aws FOR TELECOM

WHITEPAPER

Telco transformation with AWS

Introduction

The telco industry has withstood numerous challenges over the past decade. As a sector, telcos underperform many of their business counterparts — consider the 23% per annum decline of the spread between return on invested capital (ROIC) and the weighted average cost of capital (WACC) for 111 global telecom operators from 9% in 2010 to 1% 2020[1]. Correspondingly, the S&P Telecom Select Industry Total Return Index[2] has underperformed the S&P 500 Total Return index over the last ten years (7.3% vs. 12.3%).

Telco sector performance in the financial markets reflects both the perception and reality that telcos have been slow in reacting to disruptions across business models, customer expectations, and technology. Digital transformation has created new business opportunities that technology companies have jumped on — cloud applications, social media, interactive streaming, and online gaming. Seamless online experiences have up-leveled consumer expectations of other services they purchase. Likewise, cloud-scale platforms, AI and machine learning, real-time analytics, and other new technologies have unlocked new business applications and capabilities. Many telcos today struggle with transitioning to this new world and are challenged in adopting faster and leaner operational models.

Yet, the onset of the pandemic in 2020 reminds us of the critical role telcos play. Telecom operators exhibited their resilience and rose to the challenge of keeping the world connected through waves of lockdowns. Telcos continue to make substantial infrastructure investments in fiber and mobile infrastructure buildouts. GSMA estimates telcos will spend an aggregate of USD 1.5T capital expenditure (CapEx) from 2023 and 2030[3] on mobile networks.

To improve the outcomes for telcos over the next decade, telco boards and senior leaders recognize the importance of enacting transformation across multiple areas: business, operations, technology, and culture. Change at telcos is necessary to monetize 5G and fiber infrastructure buildouts, address energy use and sustainability initiatives, retain and delight customers, and attract and cultivate a new generation of employees, all while creating new revenue streams and reinventing business models. Fortunately, telcos today recognize they do not, and should not, need to achieve all these independently.

Since inventing modern cloud computing in 2006, AWS has been helping customers transform their businesses by alleviating the heavy lifting of IT infrastructure, reducing costs, and enabling a secure, reliable, and agile environment to experiment and innovate. As the numbers and types of customers have grown, we have adapted our cloud to meet the needs of new industries and diverse workloads. We have forged ahead to lay the groundwork for cutting-edge technologies across AI, analytics, containers, serverless, and secure computing, so our customers can benefit from the R&D investments we have made and continue to make. With the need for telcos to innovate and disrupt, the emergence of new technologies such as software-based networks, edge computing, and the maturing of AI and ML, we considered how we could adapt and extend AWS services and infrastructure to support telco-specific business objectives and workloads.

This paper provides an insider's view into AWS's journey to make our cloud the best foundation for telcos to build upon. We share our learnings from working with telco partners and detail how we have adapted our cloud to telco workloads while establishing our long-term commitment to the communications industry.

1 www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-telcos-can-succeed-in-launching-new-businesses-beyond-connectivity
2 S&P Telecom Select Industry Index includes equipment manufacturers that serve the telcos, along with integrated and wireless telecom operators.
3 www.gsma.com/mobileeconomy/wp-content/uploads/2023/03/270223-The-Mobile-Economy-2023.pdf

Technology trends in the telecom industry

The telecom industry has seen significant technology shifts in the past few years. Telecom companies seek to respond by transforming themselves from communication service providers to digital service providers.

Maturity of virtualized infrastructure for network functions

In 2015-2016, network vendors started unbundling software from proprietary hardware and packaging it into software network functions (NFs) running on x86 architecture-based commodity infrastructure. We see a wider adoption with most telcos in some stage of virtualization in their mobile network deployments. This trend is rapidly expanding to other network areas, such as cable and broadband networks. However, the problem of cross-domain automation and resource and service orchestration in virtualized networks remains challenging without interface standardization or the emergence of a leading ISV/solution.

Establishment of the telco network edge

Telcos are bringing computing and storage physically closer to the end-user by hosting them in their data centers and other network aggregation sites. The increased proximity serves emerging third-party applications requiring very low latency. It can take advantage of closer integration with telco network APIs providing information from the network and the ability to enforce policies on the network. This telco edge infrastructure is also the foundation of building distributed wireless and wireline access networks. The edge can host network functions such as the data plane of the 5G core, termination of broadband and cable network overlays, edge routers, and SD-WAN (software-defined wide area network) hubs.

Software-defined connectivity and security models

Telcos are moving to provisioning SD-WAN-based connectivity, complementing traditional MPLS (multi-protocol label switching) circuits. SD-WAN is a virtualized service connecting and extending enterprise networks over large distances to give remote users access to corporate applications, services, and resources. While SD-WAN is gaining ground, it doesn't include any security and access controls companies need to protect and defend their network in the cloud environment. Enterprise customers often require multiple point products to secure web gateways, support application firewalls, and secure virtual private network remote access, which creates unwieldy administrative issues, increasing costs, and complexity. To facilitate secure network access in cloud environments, telcos are implementing Secure Access Service Edge (SASE) solutions that combine the capabilities of a WAN with comprehensive security functions.

Connectivity to cloud

Telcos traditionally built infrastructure and services to connect subscribers, enterprises, and the internet. With the shift of enterprise workloads to the cloud, telcos seek to connect their networks directly to cloud points-of-presence and different cloud providers. They're automating this process through APIs with carrier-neutral co-location providers. They're also working with networking vendors or offering their in-house solutions to provide

application acceleration and security to support the enterprise transition to the cloud. Most enterprises are pursuing a hybrid and a multi-cloud approach, meaning they need connectivity across private data centers and into various cloud networks.

Wired wireless convergence

Wired wireless convergence (WWC) comprises the ability of telcos to provide subscribers and enterprises with services that use both fixed wireline cable/fiber networks and mobile wireless cellular/Wi-Fi networks. The industry has been attempting this for a few years. Still, in the legacy model, seamless service integration and interworking between wireless and wireline networks is complex because services reside in silos architecturally and from a network equipment perspective. The shift to software-based networks enables a harmonized single stack of fixed and mobile services to be implemented efficiently, which has fueled renewed interest in WWC. Additionally, there's been strong cooperation among the 3rd Generation Partnership Project (3GPP) and the Broadband Forum (BBF) bodies to lay the foundation for WWC.

Regulatory landscape affecting cloud and telcos

The European government has classified network infrastructure and services as "nationally critical infrastructure," applying data sovereignty regulations to both cloud and telcos. While Europe is leading the way, some Asian countries are adopting a similar posture. Data sovereignty can be expressed as related concerns over how foreign nations can access or disrupt customer data processed in the cloud. The focus on data sovereignty is driven by geopolitical tensions with local political and economic motives and heightened scrutiny of regulators and governments. In some countries, telcos have additional regulations around sustaining an operating network, even when disconnected from international networks or global infrastructure for extended periods.

Each of these trends underly core requirements driven by business and regulatory needs. Based on our experience building and operating large-scale infrastructure, we believe that cloud is the key to developing new telco capabilities to address these needs. At the same time, AWS's work with other market verticals has taught us that every industry brings unique requirements we have to adapt to. Our collaboration with telcos has given us insight into the enhancements needed on the AWS Cloud, and we will delve into the changes we have made after we explain the importance of cloud.

Why cloud and why AWS?

The cloud is a large-scale, global, multi-tenanted infrastructure with layers of services for developers to build applications via APIs. The nature of the cloud makes it elastic, giving users the perception of infinite, available capacity. Users can access capacity when needed and then relinquish it to the multi-tenanted, common pool for others to use – lending the infrastructure to new consumption models such as pay-as-you-go.

Since its inception, cloud has been inherently secure and available – many enterprises have migrated critical workloads from on-premises IT-controlled infrastructure to cloud. Deployment of software, upgrading, and patching is fully automated and happens 24x7 without impacting customers.

The cloud is also energy efficient because multi-tenancy, a range of consumption options, and pricing result in very high infrastructure utilization. Above all, the cloud gives users agility to innovate, develop, and experiment with new applications and services, refactoring applications very easily to quickly modify the customer experience when needed. By hosting networks and connectivity services on the cloud, we bring all these benefits to telcos, which also effectively run a multi-tenanted, large-scale infrastructure, but of a different kind. Combining this with other cloud services in databases, analytics, AI, ML, and IoT expands the possibilities for improving operations and customer value management, creating new business opportunities.

Let's dig into the specific benefits that cloud and AWS bring to telcos:

Elasticity and new consumption/pricing models

Elasticity and pay-as-you-go models, fundamental values the cloud offers, now comes to networks. Today networks are over-provisioned and rigid in their pricing models with usage slabs and caps. 5G has introduced the concept of network slicing, where you can dynamically create flexible, virtual, end-to-end network slices from a fixed physical capacity. In the cloud, resources, and capacity can be added when needed and relinquished if they're not. A network built using cloud enables the dynamic, efficient creation and management of hundreds of network slices. It allows for new, flexible consumption models of network resources that optimize physical resource use and monetization while giving customers more flexibility.

Total Cost of Ownership (TCO) reduction

Using infrastructure hosted and managed in the cloud reduces the TCO of a communications network. Telcos divert significant resources to manage and operate networks. Overall, telco OpEx is between 48% and 79% of total revenue[4] , with the share of OpEx allocated to managing networks, infrastructure, cloud, interconnect, and energy at approximately 60%[5]. This means telcos spend 30-50% of revenue operating networks and associated infrastructure. This larger number isn't surprising since network elements must be continuously monitored and maintenance windows scheduled for upgrades, requiring a large workforce and detailed planning. When telcos host their network on the cloud, managing and upgrading the network is automated without disrupting the subscribers and enterprise customers, resulting in lower OpEx. And since cloud infrastructure is multi-tenanted and shared, it benefits from significant economies of scale, reducing overall costs as telcos migrate their CapEx to OpEx through cloud services.

Agile development and deployment

The cloud brings agility to communication networks. One of the biggest telco challenges is the long timeframe required to roll out a new service. For example, the average timeframe for Bridge Alliance members (representing 34 telcos across Asia and the Middle East) to launch a new service from idea to production is 18 months[6]. Popular services like enterprise WAN or consumer broadband require coordination across multiple domains. Routers and other network elements must be configured individually, and before services can be rolled out in production, different scenarios must be tested extensively in a sandbox. With a cloud-native network, a new service can be developed and deployed in hours using simple APIs to assign or scale cloud infrastructure resources, configure cloud services and network functions, and create monitoring dashboards and management policies. Coupled with the benefits of elasticity, it also allows for creating offerings of the network-as-a-service.

Security

For telcos that process ever-increasing volumes of customer data that may include personally identifiable information (PII) and who are subject to compliance and privacy regulations for the countries and regions they operate in, security is paramount. AWS is architected to be the most secure cloud computing environment and is built to satisfy the security requirements of high-sensitivity organizations. AWS supports 98 security standards and compliance certifications, including PCI-DSS, HIPAA/HITECH/HITRUST, FedRAMP, GDPR, FIPS 140-2, and NIST 800-171, helping satisfy compliance requirements for virtually every regulatory agency around the globe. We provide a wide variety of best practices documents, encryption tools, and other guidance our customers can freely leverage to deliver application-level security measures. By using the cloud, telcos benefit from all of these, resulting in the most secure layered networks to support their customers.

Enhanced developer experience

By integrating the edge cloud into the telco network, telcos can unleash a new developer ecosystem with access to AWS's APIs and APIs from their own networks. Network APIs enable developers to extract network information (such as congestion, and location information, for example) and allow programmability of the network (such quality-on-demand, guaranteed bandwidth and latency, for example). Giving developers this capability, makes a whole set of new applications possible, unlocking new revenue pools for telcos.

Sustainability

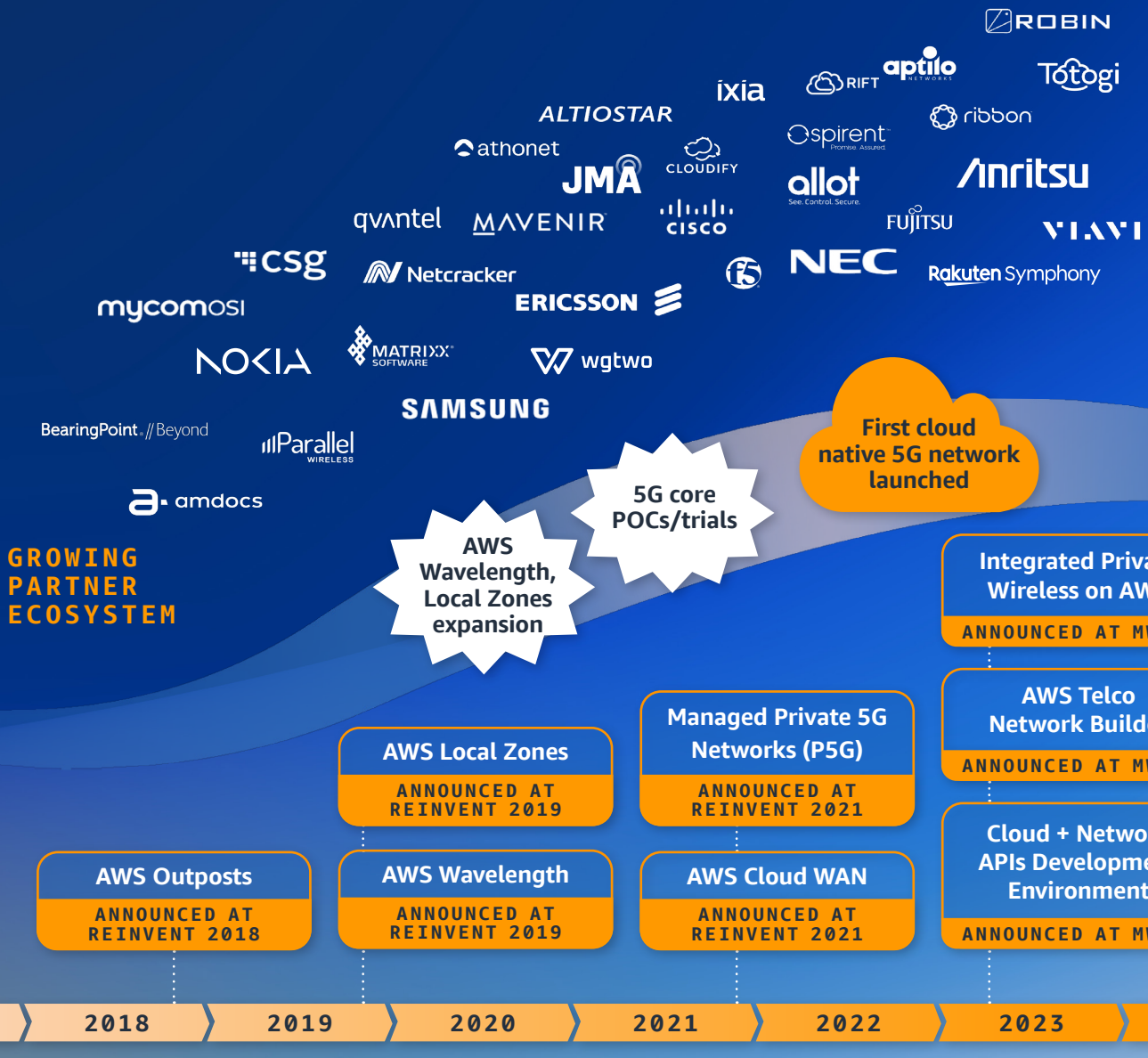
The global telco industry produced 2.6% of the total world carbon dioxide (CO2) emissions in 2020, according to a European Telecommunications Network Operators Association report. GSMA Intelligence indicates that energy consumption accounted for 15-40% of telcos operating expenditure in 2021, and that figure is expected to rise. Studies conducted by 451 Research (an S&P Global Market Intelligence research firm) show that AWS's infrastructure is 3.6 times more energy efficient than the median of US enterprise data centers and up to 5 times more than enterprise data centers in Europe and the Asia Pacific because of high utilization and other focused efforts. Additional power efficiency improvements come from our investment in chips. Graviton3 Arm-based EC2 instances use up to 60% less energy than comparable x86-based EC2 instances. Amazon is on target to power its operations with 100% renewable energy by 2025, which will further help the cause of telcos in reducing their carbon footprints. So, running networks and other workloads on AWS helps telcos achieve their sustainability goals.

4 inform.tmforum.org/features-and-opinion/the-impact-of-5g-and-cloud-on-telco-capex-and-opex
5 inform.tmforum.org/research-and-analysis/reports/telco-to-techco-capex-and-opex-implications
6 inform.tmforum.org/features-and-opinion/collaboration-opens-up-new-service-opportunities-in-asia-pacific

AWS's journey to a telco-ready cloud

AWS started in 2006 with a handful of services, notably AWS S3 and EC2, and today we have grown to over 200 services. The capabilities of the cloud have also evolved, from a self-service model to migration services for legacy applications, to pre-built solutions and managed services.

We started on the journey to use cloud for telco workloads around 2017 (reference Figure), based on our conviction that as network virtualization was accepted, cloud was next. And we were aware that cloud wasn't fully ready, so we would need to enhance our services to address any unique workload requirements.



Today, we believe the AWS Cloud is the best cloud to handle telecom workloads, and we continue to work with our telco partners to evolve it. To provide you with a deeper understanding of the changes we have made to support telco workloads, we'll explain AWS Cloud telco capabilities across four themes: extending AWS to the edge, enhancing AWS Cloud for telco network functions, transforming telco data into intelligence, and co-inventing telco connectivity services.

1 Extending AWS to the edge

A telco network spans a broad geographical area from large and regional data centers and edge sites to cell sites very close to the users. The cloud comprised a few large, consolidated data centers, unable to run an end-to-end network and bring data to the users. One of the first challenges we had to solve was how to bring the cloud to the edge of the network. In 2018, AWS set out to build an edge continuum of infrastructure. To enable multi-access edge computing as well as running end-to-end networks using cloud infrastructure, one of the critical challenges we solved was bringing the cloud to the edge. Unlike other industry verticals, the edge cloud is fundamental to enabling telco transformation. By working backward from customers and focusing on the developer experience, we realized their key concern was the consistency of APIs, tools, automation, management, and services like those used in our regions. We developed a spectrum of edge offerings as a continuum from our existing Regions to the far edge of the network to enable this.

Most workloads can run in our **Regions** and get the maximum benefits of low cost and availability of all our services. We designed **AWS Local Zones**, a type of AWS infrastructure deployment that place AWS compute, storage, database, and other select services close to a large population, industry, and IT centers. With AWS Local Zones, telcos can quickly run applications that need single-digit millisecond latency closer to end-users in a specific geography. AWS Local Zones are ideal for use cases such as media & entertainment content creation, real-time gaming, live video streaming, and ML inference. Today, we have 17 Local Zones across the US and 15 outside the US, with an additional 21 metro areas across 18 countries launching soon.

We optimized **AWS Wavelength** for mobile edge computing applications. Wavelength Zones are AWS infrastructure deployments that embed AWS compute and storage services within telco data centers at the edge of the 5G network, so application traffic from 5G devices can reach application servers running in Wavelength Zones without leaving the telco network. This avoids the

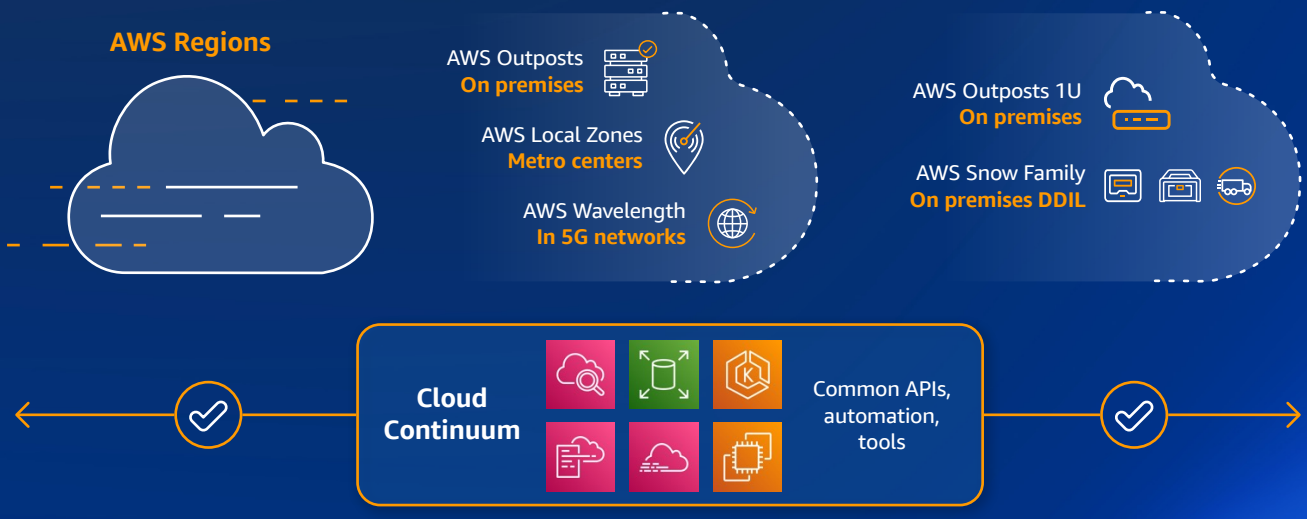


Figure 1: AWS Edge Continuum

latency that would result from application traffic traversing multiple hops across the internet to reach their destination, enabling customers to take full advantage of the latency and bandwidth benefits offered by modern 5G networks. Today, we have built AWS Wavelength Zones with Verizon (19 in the US), Bell (1 in Canada), KDDI (2 in Japan), SK Telecom (2 in South Korea), and Vodafone (1 in the UK, 3 in Germany).

AWS Outposts is a fully managed service that offers the same AWS infrastructure, AWS services, APIs, and tools to virtually any datacenter, co-location space, or on-premises facility for a truly consistent hybrid experience. AWS Outposts, available in 1U, 2U, and full 42U rack form factors, are ideal for workloads requiring low latency access to on-premises systems, local data processing, data residency, and migration of applications with local system interdependencies.

AWS Snow devices—a part of the AWS Snow Family—are edge computing, data migration, and edge storage devices. They're built for data collection, ML, processing, and storage in environments with intermittent connectivity (such as manufacturing, industrial, and transportation) or in extremely remote locations (like military or maritime operations). These devices can operate without connectivity to the AWS Regions and can be rack-mounted and clustered together to build larger installations. AWS Snow devices can migrate large amounts of data into AWS Regions either via the online AWS DataSync service or be shipped back to AWS.

Today, AWS Local Zones, AWS Wavelength Zones, AWS Outposts, and the Snow family of edge devices combine to form our edge continuum. This range of options allows us to serve OSS/BSS and the control plane of packet cores in AWS Regions with user plane core and centralized RAN functions in Local Zones (or AWS Outposts at telco data centers) and distributed RAN functions on AWS Outposts 1U at the far edge. Developers and telcos can build applications (and other services) using the principles and APIs they use for Regions — EC2, EKS, VPC, and other AWS services work the same way across the continuum.

By distributing network workloads appropriately across this edge continuum, we can build an end-to-end network managed by a single pane of glass. And with a combination of these edge cloud offerings, depending on the workloads, latency sensitivity, deployment topology, available infrastructure, and fiber assets, we can address all of the telcos network hosting needs.

2 Enhancing AWS Cloud for telco network functions

In the evolution of the cloud since 2006, we have periodically come across a vertical with a unique set of workloads that require us to delve deeper into their requirements and enhance our infrastructure and services to meet them. As an example, when AWS started looking at financial workloads from banking and insurance customers, we developed several security-compliant features needed explicitly for these workloads, and now all customers benefit from them. Telcos have similarly addressed the needs of network functions, such as the 5G packet core, radio components, virtual converged cable access platform (vCCAP), virtual broadband network gateway (vBNG), SD-WAN controllers, virtual routers, network firewalls, and IO-centric network appliances.

Providing a high performance, secure virtualization stack

One of the critical requirements of NFs—different from other workloads—was the performance expected of the virtualization stack. The stack needs to support high-speed packet processing, typically done in dedicated hardware boxes like routers with packet processing in silicon. In addition to performance, these NFs needed isolation from other processes and strict protections against unauthorized access. We have developed a virtualization stack called Nitro which lays the foundation of our AWS compute instances. The Nitro System, built using silicon chips

developed in-house, is a rich collection of building blocks that give us the flexibility to design and rapidly deliver different EC2 instance types. Nitro Hypervisor is a lightweight hypervisor that

manages memory and CPU allocation, delivering performance indistinguishable from bare metal. Meanwhile, the Nitro Cards family accelerates IO for network-bound workloads. On the security front, the Nitro Security Chip reduces the attack surface by offloading virtualization and security functions to dedicated hardware and software, and Nitro Enclaves guarantee confidential computing by enabling customers to create isolated compute environments to process highly sensitive data.

Integrating telco and cloud networking

One of the challenges we had was in the networking 'impedance mismatch' between cloud networking and telco networking. Cloud networking (constructs around VPC, gateways, and load balancers) has evolved to manage the connectivity between instances within Regions, across services needed to build an application, and for connectivity to our backbone and the internet. In cloud networking, we use layer-3 routing, hash multiple flows onto a single interface, and use a proprietary encapsulation protocol for intra-VPC traffic. We assume sufficient bandwidth is available without the need for traffic classification. Telco networking evolved to manage last and middle-mile access, and in these networks, bandwidth is provisioned and managed with QoS (Quality of Service) prioritization. Telco networks use layer 2 protocols widely. We bridged this gap by creating gateway abstractions like the local gateway in AWS Outposts and AWS Wavelength, and use of virtual routers from ISVs specifically, to allow for telco networking to interface with cloud networking.

Expanding networking capabilities for telcos

Cloud services were primarily built with compute (instances, containers) and storage (block, object, databases) as the primary resources for building applications, and networking was an enabling function to connect the resources and applications. Most enterprise applications use a lot of compute and storage capacity but have straightforward networking requirements—single or a handful of ENIs (Elastic Network Interfaces) —needing only a few IPs for attachment to an instance and bandwidth proportional to compute. However, with networking functions, the compute and storage demands are relatively simple, and the complexity is in networking. Running the user plane of a packet core takes only a few compute instances, but it requires very high throughput (to carry traffic of a large metro area, for example) and many IPs connecting to the compute instances (supporting all mobile devices connecting to that packet core). For example, we created features to assign a CIDR range to network interfaces instead of individual IPs to deal with the large scale of IPs needed. We started building instances with high networking bandwidth of up to 100Gbps.

Accelerating network function onboarding

One of the tasks in building such a cloud-native network is the assignment of compute, storage, and networking resources for the network functions, building subnets, routing tables, and access policies in VPC, connectivity between the different edge offerings, and managing all this over the lifecycle of the network functions and the network itself. We designed the AWS Telco Network Builder (TNB) to assist operators in this. TNB is a fully managed service that helps telcos deploy, run and scale cloud-native networks. The service takes as input network functions, their configurations, network topology, and other information in templates and data models used in the telco industry, such as TOSCA and YANG, and assigns AWS resources and creates service configurations automatically. When the network needs to be updated, only the template needs to be updated, and the changes are reflected in AWS reassures and services. TNB provides a single pane of glass for automating the deployments and operations of the network and orchestrating the network functions across AWS Regions, AWS Local Zones, and AWS Outposts in telco premises. It interfaces with 3rd party service orchestrators and has industry-standard northbound interfaces to enable that.

3 Transforming telecom data into intelligence

Telcos have data from multiple, varied sources – subscribers, devices that connect to the network, the applications that run on them, and the network itself. The data is spread across different systems, such as OSS/BSS, packet core network functions, call data records, and ERP systems. The data is

owned by various organizations, from network planning, network operations, procurement, and business development to customer service. It's challenging to carry out analytics when data is stuck in different silos. The silos cause multiple problems – the data needed for a given workload may be split across various sources and so be inaccessible; the silo where the data lives might not meet the price performance requirements for a given workload; the silos may require different management, security, and authorization approaches, increasing operational cost and risk.

5G architecture and underlying technology, enablers make modern data lake architecture and AI and ML applications a natural and easy fit, creating new opportunities to monetize the network that did not previously exist along with some challenges.

1. There's an **exponential increase in data** with 100x the throughput and number of devices, and modern techniques of managing and analyzing are crucial. Data also offers a richer source of business and operational value.
2. The **5G network is more dynamic** — slices can be created dynamically, and the network function lifecycle is managed through a DevOps process. So closed loop automation and optimization based on network data is essential to accomplish this efficiently.
3. The emergence of the computing edge in the telco network has created **a platform embedded in the network** that can collect, store, and analyze data.
4. Telcos are looking at **new consumption models** because they can dynamically create virtual slices, new services, and pricing plans, which requires a better understanding of customer behavior and personalization.
5. The standards bodies are defining functionality, such as the radio intelligent controller (RIC) and the Network Data Analytics function (NWDAF), which can utilize AI and ML services to **rapidly process large amounts of internal network data** for optimization, visualization, and anomaly detection.

There are three critical phases in the telco journey to becoming more data-driven using AWS. Telcos today still have on-premises data centers with relational databases and simple tooling to manage data.

Modernizing data infrastructure

The first phase is to migrate from on-premises legacy databases onto modern cloud data infrastructure, which provides a variety of databases and data services. AWS's portfolio of databases includes purpose-built database engines for relational, key-value, document, in-memory, graph, time series, wide column, and ledger databases. AWS also has a broad and deep collection of purpose-built data services that serve use cases such as interactive dashboards, log analytics, and big data processing.

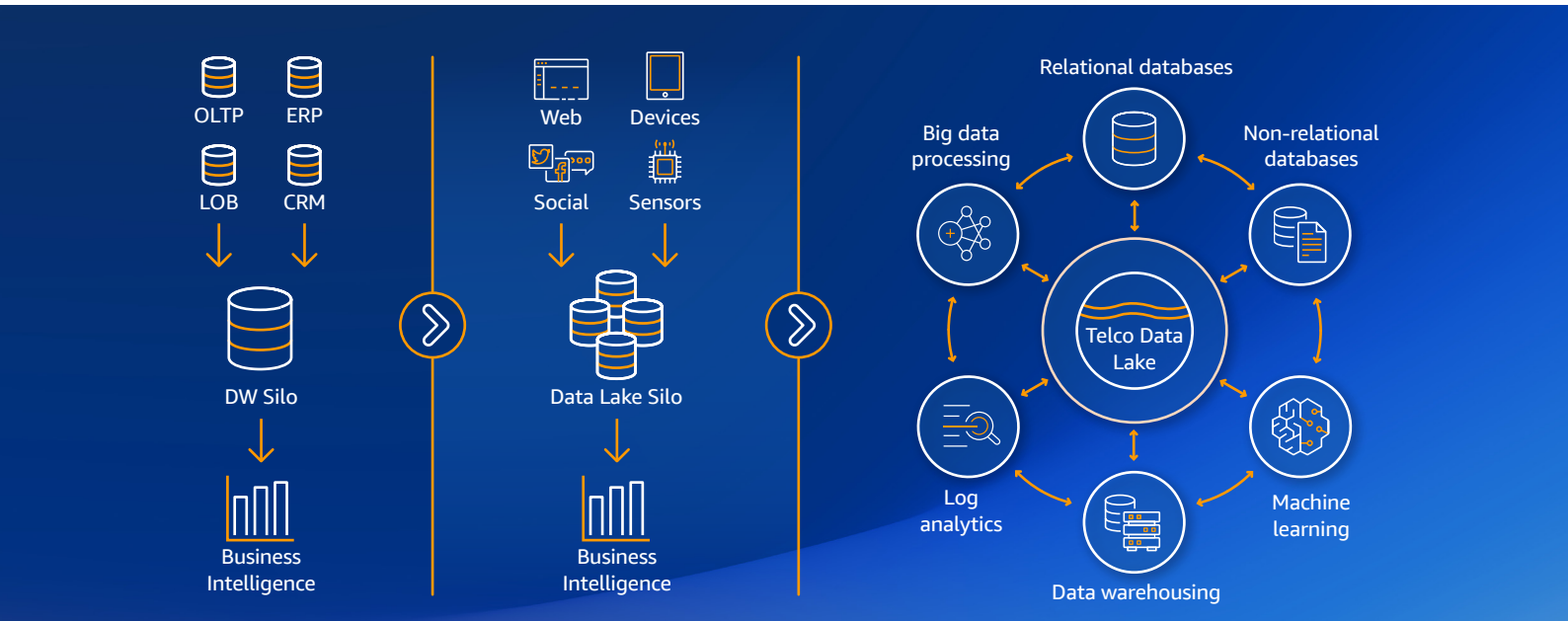


Figure 2: Evolution of Data Management Platforms

Liberating data from siloes

The second phase is about liberating the data by moving it to a data lake architecture to break down data silos and make data more accessible to all users. We have built a data lake specifically for telcos, called AWS Telco Data Lake, which includes a data mesh and MLOps pipeline that taps into the data sources and enables the data consumers to build BI and ML use cases. With Telco Lake, the data lake monolith is decoupled into individual domains to accommodate the variety of technologies, standards, and ISVs associated with telecom data. The data is managed as a product; data features and quality are owned by the producers. Relevant data attributes (quality, schema, etc.) and metadata are shared via a standard data catalog. As a result, the lines of business can quickly build use cases by searching for and subscribing to relevant, high-quality data sets via the standard catalog. The AWS Telco Data Lake also provides a common infrastructure, and platform, to centrally govern the data and user access while providing the right tools for the job at the right time. This centralized model simplifies security by sharing data versus copying or duplicating and allows for a single audit trail for data stewards.

Gaining insights with analytics and AI/ML

The third phase is innovating with that data, using AWS's AI/ML stack to gain insight and drive business and operations. At the bottom of the stack is the framework and infrastructure layer, with custom silicon such as AWS Trainium– optimized for training models, and AWS Inferentia - optimized for deriving inference. The Inf1, Inf2, and Tr1 compute instances– based on our custom silicon made available across the edge continuum– enable high performance, cost-efficient application of ML use cases at the network edge. These include third-party applications and applications for the network itself, such as network analytics and service assurance. At the middle of AWS's AI/ML stack is an integrated development environment layer called Amazon SageMaker.

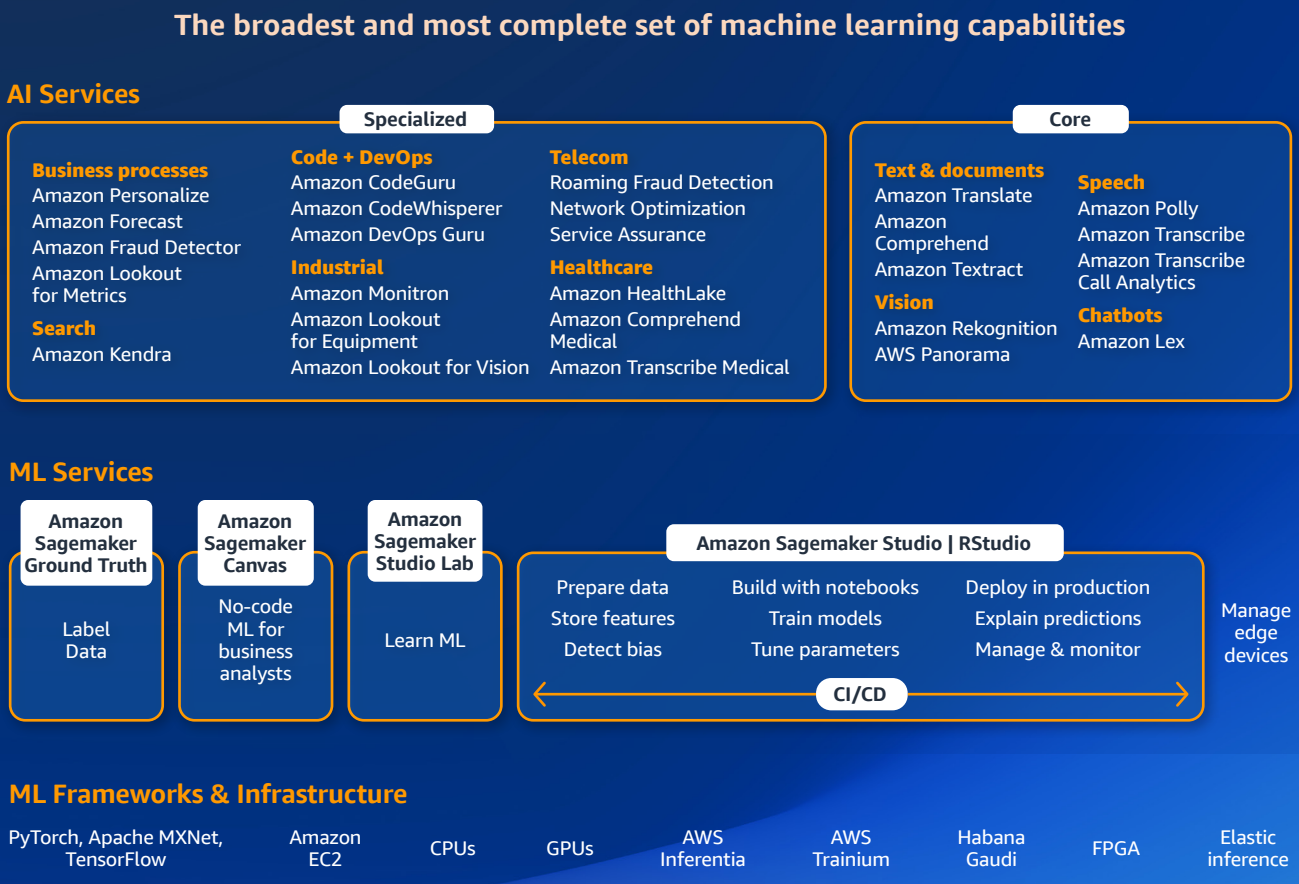


Figure 3: The AWS AI/ML stack

Amazon SageMaker helps data scientists and developers to prepare, build, train, and deploy high-quality ML models quickly by bringing together a broad set of purpose-built capabilities. The telco network edge poses some unique challenges. Compute and storage capacity at the edge is limited, so using Amazon Sagemaker, telcos and developers can trade off the size of training models with accuracy and seek to be able to do this automatically. Also, there are several thousand edge sites across which the models need to be deployed, maintaining consistency and synchronicity.

Some ISVs building ML-based tools for network optimization, service assurance, and xApps/rApps for the radio intelligent controller (RIC) in O-RAN use AWS SageMaker to develop their applications.

At the top of AWS's AI/ML stack are pre-built and trained models – purpose-built for everyday use cases such as recognizing objects automatically, transcribing voice-to-text, recommendation engines, fraud detection, chatbots, and automated call centers. We have customized these for telco-specific use cases, and nuances of telco data and customer behaviors for coupling with other telco-specific data sources such as the network.

For example, we have integrated SIP-based voice calls with AWS Transcribe (our voice-to-text processing service) and Amazon Comprehend (a natural-language processing service) for sentiment analysis to resolve customer issues. Another example is extending our forecasting service for predicting network bandwidth and capacity needs, helping telcos plan for expansion. And then, we have new dedicated services being built for telcos, such as closed-loop network optimization, identifying security threats using anomaly detection from network signals, and several others.

4 Co-inventing connectivity with the telcos

With increasing enterprise workloads moving to the cloud, we have been evaluating more efficient, cost-effective ways of connecting enterprise locations, branch offices, and data centers to the AWS Cloud.

AWS Direct Connect

In 2012, we built the AWS Direct Connect (DX) service, in partnership with telcos, to provide the shortest path for enterprises to reach AWS resources. While in transit, their network traffic remains on the AWS global network and never touches the public internet. This reduces the chance of hitting bottlenecks or unexpected increases in latency.

When creating a new connection, the enterprise customer can choose a hosted connection provided by an AWS Direct Connect delivery partner or select a dedicated connection from AWS—and deploy at over 100 AWS Direct Connect locations around the globe. The connectivity from the enterprise site to the AWS Direct Connect locations is provided in partnership with telco carriers. We have been expanding on this principle of leveraging our connectivity assets and cloud value with telco assets to create new connectivity offers for our joint customers.

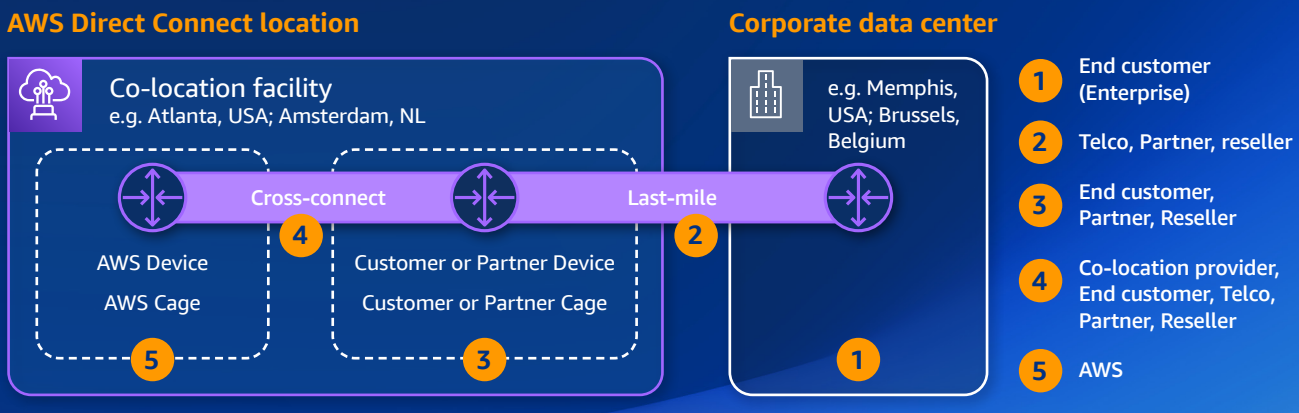


Figure 4: AWS Direct Connect

AWS Cloud WAN

SD-WAN and secure access services edge (SASE) are enterprise connectivity services experiencing rapid growth — Gartner forecasts that total worldwide end-user spending on SASE will reach $9.2 billion in 2023, a 39% increase from 2022[7], and SD-WAN will grow at a 14% compound annual growth rate (CAGR) from 2020 through 2026[8]. Our telco partners have indicated a strong desire to pursue SD-WAN and SASE opportunities with their business customers, and they can do so using our AWS Cloud WAN service.

AWS Cloud WAN enables enterprise customers to build and operate WANs that connect their data centers, branch offices, and Amazon Virtual Private Clouds (VPCs). With AWS Cloud WAN, they connect to AWS through their choice of local network providers, then use a central dashboard and network policies to create a unified network connecting their locations and network types. AWS Cloud WAN also generates a complete view of their on-premises and AWS networks to help visualize their entire network's health, security, and performance.

AWS Cloud WAN can be extended in partnership with telcos by:

1. Building an AWS-centric SD-WAN where telcos manage an AWS controller instead of a third-party SD-WAN controller.
2. Adding AWS's sizable backbone and global submarine cable networks to telcos regional capabilities.
3. Becoming the preferred provider to connect to AWS via peering points at AWS Cloud WAN hubs.

Further, we can use components like WAF (Web Application Firewall), NAT gateways, and Secure Web Gateways from APN—or the telcos preferred vendors—to integrate with AWS Cloud WAN hubs creating an AWS-native SASE that telcos can offer their customers. Our edge portfolio of AWS Outposts and AWS Local Zones plays a crucial role in hosting a network edge that converges not only 5G workloads but also SD-WAN/SASE. With WWC initiatives maturing (as described in technology trends earlier), the AWS edge can become an optimal and very efficient point-of-presence for several new connectivity services.

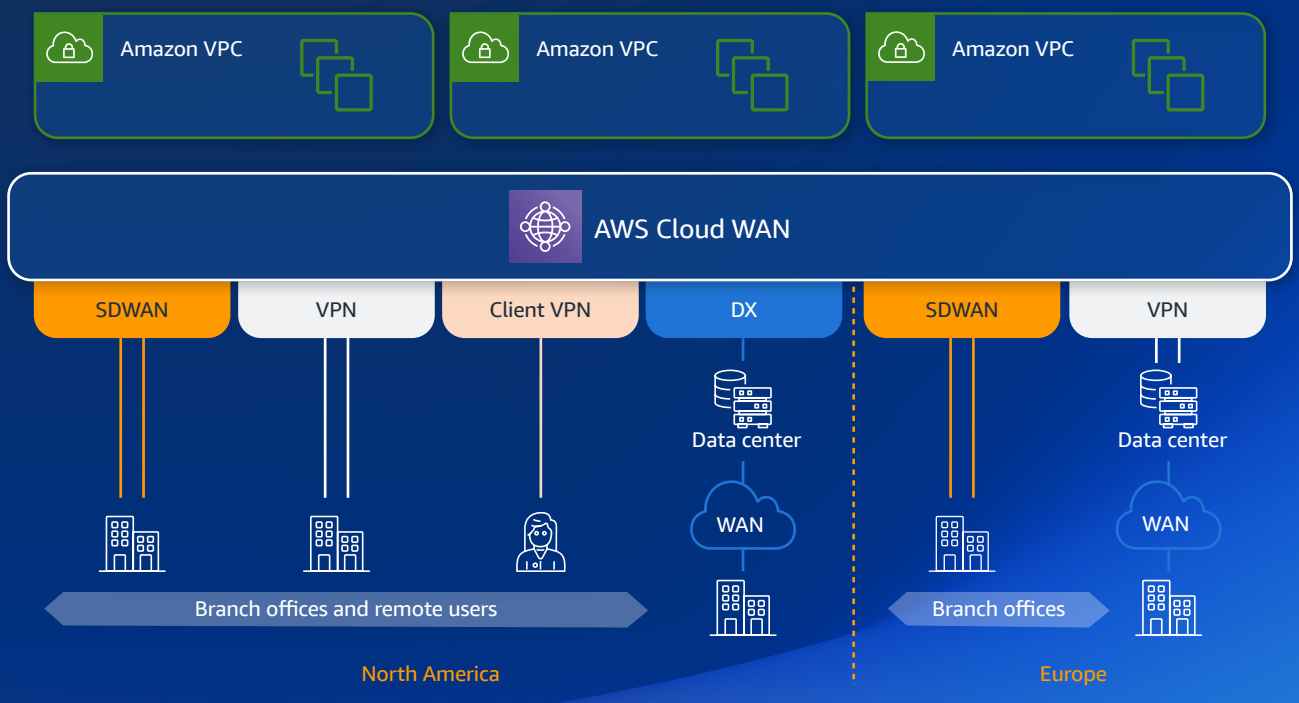


Figure 5: AWS Cloud WAN

7 Gartner Identifies the Top Trends Impacting Infrastructure and Operations for 2023 (www.gartner.com/en/newsroom/press-releases/2022-12-08-gartner-identifies-the-top-trends-impacting-infrastructure-and-operations-for-2023)
8 Gartner Magic Quadrant for SD-WAN 2022 (www.gartner.com/en/documents/4018621)

Private Wireless Networks

Private wireless offers a more secure, reliable connectivity solution compared to Wi-Fi. Private wireless can help enterprises harness AI, ML, data analytics, and IoT to enhance quality and service, drive efficiencies, lower supply chain risk, and improve customer satisfaction. These benefits have led analysts to forecast a 35.7% CAGR (2022-2026) for the private LTE/5G wireless market, achieving $8.3B in revenues in 2026[9]. To achieve those growth rates, enterprises and telcos must simplify discovering, deploying, and managing private wireless networks.

However, many enterprises have found building and operating private wireless networks too expensive, slow, and complex. To help address these challenges, we provide a range of offerings under the umbrella of AWS private wireless networks:

- **(From cloud)** AWS Private 5G – a fully managed private wireless solution delivered from AWS Cloud
- **(With cloud)** Integrated Private Wireless (IPW) on AWS – a telco partner-managed private wireless solution that combines our cloud services with telco go-to-market capability
- **(On cloud)** Customer Private Wireless — solutions offered by our ISV and network equipment partners running on AWS Cloud

Telcos and their customers can benefit from our IPW program which was co-developed with our telco partners. IPW on AWSs combines private 4G and 5G wireless technologies with AWS services across AWS Regions, Local Zones, Outposts, and Snow Family of devices. The program will also leverage AWS's dynamic community of more than 100,000 partners from over 150 countries, offering additional validated solutions that run on AWS. Customers can explore private wireless offerings from participating telcos using the Integrated Private Wireless on the AWS portal, browsing by industry or use case. Customers can contact the telco of choice from the portal, which will design, deliver, operate, and support the private wireless solution running on AWS.

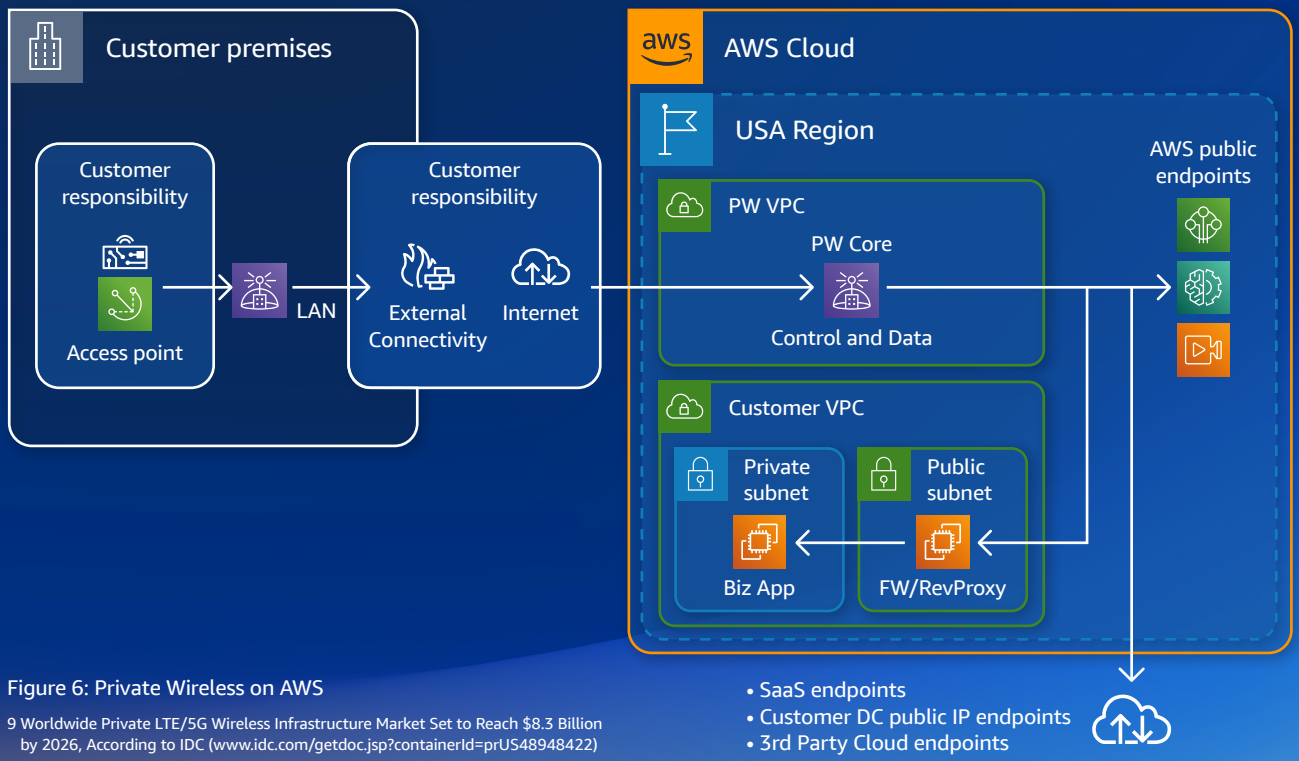


Figure 6: Private Wireless on AWS

9 Worldwide Private LTE/5G Wireless Infrastructure Market Set to Reach $8.3 Billion by 2026, According to IDC (www.idc.com/getdoc.jsp?containerId=prUS48948422)

Summary

As telcos try to reinvent themselves into digital service providers and transform their business and operational models, they are at a critical juncture in their evolution. AWS plays a central role in helping with their transformation. Over the years, we have enhanced our offerings to enable cloud to run network workloads at the performance, scale, and resiliency expected of these networks.

As a key part of this we established strong bidirectional partnerships with leading network vendors and ISVs early on, helping them refactor their solutions and embrace cloud-native principles. We have launched connectivity services in partnership with telcos to connect enterprises efficiently and flexibly to cloud resources. We are using AWS data, analytics, AI, and ML services to help telcos organize the variety of data they possess, understand their customers better, offer new customized services, and improve network operations.

While it took time and validation via proofs-of-concept, telcos worldwide have started adopting our solutions. We are seeing the increased acceptance of the value of the cloud, deployments of network functions, and migration of IT and data workloads.

aws FOR TELECOM

Ishwar Parulkar
AWS Telecom and Edge Cloud CTO

About the author

Ishwar Parulkar is the Chief Technologist for Telecom and Edge Cloud at Amazon Web Services. In this role he is responsible for setting AWS technology strategy, defining new cloud services and leading initiatives to enable AWS's edge cloud offerings and next generation telecom networks and services.

Prior to AWS, Ishwar was a Distinguished Engineer at Cisco and Chief Architect for business units responsible for mobile and wired access routing, mobile packet core, small cells and network orchestration products. He was involved in building the first LTE networks for several telecoms across the world based on these products. He was a founding member of industry-wide initiatives around mobile edge computing and 5G and looking at the technology trends emerging from those, the conviction that cloud technology can transform the telecom sector brought him to AWS in 2016. Before his stint at Cisco, Ishwar was a Distinguished Engineer at Sun Microsystems where he led the design of SPARC architecture-based data center computing technology and products, including the first multi-core processor systems and first compute virtualization platforms in the industry. He started his career at Apple, where he worked on the Mac desktop and laptop product lines and on the Newton PDA technology, which had the seeds of the iPhone revolution.

Ishwar holds an M.S. from Vanderbilt University and a Ph.D. from University of Southern California. He served for several years as a member of the Industry Advisory Board of University of Southern California. He holds several dozens of patents, has published over 25 papers in IEEE/ACM journals and conferences. Has also served as a Program Committee member and chair for IEEE/ACM conferences in networking and computing. In 2017, Ishwar was elected a Foreign Fellow of the Indian National Academy of Engineering for his excellence in and seminal contributions to the field of telecom networks and data center computing.

Read more about the four themes of the AWS for Telecom journey

1. Extending AWS to the edge

EBOOK:
Engaging the Edge - A guide for business leaders (Feb 2023)

For executives looking to transform their businesses with 5G, cloud, and the edge, learn how AWS Hybrid edge infrastructure offerings can help you identify where to start.

Download here »

BLOG POST:
Architecting multi-carrier interoperability with Edge Discovery APIs on AWS Wavelength (April 2023)

This blog post introduces a use case, an architecture, and best practices for operating a MEC application at-scale using AWS Wavelength and AWS Outposts.

View here »

---

2. Enhancing AWS cloud for telco network functions

BLOG POST:
New: AWS Telco Network Builder – Deploy and Manage Telco Networks (Feb 2023)

AWS Telco Network Builder (TNB) is a service that designed to help Communications Service Providers (CSPs) deploy and manage public and private telco networks on AWS while using existing standards, practices, and data formats.

View here »

BLOG POST:
AWS and Nokia Collaborate to Drive Operational Simplicity with Carrier-Grade Performance in Cloud RAN (Feb 2023)

This blog post describes the AWS and Nokia's Cloud RAN solution. The solution combines AWS's reliable and secure infrastructure and cloud services with Nokia's new In-Line Cloud RAN SmartNIC and Cloud Native RAN Software.

View here »

WHITE PAPER:
Open Radio Access Network Architecture on AWS (Dec 2022)

This whitepaper explores the concept of the Open Radio Access Network (O-RAN), offers a reference architecture for the O-RAN on AWS, and presents best practices for key features of the O-RAN.

Download here »

WHITE PAPER:
AWS Well-Architected Framework (Oct 2022)

The AWS Well-Architected Framework helps you learn architectural best practices for designing and operating reliable, secure, efficient, cost-effective, and sustainable systems in the cloud.

Download here »

BLOG POST:
Telco Meets AWS Cloud: Deploying DISH's 5G Network in AWS Cloud (Feb 2022)

This blog post provides details on how DISH is utilizing the AWS global infrastructure footprint, native services and on-demand scalable resources to benefit from the disaggregated nature of a cloud-native 5G Core and RAN network functions.

View here »

BLOG POST:
Ericsson and AWS Partner to Support CSPs on Their Journey to Cloud BSS (May 2021)

Ericsson and Amazon Web Services (AWS) are joining forces to deliver increased agility and efficiency by offering the Ericsson business support systems (BSS) portfolio on AWS.

View here »

WHITE PAPER:
Next-Generation Mobile Private Networks Powered by AWS (Jan 2021)

This whitepaper introduces the relevant use cases, solutions, and best practices for designing and deploying mobile private networks powered by AWS.

Download here »

---

3. Transforming telecom data into intelligence

BLOG POST:
Limiting Subscriber Churn by leveraging real-time subscribers' feedback (Nov 2022)

This blog series demonstrates a fully serverless approach to building a sentiment analytics and customer engagement solution for communication service provides (CSPs) on AWS

View part 1 here » **View part 2 here »**

BLOG POST:
A Modern and Simple Approach to Address CSP's Network Performance Analytics Challenges Using AWS (April 2021)

This blog post explains how a serverless network analytics using AWS, enables CSPs to gain insight on performance data associated with new network services.

View here »

BLOG POST:
Data monetization and customer experience optimization using telco data assets (Nov 2020)

This two-part blog post, demonstrates a working solution with an AWS CloudFormation template for how a TSPP can use existing data assets to generate new revenue streams and improve and personalize CX.

View part 1 here » **View part 2 here »**

BLOG POST:
Detecting and visualizing telecom network outages from tweets with Amazon Comprehend (June 2020)

This blog post shows how to classify tweets in real time so telecom companies can identify outages and proactively engage with customers by using Amazon Comprehend custom multi-class classification.

View here »

---

4. Co-inventing connectivity with the telcos

EBOOK:
Private Wireless and Cloud (Feb 2023)

This eBook describes how AWS and Telco Partners can simplify, accelerate, and scale their private wireless deployments for enterprises.

Download here »

BLOG POST:
AWS Teams Up with Leading Telcos to Launch the 'Integrated Private Wireless on AWS' Program (Feb 2023)

This blog post describes the new AWS private wireless program, and how it brings together AWS's reliable and secure infrastructure and services, with 5G/4G LTE networks from leading telcos.

View here »

BLOG POST:
Cloud WAN - A Managed WAN Service (July 2022)

AWS Cloud WAN is a network service that makes it easy to build and operate wide area networks (WAN) that connect your data centers and branch offices, as well as multiple VPCs in multiple AWS Regions.

View here »
