AvidThink®

Sponsored by MobiledgeX™

Orchestrator of Orchestrators — Unifying the Edge

RESEARCH BRIEF

Table of Contents

Research Briefs are independent content created by analysts working for AvidThink LLC. These reports are made possible through the sponsorship of our commercial supporters. Sponsors do not have any editorial control over the report content, and the views represented herein are solely those of AvidThink LLC. For more information about report sponsorships, please reach out to us at research@avidthink.com.

About AvidThink®

AvidThink is a research and analysis firm focused on providing cutting edge insights into the latest in infrastructure technologies. Formerly SDxCentral's research group, AvidThink launched as an independent company in October 2018. Over the last five years, over 110,000 copies of AvidThink's research reports (under the SDxCentral brand) have been downloaded by 40,000 technology buyers and industry thought leaders. AvidThink's expertise covers Edge and IoT, SD-WAN, cloud and containers, SDN, NFV, hyper-convergence and infrastructure applications for AI/ML and security. Visit AvidThink at www.avidthink.com.

Orchestrator of Orchestrators — Unifying the Edge

Introduction

Edge computing has become a vital pillar of both mobile and wireline carrier strategies. Today, many carriers are still working through their approaches for edge and telco clouds that incorporate both public clouds and private clouds, an evolution of their network functions virtualization (NFV) infrastructure. AvidThink's recent conversations with carriers firming up their plans surfaced the need for an abstraction and orchestration layer that can coordinate telco workloads across multiple clouds.

This research brief will discuss the state of the edge at carriers and why this layer, which we'll term Orchestrator of Orchestrators, is needed. We'll also touch on the essential functions this layer needs to provide. By highlighting this need in the market, we hope that the telco ecosystem can have an open discussion around how to best address this in a way that benefits the carriers and their partners.

State of the Edge

The edge market today is promising but nascent. There are early use cases and proofs of concept, but, aside from the content delivery network (CDN) as an edge application, category revenue is thin. As the edge develops, we will see distinct edge segments emerge instead of a single uniform market. This approach mirrors the development of the cloud computing market — unsurprising since the edge represents part of the distributed cloud.

This approach mirrors the development of the cloud computing market — unsurprising since the edge represents part of the distributed cloud.

Edges of Interest

There are currently two leading edges of interest:

- **Network edge** refers to edge computing resources embedded in mobile and wireline networks and situated within the figurative last mile from end-user devices. In mobile networks, edge sites include mobile switching centers (MSCs), sometimes termed service aggregation points (SAPs) or mobile switching offices (MSOs), depending on the carrier. For wireline networks, these are the central office (CO) or for multiservice operators (MSOs), the cable headend. For convenience, we'll consider edge computing clouds located in major metropolitan areas (but outside of carrier facilities) as part of the network edge. These metro edges, which are often hosted in co-location facilities, serve use cases such as cloud gaming, media and entertainment, retail and point-of-sale, and drone-related or automotive applications.
- **On-premises edge** covers computing resources located in private or semi-private spaces, including stadiums, transportation hubs, ports, factories, farms, mines, and oilfields. The critical difference between on-prem edge and legacy on-site computing servers in machine rooms is the software stack that powers the edge and a centralized management and orchestration system, typically cloud-hosted. Use cases that favor on-prem edge include manufacturing and any Industry 4.0 applications control systems, the Industrial Internet of Things (IIoT), the Internet of Medical Things (IoMT), and computer-vision powered video surveillance.

From centralized clouds to the remote edge, there are multiple edge locations on the spectrum. Some edge taxonomies include devices as edge locations. From AvidThink's standpoint, an automobile as an edge site is reasonable since today's vehicles look increasingly like mini datacenters on wheels. But treating a mobile phone as an edge location makes less sense to us. The cloud-to-edge spectrum also includes mobile base station edge sites, but these are not yet commercially viable.

Edge Ecosystem Providers

Examining the ecosystem players involved in these two edges yields a vast collection of companies with different agendas:

- **Network edge (both mobile and fixed)** — Hyperscalers like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) are involved and looking to extend their cloud reach. Commercially, the only hyperscaler with a public mobile edge in production is AWS. AWS partners with network operators including Verizon, Vodafone, KDDI, and SKT worldwide on their Wavelength product, with Bell, Singtel, and Telstra in the wings. Azure, with AT&T, and GCP have promised public mobile edge but have yet to launch. In addition, carriers have launched their public edges in partnership with third-party providers and software stack vendors. Examples include Cox with StackPath, an edge startup, and Lumen, formerly CenturyLink, with VMware. Finally, co-location providers such as Equinix own and control strategic assets in both the physical (data centers) and virtual (prime interconnects) realms.
- **On-premises edge** — Hyperscalers also feature in this space. For instance, AWS and Azure provide edge computing using AWS Outposts and Azure Stack Edge respectively, while partnering with carriers providing private 4G/5G mobile networks. Example tie-ups include AWS and Verizon, AWS and Telefonica, and Azure with Verizon. GCP touts its Kubernetes-powered Google Anthos system for edge workload hosting and cluster and container management. Anthos is agnostic to on-premises server hardware vendors, and Google showcases verified solutions with Dell, HPE, Lenovo, and other server original equipment manufacturers (OEMs). The principal server OEMs also have their own on-prem edge offerings, some providing their flavor of edge orchestration or hardware management, as well as certifying integrated edge stacks with platform partners like VMware and IBM/Red Hat. And in three-way tie-ups, we have partnerships like SKT with Dell and VMware, resulting in OneBox MEC, an integrated private 5G with edge computing. Meanwhile, major system integrators are pulling together blueprints for on-prem systems with private 5G, an edge computing hardware and software stack, and vertical enterprise applications for manufacturing, healthcare, and other key markets.

Communication Service Providers and the Edge Opportunity

The majority of communication service providers (CSPs) believe they have legitimate claims on the edge computing market. In AvidThink's conversations with major CSPs, the standard view is that because CSPs own the last mile, they should have an inherent advantage in participating in the edge ecosystem and extracting value. But, based on the complex interaction of strategies of not just CSPs but hyperscalers, independent software vendors (ISVs), system integrators, and server vendors, edge market evolution will prove more complicated. For example, even though CSPs previously dominated voice communications, they couldn't translate that dominance into the Unified Communications (UC) and UC-as-a-service (UCaaS) space. The same could happen for the edge.

Strategically though, CSPs understand the edge can play a crucial role in:

- Creating new revenue opportunities both in-network and on-premises.
- Enhancing quality of service (QoS) in their networks by distributing traffic processing and other network functions to multiple locations closer to end users.
- Improving end-to-end application performance by reducing the topological distance from users to computing resources.
- Providing leverage during negotiations with other players in the ecosystem who have strong developer or enterprise relationships, such as the hyperscalers or server OEMs.

Having proved with early rollouts that 5G enhanced mobile broadband (eMBB) services don't drive significant increases in consumer average revenue per user (ARPU), CSPs who are mobile network operators (MNOs) are looking to other sources for a return on investment (ROI) on their tens of billions of dollars spent building out 5G networks. Basic carrier connectivity is a business with eroding margins. The CSPs hope that edge computing can bring new monetizable services to drive up ARPU and business average revenue per account (ARPA).

MNOs hope to monetize applications running on mobile edges that serve their customer base — cloud gaming, enhanced event experiences, pop-up retail, etc. For wireline operators, the network edge can host enterprise applications previously on-prem, saving space and reducing headaches around server security, power, cooling, humidity, and dust management.

Based on the complex interaction of strategies of not just CSPs but hyperscalers, independent software vendors (ISVs), system integrators, and server vendors, edge market evolution will prove more complicated.

Edge as Internal CSP Platform

At a minimum, even if edge platforms can't be directly monetized, CSPs are hoping that the edge can be used to improve network performance or security services. As an example of QoS improvement, MNOs can host 5G user plane functions (UPFs) in edge locations to reduce network latency and improve throughput. Likewise, edge computing platforms can host private 5G cores (5GC) on enterprise premises, allowing local traffic breakout and integration with on-premises assets.

Both MNO and wireline operators can host security and software-defined wide-area networking (SD-WAN) gateways to provide the utmost performance to enterprise users that subscribe to managed SD-WAN or secure access services edge (SASE) products.

Using the edge as an internal platform can relieve the pressure on CSPs to prematurely unearth an edge "killer app," when the ecosystem and value chain aren't sufficiently evolved. CSPs can benefit from platform buildout and demonstrate ROI without having to immediately monetize.

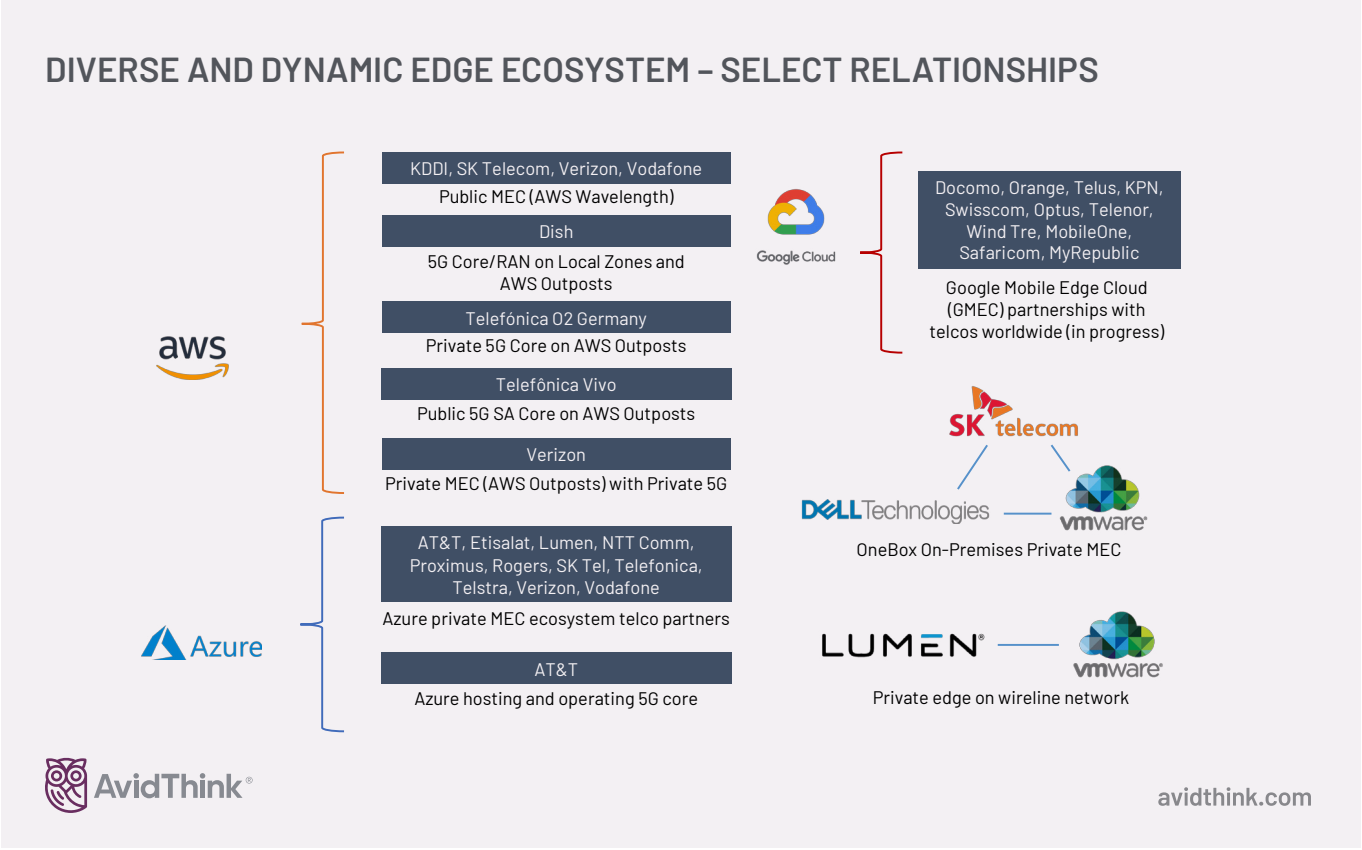
Building the Edge and Go-to-Market Options

Regardless of the initial use of edge platforms, CSPs are moving to incorporate edge computing into their 5G and next-generation service plans. We're seeing a few different approaches take shape, with many CSPs embracing a diverse set of strategies as hedging moves:

- **Build your own edge (BYOE)** — This strategy involves working closely with server OEMs and virtualization and container software stack providers — IBM/Red Hat, VMware, but also others like SuSE, Platform 9, and Rafay Systems — to develop a carrier-owned edge. This involves the CSP installing, operating, and managing the edge either independently or contracting with a system integrator (SI) for assistance.
- **Partner for an edge** — For this, the CSP partners with a hyperscaler, SI, or other edge vendor, e.g., StackPath, who operates the edge on the CSP's behalf and takes on the capital expenditure (CapEx). There may also be variants where the CSP provides the investment to create the edge, but a 3rd-party operates it.

For the go-to-market (GTM) options, the CSP can choose to front the product in both the BYOE or partner build options, selling to the enterprise developers. Or, if the CSP has picked a partner strategy, the CSP can opt for joint go-to-market with the partner or have the partner take on charging with revenue share. For example, in the Verizon and AWS relationship, AWS Wavelength public mobile edge, AWS charges a slight premium for mobile edge computing services and shares incoming revenue for the mobile edge service with Verizon.

The same partnership decisions or BYOE options are available for the on-prem edge. Carriers can choose to build a solution for enterprises through OEM servers in concert with a virtualization software stack. Alternatively, they can work with the hyperscalers and other partners who can provide a managed on-prem stack that carriers resell or partner with for GTM.



Mixing up the Edge

The reality for most CSPs today, with the choices available, is to go with a mixed strategy. There's no clear dominant strategy, so hedging allows CSPs to buy time to learn more about the edge market. Ironically, just a few years ago, CSPs were pushing back against partnerships with hyperscalers on the edge, concerned that hyperscalers would further erode CSP revenue streams. Today, partnering with multiple hyperscalers is acceptable and even necessary. By fostering multiple hyperscaler partnerships in parallel, CSPs hope that they can play one hyperscaler against another to gain leverage.

Consider Verizon, which has partnerships with AWS both on the public mobile edge and on the private 4G/5G edge. It is also partnered with Microsoft Azure on the private 4G/5G edge. At the same time, Verizon also has its edge application platform — powered by Rafay Systems technology — on universal customer premises equipment (uCPE) located in branch offices of its WAN customers. Another AWS partner, SKT, offers AWS Wavelength in Korea, but it has also partnered with Dell and VMware to build their OneBox MEC private edge computing solution.

As for hosting internal telco workloads on the edge, DISH plans to use AWS network edge sites, AWS Local Zones, to host its mobile network core and radio access network (RAN) software. Yet DISH has also revealed that they will be using Dell servers with VMware Tanzu in parallel.

CSPs are now faced with multiple cloud platforms, some owned and controlled by them, others by partners. Across these cloud platforms, complex workloads have to be orchestrated on top of bare metal, containers, and virtual machine infrastructure. In addition, CSPs may run in-house workloads or offer these platforms as edge computing services for public and private enterprise workloads.

Consequently, CSPs are faced with multiple dimensions to manage and are often overwhelmed. Recent AvidThink discussions with CSPs indicate they struggle to figure out how to work across these multiplicities. Furthermore, CSPs are seldom monolithic companies, and each division or business unit could choose a different edge platform or GTM strategy.

Introducing the Orchestrator of Orchestrators

Into this complex environment, we introduce the concept of the Orchestrator of Orchestrators (OoO). The OoO is a cloud-agnostic orchestration layer that sits above the orchestration interfaces provided by public and private cloud stacks, abstracting the underlying private and public edge and cloud platforms. Carriers we talked to have shared the need for this higher abstraction layer to manage the complexity of multiple stacks.

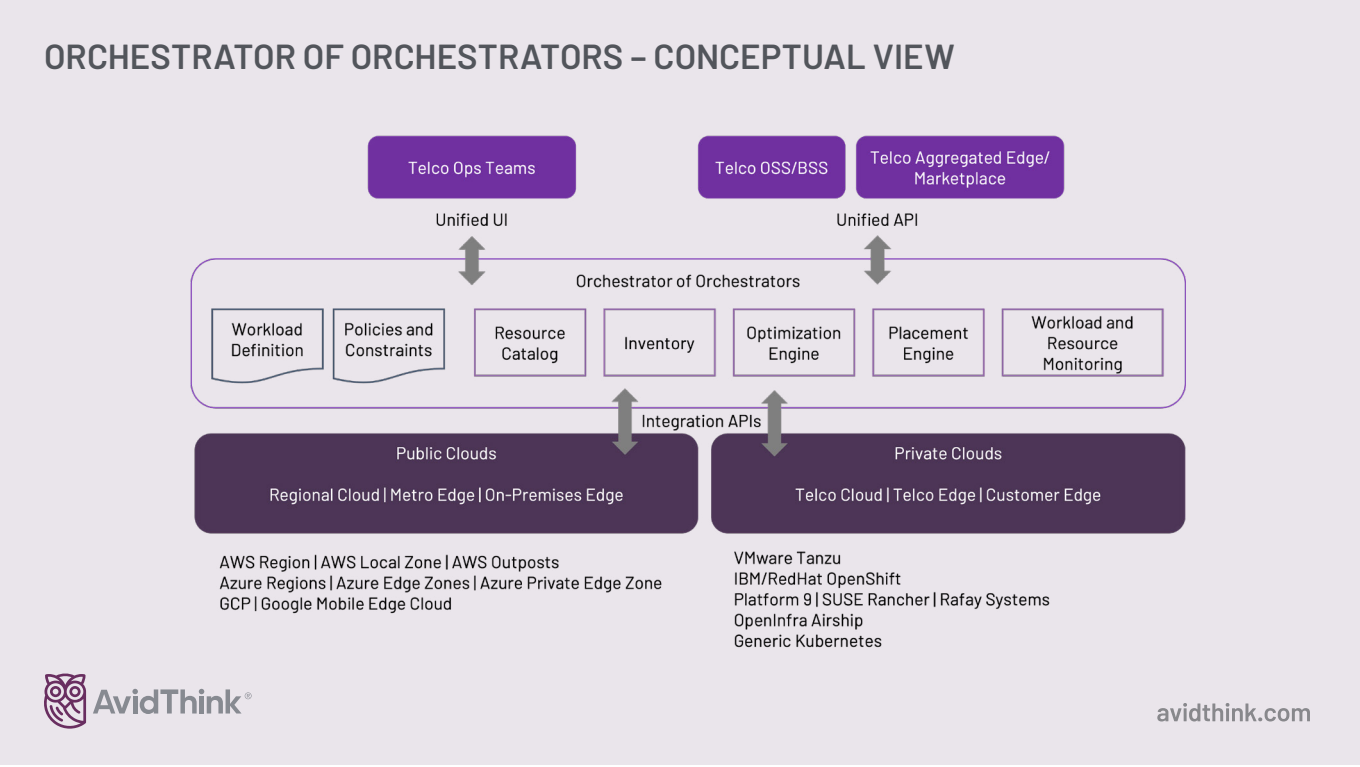
We've pulled together the high-level capabilities these carriers suggested should exist in an OoO layer, and they include:

- **Multi-edge cloud integration** — The OoO must integrate with popular edge and cloud platforms, including AWS, GCP, Azure, VMware Tanzu, Red Hat OpenShift, and generic Kubernetes. OoO must be able to unify available resources and services under a generalized catalog and tap into the inventory to extract metadata around these resources, for example, CPU, memory, storage configurations, hardware acceleration capabilities, pricing, and charging models.
- **Northbound API flexibility** — Carriers require the OoO to be easily integrated into existing OSS and BSS systems. They expect OoO northbound APIs to support provisioning, identity and access control, alerting, charging, and ticket generation. In addition, carriers would prefer the APIs to support integration with end-user-facing customer portals.

The OoO is a cloud-agnostic orchestration layer that sits above the orchestration interfaces provided by public and private cloud stacks, abstracting the underlying private and public edge and cloud platforms.

- **Workload goals and constraints specification** — A critical value-add brought by OoO is unification of how telco and end-user workloads should be specified. The common model must accommodate the spectrum from high I/O telco network functions to memory-hungry and storage-heavy enterprise web applications. The ability to convey business rules and constraints, including workload placement restriction by geography or type of datacenter (private vs. public), will be necessary to ensure compliance with government or corporate regulations, including rules around data sovereignty and privacy. Beyond those constraints, workload specifications need to include optimization goals — performance and QoS maximization or cost minimization — and service-level agreements to be met.
- **Intelligent workload placement engine** — The OoO should take the specifications described as input and combine that with the state and availability of infrastructure across the multiple clouds to make placement decisions. Workload placement is one of the critical functions of the OoO and a significant value-add to the carriers in optimizing for cost and performance while meeting workload specifications and constraints.

- **Global monitoring and remediation** — With the dynamic nature of the network and cloud resources, ongoing tracking of workloads across the different platforms will be crucial. The OoO needs to track KPIs and watch for drift. Should QoS constraints not be met, or should pricing change dramatically, e.g., with spot pricing mechanisms, the OoO may choose to dynamically reallocate workloads — to the extent the workload can be migrated or restarted without service impact. Likewise, underlying hardware or network failures should trigger the OoO to take remediation action.
- **Unified dashboard and roll-up reporting** — One of the key and immediate benefits of the OoO is a unified view across all edge and cloud resources and status of running workloads. Carriers desire a universal view that provides a global indicator of health, ongoing edge and cloud costs, and level of resource consumption.



Why do CSPs need an OoO?

With their multi-edge cloud, multi-partnership strategies, carriers today are stuck with siloed views that make it hard to understand how their overall edge strategy is performing, much less find ways to optimize their edge business. The OoO can add significant value by providing that unified view.

The OoO can help reduce the OpEx overhead and operational complexity of working with multiple edge clouds. With greenfield MNOs like DISH taking a multi-edge cloud approach for their RAN and 5G Core, using OoO to orchestrate internal 5G RAN, core, and other workloads could provide fast ROI.

In addition, the OoO can fulfill the role of a multicloud application platform that carriers can build services on top of. Carriers have intimated their hopes of building a higher-level cloud and edge container-as-a-service or infrastructure-as-a-service offering that is agnostic to and portable across different underlying clouds. At the very least, carriers believe they can provide customers with a one-stop marketplace of cloud and edge resources, allowing customers to pick and choose between providers based on instance types, prices, QoS, and SLA requirements.

Finally, having the OoO abstraction layer can provide the CSP with leverage against other edge platform providers (e.g., hyperscalers). The OoO can reduce friction as the CSP transitions their workloads across platforms and can act as a buffer against underlying changes to cloud APIs, hiding changes via its abstraction API. This increased agility and consistency allow the CSP to build value-added services instead of wasting time managing disparate edge clouds.

Follow-on Conversations around the OoO

We wrap this research brief by reiterating the value of the OoO, an abstraction layer for multiple edge clouds. The OoO is valuable given the multi-edge and multiple go-to-market approaches that CSPs are taking. By providing increased agility, reducing management complexity and cost, and improving visibility across multiple edge clouds, the OoO will become a critical layer in a telco's edge and cloud stack.

As for how such an OoO can be built and further discussions around details of its operations, you're welcome to reach out to us at **research@avidthink.com**, and stay tuned for a follow-up research brief that will delve deeper into the topic.

AvidThink, LLC
1900 Camden Ave
San Jose, California 95124 USA
avidthink.com

Rev A