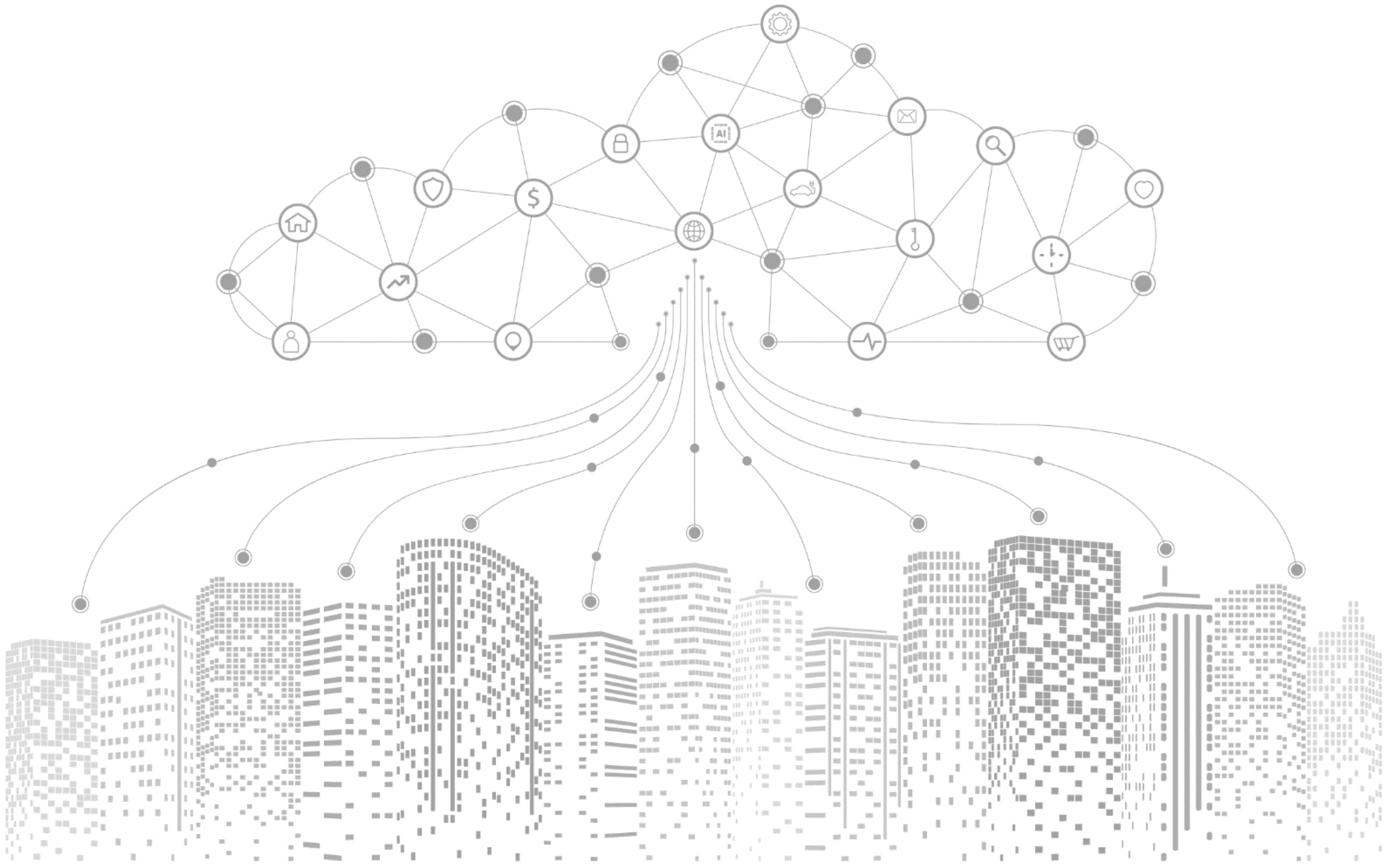
White Paper | December 2021

CHALLENGES MOVING THE 5G CORE TO THE PUBLIC CLOUD

Author: Dave Bolan | Research Director



Introduction

An extremely hot topic today in the world of 5G is whether or not to move the 5G Core to the Public Cloud. Two recent announcements, one by Dish Wireless and another by AT&T, set the industry abuzz, as the companies announced their intentions of running their 5G Cores in the Public Cloud with Amazon Web Services (AWS) and Microsoft Azure, respectively.

The big three Public Cloud Service Providers (SPs)—AWS, Microsoft Azure, and Google Cloud Platform (a.k.a. AAG collectively)—have been very vocal about convincing Telecom Communication Service Providers (CSPs) into moving their 5G Cores to the Public Cloud. The challenges involved in making this move are the focus of this whitepaper.

To support their case, or improve their ability to provide service, or expand their capabilities, AWS recently announced it would be a turnkey supplier to enterprises wanting private 5G networks and named its new service AWS Private 5G. Microsoft has acquired two 5G Core start-ups, Affirmed Networks and Metaswitch. In addition, all 5G Core vendors are making sure that their 5G Core software will run on the different Public Cloud platforms.

It sounds as if it is inevitable that the 5G Core of CSPs will run in the Public Cloud. As of November 30, 2021, Dell’Oro Group has counted 15 CSPs that have commercially deployed 5G Standalone (5G SA) networks for eMBB (enhanced Mobile Broadband) services, which requires 5G Core. None of these CSPs used the Public Cloud for their 5G Core. Rather, they went the traditional route, building their own private telecom networks, or, as known in the 5G world, the Telco Cloud.

With nearly a thousand CSPs in the world, we are obviously in the embryonic stage of the lifecycle of the 5G SA market. A lot is at stake, as billions of dollars will be invested in 5G Core. So which way should a CSP go, Public Cloud or Telco Cloud? Part of the answer depends on the target market of particular CSPs: eMBB, FWA (Fixed Wireless Access), Private Wireless Networks, or IIOT (Industrial Internet-of-Things). The rest of the answer depends on the management and engineering expertise a CSP may or may not have. One size will not fit all CSPs.

Since no CSP is yet up and running an eMBB 5G Core in the Public Cloud, what questions need to be answered before they decide to move to the Public Cloud? What are the challenges and considerations? We hope this whitepaper will give CSPs the information they need to make the move or not, while also informing Public Cloud SPs of the questions they need to respond to.

Table of Contents

5G Core Overview

Before discussing moving the 5G Core to the Public Cloud, let us look take a close look at the 5G Core. 3GPP specifications, TS23.501, identify 36 5G Core network functions. By comparison, 4G had about a dozen network functions, illustrating how much more complex 5G is than the previous generation (Figure 1).

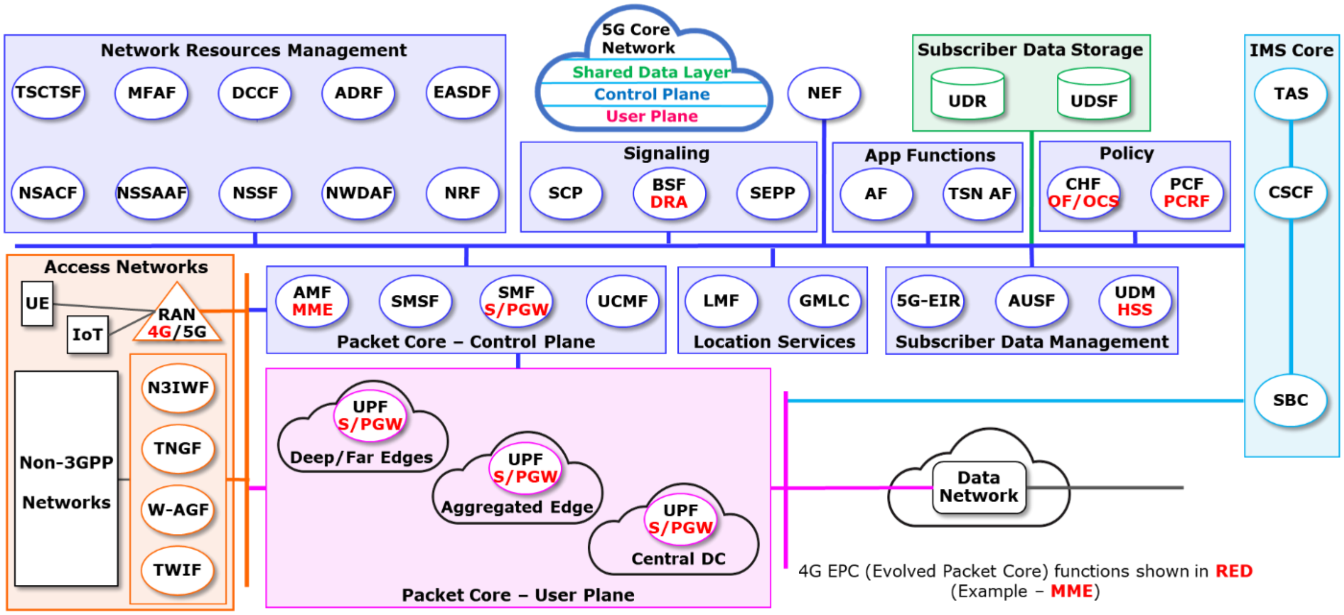


Figure 1: 5G core is more complex than 4G evolved packet core

The Network Resources Management component of the 5G Core introduces nine new network functions that include Network Slicing and Edge Computing. The Network Exposure Function (NEF) introduces a new concept in 5G, with Service-Based Architecture (SBA) rather than a Reference Based Architecture. 5G enables container-based cloud-native virtualized network functions (CNF), disaggregating the hardware from the software with a Shared Data Layer, Control Plane, and User Plane.

The 5G Core with SBA and container-based CNFs allows for the development of the Telco Cloud, which enables scaling the network to make it easier to meet demand, especially for new or temporary services. By separating the Control Plane from the User Plane, 5G makes Multi-access Edge Computing (MEC) much more feasible, as the User Plane can be distributed closer to end-user locations (Figure 2).

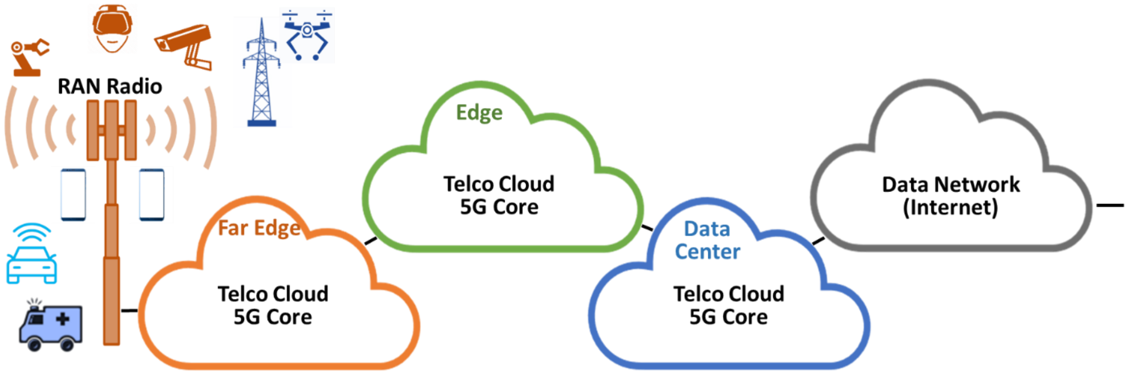


Figure 2: 5G Core facilitates Telco Cloud and MEC

The 5G Core network is also designed to manage non-3GPP networks for wireline networks, Wi-Fi networks, and non-cellular IIoT networks, making 5G Core a universal communications core for all types of traffic, real-time, near real-time, and non-real-time. As outlined in our whitepaper, *"How 5G Service Providers Can Capitalize on the Wireless Enterprise Market Opportunity,"* 5G CSPs have the opportunity to address many more enterprise applications than before (Figure 3).

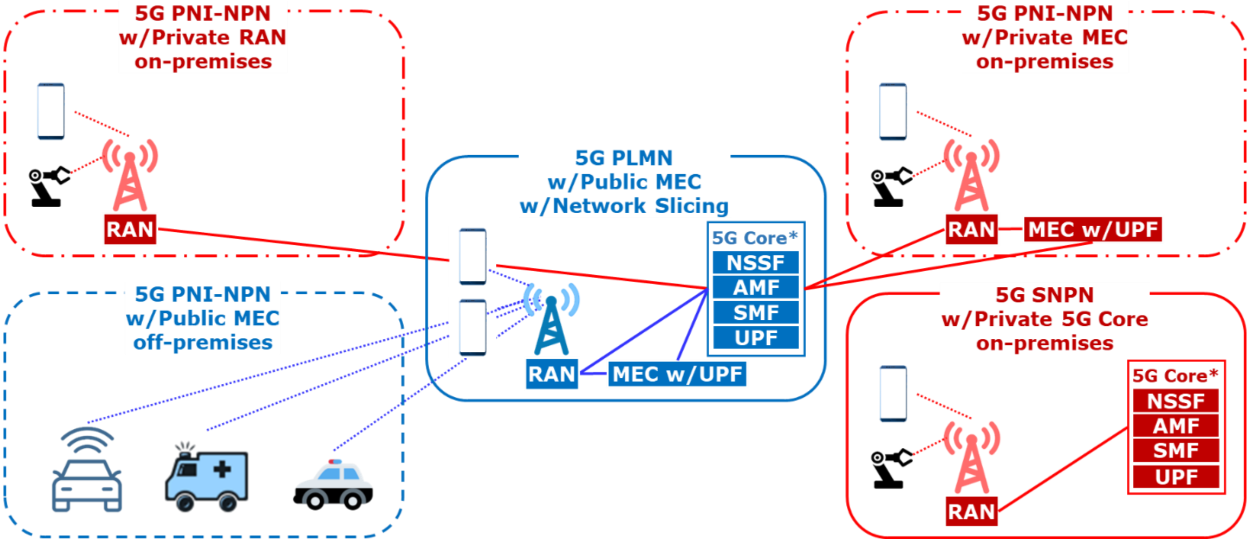


Figure 3: 5G Core allows 5G CSPs to address low-latency enterprise applications

The leading 5G Core vendors in revenue are Huawei and ZTE, the primary suppliers to the three Chinese CSPs, which began building out their 5G SA networks in 2020. Other significant 5G Core vendors are Ericsson and Nokia. These vendors are market leaders because of the breadth of their product offerings, deep experience, extensive knowledge, and comprehensive services. All four, plus Mavenir, could be considered E2E (end-to-end) vendors, in contrast to other 5G Core vendors that concentrate on certain parts of the 5G Core network.

Before 5G, traditional networks were built out by traditional E2E core vendors, which not only supplied the network functions but also the software on tightly integrated compute and storage infrastructure, usually x86-based. This system allowed the E2E suppliers to offer pre-integrated and pre-tested systems in a timely matter at a known performance level.

With the advent of NFV (Network Function Virtualization) late in the life of 4G networks, CSPs had the choice of using COTS (commercial off-the-shelf) infrastructure (a.k.a. NFV Infrastructure (NFVI), because NFV disaggregated the hardware from the software. NFV allowed a CSP to select an NFVI vendor independently of the software vendor and either integrate and test the software (network functions) itself or employ a system integrator.

SELECT 5G CORE VENDORS

- Affirmed Networks*
- Casa Systems
- Cisco
- ENEA
- Ericsson
- Huawei
- HPE
- Matrix Software
- Mavenir
- Metaswitch*
- NEC
- Nokia
- Oracle
- Samsung
- ZTE

**Microsoft companies*

Figure 4: 5G Core vendors that have announced orders in alphabetical order

In the 5G era, the 5G Core's container-based CNFs provide CSPs with new choices in NFVI vendors: Public Cloud vendors (a.k.a. hyperscalers) such as AWS, Microsoft Azure, and Google Cloud Platform (Figure 5).

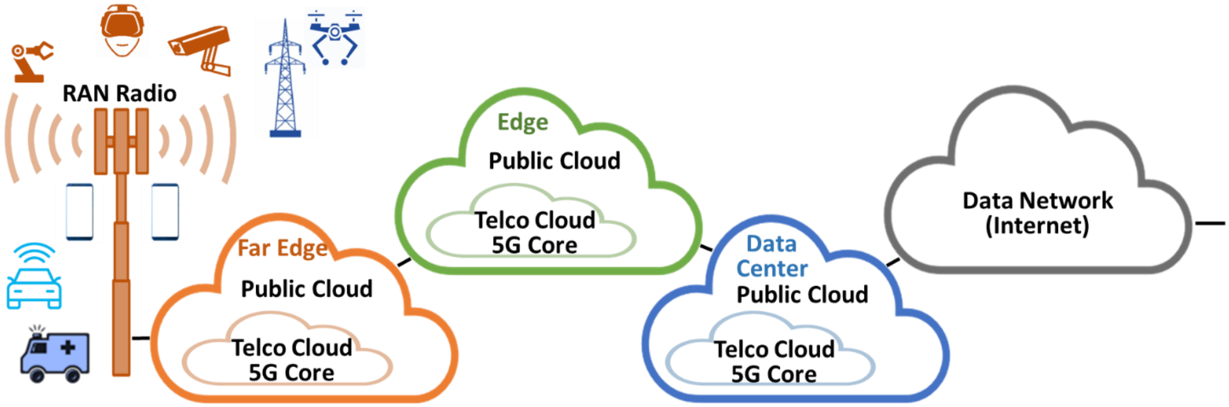


Figure 5: 5G Core architecture enables Public Cloud SPs to host 5G CSPs Telco Cloud networks

Public Cloud vendors are touting their advantages, such as robustness, resiliency, proven in-enterprise applications, security, and shift of the cost to an opex model versus a capex model, as outlined in our whitepaper, *"5G Service Providers and Public Cloud Service Providers: Ideal Partners."* For some CSPs, the advantages promoted by Public Cloud SPs are very appealing.

Challenges Moving the 5G Core to the Public Cloud

Complexity and Time-to-Market

As discussed above, 5G Core is much more complex than the 4G EPC, with three times the network functions, a new service-based architecture, network functions that are now container-based cloud-native, network slicing, edge computing with MEC, and emerging enterprise opportunities. Fifteen 5G SA networks are now up and running on 5G Core vendor NFVIs. These vendors have reported that implementing 5G Standalone networks is a very complex task in and of itself, without considering running the 5G Core on the Public Cloud. Moving to the Public Cloud would greatly increase the degree of complexity, as it requires integrating the MANO (Management and Network Orchestration), IMS Core (a.k.a. Voice Core), OSS/BSS (Operations Support/Business Support Systems), and Network Security to the mix.

Dish Wireless in the US was one of the first 5G CSPs to announce it would run its 5G network on the AWS Public Cloud. Figure 6 illustrates the numerous vendors the firm will call on, highlighting the sheer numbers of network functions and vendor software that need to be integrated and pre-tested with lab simulations, before beta trials and eventually going live. However, Dish Wireless, being a new CSP, avoids the additional complexity of interfacing with legacy 3G and 4G networks.

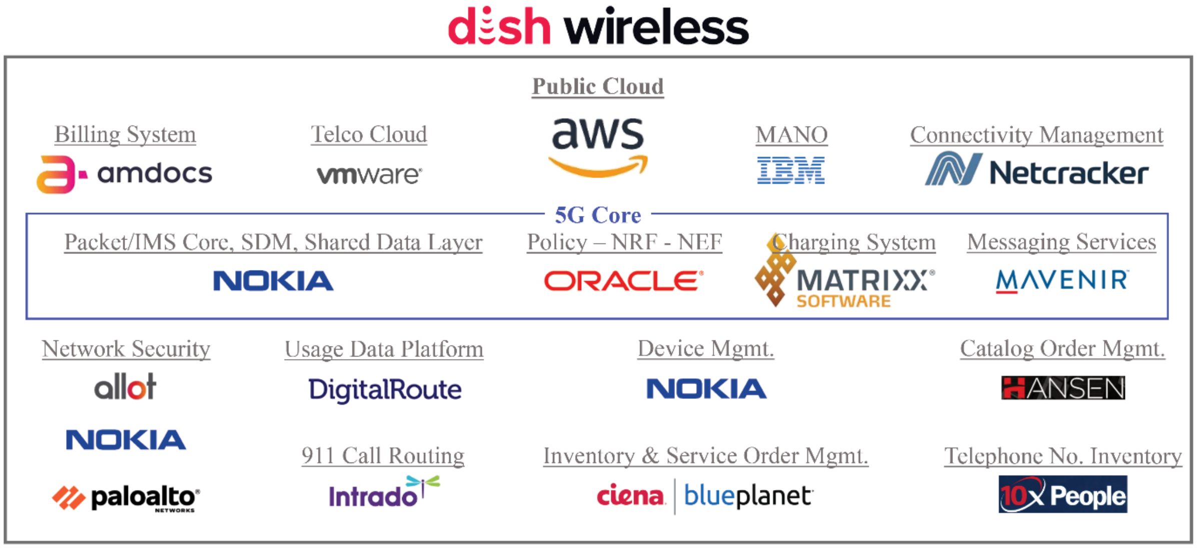


Figure 6: Dish Wireless 5G standalone vendor complexity

Dish Wireless's published timeline to market started with RFP/RFQs in July 2019, beta trails scheduled for 4Q 2021, and commercial launch in its first city, Las Vegas, Nevada, in 1Q 2022, a 33-month timeline. The firm did not publish its planning time for lab trials or field trials before it issued the RFPs/RFQs. It is safe to assume the total timeline is four to five years for the first city.

Also affecting the timeline is the buildout that might be required for Public Cloud SPs (in Dish's case, AWS) to have their Availability Zones (AZ) everywhere with the network interconnectivity required for distributed Telco Cloud architecture. These zones are mandatory to meet redundancy for network reliability, latency for edge computing, and in-country presence for data sovereignty.

Dish Wireless has a formidable list of functions and vendors for moving to the cloud, and 5G CSPs should be fully aware of what it takes to get their networks up and running on a Public Cloud, from lab simulation to beta trials to live commercial deployment at scale, before deciding to move to the Public Cloud.

Reliability and Performance

When considering moving to the Public Cloud, 5G CSPs need to understand the likely Reliability and Performance of the Public Cloud as compared to that of the Telco Cloud. Public Cloud SPs publish different network reliabilities for various parts of their networks, ranging from 99.5% to 99.99%. How does that add up for total network reliability?

Typically, Telco-grade network reliabilities are five nines (99.999%) of uptime for eMBB, and six nines (99.9999%) of uptime for demanding private wireless manufacturing applications. Can Public Cloud SPs meet Telco-grade network reliability requirements?

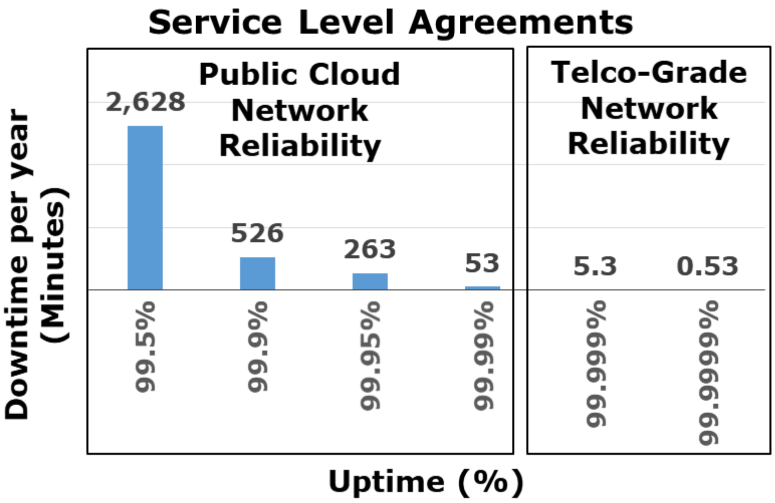


Figure 7: Public Cloud versus Telco-Grade network reliability

UPF (User Plane Function) throughput is critical to the performance of a 5G network. The technology in the NFVI (CPUs, GPUs, NPUs, ASICs, FPGAs, or software acceleration techniques) dictates how many servers and cores are required, which affects not only performance but cost. 5G CSPs must ascertain whether a given Public Cloud SP's NFVI UPF throughput is competitive with NFVI from 5G Core vendors or COTS vendors.

Another important consideration is how well the Public Cloud holds up to network stress scenarios, like busy hours, busy days (e.g., Mother's Day), and public disasters that can abnormally increase volume and create a signaling storm.

A further consideration arises in the unfortunate scenario of a network going down: what commitment does a Public Cloud SP make to the CSP to get its network back up and running? Do 5G CSPs have priority over other clients of the Public Cloud SP? The significance of a network going down can depend on its geographic coverage. For example, if a 5G CSP is multi-regional or multi-national, a Public Cloud outage could take down all regions or countries in which the CSP operates. In contrast, a CSP Telco Cloud is usually confined to one country, minimizing the impact of an outage to a smaller geographic area.

Geo-Redundancy

To meet network reliability requirements, geo-redundancy is required so there is no single point of failure that can take down the network. Because of data sovereignty rules, redundancy must be in the same country. And for edge computing, the geo-redundancy must be in the same vicinity as the primary service point to meet latency requirements. If the service point is on the premises of an enterprise, the redundancy must be on the same premises but in a separate location with different transport so there is no single point of failure.

When it comes to 5G Geo-Redundancy, the network cannot be vulnerable to a single point of failure and the backup AZ for redundancy must meet all the same network performance requirements, including latency, as the primary AZ (Figure 8).

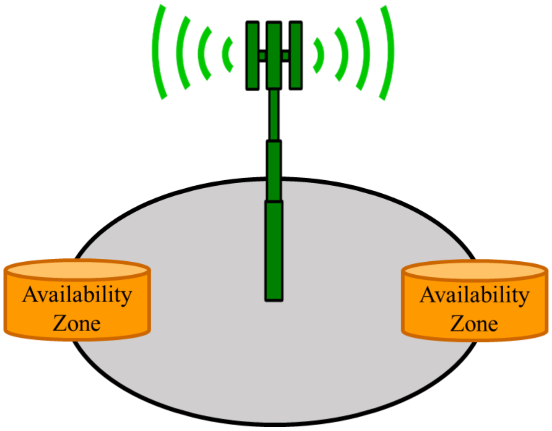


Figure 8: Public Cloud needs to meet geo-redundancy latency requirements

While it is out of the scope of this whitepaper, transport does become an important element with respect to redundancy, and which party is responsible for providing the transport, the Public Cloud SP or the 5G CSP?

Data Sovereignty and Security

Personal information, data privacy, and security are garnering a lot of attention, given that cybercrime is on the rise. Locking down a network with the best security available to protect privacy is obviously an important part of any network today.

When considering a Public Cloud SP, 5G CSPs must ask these questions about data sovereignty and security:

- Can a Public Cloud SP meet all government regulatory requirements?
- Can a Public Cloud SP keep all lawful intercept data in-country?
- Can a Public Cloud SP keep all user data in-country?
- In Europe, can a Public Cloud SP meet both GDPR (General Data Protection Regulation) and US Cloud Act requirements?

The answers to all these questions must be "yes" for both the primary AZ and the backup/redundant AZ locations.

Agility and Automation

Due to the expanded scope and capabilities that 5G offers, 5G networks introduce a level of complexity not seen previously in Telco networks. For 5G to deliver on its promises, the network has to be agile so as to introduce new features quickly, which requires automated network management.

When considering a Public Cloud SP, 5G CSPs must ask these questions about agility and automation:

- How do 5G CSPs manage their networks on a Public Cloud?
- How much network control do 5G CSPs have? Or to ask the question another way, what network controls are 5G CSPs giving up?
- How does a 5G CSP introduce new features and services on the Public Cloud? Does it take longer to introduce new features on the Public Cloud than in a Telco Cloud?
- How are 5G Core software upgrades made on the Public Cloud?
- Are Public Cloud SPs' network management planes for service and infrastructure independent of each other or centralized for more automated agility?

These questions need to be asked about all edge nodes as well. In the future, as 5G SA networks mature and use cases expand, edge nodes for low latency requirements might exist at every macro base station.

Vendor Lock-in

One of the driving forces towards a Telco Cloud with hardware and software disaggregation was to address vendor lock-in. CSPs do not want to be locked into a certain vendor, and disaggregation made it easier and more flexible to switch to another vendor's hardware or software when the need arose, for example, when costs got out of control. When considering a Public Cloud SP, 5G CSPs must ask:

- Are they trading one kind of vendor lock-in for another (i.e., locked into 5G Vendor "X" versus locked into Public Cloud Vendor "Y")?
- How hard would it be to switch from one Public Cloud vendor to another? 5G Core vendors would need a different flavor of their software for each Public Cloud SP.
- How much control would CSP have to give up; or stated another way, how much control would a CSP have to give to the Public Cloud SP?

In the future, when edge nodes are located at every macro basestation, 5G CSPs would have to consider the logistics and cost at the edge if they had to replace every location with a different Public Cloud SP.

Total Cost of Ownership (TCO)

Many factors must be considered in computing TCO. Sometimes the cost of management, training, and tools are overlooked in evaluations, especially when implementing innovative technology. In this case, 5G CSPs must learn the ins and outs of the Public Cloud.

- Determine the cost difference in network reliability between running the Telco Cloud versus the Public Cloud. Don't forget about lost revenues during downtime, including the loss of customers in response to the downtime.
- Factor in the cost of performance differences, especially UPF throughput.
- Consider the cost of redundancy, especially interconnectivity costs, under the two scenarios, especially for disaster recovery and a plethora of edge nodes coming in the future.
- Think about the cost difference in MANO between the two network choices.
- Figure in the cost of changing vendors under both scenarios. The 5G Core software will be different in one Public Cloud SP versus another: the software is not interchangeable. What is the cost difference?
- Look over the total life-cycle time horizon for unexpected costs, especially the need to upgrade hardware.
- Factor in bandwidth consumption and compute processing needs for the future, as they are only going to increase.

The Public Cloud SPs have many variables in their pricing models, and picking the right cloud pricing model is crucial for a CSP to not let costs get out of hand. The most expensive plan is on-demand compute and storage pricing. 5G CSPs would only pay for the compute and storage resources required for a given time interval (second, minute, hour). Such a plan might be advantageous if the demand for these resources is totally unpredictable. 5G CSPs must make sure that pricing includes the backup/redundancy to meet network reliability.

Compared with on-demand pricing, Public Cloud SPs offer discounts up to 50% to 75% if 5G CSPs make an upfront commitment to consuming a set amount of compute and storage resources over a one-year or three-year period. However, they do not offer a credit when the traffic dips below the committed level, and, when traffic goes above the committed level, they impose on-demand pricing. In case of a signaling storm or another event that requires excessive capacity defined by the Public Cloud SP, the Public Cloud SP might levy cost penalties, or, worse yet, might not be able to handle the additional traffic load.

The many pricing variables of Public Cloud SPs could make the cost of running 5G Core workloads over the Public Cloud unpredictable and uncontrollable, especially as compared with the costs of running 5G Core on the CSP's Telco Cloud. Calculating the true TCO is crucial to the decision of whether or not to move to the Public Cloud.

An Alternative for 5G CSPs: Integrating Public Cloud Services into the Telco Cloud

Instead of integrating the Telco Cloud into the Public Cloud, what about thinking the other way around? Why not integrate Public Cloud services into the Telco Cloud?

One reason that Public Cloud SPs are interested in the 5G Telco market is the opportunity to extend their services to enterprises using 5G connectivity for low latency requirements, as illustrated in Figure 3 on page 5. Public Cloud SPs want access to the customers and network edges of 5G CSPs to deliver real-time and near real-time communications to enterprises.

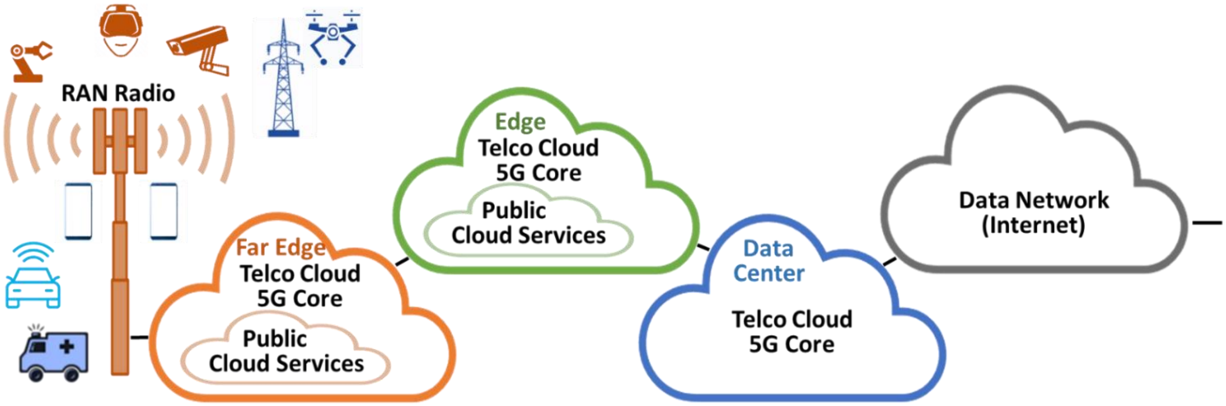


Figure 8: Integrating Public Cloud Services into the 5G CSPs Telco Cloud

Several 5G CSPs are already well down the path of integrating Public Cloud SPs services and their solutions for enterprises into their Telco Cloud from multiple Public Cloud SPs enabling a multi-cloud strategy to maximize the revenue potential from the Enterprise market.

Enterprises may already be using Public Cloud vendor "X" for their IT (Information Technology) workloads and might like to extend their OT (Operational Technology) workloads to the same Public Cloud vendor. A CSP Telco Cloud solution enables the CSP to address the preferences of an enterprise that might not otherwise be able to be addressed if the Telco Cloud is on a different Public Cloud vendors network than the preferred vendor of the enterprise. Running the 5G Core on the Public Cloud could potentially limit the addressable Enterprise market opportunity.

Integrating Public Cloud services into the Telco Cloud does not require a 5G CSP to start from scratch, building DYI (do it yourself) data centers. Data center options for a CSP for its 5G Core include using colocation data center vendors, not only for the central data center but also for edge locations with emerging colocation micro data centers. Tower companies, with their real-estate assets, are ideal for edge computing, have recognized this need and have begun partnering with colocation micro data center vendors.

Conclusion

5G CSPs should evaluate two major architectural solutions for deploying their 5G Cores:

1) Integrate the 5G CSP Telco Cloud into the Public Cloud.
2) Build the 5G CSP Telco Cloud with DYI data centers or colocation data centers.
    - Build a Telco Cloud with 5G Core vendor software and NFVI
    - Build a Telco Cloud with 5G Core vendors and COTS NFVI with a system integrator
    - Integrate Public Cloud services into the Telco Cloud

Each option has trade-offs, and careful consideration must be given to each scenario with respect to network and vendor complexity, time-to-market, reliability, performance, geo-redundancy, data sovereignty, security, agility, automation, vendor lock-in, and TCO.

5G CSPs must consider multiple factors:

- What market will be addressed: eMBB, FWA, IIOT, and/or Private Wireless? Which Enterprise vertical markets will be targeted, with Public MEC and/or Private MEC?
- Will more than one 5G Core be needed, e.g., one for eMBB, another for FWA, a third for enterprises (IIoT), public safety or mission-critical networks, and so on?
- What technical and management expertise does the 5G CSP have?
- What is the 5G CSP's financial condition? How much capital does it have access to? Is an opex or a capex strategy model better?

Answers to these questions will be different for each 5G CSP and will help determine the best alternative: build its own 5G Telco Cloud or move the 5G Core to the Public Cloud.

About Author

Dave Bolan joined Dell'Oro Group in 2017 and is currently responsible for the Mobile Core Network market research, as well as our Advanced Research report for Multi-Access Edge Computing market research programs. He previously covered the Carrier IP Telephony, and Wireless Packet Core markets for Dell'Oro Group. Mr. Bolan has written articles in industry media such as RCR Wireless. Mr. Bolan's research and analysis has been widely cited in leading trade and business publications. Mr. Bolan is a frequent speaker at industry conferences and events, including CTIA Wireless conference, NetEvents Global Summit, and SDN NFV World Congress.

Email: dave@delloro.com

About Dell'Oro Group

Founded in 1995 with headquarters in the heart of Silicon Valley, Dell'Oro Group is an independent market research firm that specializes in strategic competitive analysis in the telecommunications, networks, and data center IT markets. Our firm provides world-class market information with in-depth quantitative data and qualitative analysis to facilitate critical, fact-based business decisions. Visit us at www.delloro.com.

About Dell'Oro Group Research

To effectively make strategic decisions about the future of your firm, you need more than a qualitative discussion – you also need data that accurately shows the direction of market movement. As such, Dell'Oro Group provides detailed quantitative information on revenues, port and/or unit shipments, and average selling prices – in-depth market information to enable you to keep abreast of current market conditions and take advantage of future market trends. Visit us at www.delloro.com/market-research.

Dell'Oro Group
230 Redwood Shore Parkway
Redwood City, CA 94605 USA
Tel: +1 650.622.9400
Email: dgsales@delloro.com
www.delloro.com