2021 Telco Infrastructure Report

The Journey from NFVI and VMs to Containers and Public Clouds

RESEARCH BRIEF

Table of Contents

State of Telcos in 2021 1
- NFVI to Telco Cloud 1
- Infrastructure is Strategy 3

Key Telco Initiatives in 2021 3
- 5G Buildouts 4
- SD-WAN, SASE, and Network Edge Security 4
- Edge Computing Services — Network and On-premises 5
- Private Enterprise Networks 5
- Internet of Things (IoT) 6
- Cloudification of OSS and BSS 6
- Digitization of Customer Experience (CX) 7

Telco Infrastructure Observations 7
- Multi-platform Realities — VMs to Containers 7
- Ongoing Performance Tuning 8
- Converged and Composable Systems plus Diverse Edge Hardware 9
- Workload Placement and Orchestration 9

Role of Public Clouds 10
- Public Cloud for IT 10
- Public Cloud as NFVI 11
- Public Cloud and the Telco Edge 11
- Alternative to Public Clouds 12

Strategies for the Telco Cloud 12

Conclusion and Wrap-Up 13

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.

Hewlett Packard Enterprise

CHANGING THE WAY 5G TELCO CLOUDS ARE BUILT

Hewlett Packard Enterprise brings together a set of open, edge-to-cloud infrastructure solutions and an ecosystem of partners to help telcos accelerate transformation and realize the full potential of 5G.

Explore how

RIBBON IS READY
TO TAKE YOU TO THE CLOUD

As service providers migrate their real-time communication (RTC) workloads onto cloud native platforms, Ribbon can help.

Ribbon SBCs secure your cloud native RTC workloads and Ribbon's cloud native operations automate and optimize the deployment and management of your SBCs

Ribbon's cloud native operations:

- Enabled with automation tools packaged with SBCs
- Fully automate LCM for Ribbon workloads
- Enhance options for high availability deployment
- Simplify and de-risk software and configuration upgrades

INNOVATION YOU NEED

TECHNOLOGY YOU TRUST

PEOPLE YOU DEPEND ON

ribbon

rbbn.com

ekinops

Empowering Networks

Agile networking solutions for the digital enterprise demands

Routing

Firewall

SBC

SD-WAN

Virtualization

OneOS6

the middleware empowering your network

• Open • Simple • Future-Proof •

www.ekinops.com

2021 Telco Infrastructure Report

The Journey from NFVI and VMs to Containers and Public Clouds

State of Telcos in 2021

As we kick off 2021, the world continues to be mired in the throes of the pandemic, but the arrival of vaccines forms a beacon of hope. For telcos everywhere, the march to build out resilient, scalable connectivity continues. Carriers worldwide emerged as one of the heroes of the pandemic, enabling work-from-home (WFH) as businesses reconfigured and physical activities moved online. Communication providers powered lifelines and forged connections between friends and families as borders closed and governments curbed travel. Connectivity has become more critical than ever.

To better support enterprise WFH and work-from-anywhere (WFA) initiatives, telcos extended their software-defined wide-area networking (SD-WAN) rollouts to encompass home and mobile users. Many are now adopting secure access services edge (SASE) capabilities to provide added protection for their customers working from home. In the meantime, telcos continued fiber buildout and 5G rollouts. 5G schedules stayed on track with minimal delays. We saw concurrent interest in the edge from enterprises for both network-based and on-premises edge computing. In addition, the increased emphasis on mobile wireless networks drove companies to telcos to deploy private enterprise networks based on 4G LTE and 5G mobile technologies.

Telcos engaged in these initiatives realized the urgency of accelerating their ongoing internal efforts to digitize, automate and scale more of their processes. The evolution of underlying infrastructure platforms has kicked into the next gear. The original network functions virtualization (NFV) movement that spawned disaggregation, white box servers, and virtual machine (VM)-based NFV infrastructure (NFVI) has moved into containers and cloud-native frameworks. It is now rapidly entering the realm of telco clouds and public clouds.

The original network functions virtualization (NFV) movement that spawned disaggregation, white box servers, and virtual machine (VM)-based NFV infrastructure (NFVI) has moved into containers and cloud-native frameworks.

In this research report, we'll examine major telco initiatives and how telcos are looking to innovate their telco infrastructure to support them. We'll trace the NFV movement's evolution to today's adoption of public clouds by some telcos. Many of the telcos we've spoken with view infrastructure decisions as critical to their strategies and believe that making the right call can improve their capital efficiency, agility, and position in the market versus their competitors.

NFVI to Telco Cloud

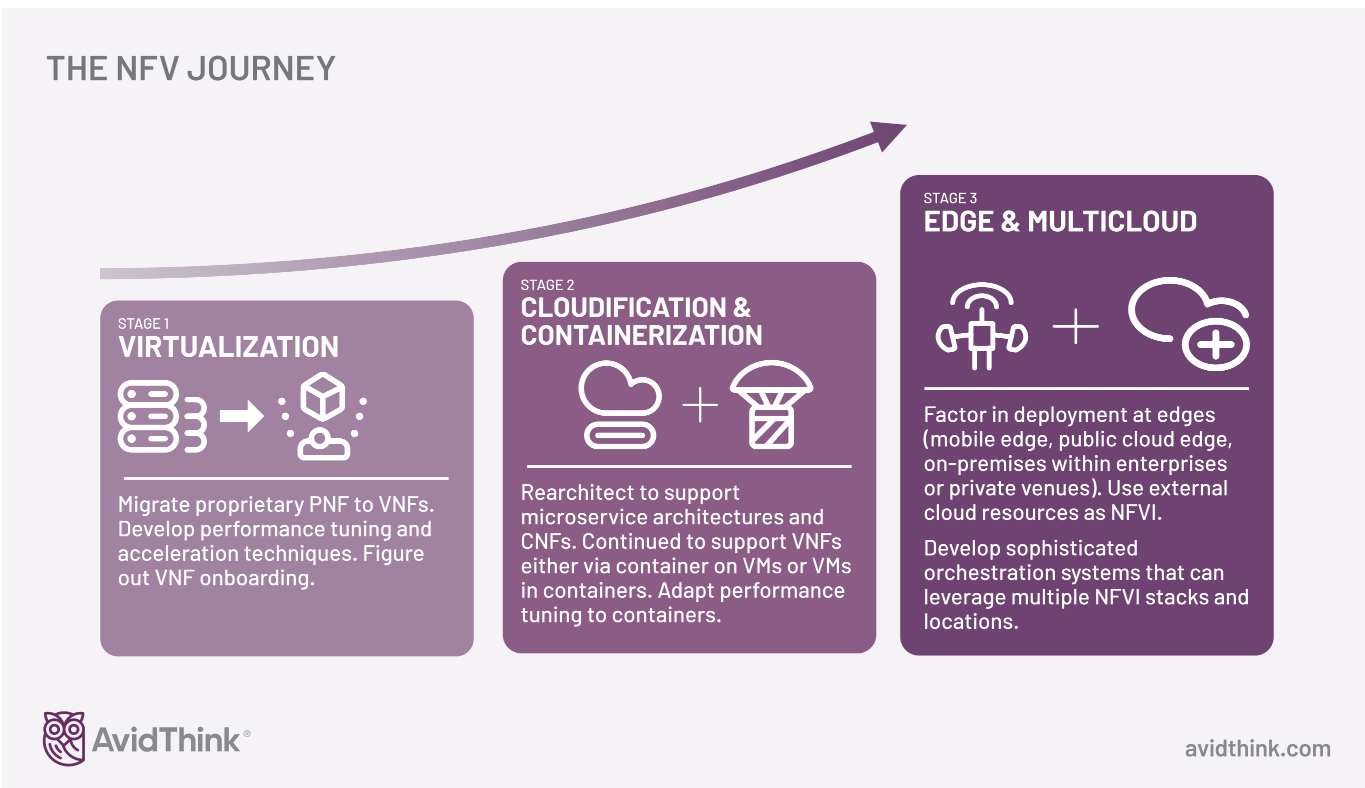
The NFV journey that started in 2012 focused on straight-up virtualization, mirroring efforts in enterprise IT data centers. It consisted of disaggregating and moving appliance-based, proprietary, physical network functions (PNFs) into virtual network functions (VNF) running on white box servers, aided by hypervisor technology (KVM and ESXi) and virtualized infrastructure managers (VIMs) like OpenStack and VMware vCloud NFV.

Telcos struggled to successfully make the PNF to VNF transition, somewhat reluctantly aided by incumbent network equipment providers (NEPs). NEPs took a while to embrace a move that reduced their market power. Challenges around VNF onboarding, incompatibilities in approaches between vendors, and intricacies of performance tweaks (managing huge memory pages, CPU pinning, DPDK software acceleration, and SR-IOV) gummed up the transition to NFV. Compounding this was a lack of consistency in pricing and business models.

A few years into the journey, cloud-native network functions (CNFs), Linux containers, and the Kubernetes orchestrator showed up, seducing many telcos and distracting them even before they had completed their NFV transitions. Today, telcos around the world tout unified visions of a "telco cloud" that spans legacy PNFs, VM-based VNFs, container-based and Kubernetes-managed CNFs on bare-metal machines, but few have delivered.

An increasing number of telcos have recognized that hyperscalers like Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and the Chinese hyperscalers (Baidu, Alibaba, and Tencent) are better than they are in managing and scaling up computing and storage infrastructure. Public clouds and hyperscaler-powered private cloud platforms are now under consideration as part of the extended telco cloud. We'll touch on this in more detail later. Going forward, the telco infrastructure decision is a more complicated choice than back in 2012.

For the remainder of the report, we'll use the words telco cloud and NFVI sometimes interchangeably. Some may argue that container-based platforms running on bare metal don't conform to the original NFV vision. Nomenclature aside, our report's focus is on the underlying hardware and software elements in the infrastructure layer that supports network functions and telco services. Any item that falls into those categories qualifies, including public cloud infrastructure offered as a service to the telcos. The one caveat for this year's report is that we'll cover software-defined networking (SDN) in our upcoming next-gen data center report, even though SDN is sometimes bundled in with NFVI.



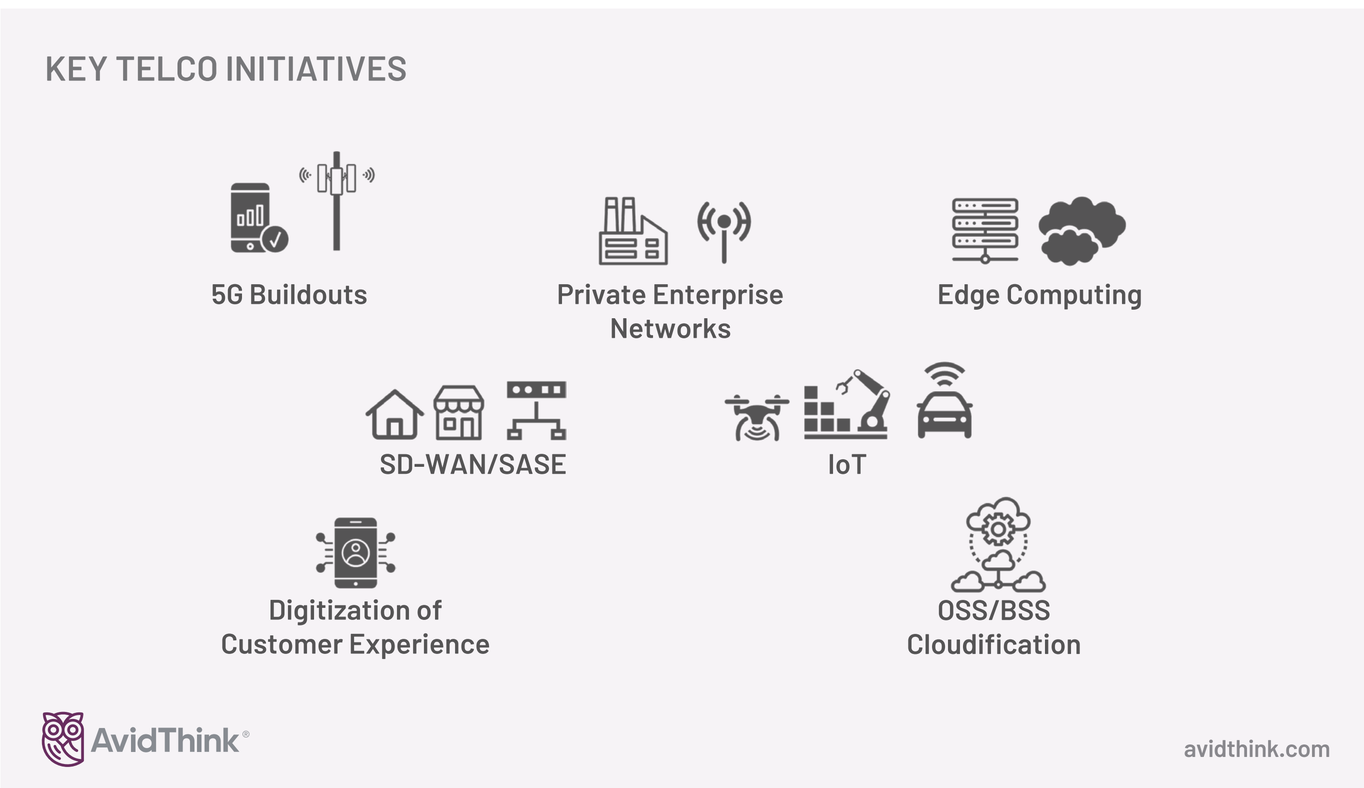
Infrastructure is Strategy

First, we need to examine key telco initiatives to understand how their requirements impact telco infrastructure choice. These initiatives center on telcos' ongoing transformation into next-gen communication service providers (CSPs) and digital service providers (DSPs). DSPs represent a potential evolution of today's communication companies, providing digital services over the communication links they sell today. These services include content, media, collaboration, security, mobility, and other applications that leverage networks and data. The DSP transformation is expected to unlock increased revenues and improved margins for today's telcos.

To support the rich set of services, telcos will need a robust, agile, and scalable infrastructure foundation. No longer are telcos faced with simple decisions of what, when, and how to virtualize and deploy VMs or embrace cloud-native and container-based offerings. Telcos are now faced with determining which parts of their infrastructure are core versus non-core, who to partner with, when and where to use public clouds, and whether and how much to extend their infrastructure to the edge and into customer premises. The telco infrastructure decision is now a strategic business decision.

Key Telco Initiatives in 2021

In reality, the telco infrastructure decision is not a single choice but a series of ongoing smaller decisions tied to critical business initiatives. For example, deciding the underlying infrastructure for a virtualized, disaggregated, and open RAN, or how to best host SD-WAN or SASE gateways at the telco edge. The telco's goal is to make those infrastructure decisions in line with an overarching telco cloud strategy. Let's examine the top initiatives that have come up in our conversations with telcos worldwide over the last twelve months.



5G Buildouts

For telcos that are mobile network operators (MNOs), the ongoing 5G buildout continues despite the global pandemic. Ericsson's bi-annual mobility report from November 2020 projected that 5G rollouts covered about 15 percent of the world's population (1 billion people) at the end of 2020. Meanwhile, the GSMA indicated that 5G activity recovered from early slowdowns during the pandemic and that 113 mobile operators have launched 5G networks in 48 countries. GSMA expects the buildout to continue with telcos spending USD 890 billion on 5G networks over the next five years. The intention is for 5G to serve not just mobile phones but also IoT devices, automobiles, and other autonomous vehicles, and even replace existing wireline connections to homes and businesses. Given the importance of 5G and the sizable share of Capex they will occupy, it's essential to understand the impact of 5G on telco cloud infrastructure needs.

Virtualized 4G and Containerized 5G Cores

The telcos we've talked to are starting 5G non-standalone (NSA) to standalone (SA) transitions. The vast majority of rollouts to date have used 5G New Radios (5G NR) in combination with existing 4G evolved packet cores (EPC). Still, to maximize the value of 5G, operators expect to aggressively migrate to 5G cores (5GC) over the next few years. Many 4G EPCs are now virtualized and increasingly adhere to control-user plane separation (CUPS) architectures, allowing the virtualized user plane of gateways to be deployed in different locations and on other hardware. With 5GCs, we're entering the realm of containerization. 5GCs employ cloud-native frameworks like micro-service architectures and require container-based deployment with Kubernetes orchestration. This, in turn, drives the underlying NFVI/telco cloud decisions.

Virtualized, Disaggregated, Cloud and Open RAN

Concurrently, the RAN side of the equation is evolving. MNOs are deploying virtualized RAN (vRAN) architectures, with increasing disaggregation of the 5G RAN into centralized units (CUs), distributed units (DUs), and radio units (RUs). To date, most of these vRAN solutions have been single-vendor solutions. With the Open RAN movement, there's a push from some MNOs for disaggregated multi-vendor solutions that conform to specifications from organizations like the ORAN Alliance. Meanwhile, the underlying NFVI that hosts these components may need hardware (HW) acceleration. For instance, the DU needs to perform real-time signal processing functions like transforms, filtering, and channel coding. Servers running the DU will likely need HW assistance in the form of digital signal processors (DSPs), field-programmable gate arrays (FPGAs), or graphics processing units (GPUs) to enable real-time handling of the RAN traffic.

Disaggregation of Cell-site Gateways and Routers

Another aspect of the RAN that is evolving underlying HW platforms is the cell-site gateway routers. These devices are migrating to white box platforms, running disaggregated software stacks that can provide the necessary performance while lowering the costs of these devices through general-purpose CPUs coupled with merchant silicon application-specific integrated circuits (ASICs) and hardware network timing and synchronization support. Organizations like the Open Compute Project (OCP) and Telco Infra Project (TIP) are defining the underlying open HW platforms for these devices through the disaggregated cell site gateway (DCSG) project.

Related to these efforts is the OCP's distributed disaggregated chassis (DDC) model, which is a disaggregated router initiative that couples ASIC high-speed switching platforms with CPU-based servers. The DDC targets a range of routers, from access routers to core routers. Recently, a Tier-1 provider, AT&T, announced the use of a DDC-based solution powered by disaggregated routing software from startup, DriveNets, on its core network.

SD-WAN, SASE, and Network Edge Security

The next hot initiative, particularly for wireline providers, is SD-WAN and SASE. With the recent pandemic and a mad scramble to connect and secure WFH employees, SD-WAN and SASE have been a boon for carrier revenue. Increasingly, the continued push towards 5G fixed wireless access (FWA) for businesses and homes, and mobile versions of SASE are also pulling MNOs into the SD-WAN/SASE arena.

Today, SD-WAN and SASE drive requirements at the branch and home office locations, pushing white-box CPEs and universal customer premise equipment (uCPE) deployments. uCPEs are usually small-form-factor servers with 4, 8, or 16 CPU cores that host a VIM or container run-time and are managed via a centralized orchestration system. There are ongoing efforts to improve the NFVI infrastructure on uCPEs to accommodate a mix of VMs and containers while maximizing performance. Carriers are also looking to use uCPEs as edge platforms that can host enterprise workloads.

Meanwhile, with the rise of SASE, telcos aim to place security gateways at the edge of their network that process incoming traffic over both wireline and mobile last-mile. These gateways can apply necessary policies to the incoming traffic and direct the packets accordingly. Packets may be sent to their final destinations at known services and safe locations, or for further scrubbing by cloud-security services from vendors like zScaler, Palo Alto Prisma, Cisco Umbrella. Even with SD-WAN, network gateways at points-of-presence (POPs) are necessary to enable seamless multi-link bonding or forward-error-correction (FEC). These gateways need to be hosted on NFVI at the network edge of their wireline and mobile networks.

Edge Computing Services — Network and On-premises

Whether you believe the edge computing market will be as big as, larger than, or a subset of today's **USD 300B+ cloud computing market** (as estimated by Gartner), it will be significant. As enterprises push to digitize more of their businesses, more data will be generated by devices and processes across the enterprise footprint. It's not always feasible or cost-effective to transport all that data for processing and analysis in regional public or private clouds. Data processing and enterprise workloads that cannot be moved into these clouds due to performance (latency, reliability), data residency, privacy, security, or cost reasons, are candidates for edge computing.

It's not always feasible or cost-effective to transport all that data for processing and analysis in regional public or private clouds.

There are two primary edges of relevance to telcos: the network edge, located where customers intersect with the telco network; and the on-premises edge, located at the enterprise (office, factory, farm, mine, sea port, and other locations). The network edge for a mobile network is at or near RAN aggregation points (mobile switching centers or MSCs), while for the wireline edge, it would be at cable headends or in central offices (CO).

It's clear why the network edge needs to be part of the telco infrastructure decision. As carriers move into value-add and over-the-top (OTT) services like streaming media, having storage and computing near subscribers can improve the quality of experience. Today, media content caches help reduce video start and load times, cloud gaming instances improve gaming latencies, and edge processing enables nascent augmented and virtual reality (AR/VR) applications. Telcos are pursuing mixed strategies for the network edge, with some doing it on their own and others partnering with hyperscalers.

For the on-premises edge, telcos also have mixed opinions. Some see the on-premises edge as a partnering opportunity with network equipment providers (NEPs) like Cisco, Ericsson, Huawei, Nokia, and Juniper, server OEMs like Dell, HPE, IBM, Lenovo, or global system integrators (SIs). Other telcos view the on-premises edge as extensions of their network edge or as a scaled-up uCPE presence. These telcos see the uCPE (or a scaled-up server stack) as a natural host for SD-WAN/SASE and managed security. Other services that telcos will want to host on these edge platforms include unified communications and collaboration (UCC) applications, IoT gateways, and even private 4G/5G enterprise network cores.

Private Enterprise Networks

Speaking of private enterprise networks, the visibility and hype around 5G has turned enterprise attention to considering 3GPP-defined 4G and 5G technologies for use in private locations. Various analysts have estimated the private 4G LTE/5G market will grow between 20-30% CAGR to between USD 5-10B by 2025. Cellular technologies can provide improved coverage, better spectrum efficiency, larger capacity handling, and smoother handover than existing WiFi-based enterprise wireless.

Private 4G/5G has less spectrum pollution and interference and provides the option for shared license schemes (like CBRS in the US) or private spectrum licenses (e.g., Germany and Japan). Proponents of private networks believe that the use of cellular technologies is inherently more secure than WiFi (it defaults to secure and has more stringent authentication and identity management). Whether as a complement to existing WiFi or an upgrade path for wired and industrial Ethernet deployments, enterprises have shown significant interest in private 4G LTE and 5G.

There are a few hundred installations of these private networks worldwide to date, most of which are based on 4G LTE. However, many of these have been deployed and managed by the NEPs — notably Ericsson, Huawei, and Nokia, not the telcos. Despite this, telcos are looking to leverage their expertise in managing large scale public macro networks to seek revenue opportunities in this fledgling market. The telco cloud's role in these rollouts is to host the private 4G EPC or 5GC. Architectural options for deployment include running the entire mobile core on-premises or running parts of the core at the network edge to reduce computing needs locally. And like in 5G rollouts in the public mobile network, virtualized, disaggregate and open RAN architectures are in play, with attendant implications for the NFVI.

Internet of Things (IoT)

IoT is an initiative that has dragged on for many years now, with many early entrants failing to monetize their investments. Carriers rolled out initiatives like NB-IoT (narrowband IoT) on their networks only to see little uptake. NTT **shut down their NB IoT network** last year, though it's expected that 5G deployments and uptake of 5G massive machine-type communications (mMTC) will bring back interest. In the meantime, the number of digitized and connected items continues to proliferate, and the promise of connected and autonomous vehicles, smart cities, smart agriculture technology, and drones mean that IoT will continue to gain importance.

For carriers and the telco cloud, IoT in industrial settings (IIoT) demands edge computing presence on-premises or at least near enterprise premises. And IoT devices on the mobile networks will require network edge processing of data to enable fast decision making and reduce uplink bandwidth costs.

Cloudification of OSS and BSS

The initiatives we've discussed are focused on external services. However, telcos are themselves going through a digital transformation process. Carriers worldwide are encumbered by legacy operation support systems (OSS) and business support systems (OSS). They need to revamp internal carrier processes and systems to become more agile and offer network and digital products as services (SaaS). It is remarkable how many carriers we interviewed were held back from offering next-gen network services not because they couldn't roll them out but that they had to wait for the billing subsystems to be reworked or the customer portals to be revamped.

It is remarkable how many carriers we interviewed were held back from offering next-gen network services not because they couldn't roll them out but that they had to wait for the billing subsystems to be reworked or the customer portals to be revamped.

When NFV first arrived, many telcos pointed to their existing virtualized IT stacks running their OSS and BSS and indicated a desire to run a unified NFVI stack for IT and production network functions. Telcos hoped that having standard hardware, VIM, and orchestration stacks could lower CapEx and OpEx through increased scale economies, consolidated purchasing, and streamlined processes. Unfortunately, few carriers have achieved those goals due to the complexity of implementing NFV for production networks. Instead, they've been left with separate islands of virtualized stacks — reaping benefits of disaggregation, commoditized hardware, virtualization per stack and vendor, but not achieving the unified NFVI vision.

In any case, OSS and BSS stacks are transitioning to cloud-native frameworks today, and telcos are contemplating migrating these into the public cloud (regional clouds and on-premises instantiations).

Digitization of Customer Experience (CX)

Telcos recognize the efforts software-as-a-service (SaaS), and cloud-based companies have made in engaging their customers directly, providing advanced self-service capabilities. Carriers recognize their poor reputation on the CX front and are looking to reinvent themselves to create superior digital service experience. Many are attempting to foster the same internal DNA as cloud companies have while they build out new customer portals and enterprise-facing services. Adopting cloud-centric micro-services architecture with containers, serverless frameworks, and continuous integration/continuous deployment (CI/CD) mindsets, carriers are attempting to achieve the agility and customer-centricity akin to that of an Amazon, Google, Microsoft, or SalesForce.

To power these experiences, telcos are moving towards leveraging public cloud platforms, which have several inherent benefits for building customer-facing web services.

Telco Infrastructure Observations

In this brief report, we won't be able to do justice to all major telco initiatives. We believe we've captured the top ones in the section above. Next, we'll examine how those initiatives will drive telco infrastructure trends over the next few years.

Multi-platform Realities — VMs to Containers

Across the various initiatives, the reality for telcos is that they will have to manage an infrastructure spectrum spanning legacy physical network functions, VMs and hypervisors, containers on VMs, and containers on bare metal machines. This is in addition to any public cloud platforms that telcos decide to leverage.

Unfortunately, we've observed a lack of unified platforms across IT functions and production networks. As we've indicated, we instead see islands of NFV platforms; for example, a virtualized EPC running on a vendor-certified VIM on vendor-approved white-box servers or vendor-supplied hyper-converged systems. Or an SD-WAN platform on a uCPE that's running a mix of containers on VMs and VNFs but has its own distinct vendor-supplied VIM with a container run-time and managed via Kubernetes. Unfortunately, because these stacks aren't converged, telcos don't have a single management plane that provides cross-domain visibility at the NFVI-level. The chance for reduced OpEx through common stacks has been limited.

Since the early days of NFV, telcos have had issues integrating multi-vendor VNFs. The Open Platform for NFV (OPNFV) project from the Linux Foundation was started as an integration initiative to help with VNF deployment on a range of NFVIs. More recently, OPNFV was merged with Cloud iNfrastructure Telco Taskforce (CNTT), initially hosted by the GSMA, to **form the new Anuket project** under Linux Foundation. The goal has expanded to include both VNFs and CNFs and provides hope for telcos looking to accelerate and ease multi-vendor deployment on cloud-native infrastructure.

> **We're also in an awkward transition point for containers and VMs. Some vendors are touting containers on VMs while others run VNFs in containers.**

Containers on VM or VM on Containers

We're also in an awkward transition point for containers and VMs. Some vendors are touting containers on VMs while others run VNFs in containers. Projects like **KubeVirt**, a Cloud Native Computing Foundation (CNCF)/Linux Foundation, are improving VNF-in-container management and performance but there's still work to be done to reach production-grade maturity. We also talked with telcos who run a VM infrastructure side-by-side with a container infrastructure as separate stacks. This fragmentation of the VIM layer has created new headaches but also opportunities.

While the open-source OpenStack (hosted by the **Open Infrastructure Foundation**) started out as the preferred VIM in the early days of NFV, management and complexity slowed its uptake. That allowed VMware to increase their penetration of the market with their proprietary vCloud NFV VIM. Recently, VMware has been aggressively transitioning into an all-encompassing VMware **Telco Cloud Platform** that uses VMware Tanzu for orchestration and includes their Telco Cloud Automation solution. VMware's unified framework now supports VNFs and CNFs and can run workloads in telco clouds in the core, at the edge, and even on public clouds. Emergent telcos like Dish Networks have bought into this vision, **betting their 5G buildout on VMware**.

In the meantime, IBM/Red Hat, as a major vendor alternative, is pushing its Kubernetes-powered **OpenShift** platform, unifying VNFs and CNFs via the KubeVirt path. Red Hat also provides an abstraction layer that allows telcos to run their workloads on both core clouds, edge clouds, and public clouds. In response to edge computing requirements, **Red Hat has recently updated their platform** to accommodate unique requirements like over-the-air updates, remote device mirroring, and intelligent rollback, in addition to supporting disconnected operations.

Kubernetes All the Way

Vendor solutions aside, telcos retain the option of using OpenStack to directly manage bare metal servers (via the **OpenShift Ironic** open-source project), which now appears to be the new favored use of OpenStack by telcos. Carriers can then use stock Kubernetes to manage their own container clusters on top of these bare-metal machines if they have the necessary staff and expertise. Many telcos tell us their strategy is to maintain enough VM capability for existing VNFs while expanding CNF hosting capacity (most likely on bare metal). For telcos who don't have the internal expertise to manage containers and who prefer smaller or more targeted vendors, there are a host of alternatives in the marketplace for managed Kubernetes, ranging from **Canonical** (providers of Ubuntu), to **Mirantis** to **Rancher** from **Suse**, **Platform 9**, and **Rafay Systems**.

Ongoing Performance Tuning

Telcos continue to struggle with the performance of both VNFs and CNFs. Telcos understand that their workloads are more I/O-intensive than enterprise or cloud applications, and with 5G and edge, real-time performance is an issue. Early NFV deployments of packet core components, including media gateways, session border controllers (SBCs), and messaging subsystems from vendors like **Ribbon Communications** and **Metaswitch** (now part of Microsoft), taught us that telco workloads are much more performance-sensitive than SaaS apps. We expect this to be exacerbated with open RAN and near-real-time RAN intelligent controllers (nRT-RIC), which manage the radio units in production networks. As open-source RAN projects like the Open Networking Foundation's SD-RAN gain popularity and RAN-specific applications are developed (analytics, radio, and spectrum optimization), the need for hardware-assisted NFVI will increase.

The NFV movement initially aimed to reduce the number of hardware SKUs. Unfortunately, with the arrival of 5G and the edge, we see greater hardware diversity.

In the past, infrastructure acceleration techniques like PCI passthrough, SR-IOV, software acceleration frameworks like DPDK, and managing operating system settings like huge pages, CPU pinning have proven crucial in maintaining VNF and CNF performance. Today, SmartNICs that provide hardware performance assistance for networking workloads are now part of most telco clouds. And the exposure of the underlying infrastructure via enhanced performance awareness (EPA) has been critical in workload placement. For more on this, check out the **AvidThink infrastructure acceleration report** covering the topic of infrastructure acceleration in greater detail.

Converged and Composable Systems plus Diverse Edge Hardware

The other complexity in telco NFVI is the increasing diversity of hardware. The NFV movement initially aimed to reduce the number of hardware SKUs, collapsing a diverse collection of proprietary hardware platforms into a set of commodity white-box servers powered by software-based VNFs. Telcos understood the need for big, medium, and small servers for different workloads — like uCPEs at customer premises versus 2RU servers in the telco core. Unfortunately, with the arrival of 5G and the edge, we see greater hardware diversity.

Edge locations have power, space, and other constraints and have driven new rugged, more diverse form factors. These systems include various hardware acceleration options like FPGAs, GPUs, and SmartNICs to provide the necessary networking performance for 5G RAN, AI/machine learning workloads in a limited footprint.

Further complicating telco hardware management is the arrival of other options for edge computing. These include hyper-converged infrastructure (HCI), which collapses computing, storage, and networking elements into Lego-type bricks of limited flavors, making them easier to order, deploy, scale, and replace. Likewise, there's a surge of interest in composable systems that can build new bare metal server configurations out of underlying pools of CPUs, storage, GPUs, FPGAs. These composable systems make more efficient use of underlying hardware and ensure the "composed" server is well-matched to its software workload.

Faced with such hardware diversity, telcos will need to figure out how to manage and monitor the platforms. This will occur either through vendor-specific HW management software or through the Distributed Management Task Force (DMTF) Redfish standards specifications. There are ongoing efforts to extend Redfish, like the Linux Foundation **Open Distributed Infrastructure Management** (ODIM) project, to scale better to large numbers of edge locations.

Regardless, we at AvidThink would like to remind telcos that the path to success in NFVI and telco clouds is simplicity. Take, for instance, Rakuten's ability to **limit their NFVI to four server configurations** despite running a virtualized cloud-native mobile network. We'll be watching to see if they can maintain the limited SKUs as they scale out to edge computing in the future.

Workload Placement and Orchestration

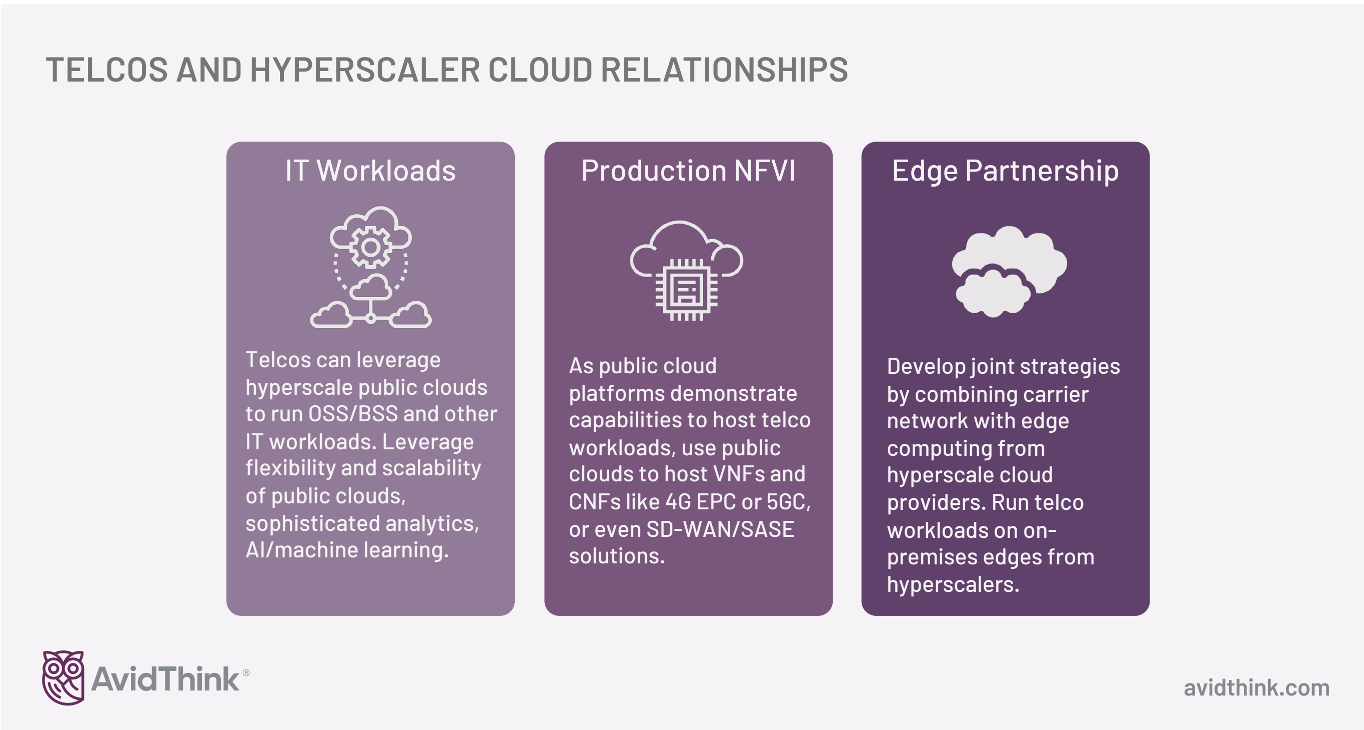
Given the multiple locations and diverse underlying hardware and software platforms within a telco cloud, we would be remiss in not highlighting the importance of intelligent workload placement and orchestration

Orchestration systems need to consider a host of parameters in making the placement decision. These include: VNF and CNF needs, service-level objectives (SLOs), real-time telemetry, resource availability, systems and business constraints. The goal will be to minimize underlying infrastructure cost, meet availability and reliability objectives, and optimize user experience. The orchestration problem itself is complicated enough within a data center, and much more when edge locations or customer premises equipment have to be taken into consideration. When public clouds are added, the problem enters the realm of extreme complexity.

Choices abound, with traditional orchestration vendors like Amdocs, Ericsson, NEC Netcracker, Nokia jumping into the fray, joined by IBM/Red Hat and VMware with their multi-cloud abstractions. Meanwhile, Google is touting its Anthos orchestration that can manage containers across private and public clouds, including those of its hyperscale competitors. Given the number of issues to examine, we'll reserve the topic of orchestration for another day. Suffice to say, telcos will need develop a robust orchestration strategy in parallel with building their telco cloud infrastructure.

Role of Public Clouds

In the last twelve months, we reached a turning point in hyperscaler and telco relationships. Previously, most telcos viewed hyperscalers with suspicion (or at least network operations teams did), concerned that hyperscalers would muscle into their territory and siphon revenue from them. Telco's experiences with OTT UCC and streaming media providers decimating the telcos' value-added services left them understandably wary. However, the telco stance on hyperscalers has now evolved.



Public Cloud for IT

Starting with back-office and IT functions, telcos have come to realize that hyperscalers are effective at running large-scale computing and storage platforms. By migrating back-office and customer-facing applications to the public cloud, telcos can benefit from a reduction in costs while improving agility, spurring innovation, and speeding up their time-to-market.

Notably, telco OSS and BSS are similar to enterprise IT workloads. Billing, enterprise workflow, customer service applications port smoothly into public clouds. As telcos look into exposing their network APIs (e.g., for network slicing or rapid provisioning) for consumption by third-party independent software vendors (ISVs), leveraging the rich set of services in public clouds enables telcos to set up API services rapidly.

Similarly, using the analytics and AI capabilities of public clouds allows telcos to tap into the enormous investment that hyperscalers have made on the AI/machine learning (AI/ML) front. Carriers have realized that leveraging public clouds to run analytics to help with targeted marketing and sales, churn reduction, or fraud detection can improve the telcos' top and bottom lines.

Public Cloud as NFVI

In the meantime, network operations teams at numerous telcos remain skeptical that public clouds can provide the necessary performance and capabilities for networking functions, especially for 5G networks. However, hyperscalers continue to invest in upgrading their platforms and have started partnering with telcos to co-develop appropriate NFVI for telco workloads. For instance, **AWS has partnered with Telefónica Germany and Ericsson** to host a mobile core for private 5G networks on AWS' Outposts infrastructure. And Microsoft Azure has similarly **partnered with Verizon** to use their edge platforms to power private enterprise networks.

We anticipate that edge platforms from the significant hyperscalers, like AWS Outposts, Microsoft Azure Stack Edge, will increasingly become viable as NFVI for on-premises deployments. This NFVI can power private networks, act as IoT gateways, and could fulfill the role of SD-WAN/SASE CPE.

Public Cloud and the Telco Edge

The hyperscalers have been active in edge deployments. AWS rolled out ten production AWS Wavelength sites for the mobile edge in the United States in 2020 (Verizon), one in Korea (SK Telecom) and another two in Japan (KDDI), with early trials in both UK and Germany (Vodafone). They recently added Telstra in Australia to the list of AWS Wavelength sites.

Microsoft Azure has announced Azure Edge Zones with Carrier with multiple telcos, including AT&T, Etisalat, NTT Comm, Proximus, Rogers, SK Tel, Telefonica, Telstra, Vodafone. However, none are in production as yet. Likewise, Google announced the Google Mobile Edge Cloud (GMEC) initiative with Docomo, Orange, Telus, KPN, Swisscom, Optus, Telenor, Wind Tre, MobileOne, Safaricom, MyRepublic, but it has yet to launch.

Meanwhile, the MNOs haven't been idle. América Móvil, KT Corp, Rogers, Telstra, Verizon, and Vodafone banded together to form the 5G Future Forum to hash out 5G and Multi-access Edge (MEC) interoperability specifications. And the GSMA pulled together its Operator Platform initiative to create a framework for network monetization with the edge.

Taking a different path, wireline provider Lumen (formerly CenturyLink) has chosen to roll out their edge platforms without hyperscaler involvement, deploying over 100 sites and partnering with virtualization vendors like VMware. And cable operator Cox has decided to invest in alternate edge provider Stackpath.

NFVI at the edge is also necessary to power telco services — like a distributed 5GC or 4G EPC, virtual broadband network gateway (vBNG) in a central office, or a virtualized cable modem termination system (vCMTS) at a cable headend.

In addition to meeting the need for enterprise edge computing initiatives, these edge offerings are important for IoT use cases, some of which require edge computing close to devices. And for use cases like connected or autonomous vehicles, having an edge platform tied to a 5G mobile network is critical.

Edge-as-a-platform versus Edge-as-a-service

NFVI at the edge serves two purposes. What we've described above is primarily edge-as-a-service. Like infrastructure-as-a-service in public clouds, it provides workload hosting for enterprises at the carrier edge. However, the same infrastructure can also be used to host 5GC/4G EPCs, or wireline termination services like virtual broadband network gateways (vBNG) or virtualized cable modem termination system (vCMTS). Beyond that, SD-WAN and SASE services can benefit from the lower latency and improved performance that gateways hosted in these edge locations can offer.

Alternative to Public Clouds

Beyond partnering with or utilizing public clouds from hyperscalers, telcos can consider alternative cloud platforms from companies like Oracle and IBM, who have been courting telcos with offers of hybrid clouds — partially on-premises, and partly in regional clouds. Telcos can also look into solutions from co-location providers like Equinix and Digital Realty. With its recent acquisition of Packet, Equinix has a bare metal offering, Equinix Metal, that's available in an increasing number of data center locations. Co-location providers may have the added advantage of being a peering point for multiple popular cloud and SaaS services, improving the quality of experience for telco customers using such services.

Another unique offering comes from the upstart mobile network operator Rakuten Mobile, whose parent is an e-commerce giant in Japan. Rakuten has packaged their in-house platform for running a 4G/5G mobile network and is offering it as a service — Rakuten Communications Platform (RCP). RCP includes an OSS solution from InnoEye, a startup that Rakuten acquired last year. Rakuten has partnered with Equinix to host the service and has signed up 15 customers for its platform.

Strategies for the Telco Cloud

Having laid out the critical telco initiatives and the current options available for building a telco cloud, we'll discuss practices from leading telcos we've spoken with, along with our recommendations.

For telcos who feel uncomfortable working with hyperscaler-powered public clouds, their options include taking a do-it-yourself approach of assembling a telco cloud from the constituent components or finding a global SI. The fundamental elements include disaggregated white box hardware, a VIM layer (or layers) that support containers and VMs, some networking fabric, and then management and orchestration capabilities. And as we've mentioned, hardware acceleration in the form of SmartNICs, or direct access to GPUS, FPGAs, may be required. Telcos will likely tap one of the major virtualization vendors, including VMware or IBM/Red Hat, for assistance or reach out to some of the other companies we mentioned earlier. This is a sizable undertaking and will take significant ongoing OpEx to maintain. For any supporting software services like databases, analytics, AI/ML, security functions, the telco will need to find in-house talent, or count on the SI, to research, deploy and maintain.

Most telcos will likely pursue hybrid approaches, using a mix of public cloud resources (or alternative providers), with local telco infrastructure from a major server vendor. Some server vendors, such as HPE, are attempting to reduce the NFVI workload on telcos by creating pre-defined and validate telco blueprints, which provide benefits of disaggregation while cutting integration costs. Likewise, VMware and IBM/Red Hat have extended their software to manage and orchestrate hybrid use cases. In hybrid deployments, some infrastructure resources are local while others are in the public hyperscale clouds.

For the hybrid situation, let's drill down one more level. We will break our analysis a couple of buckets: telco IT workloads, on-premises workloads, edge services and telco core functions.

- **IT workloads (OSS, BSS, customer-facing applications)**: Many telcos we've spoken with are expecting a natural migration to the public cloud for OSS, BSS, and other IT workloads. We do not see a strategic disadvantage to doing so and agree with their strategy. Their decisions here mirror those of enterprises who went down the same virtualization to public cloud journey many years ago. Major OSS and BSS vendors like Amdocs, Ericsson, NEC Netcracker, Nokia have partnered with the hyperscalers to ensure their solutions run in public clouds. We expect there will be some strategic or legacy workloads that remain in telco private data centers. However, there's a rich set of services available in these public clouds, from advanced databases to analytics to AI/ML to API management, and telcos should take advantage of them. In particular, for any projects involving upgrading the telco CX, we recommend taking the public cloud approach as SaaS application vendors have done. Telcos should use public cloud analytics, database, messaging, and other services, migrate to cloud-native development processes and frameworks to gain the agility, flexibility, and scalability of their OTT counterparts.

- **Enterprise on-premises workloads (uCPE/SD-WAN/SASE, Private 4G/5G, IIoT)**: Telcos tend to operate SD-WAN/SASE, managed security, and uCPE separately from their public cloud relationships. And historically, the hyperscalers have not focused on WAN offerings. We expect the uCPE and SD-WAN/SASE NFVI to remain under telco control for the immediate future and for them to depend on uCPE stack vendors for solutions. Whether the uCPE extends to host other workloads, we expect some telcos to go this route by adding more CPU cores, memory, and storage. The uCPE will eventually butt up against on-premises edge solutions. Still, we expect both the uCPE and on-premises edge to co-exist in the market due to diverse market needs. For hosting private network cores, telcos are concerned that using hyperscaler powered edge clouds like AWS Outposts or Microsoft Azure Stack Edge may disadvantage them strategically. Telcos are particularly wary of Microsoft Azure, which owns their mobile core stacks (from their acquisition of Affirmed Networks and Metaswitch). However, due to hyperscaler strength in application stacks like IoT, telcos are forced to collaborate with the hyperscalers. Some telcos are partnering with the NEPs, like Nokia, Ericsson, and Cisco, to host the 4G/5G cores, in hopes that these are more neutral or supportive. Server OEMs are also providing more comprehensive solutions on this front to offer alternatives to hyperscalers.
- **Network edge (edge computing, automotive IoT)**: As we've described, numerous MNOs are partnered with hyperscalers for mobile edge deployments. The wireline situation is less clear since the hyperscalers offer local edges that compete more directly with carriers. In general, we expect the underlying infrastructure for edge-as-a-service may eventually fall to the hyperscalers since many of these applications will be extensions of cloud services. There are other NFVI alternatives to power edge-as-a-service, but it's hard for us to imagine convincing enterprises to switch their development preferences from using public cloud solution stacks.
- **Telco core (carrier network functions)**: This is where we see the telcos hanging on to non-hyperscaler-powered solutions for some time (perhaps forever). Many telcos are not yet convinced that hyperscale cloud stacks are a fit for telco workloads. Or even if they are, they have concerns about the strategic implications of becoming fully dependent on hyperscale cloud infrastructure for all NFVI. This is where we see the OEM server vendors, along with the virtualization stack vendors, come into play. There is also a potential role for the co-location providers, who can assist in providing part of the infrastructure, perhaps even up to bare metal servers. The co-location providers are viewed to be more neutral than the hyperscalers.

In the meantime, the hyperscalers are pushing beyond offering their underlying clouds as NFVI. Microsoft Azure, leveraging their Affirmed and Metaswitch assets, is promoting Azure for Operators, a mobile core as-a-service product. Smaller tier-3 or even tier-2 telcos might be open to outsourcing their OSS, BSS, and even mobile core to a hyperscaler. MNOs can consider options like Rakuten Communications Platform as an alternative to Azure for Operators or other telco-as-a-service offerings that will likely hit the market this year.

However, outsourcing too many functions can reduce an MNO to differentiating only on marketing, sales, customer service, and spectrum holdings — all of which contribute little to technology intellectual property. Indeed, business models can be an area of innovation, as has been proven by Apple Music succeeding where others had failed, and speed of execution can differentiate. Nevertheless, telcos will have to balance the need for internal innovation versus leveraging public cloud benefits.

Conclusion and Wrap-Up

Telcos are living in interesting times. The NFVI decision, from when NFV was first adopted, has evolved dramatically. Today, it spans VMs, containers, and telco clouds to public clouds. And it covers telco core data centers, the network edge, and the enterprise premises. The plethora of choices today facing telcos is both exciting and mind-boggling.

Pick the right mix, and a telco can leverage the innovation of leading virtualization vendors, open-source efforts, and the hyperscalers to get to market faster, react quicker to customer demands, improve their customer experience, tap into new revenue opportunities, be more agile, and profitable.

Make the wrong choices, and the same telco could be burdened managing and scaling hardware and software infrastructure they don't have the skills or know-how to. The telco would burn unnecessary OpEx and CapEx trying to keep up, resulting in slow product launches and low profitability. Alternatively, the telco could become dependent on hyperscale cloud provider platforms by outsourcing too much, leaving little room to differentiate and innovate in the market, making them less competitive.

There's no simple answer. Each telco will need to evaluate their customer needs, their regional competitive environment, the available offerings from software vendors and hyperscalers in their geography. They will need to candidly and comprehensively assess their internal capabilities and culture as part of this decision. In the end, telco infrastructure is a crucial strategic decision and one that requires utmost contemplation and consideration.

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