Edge: Where's the Money?

Introduction

By 2030, analysts estimate that global demand for edge services will reach $445 billion.[1] Yet, despite its enormous potential in the near future, the revenue opportunities for edge services remain partially obscured as new applications for the edge evolve. Communications Service Providers (CSPs) are in a position to capture a significant portion of that revenue, provided they adopt a strategic "lean-in" approach. Competition from hyperscalers, private cellular network players, systems integrators, and neutral host networking players will be fierce. Differentiation will depend on two key developments: the emergence of a new breed of highly targeted edge services and a strategic shift in monetization strategies.

If CSPs are to capitalize on the overall edge opportunity, they must pivot from the "network-out" model of service delivery that has served them for the last 20+ years to a customer-focused, "value-in" model. This means moving beyond monetizing the 3 M's (messages, megabytes, and minutes) and focusing on the 3 C's of content, context, and convergence.

1. STL Partners, "Edge computing market forecast: Second release," December 2022.

The Edge Isn't a Place, It's a Revenue Play

The edge isn't an intrinsic benefit so much as a requirement. What happens at the edge needs to happen at the edge. Migrating existing, centralized applications and workloads to the edge is often unnecessary and has no tangible business benefit. The real revenue opportunity lies in emerging edge applications and services that are tailored to a specific customer segment, vertical industry, or application type. These edge services will have underlying deployment requirements that are well suited to the edge, including:

- True low latency below 30 milliseconds round-trip (as opposed to 100–150 milliseconds for centralized deployments)
- Localized security where an application or data must remain on-site at a physically identifiable location
- Control plane traffic offloading when an excessive number of devices requires distribution of key control plane functions (e.g., AMF, SMF, and UPF in 5G) to prevent overloads and outages
- User plane traffic offloading where a need exists to localize application delivery (e.g., content delivery network caching)

This new breed of applications won't be delivered from a single type of edge location. There will need to be multiple edge architectures to support them, based on specific use cases, monetization drivers, and industry requirements. Below are several examples of the different types of edge deployments likely in the near future.

The Offload Edge

Here, key control plane functions are distributed to support high-volume, highly interactive IoT networks such as those used for smart cities or densely populated IoT areas. Localized Internet breakout may be deployed as part of this. A variation of this model would be a user plane offload edge for caching in content delivery networks.

The Satellite/Non-Terrestrial Networks (NTN) Edge

This edge architecture is designed to minimize the latency inherent in satellite/Non-Terrestrial Networks (NTN) communications and place key applications as close as possible, latency-wise, to the end user without adding further terrestrial network latency.

The Federated Edge

These locations require closely coupled synchronization between primary and secondary (subordinate) applications to ensure equal performance characteristics. Typically, this model will be based on a multi-operator, inter-edge deployment. Online gaming is a good example of a federated edge application.

The Private Edge

In use cases where data security is paramount, a private edge is ideal. These can be created using a standalone private cellular network, a fixed/Wi-Fi network, a fixed VPN, or a network slice from a public land mobile network (PLMN). Good examples of private edge use cases include automated, AI-enabled factories or machine-learning applications that require confidentiality and protection of intellectual property.

The Neutral Host Network Edge

This deployment model is best for use cases where superior indoor cellular network performance is needed and a neutral, single-owner architecture is preferred. Good candidates for a neutral host network edge could be large office buildings, airports, shipping ports, and sports stadiums.

The SASE/Branch Offload Edge

These edge deployments are typically part of an enterprise VPN with in-house or managed applications (e.g., security, UCaaS, VNFs) and are located at either the branch location or a closely coupled cloud location. The goal here is to improve application performance and reduce network bandwidth costs.

The Dynamic Edge

This deployment model simultaneously supports dynamic traffic management, premium offers based on geo-location, dynamic application loading based on VIP customer needs, and dynamic pricing. The dynamic edge is designed to provide a customized experience for a wide range of customers across B2C, B2B, and wholesale segments.

Show Me the Money

In order to deliver a differentiated user experience for customers, CSPs need to look into wider adoption of digital channels, online self-service capabilities, and flexible payment models. This holds true for B2B, B2C, and wholesale customers. The decades-old model of postpaid "revenue collection" billing offers no experience differentiation and is ill-suited to the dynamic, "value-in" environment that will drive edge services revenue in the future.

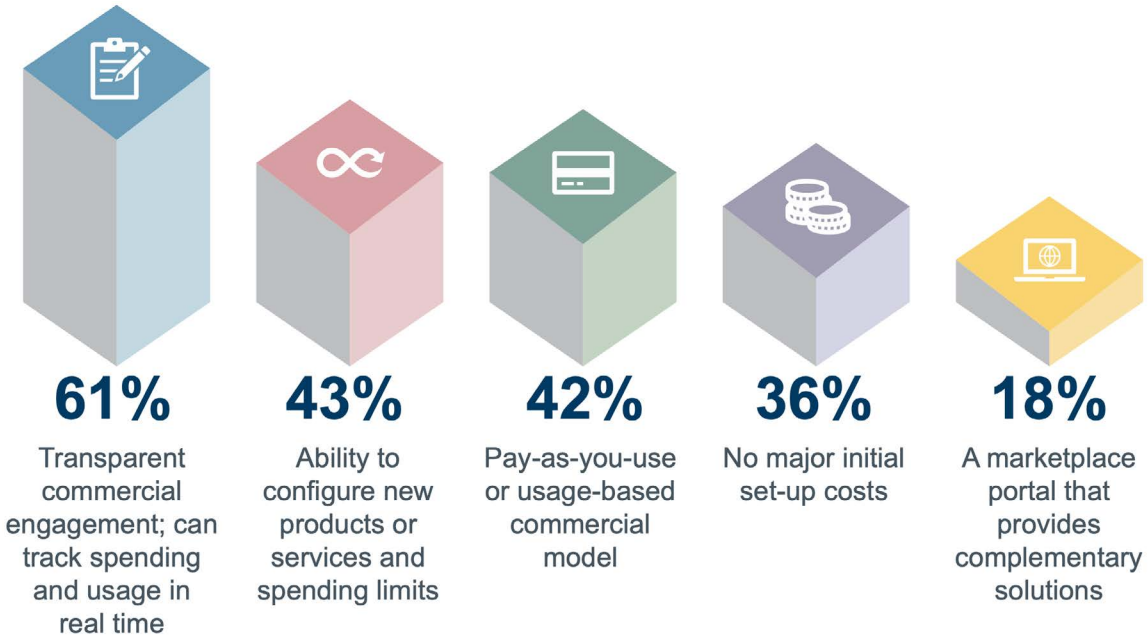

Figure 1. Analysis Mason enterprise survey response to the question, "What key capabilities would you like to see delivered as part of your commercial relationship with your CSP?"

If real-time, agile, digital experiences are to become the foundation for the future of edge monetization, the commercial models enabled by that foundation need to be equally innovative in order to deliver on true differentiation. Equally important, these models need to be adaptable across the various edge scenarios previously outlined.

There are four key edge monetization models to consider:

1. Network-as-a-Service

This replaces flat-rate tariffs with a real-time consumption model. It could be the consumption of compute, storage, or network resources and might include considerations such as latency and security as edge dividend components. Importantly, pricing information would be presented as real-time usage and spend data from which enterprises could make informed decisions. The NaaS model is still something of a "network-out" model of monetization, but it could serve as the foundation for the next model, outcome-based pricing.

2. Outcome-Based Pricing

A 180-degree pivot from network-out pricing models, outcome-based pricing takes a customer-centric, "value-in" perspective. Given its tight coupling to perceived customer value, this pricing model works best when it is applied to a specific vertical industry or business metric such as those typically found in annual reports (see Figure 2 on the next page). Read more Outcome-Based Pricing examples in the white paper.

Compliance	Delivery of vertical compliant service offering (such as HIPAA II in Healthcare)
Security	Clean-Pipe delivery across all access types, agreed response time (<4hr?) to zero day attacks
Productivity	Reduced time to service activation/adds/moves/changes through self-help digital marketplace automation
Transparency	Zero Bill shock – provision of real-time spend control, usage insights and flexible payment terms via digital marketplace
Availability	Service uptime benchmark independent of location (office, telecommuter, mobile)
Access	Consistent 'no loss' service availability across cellular/Wi-Fi services indoor and outdoor
Performance	<2 second application load times independent of location
Consistency	Any service, any device, any location

Figure 2. Examples of outcome-based pricing metrics.

3. Congestion-Based Pricing

This is a consumer-centric model based on low-volume, adaptable pricing plans with a strategic goal of taking market share from competitors. Congestion-based pricing starts with an entry-level plan for consumers and seeks to upsell them to a premium plan over time. For example, the starting plan may feature 5 GB of data per month dynamically priced at 50 cents per hour under normal traffic conditions, and dropping to 30 cents per hour during high-congestion periods but with a commensurate drop in throughput from 1 Gbps to 250 Kbps. This throttling would be applied at the edge to prevent core network congestion. Plan participants would be identified by assigning their IMSI to a congestion group within the converged charging system. During periods of high congestion, these users could then be targeted for upsell opportunities to increase their throughput during congestion periods as part of a premium plan.

Details	Group A Premium	Group B Congestion
Plan	Data Volume – 40GB	Data Volume – 5GB
Price	$30/month	Dynamic – 50c/hour (~$15/month), reduces to 30c/hour when congested
Throughput	Best available/5G	1GBps/4G
Congestion Action	Reduce multi-carrier aggregation	Throttle to 250Kbps, Upsell Offer
Non-Congestion Action	Multi-carrier aggregation	1Hr temporary 'Group A' experience offer FOC bi-monthly

Figure 3. A comparison between premium and dynamically priced congestion-based mobile plans.

4. Dynamic Edge Applications

Dynamic edge applications represent a kind of "VIP lane" for consumer traffic that reflects the ability to treat specific edge applications with special properties such as low latency. For nomadic cloud-gaming users who enjoy high-throughput, low-latency access to edge-based, cloud-gaming apps via fiber-to-the-home, they could be offered a high-performance mobile plan to ensure the same high performance and low latency on the road. Such a plan, of course, would require targeted 5G support through a specific QoS class or be enforced through an ultra-reliable, low-latency communications (URLLC) network slice.

Critical to delivering a mobile "VIP lane" would be the ability to service these users when they connect to an edge location where the gaming application is not present. This scenario could lead to excessive round-trip latencies of greater than 50 milliseconds. Using information about the user stored in the converged charging system, a multidomain orchestration system might be instructed to create a copy of the gaming app at that edge site to deliver lower latencies.

Plan	Details
Target gaming app latency	<50ms RT
Location	Urban Cell Tracking Areas
Performance	5G QCI/URLLC slice defined 100Mbps
Pricing	$40 for 60GB plus $25 for VIP Lane annual subscription
Extras	Add additional devices, grant 1hr gaming experience to friend, home femtocell for 'always mobile'

Figure 4. An example of a dynamic edge application pricing plan.

Further Considerations

The edge and pricing models provided here, while extensive, are far from exhaustive. As edge services and 5G applications continue to evolve, a host of new models will emerge such as geo-location models for AI-based retail applications, neutral host models for augmented reality applications in sports stadiums, IoT applications in ports, federated edge models comprised of multiple edge providers and vertical applications such as smart cities.

The developments at the edge will take place there because they must. As outlined earlier, what happens at the edge needs to happen at the edge. Similarly, new monetization models will need to be created for those new edge services. Using existing monetization approaches and trying to retrofit them to the edge is futile. With billions of dollars of revenue in the balance, CSPs need to stop thinking about revenue from the network out and start thinking about it in terms of customer value in. Only then will they have a true competitive edge.

About MATRIXX Software

MATRIXX Software delivers a modern converged charging and commerce solution proven at scale. Its cloud native Digital Commerce Platform provides network-grade, mission critical software that unlocks new network monetization opportunities. With its no-code configuration capabilities, MATRIXX empowers service providers with the agility necessary to easily develop, deploy and monetize new products and services. MATRIXX is the platform of choice powering many of the world's leading communications companies, IoT players and emerging network infrastructure providers. MATRIXX makes it possible to harness commercial innovation and on-demand customer experience to better compete and drive new revenue and growth opportunities across markets and verticals.

matrixx.com