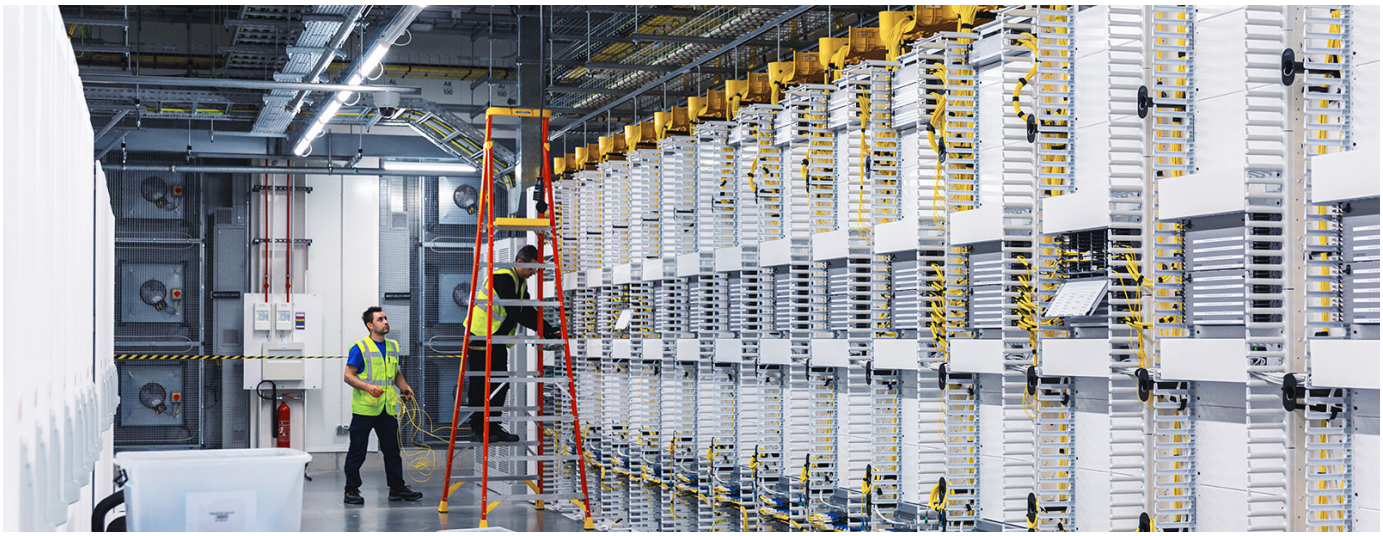
Moving 5G network functions to a hybrid cloud environment



Introduction

According to analysts, the telecommunications industry will spend over $25 billion on 5G network equipment by 2022. This level of investment intensity represents an inflection point for the industry and an ideal opportunity for mobile network operators (MNOs) to consider the significant financial benefits of shifting network investments to the cloud. Still, many MNOs are challenged by the questions of how to transition 5G network functions strategically and efficiently to a cloud ecosystem that includes multiple options across public, private, and hybrid cloud platforms.

While the 3rd Generation Partnership Project (3GPP) has defined 5G network functions, the deployment of those functions has yet to be fully implemented by MNOs and their vendors. Hosting network functions on a private telecom cloud, hyper-scale public cloud, or a hybrid combination of the two are all equally valid options that carry unique benefits and challenges.

The new generation of 5G applications requires a network architecture that supports flexibility, agility, openness, and scale. These characteristics are ideally suited for the cloud, particularly public clouds designed to speed up service deployment and significantly reduce costs. Yet, most 5G services today run on private cloud networks managed by telecom operators. Therefore, an attractive path forward for 5G is to create a hybrid cloud environment that leverages the best capabilities of the public and private cloud to host network workloads and applications.

As the transition to hybrid environments grows, we recommend looking beyond the traditional total cost of ownership model for measuring return on investment and adopting a total value of ownership approach for better savings and faster time to market, ensuring a more value-driven network that enables new revenue-generating opportunities.

This white paper explores the drivers of 5G investments in the cloud, how MNOs can strategically move 5G network functions to a hybrid cloud environment, and how operators can leverage a hybrid cloud architecture to drive innovation and create new sources of revenue generation.

Cloud market drivers

The road to the telecommunications cloud began a decade ago with the industry's adoption of network and hardware virtualization. Virtualization and, later, Network Functions Virtualization infrastructure (NFVi) provided the first wave of significant reductions in capital expenditures (CapEx). Achieving these CapEx reductions meant moving infrastructure investments from proprietary, hardware-based appliances to open, software-based network functions running on standards-based, virtualized hardware.

Cloudification, the idea of converting and deploying network functions as purely software-based applications to use cloud computing, is the next logical step in this network transformation. It unlocks additional opportunities for reducing CapEx by eliminating the need to manage and maintain hardware while also improving utilization, redundancy, and elasticity. Many telecom operators are already experiencing the savings from moving their IT workloads to the cloud and are now looking to other parts of their network to achieve similar levels of cost reduction.

The timing of cloud investments is not coincidental. 5G applications require much higher flexibility and performance at a significantly larger scale than 4G services. As MNOs weigh their investments in 5G, several key market factors are driving their future investments toward the cloud.

Operational efficiency
MNOs need to maximize their 5G investment dollars, particularly as they wait for 5G use cases to expand. In the short term, this means focusing on cloud technologies that can help them operate the network services they have today more efficiently and test early use case opportunities such as private mobile networks for enterprises.

Toolsets available in the public cloud provide new capabilities to MNOs, including automation and enhanced network intelligence. In the case of private mobile networks, deploying these services in a public cloud environment can significantly reduce time to market, risk, and cost while still meeting the critical requirements for ultra-low latency, high performance, and mobility.

Scalability
Traditionally, there has been a hard cost associated with scaling networks due to reliance on physical resources. However, with a diverse set of 5G applications and use cases on the horizon, the operator's network must be inherently scalable. Combining public cloud infrastructure and the cloud-native architecture of 5G workloads can enable operators to build a network that can scale on demand.

Mobile call roaming, for example, is costly to subscribers because of the complexity of network interworking connections, a cost that rises as mobile roaming services increase. We predict this level of complexity to grow with 5G deployments as the

> "Adding capacity and scale is no longer a multi-year CapEx planning and purchasing cycle."

number of roaming users and devices increases. Cloud interworking can significantly reduce the cost and complexity of mobile roaming, allowing MNOs to scale more efficiently. In addition, adding capacity and scale is no longer a multi-year CapEx planning and purchasing cycle; instead, it is dynamic and elastic with the on-demand increase or decrease in network capacity.

Technology in harmony

The telecommunications industry may well take many years to fully move to 5G. Thus, it is reasonable to expect that 5G and 4G (and even 3G) technology will co-exist in the operator's network, both in their private network and in the cloud. As such, MNOs must manage these systems concurrently, consistently, and ideally through a single management and operations interface.

The move to the cloud is more than a CapEx consideration. In fact, it is driven equally by the need for efficient scalability, simpler network management, and service acceleration. If we accept that the cloud is the future, the question becomes, how do we get there efficiently? What should be moving into the public cloud today, and what should move there tomorrow? We examine these questions in the following sections on cloud migration.

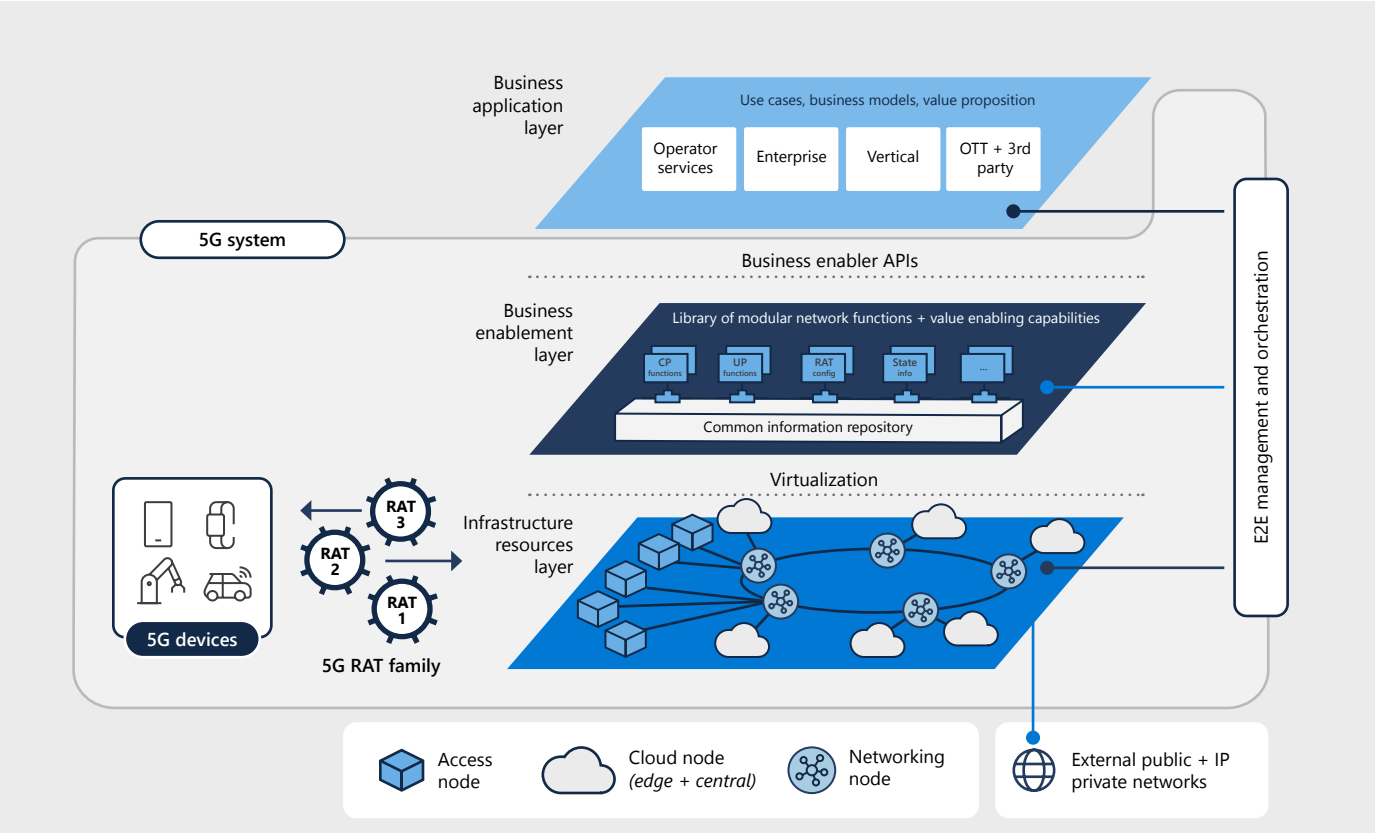


Figure 1: NGMN 5G architecture

Migrating network functions to the cloud

Let's begin by looking at 5G network architecture. Today, it consists of a mobile core, radio access network (RAN), edge infrastructure, transport, and services to support various specialized applications. For this paper, we'll focus primarily on the mobile core, as this is where the majority of 5G network functions reside. However, even from this simplified perspective, the 5G network is already a complex ecosystem of multi-vendor solutions (*see Figure 1*).

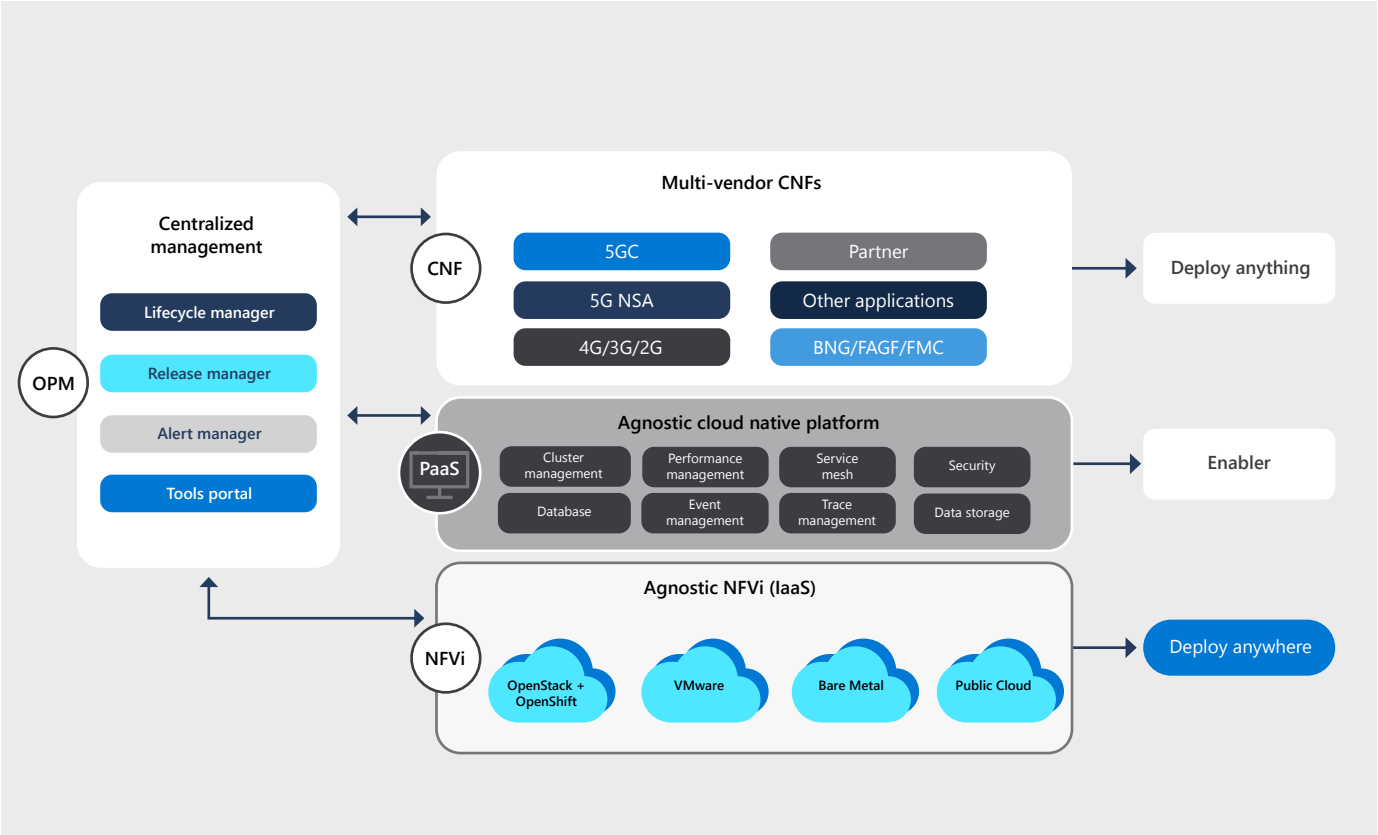


Figure 2: Multi-vendor 5G packet core architecture

As shown in *Figure 2*, four distinct layers make up the core of 5G networks.

- The network functions layer featuring the cloud-native 5G network functions (CNFs)
- The Platform-as-a-Service (PaaS) layer, including various cloud and container technologies
- The NFVi layer of virtualized hardware resources from multiple vendors
- The network operations layer

These layers can eventually shift into a hosted public cloud environment for additional cost/feature benefits. However, Microsoft recommends a phased migration to the public cloud that begins with short-term, high-yield returns centered on CapEx reduction. As MNOs turn their attention to the *value* portion of their network (i.e., 5G services), operational processes can also move into the public cloud to leverage enriched capabilities, such as automation, NoSQL database technologies, and advanced network telemetry analysis.

The implication is that 5G networks will initially—and in some cases indefinitely—co-exist in a hybrid deployment state between premises-based and public cloud environments. What this hybrid state looks like and how it evolves are covered in the next section.

Understanding private, public, and hybrid cloud models

The principal driver for moving 5G network functions into the public cloud is cost reduction. This is a strategically sound move as the cost savings of the cloud can help fund future investments in 5G revenue generation.

There may be unique requirements whereby an MNO would choose to keep some network functions on-premises. For example, MNOs may:

- Need to first recoup their investments in NFVi infrastructure
- Have security or latency requirements that are well-suited to a private cloud environment
- Need to meet client or regulatory compliance requirements

Keeping the above considerations in mind, Microsoft has developed a multi-phase hybrid cloud migration approach.

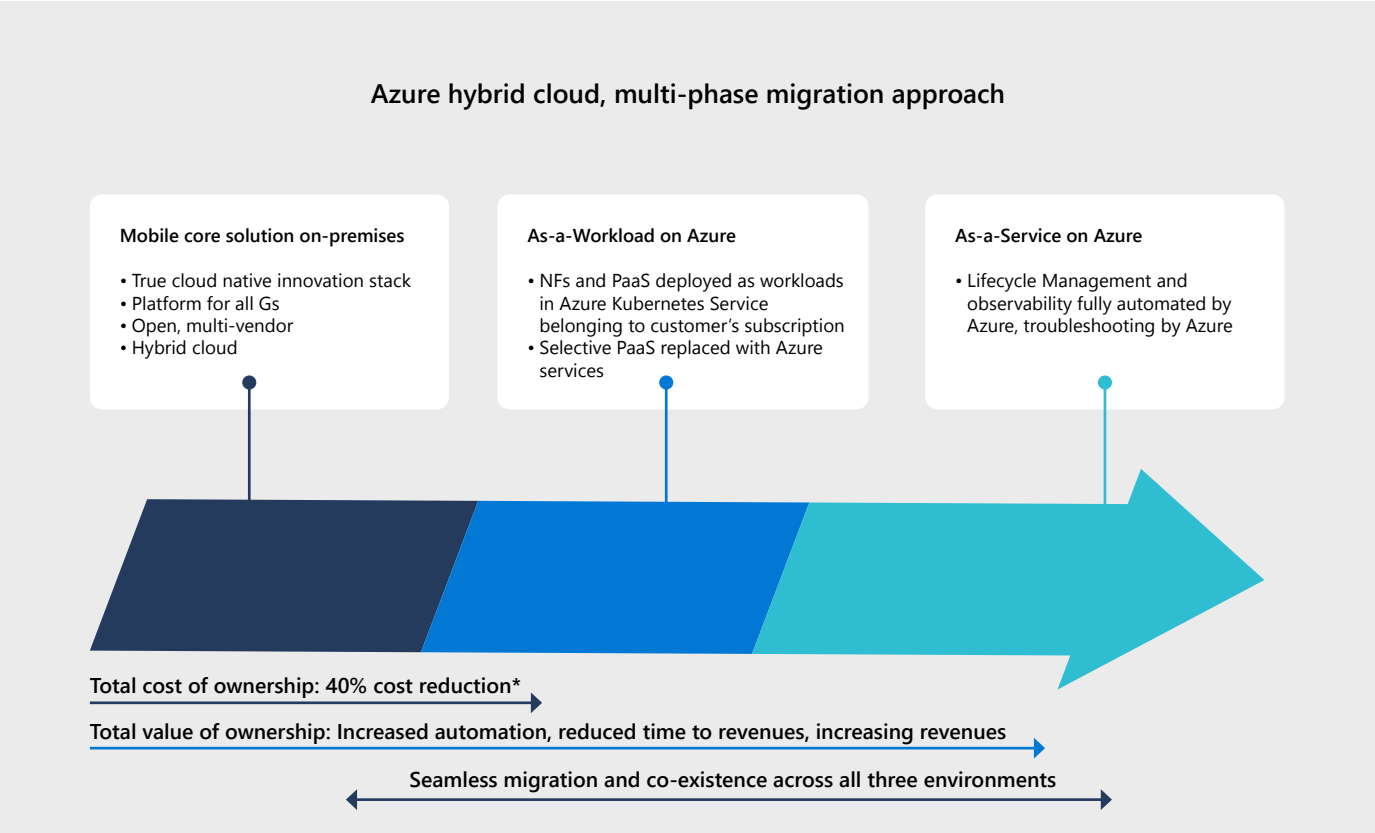


Figure 3: Microsoft's hybrid cloud, multi-phase migration approach

In the first phase, an operator runs its 5G workloads on a cloud-native 5G mobile core solution, using a private on-premises NFVi infrastructure. A private cloud deployment in this phase already provides operators with a host of cloud-native benefits, such as:

- Seamless operation across 3G/4G/5G services
- An open, multi-vendor interoperable platform that leverages open-source innovation to support a vast number of partner solutions
- A private NFVi platform capable of running applications, services, and network functions

In the second phase, 5G control plane workloads shift from the private cloud to Microsoft's **Azure for Operators** public cloud environment. Azure hosts the CNFs and the PaaS layer components using Azure Kubernetes Services. At this point, the "hardware" platform now consists of premises-based private NFVi resources, resources hosted in the Azure public cloud, and Azure Edge locations. In addition, Azure Marketplace tools enrich these services and, in some cases, replace third-party PaaS components.

In phase three, the entire 5G mobile core is deployed and operated as a service in the hosted public cloud environment. Lifecycle management of the CNFs, observability, and telemetry are fully managed and automated by the public cloud provider. In this case, Microsoft Azure Resource Manager (ARM) and Azure Arc assume responsibility for the operations layer.

Successfully deploying cloud network functions

The 5G transformation evolution will likely extend across several years. This time horizon reveals the necessity of managing hybrid cloud environments effectively. Any network transformation must also minimize disruption to existing services, which means ensuring support for legacy services and leveraging existing network operator skillsets as much as possible. Mitigating these risks to smooth service delivery opens the door to significant benefits.

Hybrid cloud economics
Telecommunications operators have significantly different operating models from other industries that have achieved benefits with the cloud. Therefore, a fair question is, "what level of cost reduction may MNOs expect by migrating to a hybrid cloud environment?"

Microsoft analyzed a global tier-one operator who had recently completed phase one of moving 5G workloads into the public cloud to answer this question. The operator in question had over 30 million smartphone subscribers, around 15 million each for high-bandwidth and low-bandwidth IoT users.

Before the transition, the operator's annual operating expenses (OpEx) consisted of servers, switches, NFVi software licenses, power, cooling, and storage. Additionally, the operator incurred depreciation and amortization as non-cash expenses for these expenditures on its income statement. After moving 5G workloads to the cloud, annual OpEx consists of Azure IaaS and Azure Egress, which includes the compute infrastructure number of virtual machines required to host the workloads and the cost of connecting network traffic through a private connection via **Azure ExpressRoute.**

> "Any network transformation must also minimize disruption to existing services, which means ensuring support for legacy services and leveraging existing network operator skillsets as much as possible."

Tier 1 operator TCO
On-premises costs vs. Azure pricing

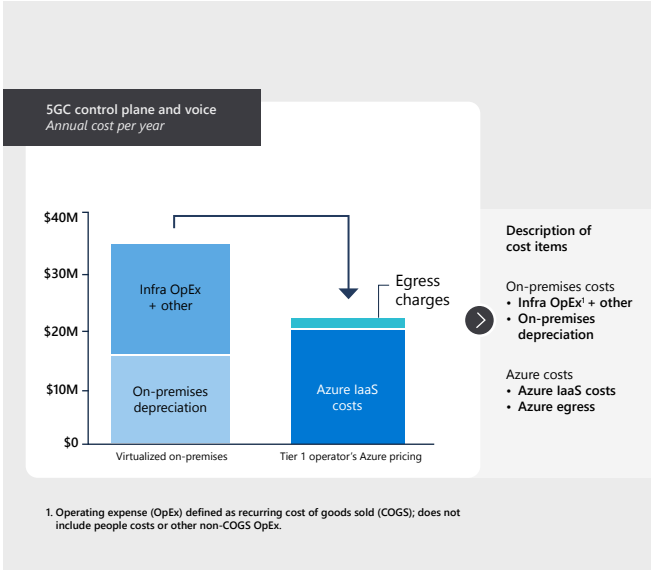


Figure 4

Tier 1 operator
On-premises vs. Azure resource requirements

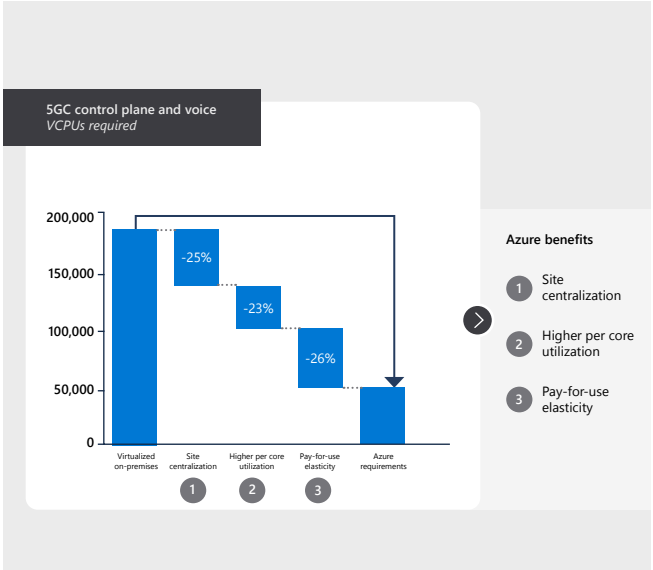


Figure 5

The result of our analysis is shown in *Figure 4* above. As we can see, the operator saved nearly 40% in annual operating expenses. This represents a reduction to the operator's TCO of approximately one-third per annum.

Cost reduction factors

By moving 5G workloads into the Azure cloud, the operator reduced their annual costs by a third. This cost reduction was the result of three factors: site centralization, higher per core utilization, and pay-for-use elasticity.

Site centralization

Site centralization reduces the number of resources needed for operations. For example, in our analysis, the operator was able to reduce deployment sites from ten to three by leveraging the public cloud's built-in geo-redundancy. Similarly, in a private NFVi architecture, control plane resources are often co-located near the user plane at regional sites. This decentralized approach results in significant cost overhead. In a public cloud deployment, a centralized control plane can support multiple distributed user plane instances.

CPU utilization

The public cloud also enables higher resource utilization, which provides a cost advantage in capacity planning. For example, instead of running servers at 50% capacity to account for peak utilization, the operators' model could assume a conservative peak capacity of 70%.

Pay-for-use scalability

By moving to the public cloud, the operator can more effectively address peak demand due to the public cloud's natural elasticity, which scales up and down as needed on a pay-per-use basis. Since resource demand isn't uniform throughout the day, cloud-native network functions can scale down during off-peak hours, which reduces the incurred cost. This isn't possible in on-premises deployments, which must be scaled to support a future peak busy hour capacity at all times. For example, before cloud migration, it is necessary for operators to plan out resource capacity a year in advance to allow for enough time to procure, install, and deploy resources. After moving to the public cloud, however, resources can be added (or reduced) as needed. This dynamic scaling of the network means operators only pay for actual resource consumption instead of buying excess capacity at the beginning of the year that they may or may not need and based on a demand forecast of limited accuracy.

Best practices for deploying cloud network functions

We introduced the 5G mobile core framework for a hybrid cloud, multi-stage migration in previous sections. In this section, we will go beyond this framework by examining best practices for deploying cloud network functions.

Centralize network functions at renewal or end of life to cloud
Network functions approaching renewal or end of life are good candidates to consider deploying in the public cloud. Examples include the access and mobility function and the session management function that serves the control plane of the packet core. Also, resource management and observability functions are good candidates to centralize and move to the cloud to gain efficiencies of scale.

Move container environment to AKS
Another best practice to consider is for the operator to move the container environment into the Azure Kubernetes Service (AKS). This move allows operators to store the container state and configuration information in the Azure cloud database and manage them as AKS clusters for better performance and efficiency. From this point, the operator can continue to incrementally leverage the existing 5G solution components in combination with Azure components, providing richer functionality and features, as shown in *Figure 6* on the next page.

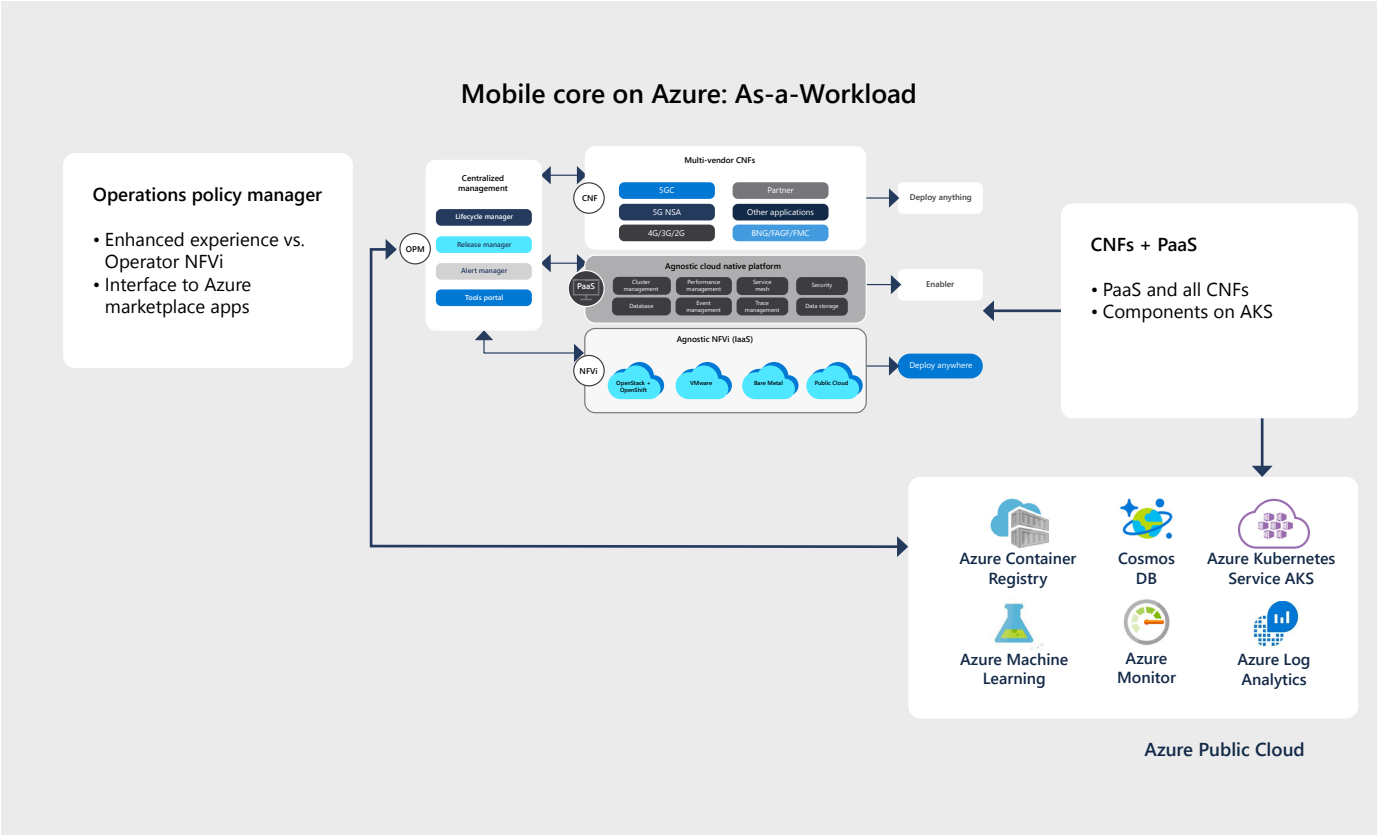


Figure 6

Advanced telemetry and monitoring to optimize network operations

Another opportunity for operators is to feed the rich network telemetry data collected by the 5G mobile core solution into the Azure cloud for further analysis. The **Azure Monitor** tool's advanced analytics engine utilizes machine learning and artificial intelligence capabilities to deliver deep insights for improving network operations and the creation of new services. Additionally, the operator can also replace some existing database functions with Azure Cosmos DB, which reduces costs while also providing a more resilient, secure, and highly scalable database platform for 5G applications.

Cloud migration benefits by phase

This first phase of hybrid cloud migration brings the 5G network into an "as-a-workload" state through a transparent migration of the basic container environment, network analytics, and some database requirements. Operationally, the user interface remains the same, and operators can continue to manage on-premises resources alongside Azure resources in the same manner. In the second phase of migration, all 5G operations components move into the public cloud, and the 5G mobile core is now delivered as a hosted service (see *Figure 6*).

In moving to an "as-a-service" state, the operator moves the entire 5G operations stack into the Azure cloud as a managed, hosted service. Azure Resource Manager (ARM) serves as the management layer for the 5G network, providing complete lifecycle management of the 5G network functions and API exposure points. In this stage, the operator makes full use of Azure's toolset, including Azure Kubernetes Service, Azure Cosmos DB, Azure Container Registry, Azure Monitor services for cluster monitoring, and additional services from the Azure Marketplace. Here, the cloud service provider is responsible for the management, automation, and service level assurance of the 5G cloud infrastructure and mobile core. Even the delivery of new software is managed in the public cloud through Azure DevOps.

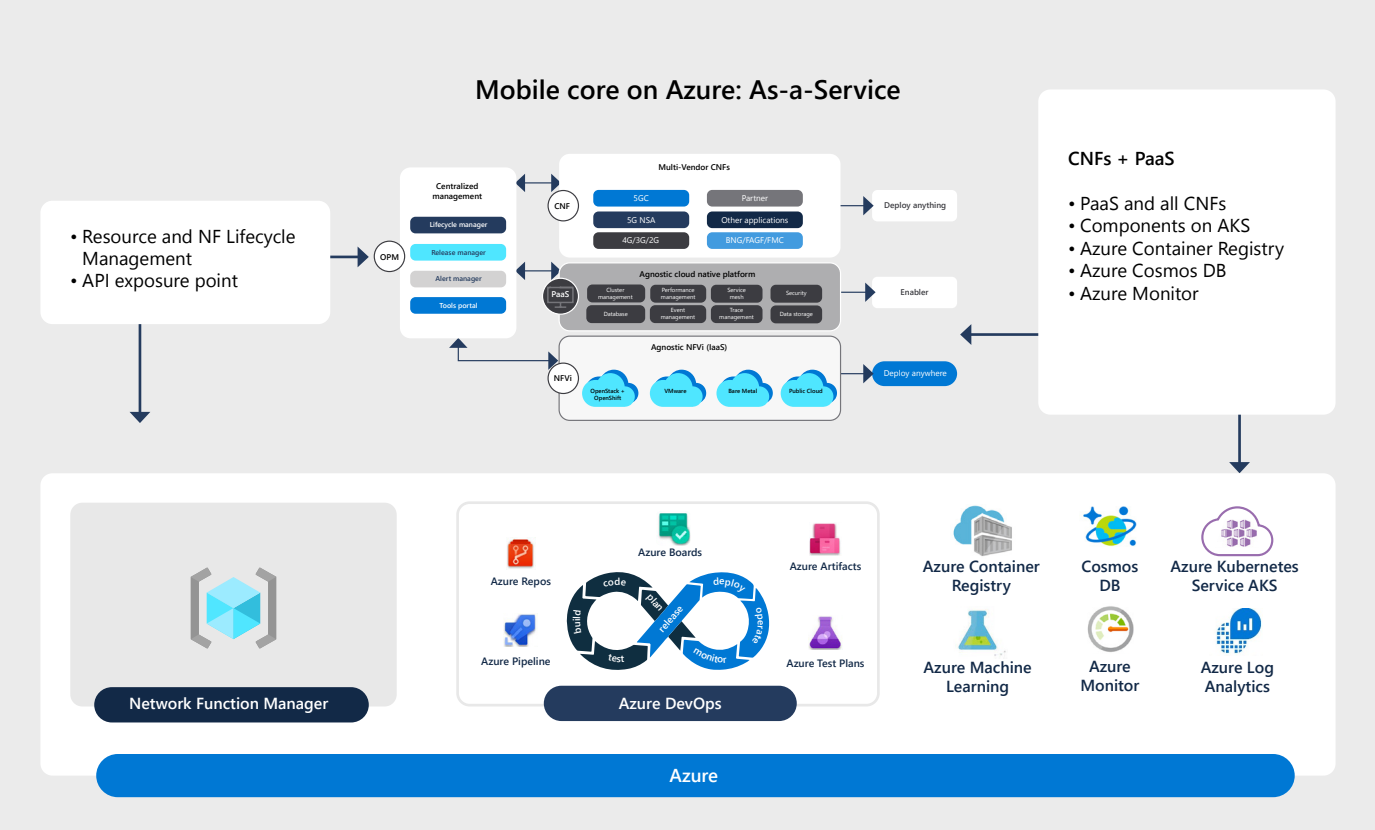


Figure 7

Benefits of effective cloud deployment

The transition from a private cloud 5G network to a public cloud 5G network is a significant change to an MNO's operating model. As such, let's now turn our attention to how the transition impacts 5G network operations. Specifically, the benefits MNOs can expect with respect to user experience, resource management, automation, and new functions.

Common user experience
In a hybrid cloud 5G network, the 5G control plane is centralized in the cloud. In contrast, the 5G user plane remains hosted primarily on NFVi hardware, either through the operator's infrastructure or via Azure edge locations. In either case, managing the user and control planes must be a seamless experience for operators to achieve operational efficiencies. Azure Arc provides this standard management interface by abstracting NFVi resources as Azure-manageable resources. As a result, public cloud resources and private cloud resources appear the same for the operator and can be managed together seamlessly over a single, familiar interface.

Resource management and SLAs
As 5G network functions move into the public cloud, MNOs add new capabilities, such as lifecycle management and service assurance, further simplifying and expediting service creation. Additionally, within a public cloud like Azure, SLAs can be managed by geography or availability zones to meet regional requirements for availability and performance.

Cluster and function automation
One of the key benefits of moving 5G network functions into the public cloud is automation. For operators, this means that control plane functions are no longer managed internally on Kubernetes but are automated and managed more efficiently as Kubernetes clusters via AKS. For user plane functions, ARM templates automate the Kubernetes application management previously performed by the operator using tools such as Helm.

Integrated tools
In the Azure environment, operators can take advantage of a vast ecosystem of Microsoft and third-party tools to enhance and improve network operations. These can seamlessly integrate and be part of the 5G network deployment. Advanced capabilities like analytics, machine learning, AI, cloud observability, and hyperscale can easily be part of the solution. Moreover, the public cloud also affords operators a virtually unlimited amount of cost-effective storage: storing large amounts of network telemetry data that currently sit on more expensive, disparate storage systems. The cloud has a distinct advantage in this respect as it allows an operator to expand and scale quickly at a lower cost.

> "The public cloud also affords operators a virtually unlimited amount of cost-effective storage."

5G mobile core

Microsoft is unique among cloud service providers to offer a completely integrated 5G mobile core platform to operators. Affirmed Networks and Metaswitch, now part of Microsoft, provide the 5G mobile core, which can be deployed in a private cloud, hybrid cloud, or entirely in the public cloud as a hosted, managed service. It also contains a comprehensive set of 5G core cloud-native network functions, including many partner components defined by 3GPP:

- Access and Mobility Management Function (AMF)
- Network and Service Capability Exposure Function (NEF)
- Network Repository Function (NRF)
- Network Slice Selection Function (NSSF)
- Non-3GPP Interworking Function (N3IWF)
- Security Edge Protection Proxy (SEPP)
- Session Management Function (SMF)
- End User Plane Function (UPF)

In addition, the solution features a fully integrated and automated PaaS that includes open-source, third-party tools such as Kubernetes, Prometheus, Jaeger, Envoy, and many others.

Features and benefits of 5G mobile core
5G mobile core delivers exceptional levels of innovation, automation, network assurance, and reliability. Let's now look closer at the benefits in each of these categories.

Innovation

- Finely grained network slicing capabilities that allow operators to support up to 1000s of network slices easily
- An advanced DevOps environment with continuous integration/delivery to support the rapid development, testing, and monitoring of new services
- A fully cloud-native architecture that supports the independent management of microservices, as well as the ability to create "combined" network services from 4G and 5G network functions/nodes

Automation

- Multi-vendor comprehensive lifecycle management, including spin-up and upgrade of network functions as needed, accelerating the time to deploy and grow the network
- Dynamically orchestrated and instantiated microservices based on real-time traffic conditions enable auto-scaling so that operators can rest assured that resources are available to satisfy demand in critical times of need

Network assurance and reliability

- Built-in monitoring, alerts, tracing, and logging to meet operational insights to maintain network performance
- Greater than five nines of availability helps deliver a superior customer experience
- Canary upgrades with no service interruptions and no maintenance windows required ensure that the latest features are available in the network through controlled, atomic upgrades of individual microservices

Why Microsoft's approach is the right choice for mobile operators

Our approach brings the most advanced, flexible, and operator-friendly 5G core solution available on the market. It has been designed specifically to meet MNOs where they are today—transitioning from 4G to 5G networks—and where they need to be tomorrow: creating and deploying revenue-generating 5G services.

With our solution, MNOs have a clear path forward to the future. The benefits of the solution include:

- **Lower total cost of ownership** through an NFVi-agnostic platform that is designed to deliver industry-best performance on commodity x86-based servers or in the cloud (or on Azure)
- **Build-once, deploy-anywhere design** that enables the deployment of 5G services wherever they are needed: in the core, in the cloud, at the edge
- **Reduced management complexity** that allows operators to manage on-premises and public cloud resources as a single, seamless solution or shift management of the 5G network functions to Microsoft
- **Accelerated service creation** that leverages a microservices-based architecture and network slicing capabilities to quickly create, deploy, and monetize new 5G services

Use cases for 5G hybrid cloud networks

The business case for 5G network transformation is different from previous generations of mobile technology. Consumer services have been the primary revenue model in the past, but 5G emphasizes revenue generated through new enterprise services. The ability to deliver high-bandwidth, low-latency mobile services to enterprises in a secure and SLA-guaranteed fashion opens up a variety of new 5G use cases.

5G provides MNOs with new opportunities in both consumer and enterprise networks. Microsoft's approach provides MNOs with the opportunity to create slices for specific applications to optimize bandwidth and traffic and also provide private networks for diverse use cases.

The need for private mobile networks is driven by new applications such as AR/VR, IoT-connected cities, and autonomous vehicles, along with the need to deliver on the unique requirements of various industry segments, including manufacturing for Industry 4.0, oil and gas for remote platform, airports and ports for traffic, and cargo management.

Manufacturing
Industry 4.0 is about leveraging a collection of new technologies in modern manufacturing environments, including analytics, automation, and wireless broadband connectivity. As manufacturing floors generate more data than ever before through IoT devices, video cameras, and sensors, this data must be collected and analyzed to drive decision-making and process optimization.

MNOs can partner with manufacturers to support these future factories via dedicated 5G network connectivity and multi-access edge computing. With Affirmed 5G mobile core deployed on Azure, MNOs can quickly stand up a dedicated 5G mobile core close to manufacturing sites and host additional data processing capabilities in the manufacturer's enterprise via Azure edge locations. This scenario would deliver the dedicated, secure 5G connectivity needed to support high-bandwidth video and telemetry data. Additionally, dedicated edge-processing capabilities leverage Azure tools, such as machine learning and AI, without the need to backhaul large amounts of data through the MNO's network.

Private mobile enterprise networks
Beyond manufacturing, many enterprises are looking at private 5G networks to replace their existing Wi-Fi networks. With 5G mobile core provided by Affirmed, these services can be delivered as multi-tenant network slices on the operator network or at the enterprise edge. In addition, MNOs can opt for a centralized control plane in a hybrid cloud model to manage these communications more efficiently with a distributed user plane at the edge to support on-site processing and lower latencies. These private 5G networks could support many enterprise applications, including real-time remote healthcare, secure banking transactions between branches, live sporting events, and secure wireless connectivity for airport kiosks.

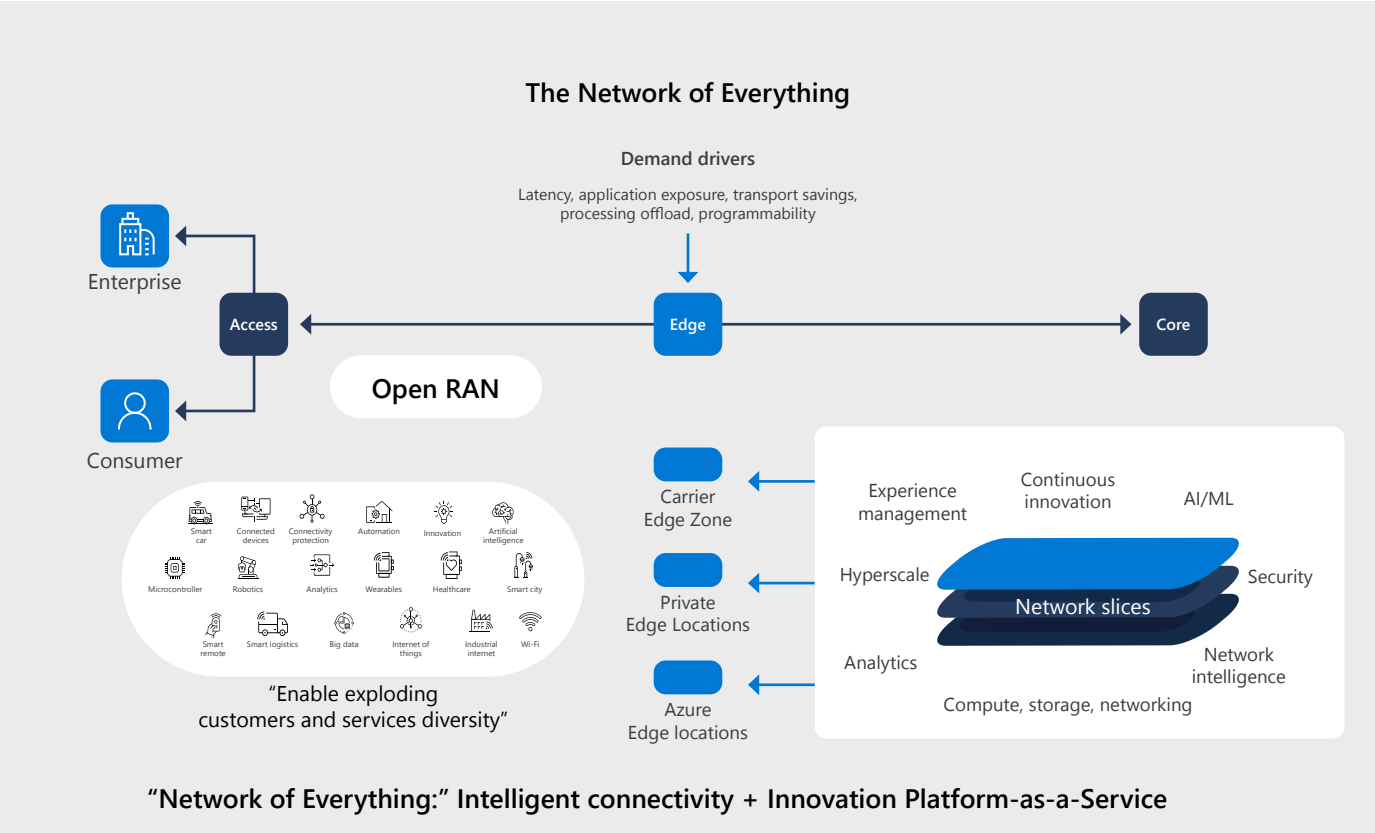
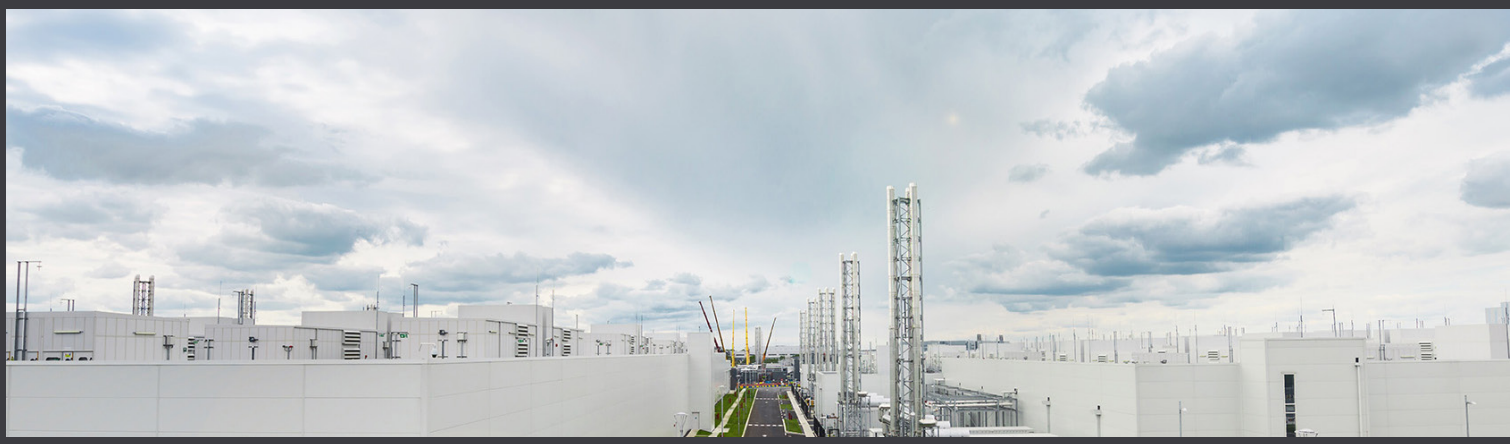


Figure 8

The future of the 5G hybrid cloud

For mobile network operators, the cloud is the future. Successful web-scale companies have utilized cloud-native technologies for years, successfully growing and scaling their business. With 5G mobile core by Affirmed, Microsoft brings the same powerful cloud-native technology to the telecommunications industry. The result is a combination of the best of intelligent core, intelligent cloud capabilities, and features to transform the telecom industry. Network operators can now fundamentally define how the network can run at webscale, create new innovative services leveraging the power of the cloud, and accelerate time to revenue at the speed of business.



Looking forward

Summary and conclusion

Unlike other 5G core solutions providers, Microsoft provides an approach that offers everything that MNOs need to deliver 5G services: the core network functions provided by Affirmed, hybrid cloud infrastructure, operations platform, container platform, and enriched toolsets (AI, automation, etc.). At the same time, 5G cloud components are open and interoperable, leveraging open-source innovations to ensure that MNOs can build their own best-of-breed hybrid cloud solution for 5G. From there, the future resides in the 5G hybrid cloud that enables "The Network of Everything."