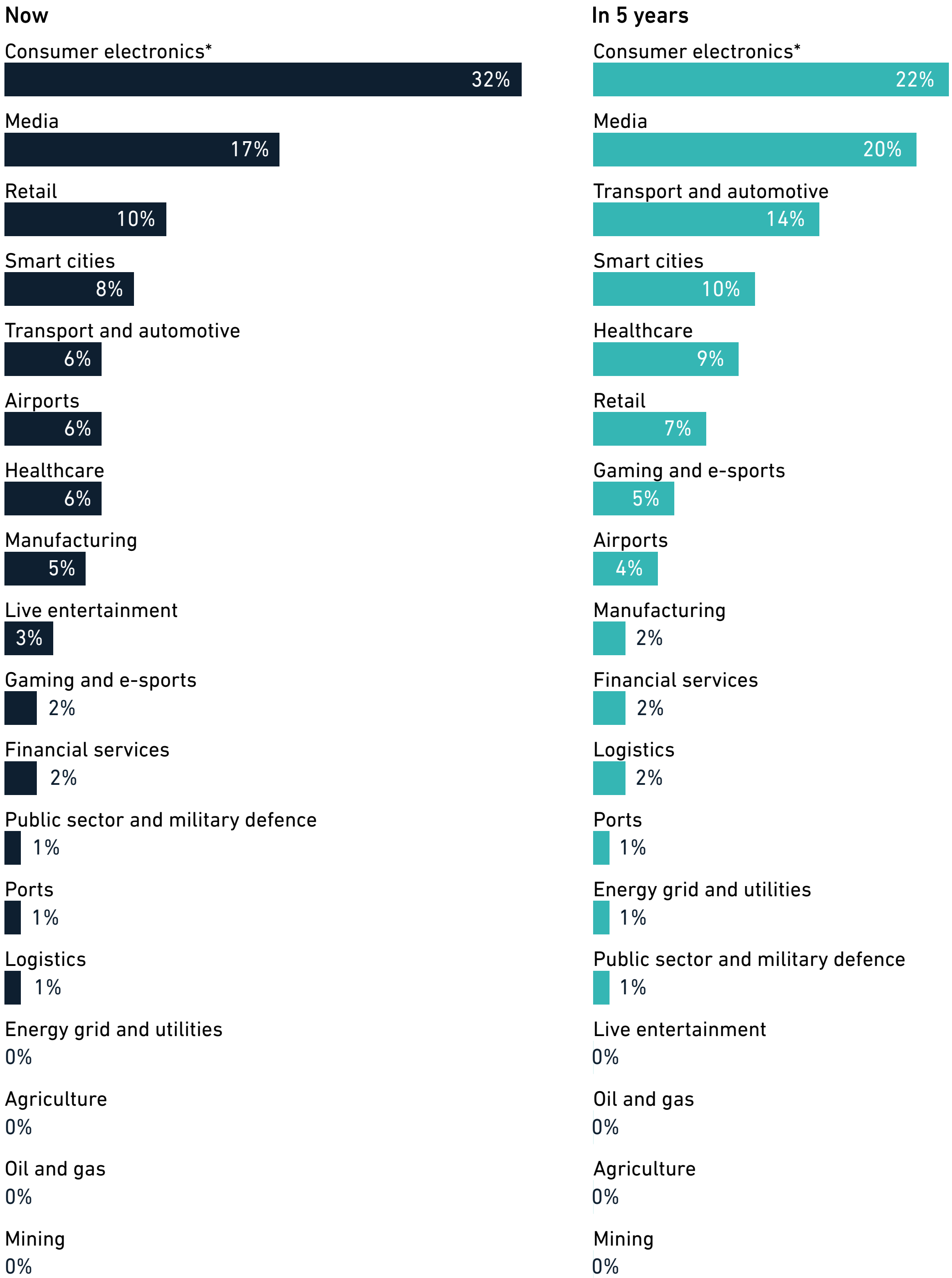
Note: commercial deployment refers to deployment in one or more countries. RoI score calculated as (Ranked 1st*1)+(Ranked 2nd*0.5)
Source: GSMA Intelligence

As the use cases for edge computing continue to expand, players in the edge computing industry are increasingly seeing an opportunity to capture more value in the B2B segment. There is growing interest in edge computing across a range of verticals, particularly consumer electronics, media, retail and smart cities.

Figure 6

Top industry verticals for edge

What do you see as the top industry verticals that will drive the highest demand for edge computing? Percentage of respondents rating sector as No.1. (All survey respondents)



*includes products such as smartphones, laptops, tablets, PCs, XR/VR devices, wearables and drones
Source: GSMA Intelligence

Consumer electronics

The adoption of edge computing in consumer electronics is gaining momentum due to the growing popularity of the metaverse, remote collaboration and operation, e-sports and gaming. The technology underpinning certain consumer devices enables compute-intensive workloads to be offloaded from the end device to an edge node, which processes the data and transmits the results back to the device. This approach is particularly beneficial for devices with applications that require high processing power, such as AR/VR.

Edge computing technology is proving to be a significant enabler of advanced consumer electronics devices, including smartphones, smart home solutions, laptops, drones and XR/VR devices. These are capable of delivering immersive experiences that were previously unattainable due to limitations in processing power and energy consumption.

EXAMPLE USE CASE

Education

Educational institutions are seeking new, more flexible and engaging ways to enable learning. They are looking for a cost-effective means of enabling students to use VR applications to participate in lessons remotely and engage with 360-degree videos and 3D models.

Using low-cost cardboard headsets, 5G connectivity and edge computing, Telefónica, the IE University and Nokia have enabled teachers and architecture students to interact in a virtual space. Employing edge computing to handle the computation associated with rendering virtual spaces reduces hardware requirements (and costs) for the end user. The combination of 5G, edge computing and VR enriches interaction between teachers and students and allows for further development of educational methodologies.

Media

Digital media content has traditionally been delivered over the internet from cloud-based servers. These are typically located at central data centres across the globe, operated by hyperscalers such as AWS, Microsoft and Google, or other internet hosting companies. Although content delivery networks (CDNs) have been used to distribute static media content closer to end users, they have not always been suitable for interactive experiences with complex user interactions requiring server processing. In such cases, reducing end-to-end latency becomes crucial. Serving high-bandwidth interactive content to users located far away can lead to network congestion, a poor user experience and inefficient data transit. Edge computing offers a solution to this challenge by offloading certain workloads to an edge node closer to the user.

EXAMPLE USE CASE

TV production

The production of TV programmes requires filming at indoor and outdoor locations, making it expensive and time-consuming to install wired networks to transfer video images and audio to the editing and production teams. To help solve this problem, a private 5G network connecting cameras and microphones to a portable base station can be implemented. With edge computing, the system transmission delay in the 5G private network can be halved.

Smart cities

While technology innovation continues to drive smart cities forward, sustainability's rise up the agenda has seen a shift from smart cities to smart and sustainable cities. This involves using ICT and other emerging technologies such as AI and edge computing to improve quality of life, efficiency of urban services and competitiveness, but also to ensure smart cities meet the environmental and social needs of current and future generations.

The technological capabilities required to deliver smart city strategies are diverse and multidisciplinary. Some areas (such as IoT, connectivity, cybersecurity and data management) have become essential. Others (such as AI, digital twins and edge computing) offer opportunities for more advanced smart city applications. As the quantity of data grows in the decision-making process for smart city applications, edge computing can contribute to the analysis of the data at the gateway or even device level, closer to where the data is generated.

EXAMPLE USE CASE

Smart cities

The SmartCityPHL project in Philadelphia (US) is a collaborative initiative by the city, Comcast and US Ignite. It involves the collection of real-time data related to air quality, weather, transport and more. As part of the programme, SmartBlockPHL is a pilot project involving the deployment of 14 smart streetlights in Philadelphia's Midtown Village. The data collected and processed with the support of edge computing technology can help the city respond more effectively to emergencies and better understand the on-the-ground conditions of the neighbourhood in terms of environmental quality, safety and traffic/people flow.

Transport and automotive

Cloud, AI, 5G and edge computing are revolutionising the transport and automotive sectors. The primary processing of connected car workloads can occur on-board the vehicle (user edge), at the service provider edge (i.e telco base station) or in the centralised cloud.

EXAMPLE USE CASE

Automotive sector

Advances in mobile connectivity and edge computing are enabling new use cases for the automotive sector, including those highlighted in Table 1.

Table 1

Edge use cases in the automotive industry

Multifactor authentication for keyless entry	Machine learning for intuitive infotainment systems	Autonomous driving and smart infrastructure	Protection of sensitive data	Road safety (early detection)
Frictionless car entry can be enabled based on multiple security factors. This can be achieved using a camera for facial recognition, an infrared camera for spoofing detection, and a Bluetooth sensor to detect the proximity of the driver's mobile phone. This can involve significant data transfer and processing; edge computing has made this possible.	The infotainment system is a crucial user interface. To determine which functions and applications people are actually using and where the user interface needs to be improved, machine learning can be an essential tool, deriving insights from an abundance of data available. Edge computing facilitates the deployment of machine learning models (developed in the cloud) to the device effortlessly.	To enable autonomous driving, the control unit in the vehicle needs to have sufficient computing power to execute driving algorithms in real-time. Smart infrastructure such as 5G base stations, traffic lights and intermediate smart routers located at the 'fog edge' can help increase efficiency and flow at road junctions. If an edge node is installed at the junction, with vehicles connecting when approaching it, the edge node can orchestrate all nearby vehicles' trajectories instead of separately computing each vehicle's path. This significantly reduces waiting times and increases the overall efficiency of the junction.	Many connected devices have built-in sensors that generate a large amount of data; connected vehicles are one example. While the car can process most of the data, certain applications, (such as alerts) require data to be sent to the cloud. Edge computing reduces the amount of data that is sent out, reducing data transmission costs and limiting the amount of sensitive data leaving the vehicle.	V2X early detection helps when a vehicle in front brakes abruptly. The driver of the vehicle behind may respond too late, causing them to brake hard to prevent an accident. This can create a chain reaction, impacting vehicles further behind the incident. With V2X, the vehicle's advanced driver-assistance system (ADAS) detects the event earlier, and can moderately decelerate, increasing the gap between vehicles and avoiding hard braking.

Source: GSMA Intelligence

2

Competitive landscape

2.1 The edge computing landscape is highly fragmented

The ecosystem for edge computing is diverse, with multiple stakeholders jostling for position. It is a fragmented landscape with network equipment vendors, operators, service providers, hyperscalers and others with product and service offerings. No single participant in the edge computing value chain can boast a true end-to-end offering or service, with many controlling or contributing to different parts of the technology stack.

Major participants in the edge computing ecosystem include the following:

- **Semiconductor vendors** – The likes of Qualcomm, Sony Semiconductors and Xilinx provide silicon for end points and edge compute servers.
- **Connectivity infrastructure providers** – This group includes fixed and mobile operators, as well as service providers including neutral hosts, IoT connectivity providers, tower companies and related digital infrastructure players.

- **Cloud providers** – Cloud service providers, or hyperscalers, build and operate distributed cloud infrastructure to run enterprise workloads. They also have established developer ecosystems and marketplaces. Hyperscalers are at the forefront of edge solutions for enterprises currently, as many businesses look to extend their current cloud-based offerings to the edge and maintain the same interface and experience.
- **Edge platform and app providers** – The edge of the network typically overlaps with equipment and assets that fall into the category of operational technology (OT), as opposed to IT. Specialised vendors have developed platforms for the management and operation of these assets, beyond connectivity.
- **Software vendors** – A number of companies create specialised software and applications for edge computing infrastructure.
- **Hardware vendors** – These create and deploy a range of devices and end points for enterprises. In certain industry verticals, these can be highly specialised.
- **Systems integrators** – These offer integration and other professional services to enterprise end users. Some are cross-industry, while some work in specific industry verticals.

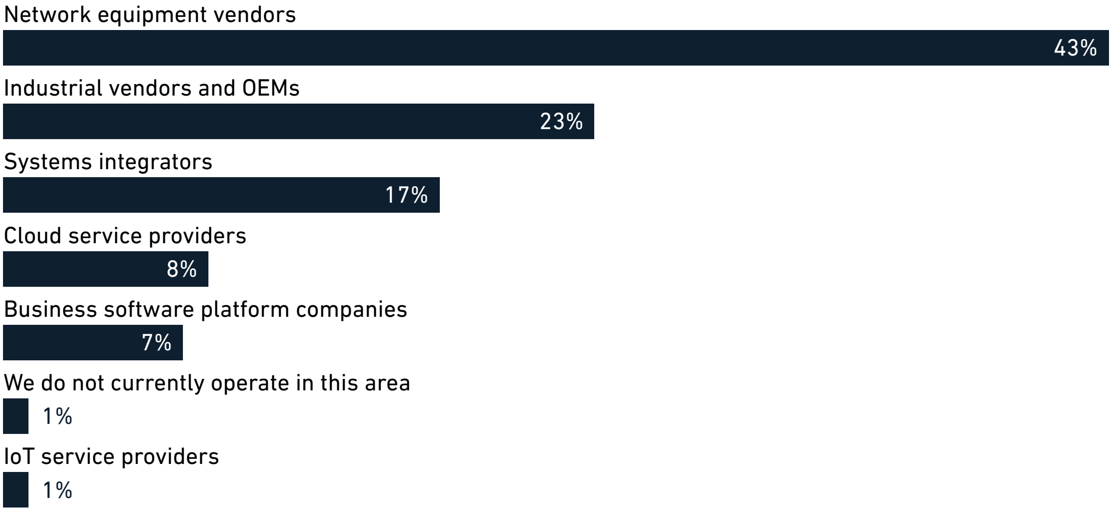
2.2 Fragmentation spurs co-opetition and partnerships

As core connectivity providers, operators are at the centre of the emerging ecosystem for edge computing. However, they do not necessarily have the requisite domain knowledge for all industry verticals or the relevant network infrastructure. Invariably, they will therefore have to form partnerships with multiple stakeholders to create solutions for end users. The potential partnerships and operator preferences are shown in Figure 7.

Figure 7

Network equipment vendors are operators' preferred partners for edge

Edge networking: rank 1st as preferred partner (operators only)



Source: GSMA Intelligence

Many of the operators interviewed in the GSMA Intelligence survey indicated a preference to partner with network equipment vendors (such as Nokia, Ericsson and Huawei). This is unsurprising as there is strong familiarity with these players and sustained engagement on the network engineering side. However, operators are increasingly aware that they need to engage with vendors and OEMs that produce industrial equipment as well as specialised products for verticals. There is increasing engagement with systems integrators too; these can provide the crucial domain knowledge for specific industry verticals.

There are also increasing partnership opportunities with cloud providers at the edge of the network. While cloud providers do not offer the connectivity part of the solution, they already have mindshare within the enterprise for managing workloads, as well as significant tools and developer marketplaces – something traditional telecoms operators cannot match. Recent collaboration examples include the following:

- Telefónica partnering with AWS to leverage private multi-access edge computing (MEC), aiming to improve operational efficiencies and deliver value to customers
- Verizon partnering with AWS Wavelength to provide edge cloud and compute service at the edge of Verizon's 5G network
- Google Cloud, Ericsson and TIM partnering to pilot 5G cloud solutions for telco edge enterprise use cases in automotive, transport and other sectors
- Microsoft partnering with BT in the UK to develop new private 5G and edge technology, targeting sustainability, safety and productivity in a range of enterprise verticals.

Operators also have opportunities to partner with hardware manufacturers. For example, in the automotive industry, major car manufacturers are producing increasingly 'digital' vehicles that require connectivity for wide area mobility applications and services (e.g. NTT has collaborated with Toyota).

The broad, emerging ecosystem of partnerships is captured in Table 2, with examples to illustrate.

Table 2

Examples of partnerships in edge computing

Telecoms operators	Cloud providers	Systems integrators and IoT service providers	Network equipment vendors and OEMs	Data centre providers	Car manufacturers
Telefónica	AWS				
	AWS			EdgeConneX	
Verizon		Wipro			
Verizon	AWS				
NTT		Cisco	Intel, Ericsson		Toyota
Telecom Italia	Google		Ericsson		
AT&T	Microsoft				
SK Telecom	Microsoft				
T-Mobile	Google				
Liberty Global, SoftBank	Hitachi			Arrcus	
Vodafone	AWS	Accenture			
	AWS				BMW
SoftBank				F5	
BT	Microsoft				
Vodafone	Microsoft				
Telstra	Microsoft				

Source: GSMA Intelligence

2.3 Evolving partnerships and ecosystem present new business models

As the edge computing market continues to evolve, it is creating new partnership opportunities and business models. Operators and specialist service providers will continue to be core to emerging partnership models, as connectivity infrastructure and services become the 'glue' enabling the extension of enterprise workloads to the edge of the network. In some cases, these partnerships and business models will be driven by new use cases defined by proximity to end points; in others, they will be driven by existing use cases deployed across the enterprise's workforce and locations.

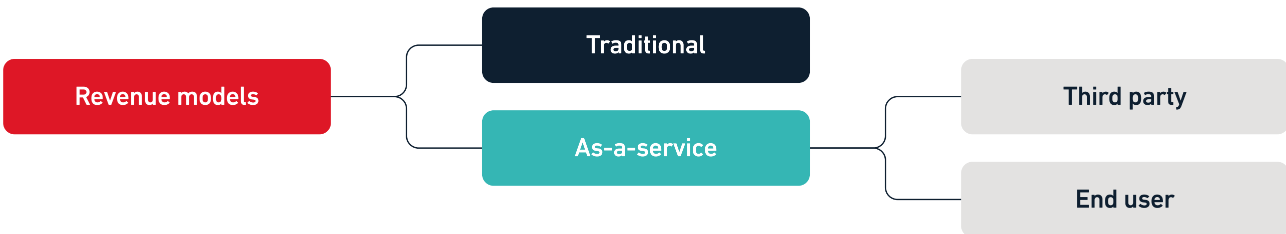
As edge computing is deployed, a crucial question concerns the funding of this new category of digital infrastructure, and the monetisation of the assets deployed. Funding can be done from multiple levels of the value chain – from IoT device makers to hyperscalers – but telecoms operators continue to be seen as the group that should bear primary responsibility for investment. This is unsurprising, as operators have the most distributed network infrastructure and close proximity to end points, making them the ideal candidate to build and operate edge computing infrastructure. Operators can be contrasted with hyperscalers, which deploy large-scale but highly centralised compute power in large data centres.

Regardless of who deploys the edge computing infrastructure, there are a few primary models for monetisation. These can be divided into traditional and as-a-service models.

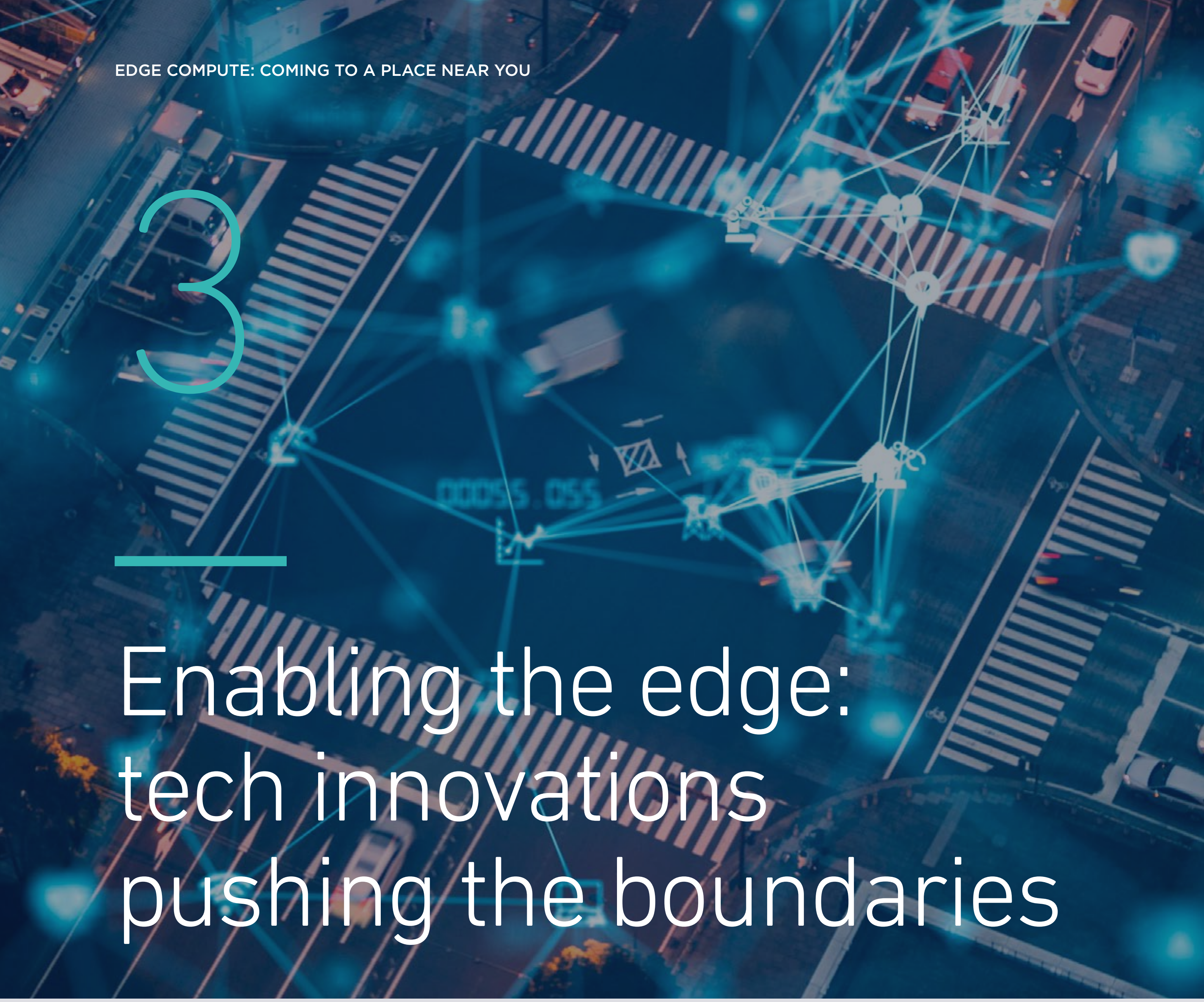
- With the **traditional monetisation model**, the costs of hardware are borne by end users, and fees are paid for consulting, systems integration and professional services. Other scenarios can include rental costs for the facilities used for deploying edge compute servers. In many cases, these facilities could reside within operator network assets, but could also be in a neutral host or commercial facilities. A further monetisation option is recurring charges for the connectivity provided for connected end points.
- Under the **as-a-service** model, developers can create applications for end users that tap into the edge computing resources provided at the network edge through APIs. These could be third-party APIs or those provided by end users themselves.
  - **Third party** – Monetisation can come from two main areas. The first is selling data derived from edge computing to third parties, who can use the data for business intelligence as well as monitoring assets in the field. The second option involves owners of the edge compute platforms charging developers for access to the resources, which can take the form of commission or recurring fees through API calls.
  - **End users** – Monetisation of end-user access to edge computing resources can take several forms. The most obvious is a direct revenue-share agreement, which can be seen in some of the recent telco-cloud partnership models that have emerged, with the telcos leveraging their large customer base. The second option is to charge end users for the service, either indirectly through a bundled service offering, or directly through a distinct subscription or pay-per-use fee. Finally, hybrid models are also emerging, with a baseline of services bundled in but several value-added services offered on a chargeable basis. The automotive industry is a good example of this approach, with new cars coming with premium services that are charged for.

Figure 8

Monetisation models for edge computing infrastructure



Source: GSMA Intelligence



3

Enabling the edge: tech innovations pushing the boundaries

3.1 Private 5G spurring new applications and services

5G unleashes the true potential of edge computing. Specifically, 5G private networks and edge computing together can lead to faster, more responsive and efficient applications and services across a range of industries.

Private 5G networks are dedicated, independent networks deployed and managed on an organisation's site or sliced off publicly deployed 5G. Private 5G has capabilities that distinguish it from Wi-Fi technology, such as secure access through SIM-based controls and encrypted network slices, radio access quality-of-service (QoS) controls on a per-application basis, and seamless handoffs between access points for improved mobility.

Private 5G will not replace Wi-Fi, but when coupled with IoT and edge computing it enables real-time use cases previously not possible due to a lack of reliability, security, low latency and high speed. Private 5G is expected to play a significant role in enhancing the efficiency and performance of autonomous vehicles, robotics and other automation technologies in the manufacturing industry. The technology can be used to develop applications that can improve efficiency, such as untethered robots or autonomous vehicles performing repetitive tasks with minimal downtime and fewer errors.

The edge platform is where value-added services or smart applications delivered over private 5G are run. While edge computing pushes computing resources close to the end user, private 5G brings the reliability, security and low latency that mission-critical applications need.

3.2 Edge computing and AI for automation

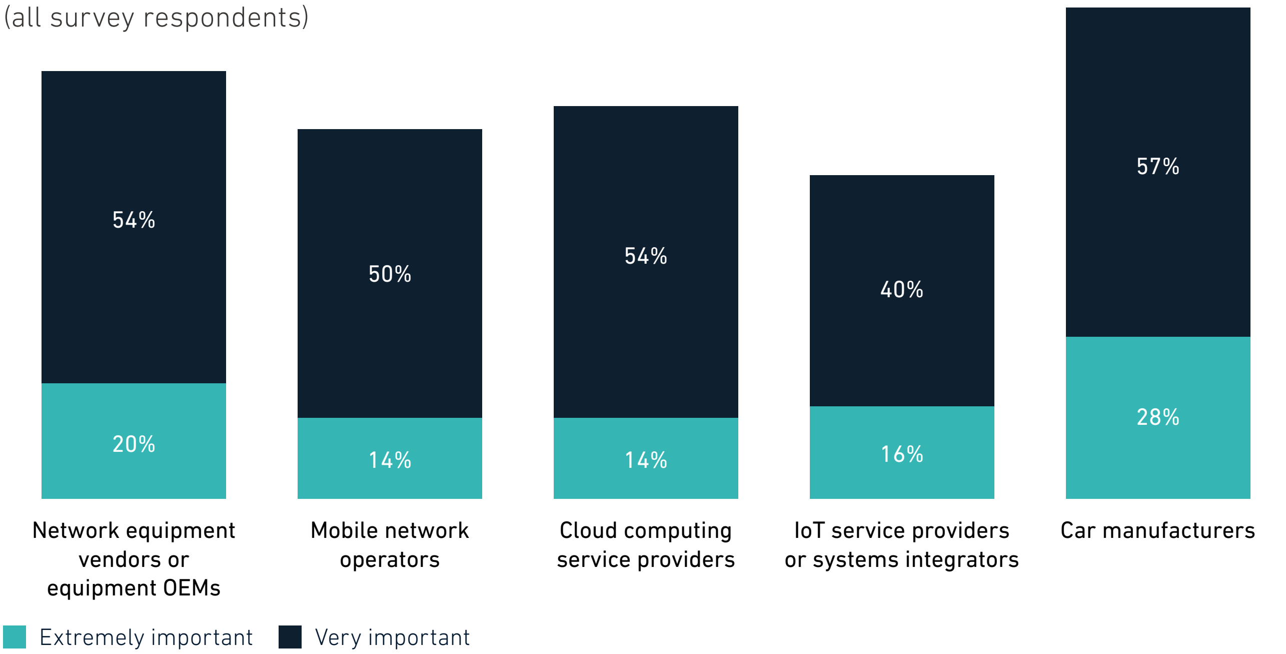
The integration of edge computing and AI allows for automated analytics. With 5G-Advanced and RedCap technologies, edge device processors offer the capability to integrate and run AI/ML models without relying on the cloud. AI algorithms running on edge devices offer real-time insights. Edge AI allows devices to make smarter decisions without connecting to the cloud or offsite data centres, enabling faster response times and decision-making. This is particularly important for applications that rely on large volumes of data, such as those based on cameras or sensors. Privacy and security concerns can also be addressed, as sensitive data can be processed locally.

The combination of edge and AI offers new opportunities for businesses and industries to develop innovative applications and services. Edge AI/ML algorithms can continue to run even if access to the network is disrupted. This capability is key in areas such as intelligent driverless vehicles and medical equipment.

Figure 9

More than half of survey respondents see AI and ML as extremely or very important to future edge success

How important do you think AI/ML is for the future of edge success? Breakdown by group (all survey respondents)



Source: GSMA Intelligence

3.3 Ground and space network integration stimulating a rise in edge

The expansion of mobile networks has been rapid. By the end of the first half of 2023, 238 operators in 94 markets had launched 5G mobile, with more people connected to the internet via mobile than ever before. The number of people using mobile internet reached 4.6 billion, or 57% of the global population.

Despite this progress, internet coverage is still not universal. Many users in remote areas and low- and middle-income countries (LMICs) do not have access to the internet due to the high costs associated with expanding terrestrial networks into areas of low population density. Meanwhile, climate-related disasters can pose a significant risk to terrestrial networks, making global internet coverage difficult.

To achieve global broadband coverage and highly reliable data transmission, satellite-terrestrial communication networks have garnered attention. Direct-to-device (D2D) communications has seen renewed interest, with low Earth orbit (LEO) satellites closer to terrestrial users than geostationary (GEO) satellites, resulting in lower latency and improved reliability. A lot of attention has centred on LEO since non-terrestrial network (NTN) integration became part of the cellular standards promulgated by the 3GPP. Release 17 (finalised in 2023) provides support for mobile broadband IoT integration between satellite and terrestrial networks.

NTNs use both satellites and high-altitude platforms such as balloons, airships and unmanned aerial vehicles in the stratosphere to ensure coverage and reliability. LEO satellites serve as relay nodes to complement the ground network. However, the large volume of data captured by satellites makes it impractical to transmit all the data generated to the ground and perform manual analysis. The rise of edge computing has sparked solutions; by deploying edge computing servers, data can be processed on satellites, which reduces delays from computing results, improves user experience and enables new use cases.

3.4 Convergence of telecoms and computing

The convergence of IT and telecoms is nurturing edge computing networks, which unite the two worlds, providing cloud computing capabilities at the edge of the telecoms network, near end users. This convergence is driving the rise of edge nodes. The GSMA defines an edge node as the point of presence in an edge computing architecture that hosts computing, storage and networking resources used for customer applications and data. According to the distance from the end user, there are two main categories of edge node: user edge and service provider edge. See Figure 10.

Figure 10

The edge computing network

USER EDGE				SERVICE PROVIDER EDGE			
Constrained device edge	**Smart device edge**	**Enterprise/ on-premise data centre edge**	**LAST-MILE NETWORKS**	**Access edge**	**Regional edge**	**INTERNET EDGE**	**Centralised data centres**
Microcontroller-based devices that are highly distributed in the physical world	Includes IoT (headless) and end-user client computers (smartphones, tablets, PCs, gaming, smart TVs) in accessible locations	Server-based computing (i.e. edge platforms) in secure locations (e.g. offices, smart factories)		Server-based computing at telco network and edge exchange sites	Server-based computing at the regional telco and direct sites		Server-based computing in traditional cloud data centres

EDGE NODES

Source: GSMA Intelligence

4

Outlook

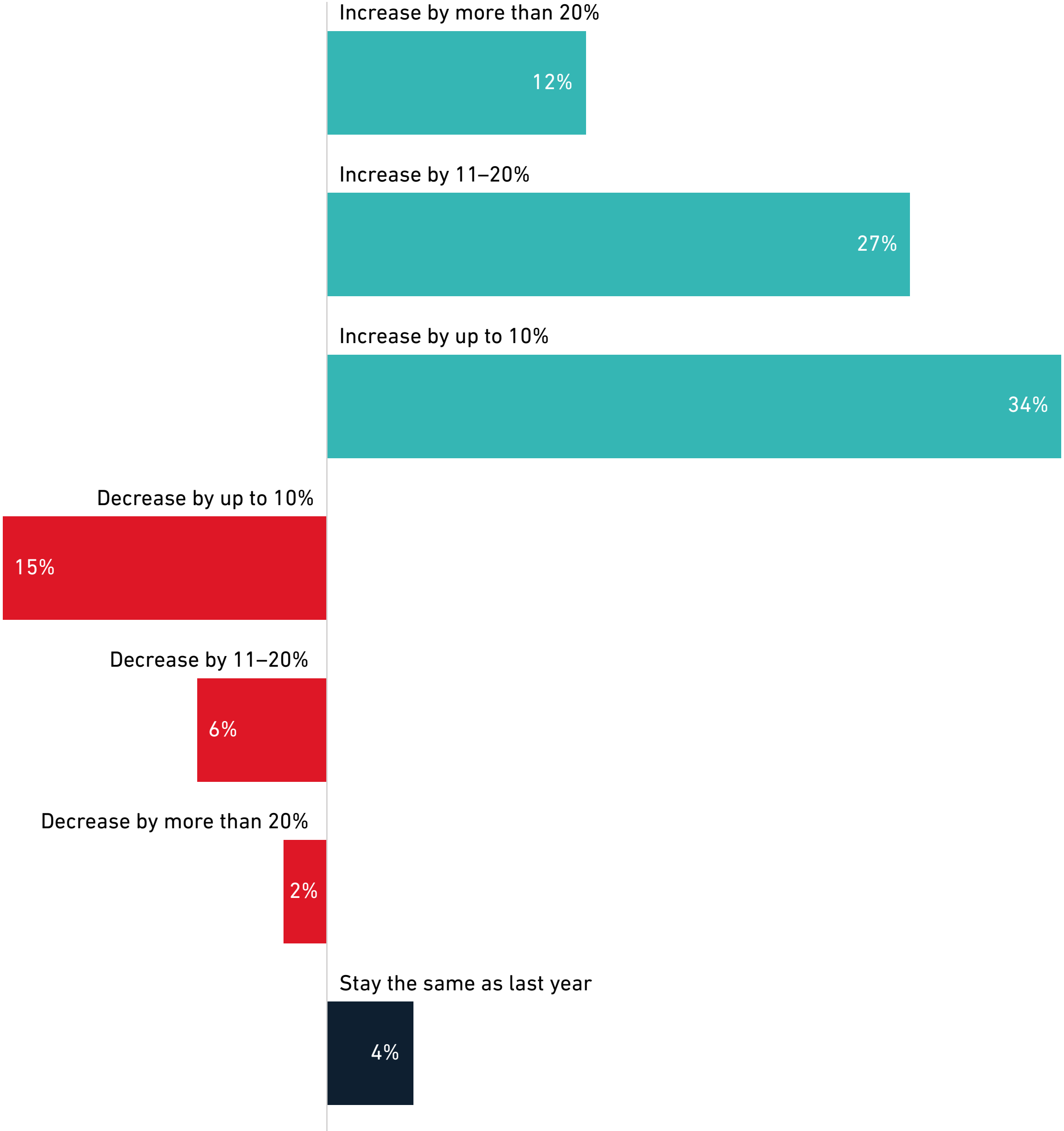
4.1 Positive investment sentiment for edge

Edge offers a multitude of benefits and can unlock new opportunities. This is reflected in the outlook for edge, with investments in the technology set to increase. Some 73% of respondents to the GSMA Intelligence survey claim they will increase their year-on-year spend on edge computing network investments, spurred by increasing edge demand, expanding use cases and the potential for high returns.

Figure 11

The majority of respondents plan to increase year-on-year investments in edge computing

Compared to last year [2023 versus 2022], how do you think the year-on-year expenditure of your edge computing network investment will change? (All survey respondents)



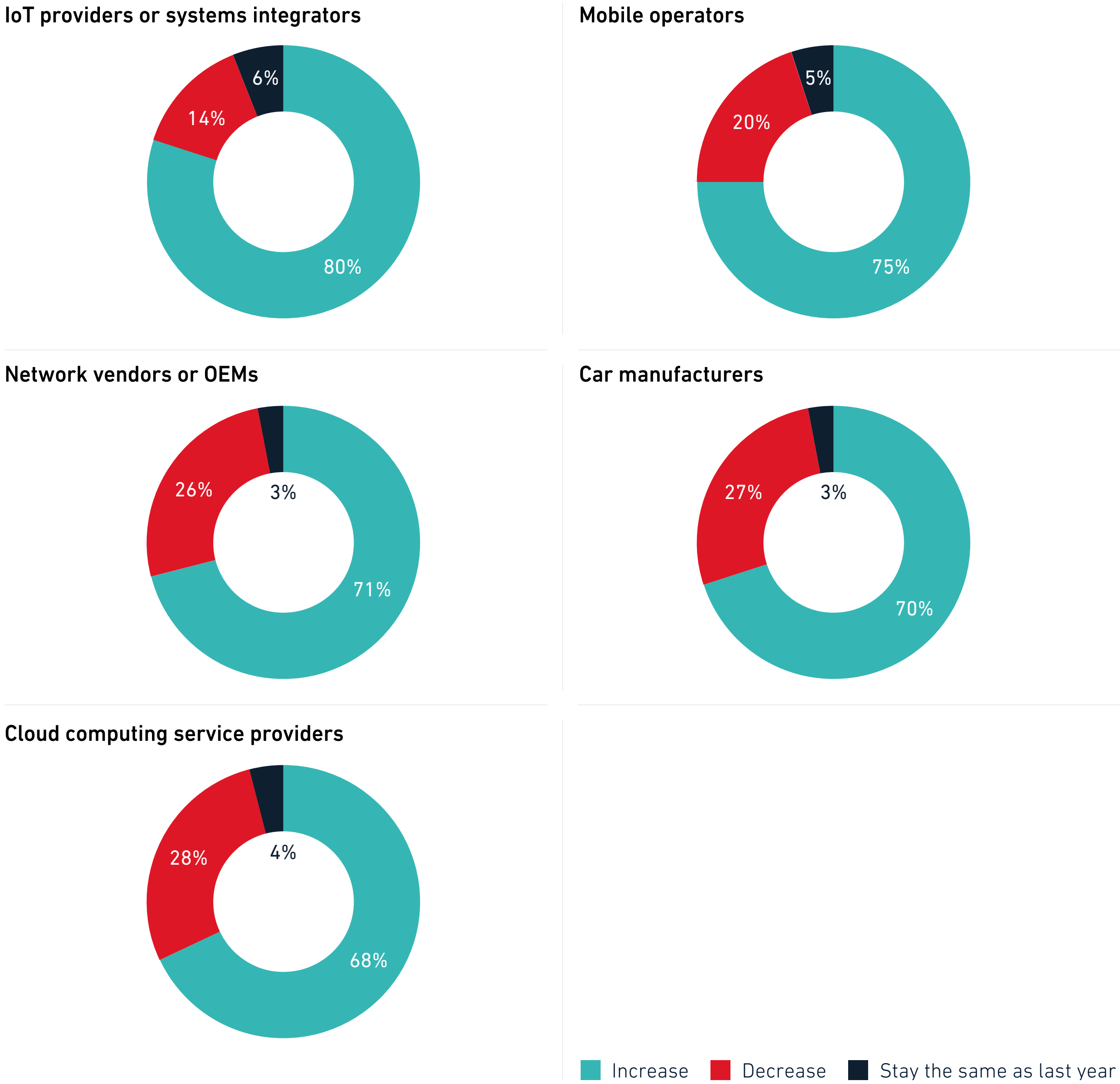
Source: GSMA Intelligence

The positive sentiment is consistent across different parts of the ecosystem, indicating a genuine rise in demand for edge-enabled services and suggesting the Covid-19 headwinds affecting capital budgets are reducing.

Figure 12

Positive investment sentiment holds across the value chain

Compared to last year [2023 versus 2022], how do you think the year-on-year expenditure of your edge computing network investment will change? (All survey respondents, by group)



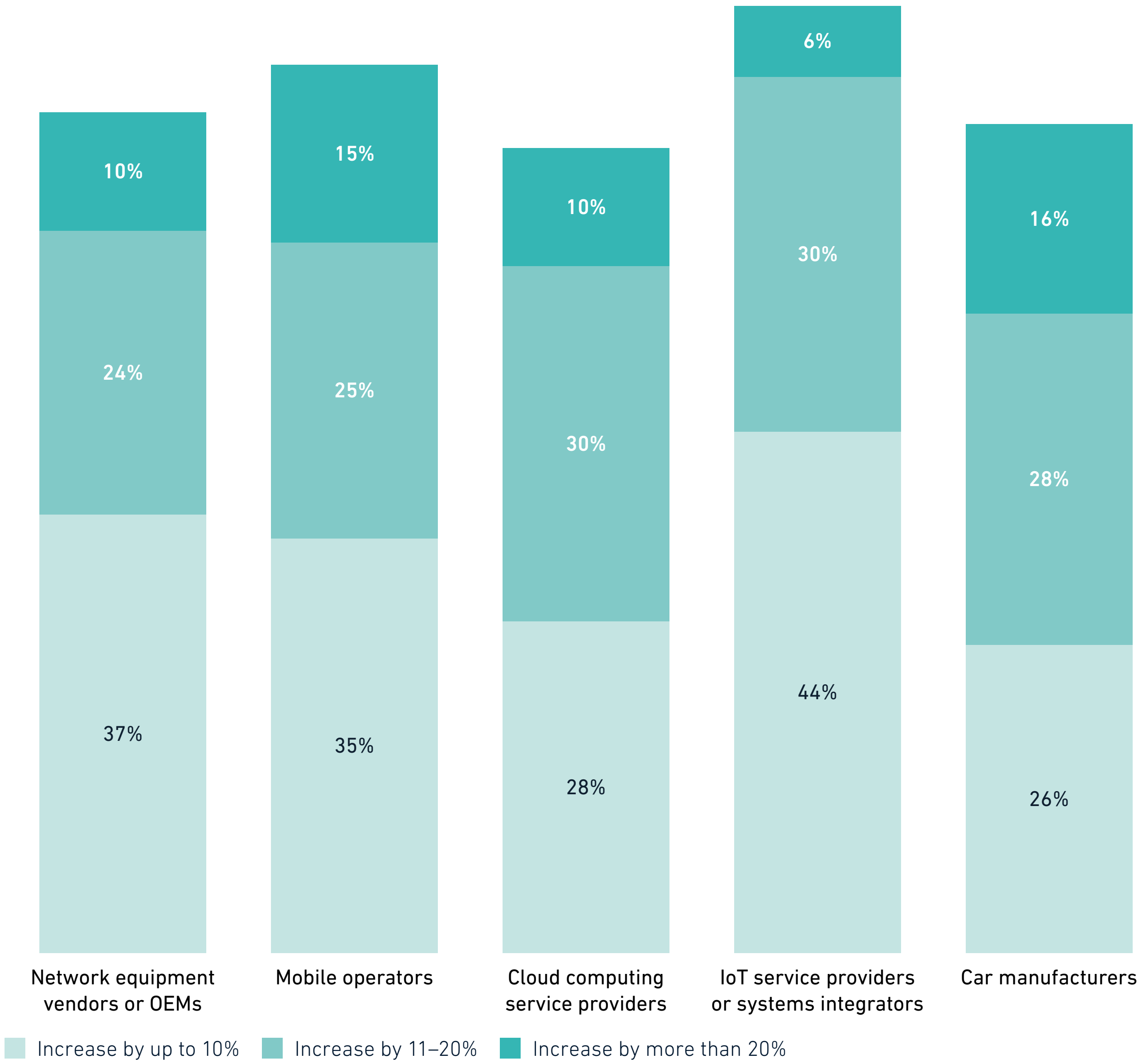
Source: GSMA Intelligence

Mobile operators and IoT service providers/systems integrators are expected to increase their spend on edge computing infrastructure the most year-on-year. The consistent response on investment plans could offer opportunities for partnerships. Expectations are for a median increase of around 10–15% compared to the previous 12 months. Of course, survey responses should be treated with caution. However, the consistency of positive responses across various sectors supports the credibility of an anticipated rise in edge network investment.

Figure 13

Investment plans could offer partnership opportunities

Compared to last year [2023 versus 2022], how do you think the year-on-year expenditure of your edge computing network investment will change? (All survey respondents, by group)



Source: GSMA Intelligence

4.2 Integration with legacy IT: the number one barrier

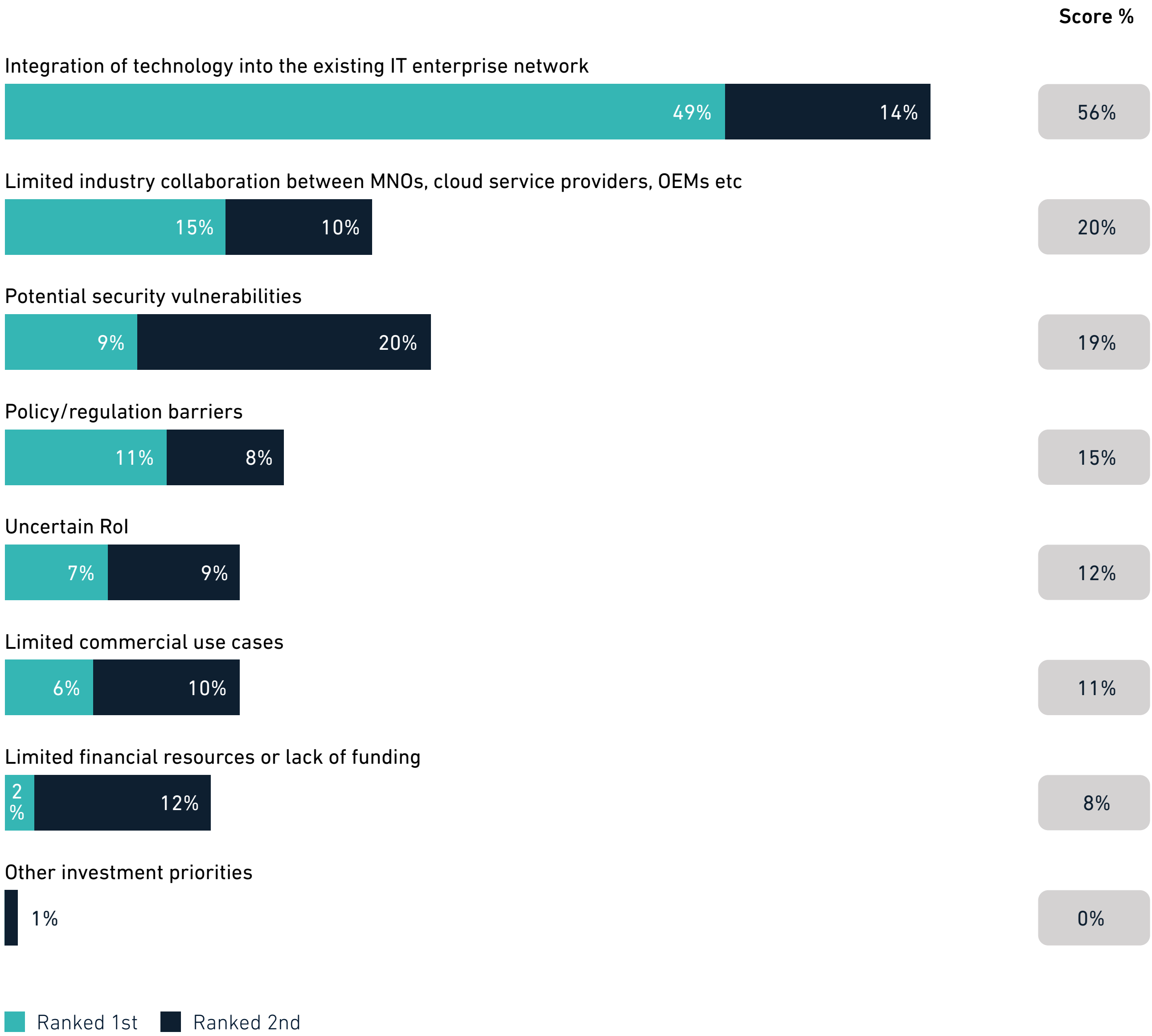
As edge computing is a developing technology, its deployment across the ecosystem faces significant obstacles. The most challenging is the integration with existing IT enterprise networks. To fully harness the potential of edge computing, it must be integrated with other technologies.

Organisations need to consider factors such as security, scalability and performance when integrating edge computing with their existing IT infrastructure. They need to ensure they have the necessary tools and processes to manage and monitor edge computing resources effectively. APIs have emerged as a critical tool in facilitating the integration of disparate IT systems and edge computing.

Figure 14

Primary obstacles to the deployment of edge computing infrastructure

What are the primary obstacles, at this moment in time, to the deployment of edge computing infrastructure? (Rank the two top obstacles)



Score % calculated as (Ranked 1st * 1) + (Ranked 2nd * 0.5)

Source: GSMA Intelligence

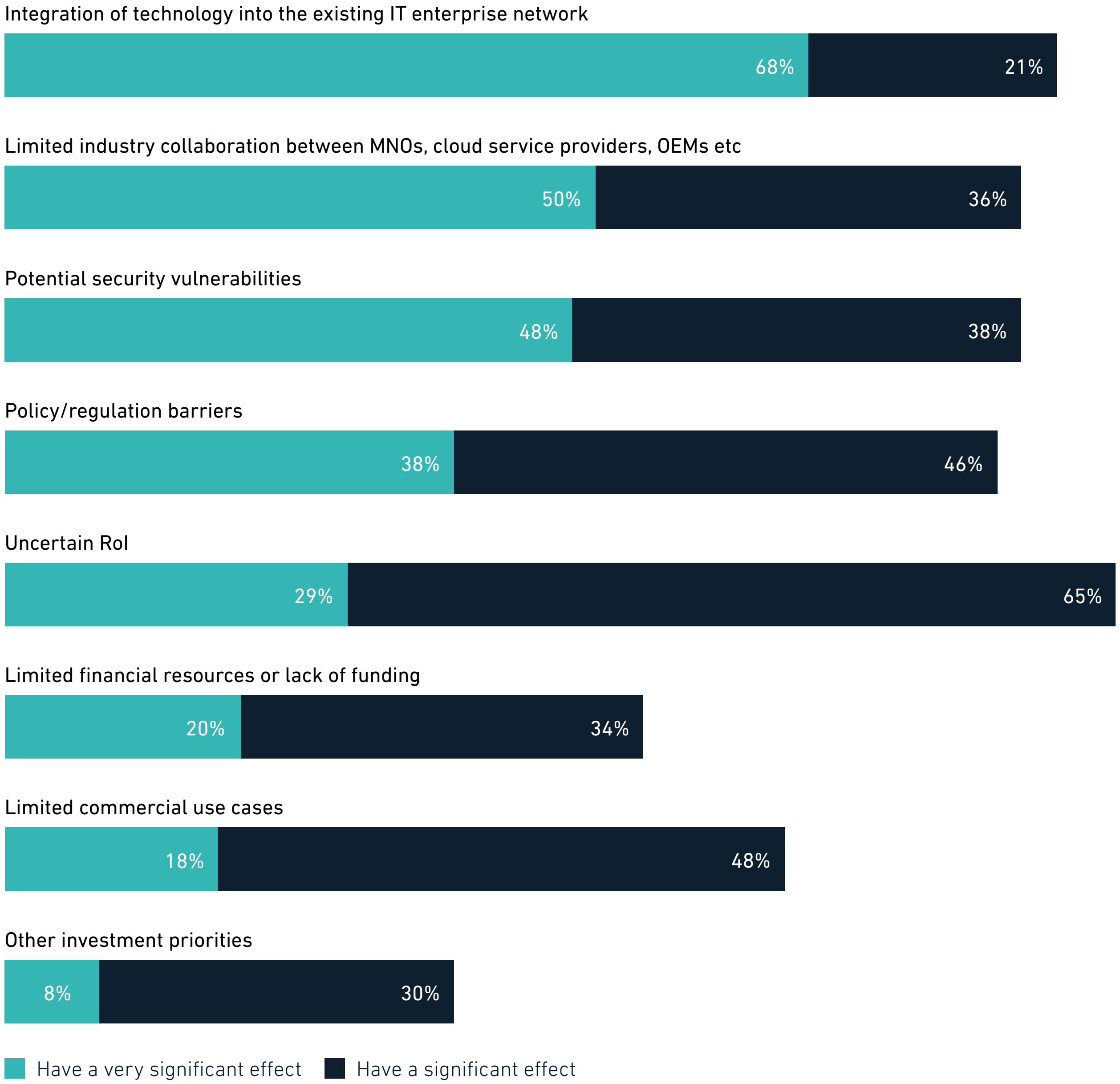
Limited industry collaboration between key players in the ecosystem was ranked as the second biggest obstacle, with 50% of respondents indicating that it could have a very significant effect on their company's edge computing deployment status.

Potential security vulnerabilities and policy/regulation barriers will also need to be factored in. Policy/regulation barriers can be seen across different aspects of edge computing, including data privacy and sustainability targets. The exponential growth of data processed at the network edge will make adhering to data privacy policy measures essential. Data centres and edge devices also consume a considerable amount of energy. If players are to meet their sustainability goals, it is imperative to factor in the sustainability impact of data centres, edge nodes and devices.

Figure 15

Integration of technology has the greatest impact on deployment of edge computing

How do you expect each of the following obstacles to affect your company's edge computing deployment status? (All survey respondents)



Source: GSMA Intelligence

4.3 Bridging the gaps: use cases and proof points

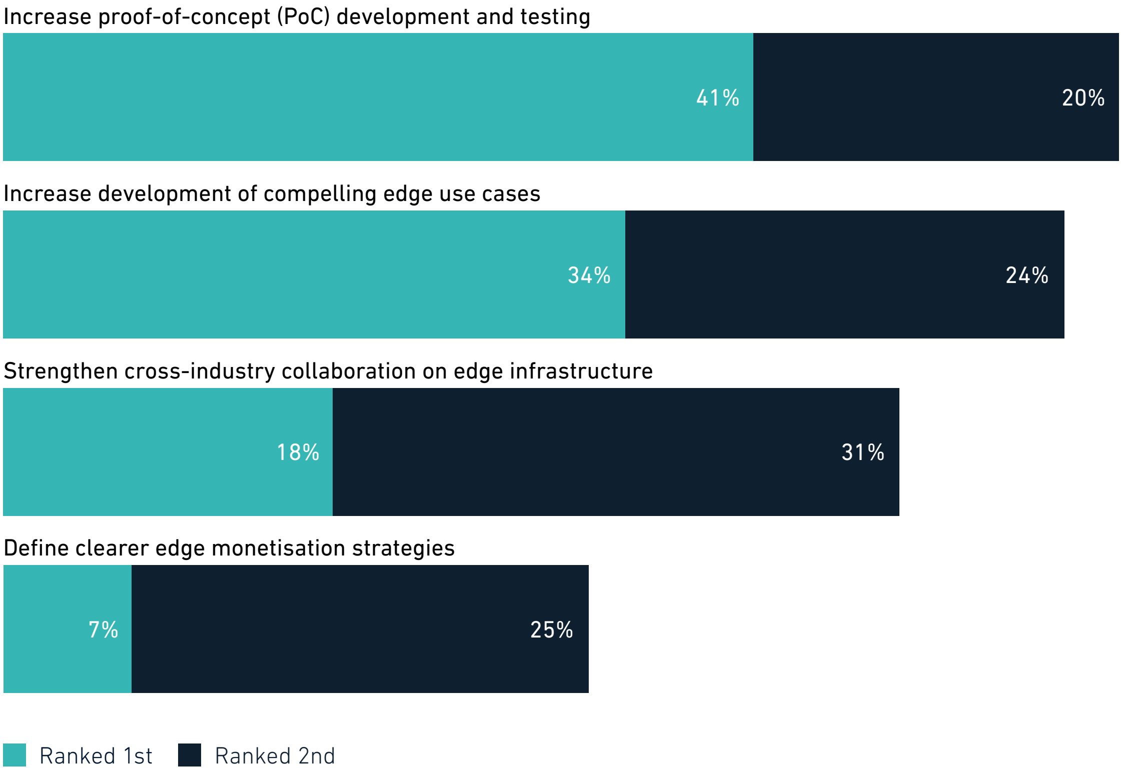
To scale edge computing technology, stakeholders must focus on two main aspects: increasing the development and testing of proof-of-concepts (PoCs), and creating more compelling use cases for edge computing. PoC development and testing allow for the creation of standardised software that can be deployed across different verticals to meet the varying demands of enterprises. This helps ensure edge computing is developed in a way that is both efficient and effective.

The development of compelling use cases is essential to ensure edge computing is embraced by a wider audience. Use cases play a significant role in showcasing the benefits of edge computing to potential users, demonstrating how the technology can be used to solve real-world challenges and create new opportunities. By spurring more use cases, businesses can help build momentum around the adoption of edge computing, which will ultimately drive further growth and innovation.

Figure 16

Increasing PoC development and testing will be essential to scale up edge computing

Thinking about edge computing, which of the following needs to be achieved to scale it? (All survey respondents)



Source: GSMA Intelligence

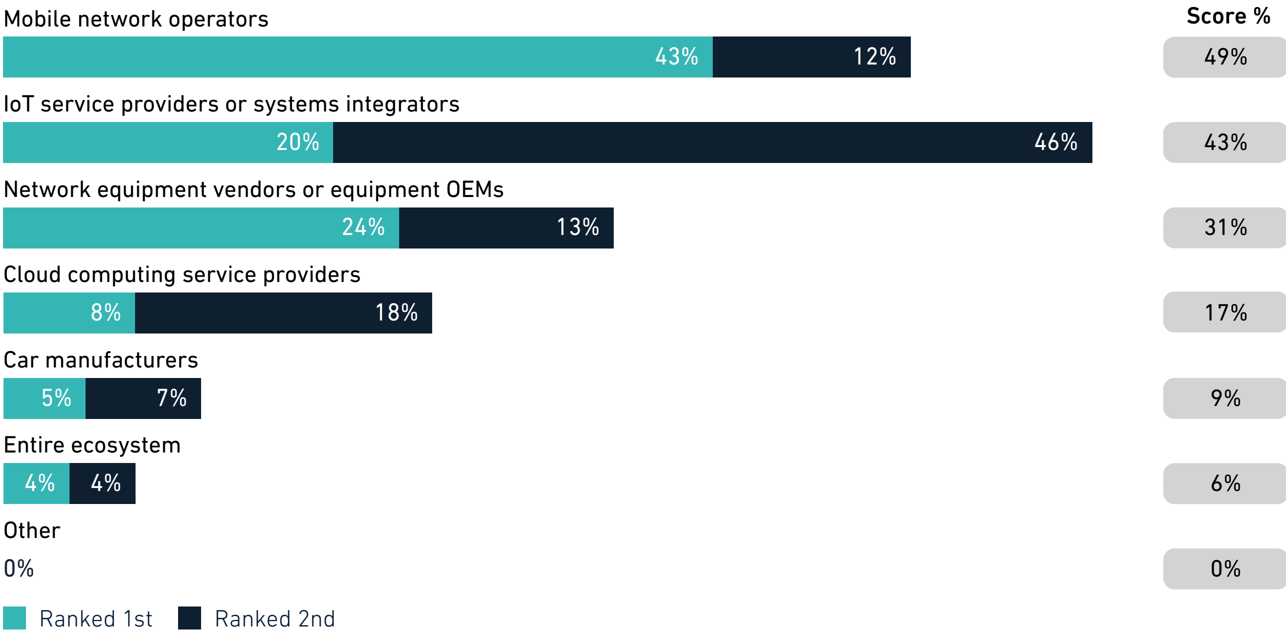
4.4 Sharing the investment burden

While edge computing infrastructure can be monetised by companies at different levels of the value chain (from IoT device makers to hyperscalers), telecoms operators continue to be seen as the group that should bear primary responsibility for investment. More than 40% of respondents surveyed rate operators as having primary responsibility.

Figure 17

Mobile network operators are expected to be primarily responsible for edge investments

Thinking about edge computing, who should bear the costs of network infrastructure deployment? Rank the top two (all survey respondents)



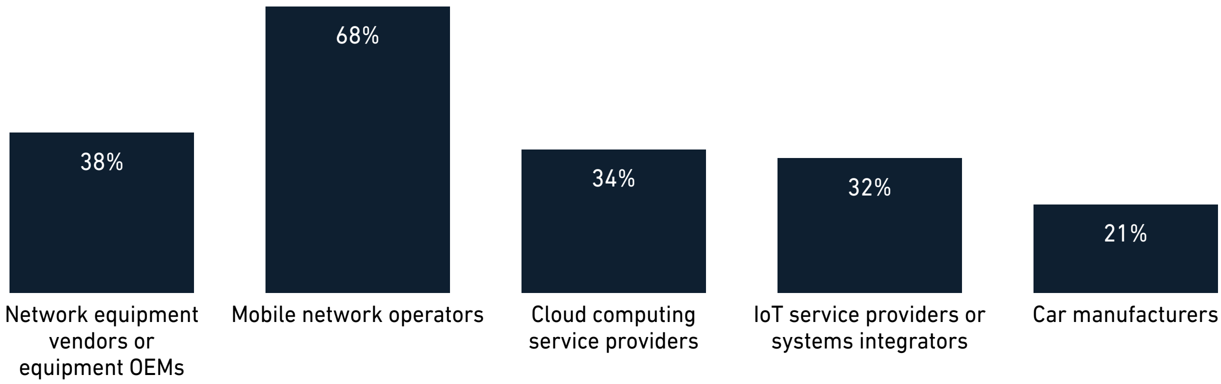
Score % calculated as (Ranked 1st * 1) + (Ranked 2nd * 0.5)

Source: GSMA Intelligence

Figure 18

Around two thirds of operators rate themselves as having the most responsibility for edge investments

Share of respondents that ranked their own sector first: breakdown by group



Source: GSMA Intelligence

That 68% of operators rate themselves as having the greatest responsibility reflects the fact that most edge costs are related to infrastructure, including active and passive equipment. When it comes to ICT, network operators play a vital role in keeping users connected, ensuring data traverses connected devices, routers, switches and data centres. However, operator revenue growth is still mostly in the low single digits, and the cost of capital is above net income margins – both of which are headwinds to investment. Co-investment models – either between operators and other suppliers or with enterprise clients – can help mitigate this pressure.

Edge computing offers the promise of new revenue opportunities, using the data generated by connected devices, as well as the advancement of use cases, particularly when coupled with other technologies, including 5G, slicing, AI and private networks.

Telecoms operators continue to be seen as the group bearing primary responsibility for investments in edge infrastructure according to the GSMA Intelligence survey. IoT providers and systems integrators are next, followed by equipment vendors. This reflects the RAN and localised data centre outlays inherent with the edge. Around two thirds of operators surveyed expect themselves to carry the main financial responsibility, but this situation is not sustainable. Operators' revenue growth is still in the low single digits, while the cost of capital exceeds net income margins, creating headwinds for investment. Edge computing opens up new opportunities, including the sharing of costs, revenue and expertise between stakeholders. Collaboration across all aspects and stages of edge computing deployment will be essential to unlock the opportunities ahead.

gsmaintelligence.com

GSMA Intelligence

gsmaintelligence.com