A Converged Network Design for Flexibility and Service Evolution

Reduce network complexity to save time and money

Contents

Services are outpacing current network capabilities

With service revenues flattening from price compression, Communication Service Providers (CSPs) are faced with a difficult balancing act. They need to design and build a network that supports in demand, Service Level Agreement (SLA)-based service offerings while still maintaining network support for existing services. With these new services, networks need to be more dynamic by integrating real-time analytics to monitor network pathways for congestion and optimally utilize network assets to support the SLAs. A new network architecture is needed to simplify the delivery of mobile, business, and residential services on one platform to give CSPs one network to operate.

Just getting by is no longer good enough

Over the past 20 years, Communication Service Provider (CSP) networks have scaled to address an insatiable desire for bandwidth to support streaming services, online gaming, and cloud-based applications. During 2020, the impact of the COVID-19 virus turned a general appreciation of the Internet into a critical dependence, as the collective global community transitioned to remote connectivity for work, education, essential goods, and healthcare. Even if their networks experienced a significant increase in traffic as well as a change in traffic patterns, network providers successfully delivered all these new digital experiences at mass scale, making them the heroes of this story as they were able to maintain critical services for all of us across the globe.

While connectivity did not break under this extraordinary demand, managing and operating the network became more difficult and expensive. Years of piling new technologies on top of existing technologies created a complex operational model and any infrastructure added in support of the demand growth only exacerbated the complexity, increasing the chance for errors or failures and placing additional strain on network operations. Moreover, with the uptake of telco cloud services, the network must be highly flexible to support SLA-based services delivered from distributed locations closer to the end users. Placing the services closer to the end user is optimal because it reduces total network traffic and packet latency to offer an unprecedented customer experience. This distributed design creates a strong dependency between the transport network and the telco cloud hosting locations and drives the need for access and aggregation platforms to support advanced capabilities such as network slicing to provide differentiated performance tiers across the network.

Using a Converged Software-Defined Network Transport (Converged SDN Transport) architecture will help CSPs reduce the overall complexity of their network and operations. It also allows them to build the necessary analytics and automation tools into their network to satisfy the stringent performance needs of today's new service offerings.

What is Converged SDN Transport?

This architecture focuses on converging network infrastructure in multiple dimensions to change the way networks are built. Simplifying the network starts by reducing the number of 'moving parts' as much as possible. There are five key areas where convergence efforts can be focused:

- Converge Access: Bring residential, mobile, and business services on to a single infrastructure platform for an access-agnostic connection to services and converge distinct subscriber management solutions under a unified, cloud native platform.

- Converge Software: Leverage an end-to-end network operating system for high-quality, agile, and consistent operations.

- Converge Protocols: Reduce network protocols down to segment routing for the transport protocol, and to Ethernet Virtual Private Networking (EVPN) for the service protocol to provide advanced capabilities such as network slicing, improved resiliency, and service chaining

- Converge Network Operations: Simplify operations with an automation platform to support service delivery across the entire lifecycle that will reduce time spent on planning (Day -1), activation and implementation (Day 0), and ongoing support during operations (Day 1+).

- Converge Optics and Routing: Integrate 200G/400G coherent optics directly in the router to simplify operations and improve capital efficiency. Traditional optical services can be emulated over a single IP transport network grounded on a hop-by-hop network architecture to improve existing service profitability.

Converged SDN Transport benefits

Use convergence to lower complexity and cost

With the quantum leap in capacity and density available in new routing systems, CSPs can build a network that will support a convergence of services (residential, cable, mobile, enterprise, or wholesale) over a unified infrastructure. Additional convergence comes from advancements in coherent optics to place transponder functionality into a pluggable form factor. This allows CSPs to begin removing transponder equipment and reduce the number of physical devices needed, removing operational complexity with a more efficient network design.

To build a consolidated infrastructure capable of meeting diverse SLA requirements for their services, CSPs need to support end-to-end network slices that will satisfy these SLAs. Using telemetry data collected from the network, latency performance can be monitored and if there's any change due to higher congestion, closed loop automation will automatically reroute traffic without the need for human oversight to manually configure the changes.

Segment routing is foundational to implementing network slicing at scale as it can steer traffic on stateless, flexibly defined network paths and has the benefit of being programmed either by an SDN controller or locally by the head-end, or source router. Since it's fully stateless and scalable, many SLAs can be simultaneously enforced across multiple access types with wired, wireless, and business services all converged on the network. Automated provisioning of network slicing powered by segment routing offers the ability to logically instantiate a private network service within the larger communication service provider network and enables CSPs to deliver secure, high-speed, low-latency content to their customers.

Likewise, EVPN provides an optimal user experience by

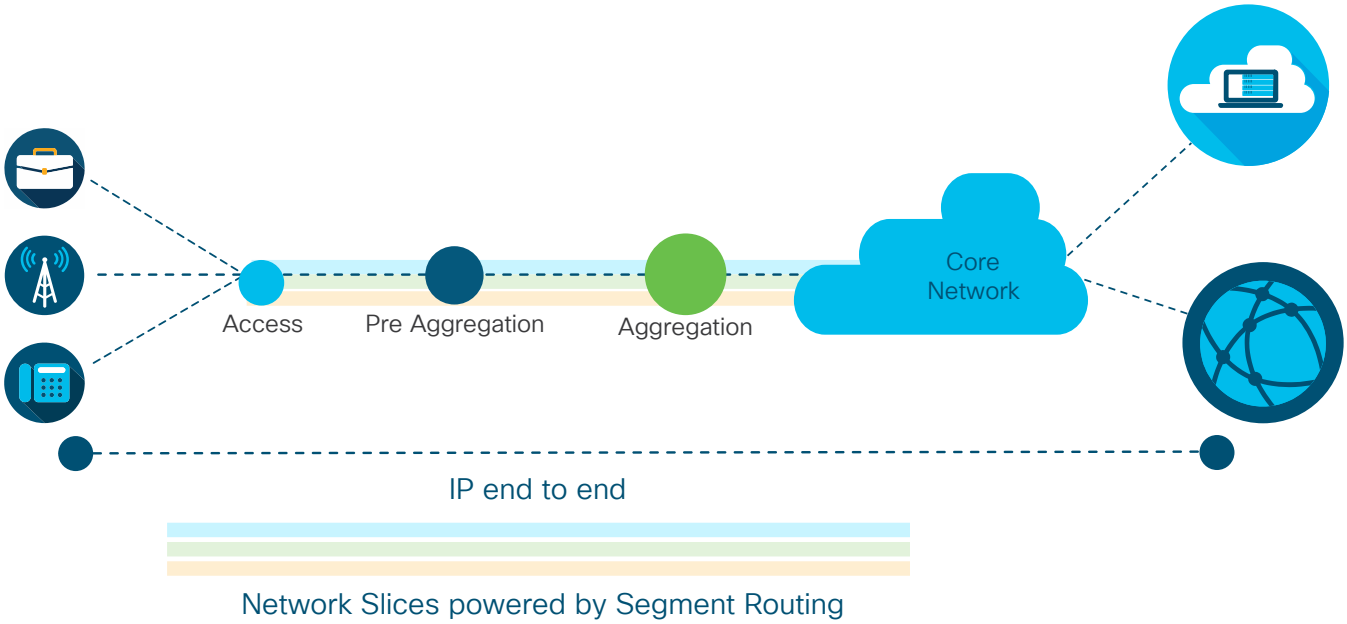

Figure 1. End-to-end IP allows you to build specific experiences for clients with network slicing and segment routing.

using SLA-based forwarding to minimize latency and maximize bandwidth. EVPN policies can extend from the data center down to the metro/access level, further simplifying the networking stack. From a management and operations perspective, EVPN speeds up service recovery and restoration and reduces day-to-day configuration and operations overhead, using a turnkey network automation platform, allowing CSPs to fast track the provisioning and delivery of Layer 2 and Layer 3 VPN services.

Distributing services closer to end users requires network operators to evolve from a traditional Broadband Network Gateway (BNG) design running on a single router, to a disaggregated solution based on Control and User Plane Separation (CUPS) design. Using CUPS, a centralized subscriber management policy runs on a scalable cloud native control plane and the forwarding functionality is delivered by the user plane. This distributed design enables simplified operations, efficient scaling, and paves the way for the convergence of wireless and wireline services with unified CUPS architecture. A disaggregated BNG architecture has several key advantages such as independent user plane and control plane scaling, independent control and user plane life cycle management, and a centralized control plane for unified configuration.

Built on a cloud native infrastructure stack, a cloud native BNG control plane is hardware-agnostic and compatible with virtualized as well as bare metal systems and portable across multi-cloud environments. The cloud native BNG architecture is targeted for faster rollout of features and upgrades, with a view of operators moving to Continuous Integration and Deployment (CI/CD) model. The cloud native BNG Control Plane architecture simplifies the Session Redundancy Group (SRG) mechanism and supports multiple redundancy mechanisms – within the cluster as well as across clusters with geo-redundancy.

In CUPS, the user plane is freed up of subscriber state management overhead and the architecture gets simplified, allowing user planes to be more distributed, with minimal integration efforts. A Service Provisioning Agent (SPA) is used to encapsulate and decapsulate the communication messages for the standardized interfaces between user plane and control plane. SPAs are generally bundled with network operating system images so that there is a minimal configuration requirement to transform from an integrated BNG router into a cloud native BNG user plane. This serves as investment protection for the already deployed integrated BNGs in the network.

Remove layers of complexity

Existing networks rely upon IP and optical network layers that are commonly operated separately by different teams. Despite several attempts to implement multilayer optimization solutions, CSPs are still grappling with Operating Expenditures (OpEx) increases and inefficient use of Capital Expenditures (CapEx) exacerbated by different upgrade cycles for each layer.

But the situation is changing rapidly as functions traditionally delivered in separate chassis-based

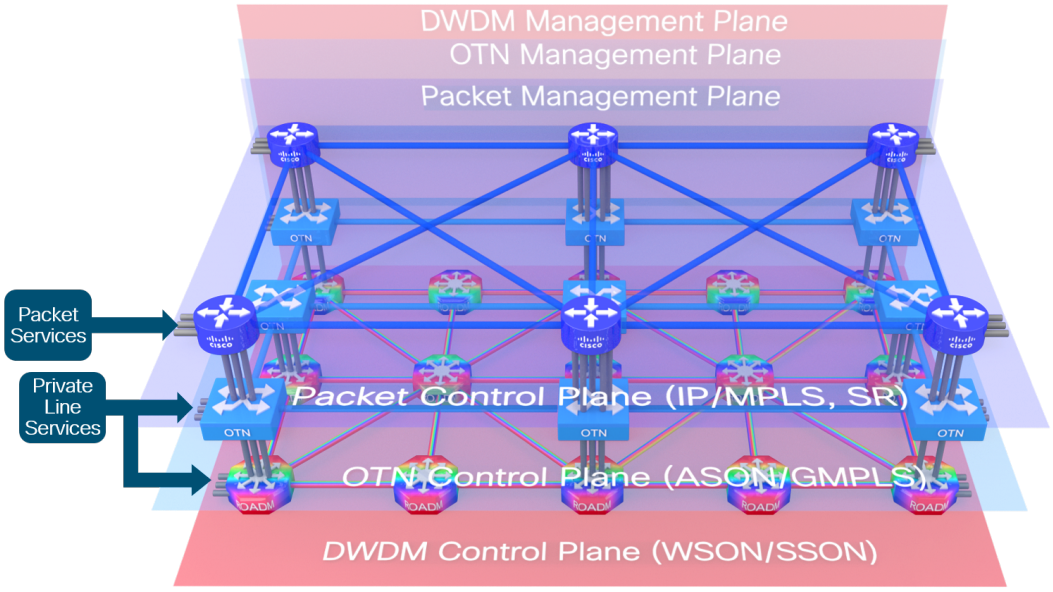

Figure 2. Traditional networks have three layers to cover the IP to optical conversion.

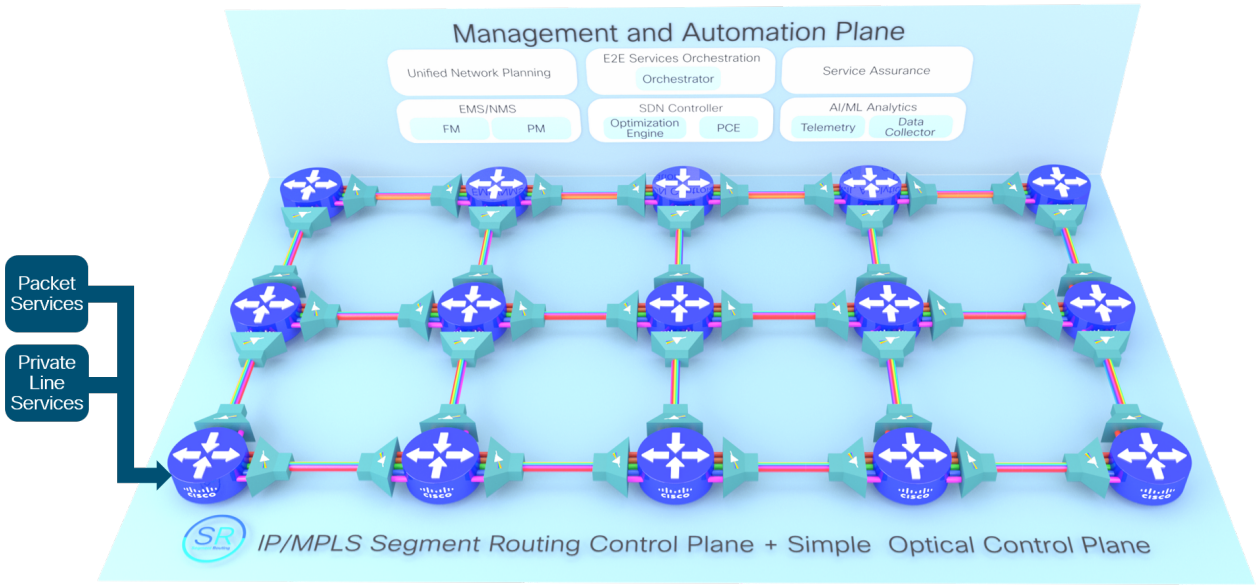

Figure 3. Routed optical networking convergence brings the IP-to-optical conversion into a simplified structure.

transponder solutions can now be delivered in a pluggable form factor embedded into routers, enabling CSPs to consider collapsing network layers to reduce operational complexity.

By bringing standardized 400G ZR/ZR+ optics and massive routing scale together with pluggables, the Routed Optical Networking solution makes the network leaner by removing transponders and replacing them with 100G, 200G, and 400G ZR standardized coherent optics. As a result, the IP and the optical layers – that have always operated separately in networks – can be unified, greatly improving network economics.

A phased approach starts by rolling out digital coherent optics that can directly plug into routers to save on transponders. Later, operators can migrate legacy Time-Division Multiplexing (TDM) and wavelength services to a full IP infrastructure with Private Line Emulation (PLE) solutions resulting in less of a need for Optical Transport Network (OTN) switches. Finally, operators can fully collapse IP and optical networks by entirely removing Reconfigurable Optical Add-Drop Multiplexers (ROADMs) and have routers connected hop by hop over point-to-point simplified Wavelength Division Multiplexing (WDM) links. This overall simplification reduces the number of moving parts, removes management complexity, and serves to drive costs out of the network.

Maintain control and improve security response times with automation

Migrating to a simplified, converged framework enables CSPs to improve operations by further integrating automation tools across their network. With an end-to-end IP network, telemetry data can be consumed from anywhere, helping to detect and respond to congestion or other network events before they become a major issue. Network operations teams can take early actions to achieve faster remediation and ensure service SLAs are guaranteed, resulting in a better end-user experience.

Using automation to help manage your network architecture has inherent operational cost benefits:

- Improved network resiliency with congestion management threshold controls to shift loads during peak usage times or maintenance and outage events
- Network maintenance is simplified because nodes can be programmed in or out of service without disrupting traffic
- Cloud-based test environments allow engineers to stress-test new software patches and configuration designs before deployments
- Faster and more consistent network updates through orchestrated rollouts and zero-touch provisioning

In addition to the benefits listed above, automation can be used to improve a CSP's network security. To support the scale and agility required to manage the dynamic network and maintain a proper security posture, security teams need complete and continuous visibility into the network. The automation of management tasks such

as patch and policy updates reduces costs and has a positive impact on operational experiences. Providers must integrate automation solutions to simplify network security with data collection and analysis to address the most critical security concerns.

Base your network on a trustworthy infrastructure

Today's network backbones are the critical infrastructure for the successful operation and growth of a nation's economy. Without access to high-speed connections and trusting those connections to be secure, business operations and transactions would halt. The only way to deliver a secure network is to establish trust in the core infrastructure. A trusted system is based on clear criteria which can be measured, verified, reported, and audited.

A trusted network component must provide assurance that:

- The hardware running your network is authentic and that processes are in place to protect network devices from taint and modification.
- The network operating system and its associated services are trusted and haven't been modified.
- The management of the network device is properly authenticated and audited.
- Any secrets stored within the network device can't be extracted or changed without authorization.

Because software can lie

Software relies on its firmware and hardware and reports back whatever it's told by that underlying architecture. Existing practices of verifying software through hashes and code signatures are an important part of establishing a trusted system, but software systems can be affected or compromised by the underlying architecture. For instance, a known-good application can't be trusted if the underlying operating system doesn't maintain effective protection and sandboxing of that process. Similarly, an operating system can be compromised through attacks based in the firmware or the underlying hardware.

Attacks on the underlying hardware or firmware of a system are commonly used to establish persistence in compromised systems after a breach. Often, they're designed to persist even after an operating system has been reinstalled. Examples of these types of attacks include hypervisor-based attacks such as "Blue Pill", which can invisibly compromise the running operating system kernel. And attacks such as "ThunderStrike" have already demonstrated persistent compromise through firmware.

Trusted infrastructure must be based on trusted hardware

Because software alone can't prove its integrity, truly establishing trust can only be done in hardware, using a hardware root of trust. To be effective, this root of trust must be based on an immutable hardware component that establishes a chain of trust at boot time. Each piece

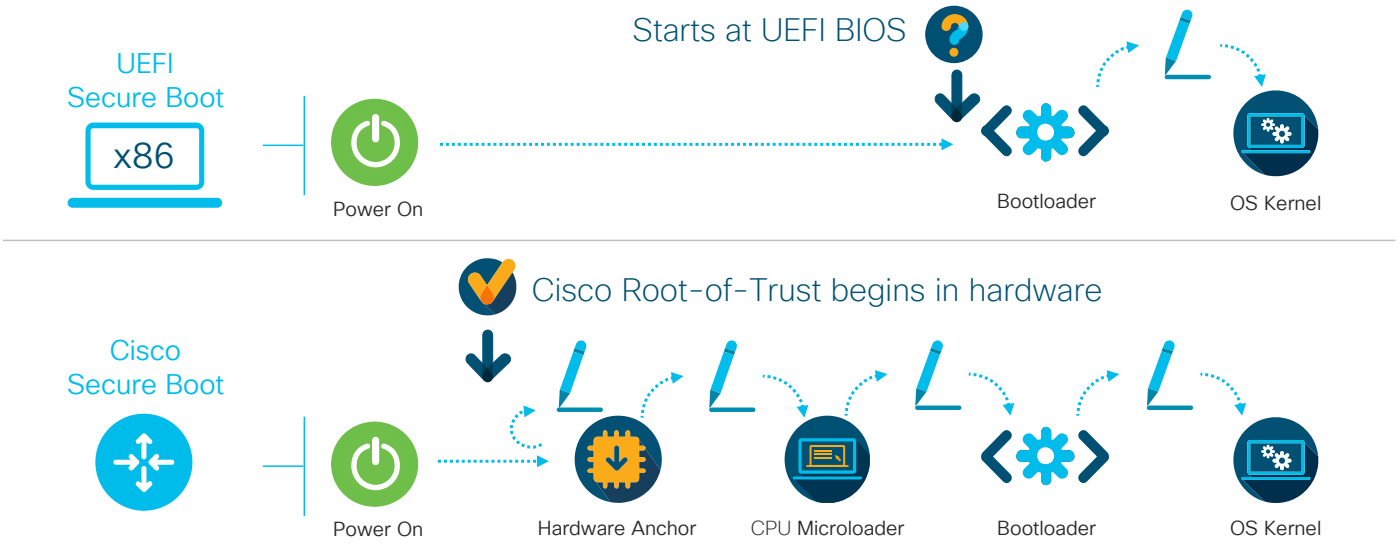

Figure 4. Comparison of a traditional server OS start to the Cisco secure boot process with a hardware-anchored root of trust

of code in the boot process measures and checks the signature of the next stage of the boot process before the software boots. Without a hardware root of trust, no amount of software signatures or secure software development can protect against a compromise of the underlying system.

Hardware root-of-trust and secure boot

Many Cisco service provider routers utilize signed images and hardware-anchored secure boot to prevent inauthentic or compromised code from booting. Anchoring the first code in the boot sequence in hardware establishes a chain-of-trust and is the foundation of the Cisco secure boot process.

Secure boot is already a familiar term in the world of x86 servers. It's used to cryptographically verify the authenticity of the OS boot loader and the OS kernel as part of the boot process. Secure boot is commonly used to protect against boot-kit attacks in server operating systems like Linux or Microsoft Windows Server. In traditional x86 server systems, this secure boot process begins in UEFI BIOS and doesn't have a hardware trust anchor.

Cisco IOS XR includes a more extensive boot process that's designed around a hardware trust anchor. This process begins before the CPU is allowed to boot and offers significant protections against hardware or firmware compromises. The Cisco secure boot process establishes an extensive chain of trust that begins in hardware. The hardware anchor implements self-measurement, followed by measurement and signature verification of the CPU microloader. It then verifies the signature of the bootloader and the OS kernel.

Secure hardware for strong cryptography

Cisco also uses a Trust Anchor module (TAm) which provides a Secure Unique Device Identity (SUDI) and other services. Within IOS XR, the TAm also provides critical cryptographic services such as:

- Unique cryptographic hardware identity (SUDI)
- Designed to be compliant with NIST SP 800-90B entropy source via hardware random number generator
- Secure generation and storage of secret key material

Trust, but verify

'Trust, but verify' isn't just a proverb. It's critical to verify hardware, firmware, and software components of a system at boot time to establish trust, but it's equally important to provide visibility and audit capabilities for trusted systems. Secure boot processes can provide valuable protections against compromise of the OS or underlying components, but once it's complete there needs to be an external reporting system that can prove a non-compromised environment. Trusted systems require effective measurement and trust posture reporting from external systems, where measured values from the system can be compared against "Known-Good Values" (KGV) for that system.

Secure measurement

Hardware root-of-trust is critical to establishing trust in a critical system and operators should require that networking devices include access to external mechanisms to securely record and store measurements taken during the boot process. Values measured at boot time in hardware should be securely recorded into configuration which can provide cryptographic proof that values have not been altered and accurately represent what was measured at boot time. Operators can enhance their value proposition to potential subscribers by leveraging their critical infrastructure with externally validated trust anchor measurements proving the integrity of their infrastructure systems.

In addition to recording these values at boot time, secure measurement can be extended to measuring individual processes within the OS at run-time. It's important to extend trust measurement to the runtime environment as well to allow for comparison against baselines. IOS XR will be implementing advanced capabilities to measure processes and critical files at runtime and compare these measurements against KGVs from Cisco's IOS XR build process.

A trusted infrastructure needs an external service for comparing all measurements to KGVs for boot integrity or individual process measurements. By providing secure measurement and reporting of measurements, a trusted system can enable remote attestation, or the ability to cryptographically prove that measured values are accurate and compare these to KGVs on an external system. This approach is the most effective way to provide visibility, trust posture assessment, and auditing of trustworthy systems.

Outcomes to expect

A Converged SDN Transport architecture will help CSPs

Learn more

To learn more about these solutions, visit the Cisco Converged SDN Transport page.

For additional information on the economic benefits of network automation, network slicing, and cross domain orchestration for 5G, please use the following resources:

ACG Research Paper: Economic Benefits of Network Automation

Analysys Mason: 5G network slicing: Cross domain orchestration and management will drive commercialization

migrate to a scalable network prepared to meet the stringent bandwidth and performance demands for the 5G era and beyond. At the most basic level, the primary goal for a Converged SDN Transport architecture is to drive simplification into the service provider's network and operations.

Implementing the design will allow service providers to realize the following business benefits:

- Increased service agility with a simple, resilient, and secure network platform that can fast-track the delivery of innovative services

- Unlock investment efficiencies through industry leading technologies that enable real-time network adaptations to optimize asset utilization

- Achieve operational excellence by integrating closed-loop automation tools that will reduce mean time to remediation and time to value

- Establish a trustworthy infrastructure that can be audited and verified, ensuring the integrity of the transit routes for sensitive traffic

This pandemic could eventually serve as a blessing for network providers as the catalyst to reimagine their network architectures with a different mindset where simplicity, trust, and value creation are key architectural tenets. No one can predict Black Swan events and as a result, the network needs to be ready to cope with the unknowns. The network must be more scalable, cost-effective, power-efficient, agile, resilient, automated, and trusted than ever before. The status quo won't work, but with the Converged SDN Transport architecture serving as a blueprint, network providers will remove complexity from their network and focus on enabling the next service revolution.

Moving forward

Cisco works to improve network platforms and assists CSPs in operating critical infrastructure needed by today's global connectivity. Flexibility, resilience, and trustworthiness are essential to modern, critical service provider networks, but doing so without adding complexity into operations or management is key. Current networks need to be simplified to help operational teams support legacy services while providing advanced technologies to build new, high performing offerings. Using the Converged SDN Transport framework will enable CSPs to remove complexity from the network and converge multiple services onto a single network infrastructure that's prepared to scale for future growth.

Footnotes: ACG Research paper: TCO Benefits of Converged 5G Ready IP Transport, https://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/mobile-internet/acg-tco-benefits-of-converged-5g-ready-ip-transport.pdf