The Driving Factors behind the Telecommunications Shift to Cloud Metro Networks

Peter Fetterolf, Ph.D.

EXECUTIVE SUMMARY

The telecommunications landscape is undergoing a transformative shift with the emergence of cloud metro networks. This whitepaper discusses the driving factors behind this evolution, emphasizing the migration of services and applications to the network edge and driving the need for cloud metro networks. The rapid growth in metro network traffic and bandwidth necessitates an architecture capable of scaling both up and out, which is where the cloud metro architecture plays a pivotal role.

At the core of this transformation is the need for high scalability, flexibility, and sustainability. Edge network capacity is increasing, creating a demand for service and application intelligence. Innovations such as IP over DWDM (IPoDWDM) and SFP-Formfactor Coherent DWDM pluggables (ZR/ZR+) are making packet-optical convergence more efficient than ever. Additionally, the integration of AIOps and experience-first, intent-driven automation aligns with the principles of hyperscaler-based design, advocating for resource pooling, automated operations, and a flexible, scalable spine-leaf architecture. Intent-based automation allows for a shared IP-service fabric that allows communications service providers to design, build, and provision new services faster over a heterogeneous network. Service agility is the key to new revenue growth. Juniper's Paragon provides a good example of intent-based automation, which is discussed in more detail in this paper.

The cloud metro provides a migration path for traditional Ethernet-based ring networks, introducing a new intelligence layer to support next-generation (NG) services and new revenue streams. It champions the support of both legacy and NG protocols, enabling operators to transition at their pace while tapping into innovative opportunities offered by 5G and hyperscaler methodologies.

The whitepaper further explores the new edge services and revenue opportunities enabled by cloud metro networks: video surveillance, network slicing for enterprises, AR/VR applications, cloud gaming, and the burgeoning sector of e-sports as they stand to benefit significantly from the low latency, high bandwidth, and enhanced security provided by cloud metro architectures.

The network requirements that drive the cloud metro network architecture will be outlined, reflecting on how these networks are not merely a technological upgrade but a strategic foundation for future-proofing business models and service offerings in the fast-evolving digital economy.

Key Drivers of Cloud Metro Network Evolution

The concept of the edge in network architecture is evolving. Traditionally, the edge was considered the access layer, but with the advent of cloud technologies and the explosion of data, the metro network is now becoming the new edge. This shift is driven by several factors:

- The decentralization of computing resources, necessitating closer proximity to end users to reduce latency.
- An increase in mobile devices and the Internet of Things (IoT), which generates vast amounts of data that need to be processed quickly and efficiently.
- The push for greater bandwidth and the proliferation of fiber optics in metropolitan areas.

As both mobile and fixed broadband network functions become disaggregated and virtualized, many network applications and services are moving closer to the edge.

Some of the key applications and services driving the need for edge computing are:

- **Cloud RAN:** Cloud Radio Access Networks (RAN) are transforming the telecommunications landscape by disaggregating the RAN BBU and allowing both CU and DU functions to be distributed to the cloud metro edge network.
- **Distributed UPF:** The distributed User Plane Function (UPF) is essential for supporting the massive data throughput and low latency required by modern applications. UPF functions have traditionally been centralized. However, now that the packet core is virtualized and disaggregated, the UPF function that forwards users' traffic is being moved to the network edge.
- **CDNs at the edge:** Content delivery networks have traditionally been implemented at the edge to provide faster content delivery and improve users' experiences.

- **New edge services:** This refers to the various new applications that can be enabled by edge computing, such as video surveillance, network slicing, and AR/VR, which will be discussed in the section of this paper focusing on new service revenue.

Changes in Network Traffic

The emergence of edge applications is having serious implications for network traffic patterns. More traffic is staying in the metro due to the distribution of virtual service instantiations across an increasingly distributed cloud architecture. This leads to more traffic terminating in the far-edge or regional-edge data centers while, at the same time, traffic patterns are becoming more bursty and unpredictable.

ACG Research has forecasted network traffic at the metro, regional, and national levels. Specifically, we have forecasted mobile traffic per user and residential broadband traffic per household for metro, regional, and national networks. These monthly traffic projections are presented in Figure 1 and Figure 2. This analysis shows that over the next five years, traffic will continue to grow, but metro traffic will grow at a higher rate than regional traffic, and regional traffic will grow at a higher rate than national traffic. This is because the key drivers of network traffic are the edge applications described previously in this paper. These changing traffic dynamics have serious implications to metro edge network architectures.

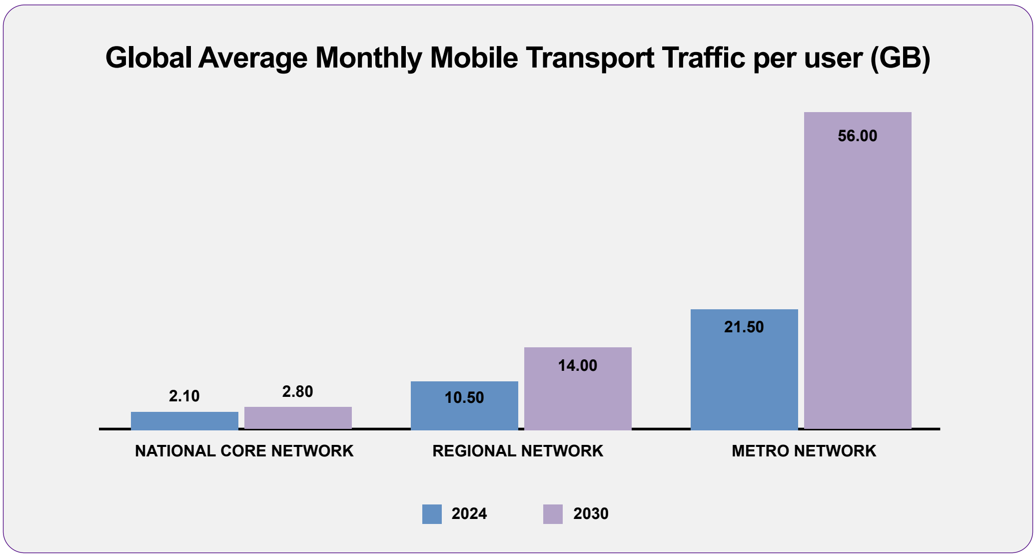


Figure 1. Global Average Monthly Mobile Traffic per User (GB)

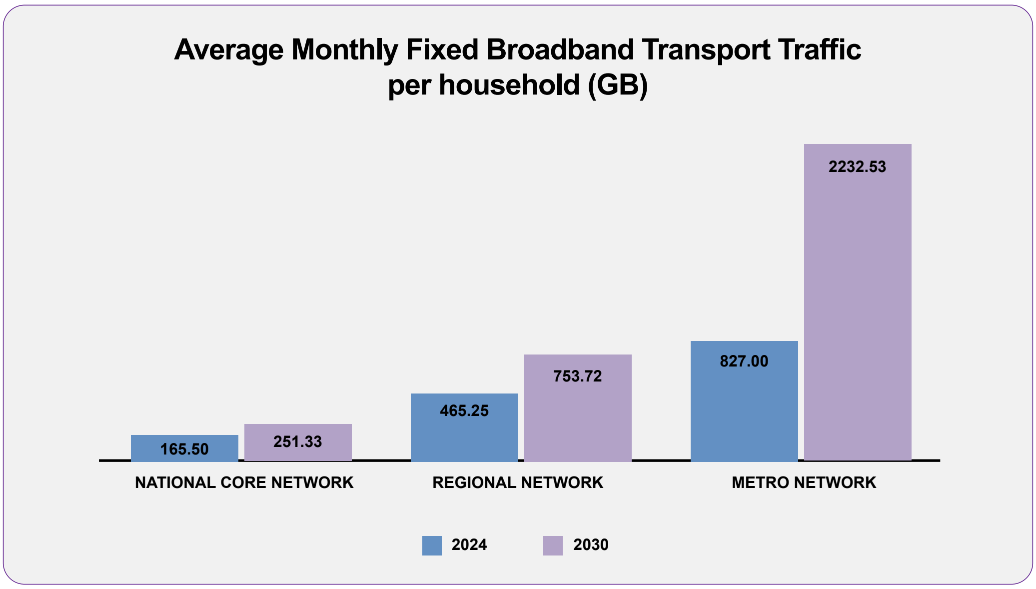


Figure 2. Average Monthly Fixed Broadband Traffic per Household (GB)

Evolution of the Metro Network Architecture

As traffic volumes and bandwidth requirements grow, there is a pressing need for a metro network architecture that can scale. The cloud metro architecture offers the solution, allowing networks to scale up (in capacity) and out (in reach) to meet current and future demands. Networks also need to be flexible enough to allow for the rapid introduction of new services. Many of these requirements are being driven by the deployment of 5G networks, which are driving new network architectures and new services. Some key requirements for the new cloud metro network architecture to enable these capabilities are:

Scalability and flexibility: The metro network must support a wide range of services and adapt to varying traffic patterns. Cloud metro networks are designed to be highly scalable and flexible, meeting the needs of different applications and services.

Sustainability: With increasing attention on environmental impacts, sustainability is a critical consideration. Cloud metro networks contribute to sustainability goals through energy-efficient designs and operations.

Increased edge network capacity: As services move to the edge, the capacity of metro networks must increase. Cloud metro provides a flexible architecture for scaling up and out to allow for future network growth.

Service and application intelligence: Intelligence at the edge is not just about processing power; it is about making smart decisions. Cloud metro networks facilitate intelligent services and applications by providing a flexible network and cloud architecture at the edge to enable faster deployments of edge applications and services.

IPoDWDM innovations: Technical advancements in IPoDWDM, particularly the integration of SFP-Formfactor Coherent DWDM pluggables (ZR/ZR+) and their role in efficient packet-optical convergence, are bringing the long promise of more efficient packet-optical convergence to reality. This is helping increase network scalability.

AIOps: The implementation of AI in network operations (AIOps) is revolutionizing network management. AIOps dramatically reduces operations labor and human errors while improving network reliability and quality of experience.

Experience-first, intent-driven automation: Automating network operations with an emphasis on users' experiences and business intent is critical for modern networks. Intent-driven automation focuses on provisioning services and functions based on users' intent, not detailed technical specifications.

Service agility: Intent-based automation allows for a shared IP-service fabric that allows communications service providers to design, build, and provision new services faster over a heterogeneous network. Service agility is the key to new revenue growth.

Spine-leaf cloud architecture for metro networks: The spine-leaf architecture used in modern data centers is essential for creating scalable and robust metro networks. Cloud metro networks apply spine-leaf architectures and other cloud data center concepts to the metro network.

Trends in Network Automation

The evolution of metro network architectures and the requirements for scalability and flexibility in increasingly complex networks drives the need for next-generation network automation. Most network operators believe automation is necessary to remain competitive and support the increasing scale of operational demands. However, most operators still resort to mostly manual and semi-automated network operations. Next-generation network automation should cover all aspects of the network life cycle within a common platform architecture, with use cases spanning Day 0 through Day 2, such as those depicted in Figure 3.

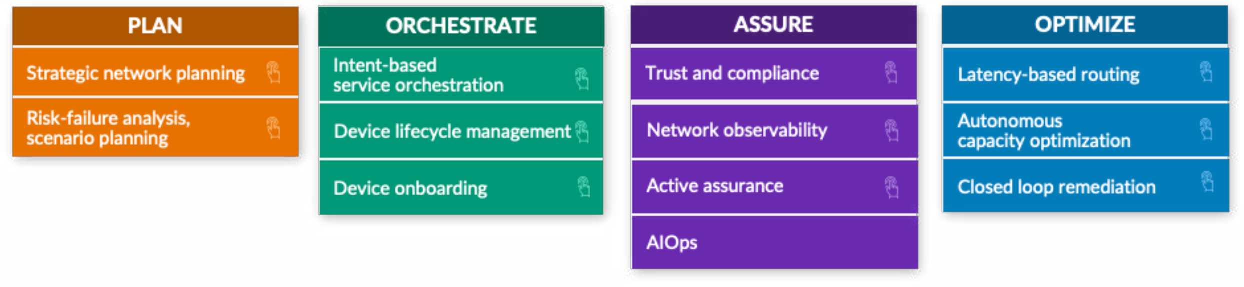


Figure 3. Life-Cycle Network Automation

Network planning and engineering includes the design of optimized network architectures and topologies, supported by network simulation for performance and risk analysis to understand how and where to deploy capacity to deliver the optimal combination of resilience and CapEx efficiency.

Network orchestration requires the adoption of intent-based networking principles, both in the resource layer (the onboarding and management of physical equipment) and the network service layer (the design, provisioning, and management of network services over the physical infrastructure). Today, service orchestration tends to rely on difficult to find expertise, and greater programmability in networks has meant that the same network service can be delivered in many ways. Knowledge is therefore limited not just to a few domain experts, but to those familiar with how a specific service was designed. A clear symptom of this status quo is that human error continues to lead to a large proportion of service outages and brownouts today.

In intent-based orchestration systems, conversely, the user specifies at a high level the service parameters, such as locations, redundancy, speeds, and capacity (the declarative statement of intent). The orchestration engine uses predefined service models to allocate, translate, configure, verify, and automatically ensure that service. Intent-based service orchestration is:

- Model and API driven
- Provides user friendly GUIs
- Abstracts network complexity, reducing the requirements for expert skill sets
- Repeatable, predictable, and error free
- Provides integrated service testing and monitoring

When services are running in production, active assurance becomes a critical component of guiding network automation to maintain service intent. Active assurance should provide:

- Both active testing and monitoring based in synthetic traffic, approximating real traffic
- A broad set of deployment options, including through test agents integrated into routers
- Support for remotely controlled or managed testing with no physical equipment or field technicians
- Prebuilt tests and monitors, configured based on the business intent for a given service
- Support for simple self-service creation of tests and monitors

Continuous monitoring combined with advanced traffic engineering allows for both closed-loop and open-loop remediation use cases, including:

- Capacity management
- Congestion management
- Latency management
- Routing optimization (for example, to minimize power consumption)

With optimization, the systems should allow for the recommendations and implementation of changes based on users' feedback. An example of a fully automated network life cycle as provided by Juniper's Paragon is depicted in Figure 4.

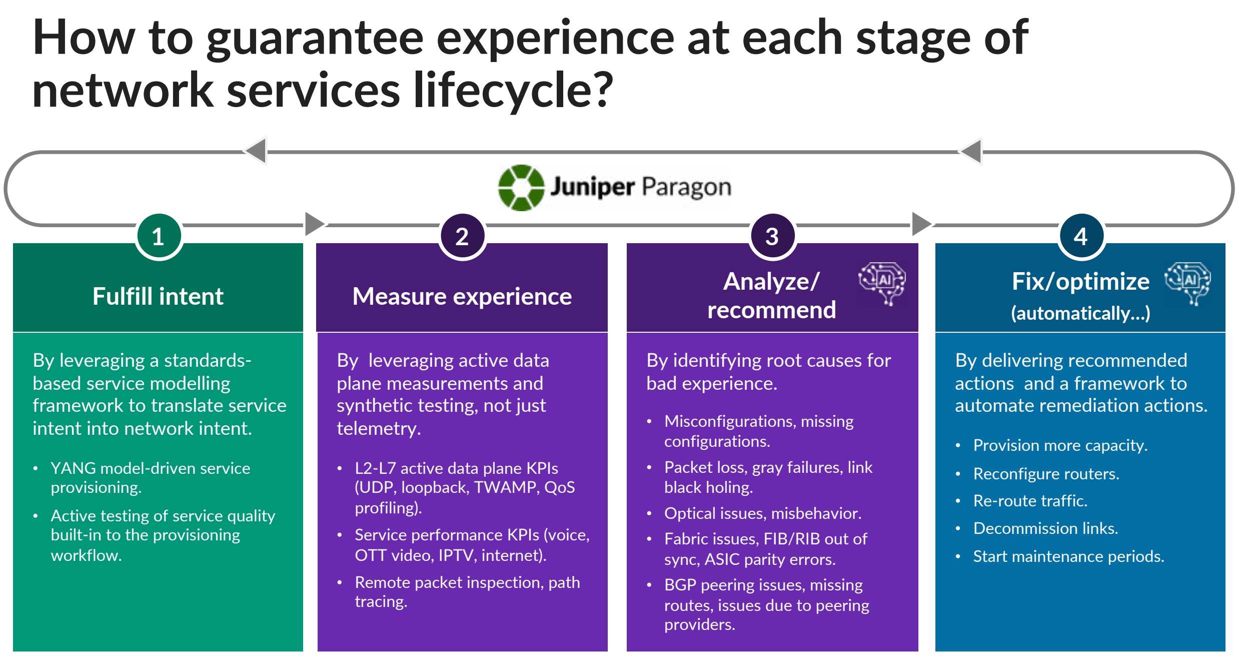


Figure 4. Guarantee Users' Experience at Each Stage of the Life Cycle

Implications for the Metro Network

Most metro networks, today, are built using rings and are designed to carry traffic from access nodes across an aggregation network to regional data centers. Traffic is typically best-effort internet or MPLS business traffic. Traffic is sometimes throttled when monthly traffic volume limits are exceeded, but quality of service (QoS) is not typically applied at the application and service levels.

Future metro networks need to evolve from fixed, ridged networks to flexible, service-aware cloud networks. Metro networks need to have cloud scale service agility. Edge data centers support resource pooling for multiple edge applications, including vDU, vCU, UPF, MEC, and others. Networks need to have elastic scalability to support flexible edge resource pools. This elastic scalability means that new routers and network architectures with scalable bandwidth, real-time monitoring and control, and end-to-end automation must be implemented.

Most access rings are 10 Gbps. Access rings must grow to 10/25/100/400 Gbps and aggregation nodes need to support 400/800 Gbps. This is a dramatic change from the metro networks in place today. To support this dramatic growth in traffic, we see two types of architectures coexisting:

1. Metro ring networks
2. Metro spine-leaf networks

These network architectures are depicted in Figure 5. In some cases, rings will provide adequate bandwidth and scalability to meet emerging edge and traffic requirements. However, in some dense urban metro areas, network traffic will drive the need for a metro network with a spine-leaf topology. The spine-leaf architecture provides a higher level of scalability, flexibility, and resiliency—an approach that has been used in cloud data centers for many years. These network topologies can and will coexist, and service providers need to use routers that have the flexibility to support different network topologies.

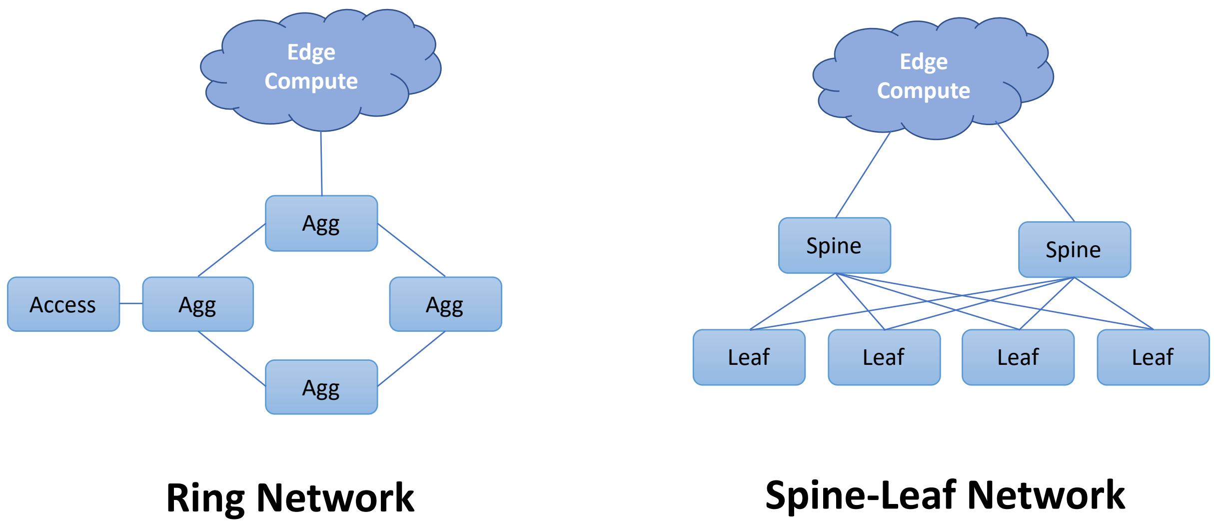


Figure 5. Two Network Topologies: Rings and Spine-Leaf

Routers must also be able to support network slicing, which requires a combination of virtual private networks (VPNs) and QoS for different services and applications. For example, front haul service between an RU on a cell site tower and a DU in a far edge node will require guaranteed bandwidth and latency. Similarly, real-time control applications used for robots and drones will also have strict latency requirements. Premium video traffic will need guaranteed bandwidth while many other internet applications will use best effort. Some of the key routing technologies required to support these requirements are:

- Segment routing
- IPV6
- L2VPN
- L3VPN
- Weighted fair queuing

Next-generation cloud metro network platforms must:

- Support legacy Ethernet protected rings, as well as next generation spine-leaf architectures
- Add service intelligence required to support next-generation services and new revenue streams
- Support existing and next-generation network protocols, allowing operators to upgrade network elements in their service fabric immediately while enabling them to evolve to next-gen protocols and capabilities selectively at their own pace

Hyperscaler-inspired design principles

It is important to clarify that cloud in cloud metro does not imply a wholesale migration of metro functions into the cloud, a concept that may concern operators. Instead, the term signifies the adoption of hyperscaler best practices within metro networks. These practices include resource pooling, automated operations, and flexible, scalable architectures such as spine-leaf designs, as well as measured services for proactive monitoring and active assurance. The cloud metro takes cues from the efficiency and simplicity of hyperscaler and 5G designs. This includes simplifying protocols to remove unnecessary complexity, streamlining data flows for better performance, and other enhancements that will be detailed in subsequent sections.

New Edge Services and Revenues Enabled by New Cloud Metro Networks

Cloud metro networks provide a flexible, scalable, and secure edge infrastructure that will enable a new generation of edge services. Many of these services, especially those targeted at enterprise customers, will allow service providers to grow revenue beyond the existing consumer and enterprise revenue for the current generation of services. Some examples of these new services are:

Video surveillance and smart cameras with event recognition at the edge: Video surveillance is evolving from passive recording to active event recognition. With cloud metro networks, video event recognition and processing can occur directly at the edge, reducing the need for continuous long-distance data transmission and enabling real-time responsiveness. CSPs can collaborate with partners to offer video edge services to enterprise customers, using network QoS guarantees and edge computing hardware and software to instantiate video edge services.

Smart cameras and video surveillance systems can store footage locally for a short period, reducing the burden on central data repositories and enabling quicker access for local processing and analysis. Only event-triggered or relevant video data is transmitted to central data centers for long-term storage, optimizing network and storage resources.

Network slicing for enterprises: 5G network slicing is set to revolutionize enterprise networking by providing end-to-end VPNs that span both wireless and wired domains, offering unparalleled quality of service and customization. Like how MPLS transformed enterprise networking, network slicing has the potential to be a significant revenue generator for CSPs by offering tailored network solutions for various enterprise needs. Enterprises demand privacy, security, and guaranteed bandwidth for their critical applications. Network slicing can elevate these capabilities to a new level, facilitated by the underlying cloud metro architecture. The cloud metro network is crucial in supporting network slicing by providing the necessary infrastructure to implement and manage slices effectively.

AR/VR for business applications: Augmented reality (AR) and virtual reality (VR) are finding new applications in business, such as visualizing underground cables, maps or equipment diagrams, enhancing both operational efficiency and safety. AR/VR applications require low latency and high bandwidth—conditions that are innately met by cloud metro networks. The use of AR/VR in business applications opens new revenue streams for CSPs by providing the necessary network infrastructure to support these advanced technologies.

Cloud gaming: The popularity of gaming on smartphones and tablets necessitates a network architecture that can support high-performance, multiplayer, and low-latency gaming experiences. Cloud metro networks provide the low-latency edge architecture required to enhance gaming performance, thereby supporting the growth of cloud gaming services. CSPs can leverage cloud metro networks to offer high-end 5G gaming services, creating additional revenue opportunities.

For example, the popularity of esports is rapidly growing. Universities are now using esports for recruitment, which is like other college sports (football, baseball, etc.). Some universities are investing significant funds to create e-sport auditoriums and adding the resources necessary to boost the ability of their network to support gaming activities and offering scholarships to attract new students. Esports gaming universities compete against each other, and esports is enabling smaller universities to compete with the larger universities as a lower price point. The network requirements focus on bandwidth, low latency, and security, all of which are addressed by cloud metro networks.

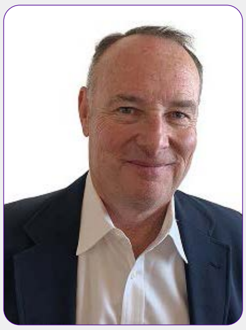
IoT use cases: The IoT encompasses a wide range of devices requiring different network capabilities. Cloud metro networks can facilitate the diverse needs of IoT applications by providing scalable, secure, and reliable connectivity.

[1] https://globalsportmatters.com/business/2020/09/24/collegiate-esports-much-more-than-a-game/

Conclusion

The whitepaper concludes that the cloud metro network is the future of telecommunications, offering the necessary scalability, flexibility, and intelligence to meet the burgeoning demands of edge computing and 5G networks. By integrating hyperscaler best practices, adopting new routing technologies, and embracing innovations such as AIOps, the metro network can support a diverse array of services and applications. A key enabler of life-cycle management is intent-based automation integrating planning, orchestration, service assurance, and network optimization. This evolution also promises to unlock new revenue streams for service providers through advanced services such as network slicing, AR/VR applications, cloud gaming, and IoT use cases. The cloud metro network, with its enhanced capabilities and sustainable design, is not just a response to current requirements but a forward-thinking adaptation for future technological landscapes.

Peter Fetterolf



Peter Fetterolf, Ph. D. is an expert in network technology, architecture and economic analysis. He is responsible for financial modeling and whitepapers as well as software development of the ACG Research Business Analytics Engine. Dr. Fetterolf has a multidisciplinary background in the networking industry with over thirty years of experience as a management consultant, entrepreneur, executive manager, and academic. He is experienced in economic modeling, business case analysis, engineering management, product definition, market validation, network design, and enterprise, and service provider network strategy.