WHITE PAPER

Revolutionizing data center management with PON technology

A technical paper for data center management

The cloud market continues to grow in all geographic regions, with the U.S. providing much of the revenue and data center (DC) capacity. New and growing demand for cloud services and AI is reshaping the industry landscape, presenting hyperscalers with both challenges and opportunities.

Many of the largest hyperscalers continue to build their own DCs. However, when necessary, they collaborate with co-location providers to meet demand and address their needs. Scale is the primary reason hyperscalers build DCs, closely followed by the need for customization and design control.

While the cloud market is propelled by digital business growth, generative AI (GenAI) applications are creating a unique and heightened computing-power demand on DCs to analyze vast and complex datasets.

Hyperscalers are responsible for constructing the majority of DCs worldwide, and Meta is one of the four largest hyperscale platforms. So, when Meta and Ciena investigated drastically simplifying DC management, including reducing the space, power, and cabling of their data plane management, the two companies took a novel approach that differs from legacy hyperscale management methods.

Meta challenges and opportunities

Simplified DC management infrastructure: Legacy switch and console servers had to be individually managed, including their entire network management lifecycles. Meta was looking for ways to reduce the complexity that comes with individual management, which includes software upgrade, security vulnerability management, IP address management, physical connectivity, and so on. It's important to highlight that console servers, regardless of the vendor, are generally more challenging to manage compared to Ethernet switches.

Prepared by

Mike Joseph
Technical Lead / Manager of Infrastructure Network Engineering
Meta
mikejoseph@meta.com

Furqan Haq
Principal, Product Line Management
Ciena Systems, Inc.
fhaq@ciena.com

Wayne Hickey
Senior Advisor
Ciena Systems, Inc.
whickey@ciena.com

Prabhakar Nagral
SVT/PV Engineer 5
Ciena Systems, Inc.
pnagral@ciena.com

Optimized DC design by eliminating copper trays for cost, complexity, and space efficiency: Management connectivity was the primary reason Meta installed copper trays in legacy DCs. A typical DC includes substantial amounts of copper between rack positions and the end of the row to support management Ethernet and legacy console server connectivity. Copper trays cannot share space with fiber in a DC and are larger and more complex than fiber trays.

Dedicated copper trays result in additional rack layers, and each additional layer introduced into the infrastructure increases both cost and complexity. Copper cables, especially in large-scale deployments (~15,000 cables per DC) represent a significant cost burden.

Meta aimed to reduce both factors by entirely eliminating copper trays and designing a DC that exclusively used fiber, thus negating the need for copper accommodation.

Ciena's DC passive optical network (PON) eliminates the need for copper trays by providing per-rack port management Ethernet and legacy console server connectivity from the uniform demark point.

Removing copper trays not only reduces cost and complexity, but it also allows for greater integration of fiber into the infrastructure. This feature enhances the data plane traffic capacity and enables the construction of DCs with lower ceiling heights. It also optimizes overhead and vertical space, enabling more efficient use of real estate.

Lack of fungibility: Meta was navigating various challenges that presented opportunities for optimization when managing assets, determining the appropriate deployment increments, deciding how much copper to install, and handling the lack of fungibility. The ability to maintain uniformity in the infrastructure layer provides Meta with a significant advantage.

Ciena's DC PON limits the complexity within the rack, with a splitter placed either on top of the rack following the enhanced Open Rack v3 (ORv3) standard—a specification within the Open Compute Project—adopted by Meta or integrated within the rack with ORv2. This approach ensures a consistent handoff across the DC, offering virtually limitless port combinations. While Meta has chosen to support up to 80 optical network units (ONUs) per pluggable optical line terminal (OLT) with four ports per ONU, larger ONUs can also be deployed.

This structure enhances fungibility within the DC, allowing for greater flexibility and scalability.

Uniform technical demark for every rack position: Meta also wanted uniformity in the infrastructure layer with a self-contained management solution for each rack position.

Ciena's DC PON offers Meta fixed fiber demarcation. Every rack position gets two fibers. Those two fibers represent a single demarcated handoff, streamlining the deployment process.

Capacity planning: One of the key challenges Meta wanted to address was determining the appropriate capacity to allocate without knowing the specific workloads that would be deployed in the DC and their respective locations. This problem is even more prevalent in the AI DC. Five years ago, in hyperscale DCs, every rack position was predetermined. Now, however, it has become highly variable because of power, cooling, and DC connectivity.

Ciena's DC PON offers practical deployment increments, allowing operators to deploy one port, four ports, 12 ports, or any required amount per rack position based on demand, with the option to add more capacity as needed. This approach offers Meta more flexibility than traditional switching and console server models.

Device consolidation to lower cost: Console servers are not only more expensive than Ethernet switches, but they are also more difficult to manage. Meta wanted to consolidate Ethernet and console server access while reducing OPEX and CAPEX.

Ciena's DC PON offers a consolidated solution, self-contained in each rack position for Ethernet and console management access.

Centralized rich and scalable streaming telemetry: Meta wanted modern and programable northbound interfaces with rich streaming telemetry capabilities for real-time monitoring and machine-to-machine (M2M) automated provisioning.

Ciena's DC PON offers comprehensive alarming and monitoring capabilities using standards-based programable interfaces such as NETCONF, gNMI, and YANG models.

Modularity, scale-out, and flexible ratios between Ethernet and console: Meta sought a modular solution capable of fanning out to "n," where n is the number of Ethernet and/or console ports within each rack. Additionally, Meta wanted flexible ratios between Ethernet and console ports to address the application demands at each rack position.

Ciena's DC PON uses low-cost passive splitters to connect the OLT port to n x ONUs. Passive splitters also simplify the network by reducing the number of actively managed devices and overall power consumption requirements. Meta chose to deploy eight splitters per rack, but PON technology allows hyperscalers to customize the split ratio based on their specific requirements.

Uniform management layer: The absence of a unified management layer in the networking infrastructure leads to inefficiencies, complexities, and higher costs. There is a critical need for a centralized platform to integrate network components, automate tasks, provide real-time insights, and streamline operations for improved performance and reliability.

Using PON technology within the DC eliminates the need for discrete management of equipment providing Ethernet and console access. This is achieved by integrating them and implementing the ONU management and control interface (OMCI) as the uniform management layer.

Native QoS with PON and HE-QoS on the host switch: Meta requires quality of service (QoS) per host device connected to ONUs. To meet this need, a combination of PON and host switch QoS is necessary due to the 10G port's capacity for up to 80 ONUs. Precise egress QoS control per host port is essential in the downstream direction before traffic reaches the OLT.

In PONs, a transmission container (T-CONT) is a group of logical connections within an ONU that is treated as a single entity for upstream bandwidth allocation. Each T-CONT carries a specific type of service traffic with its own QoS characteristics.

T-CONTs are uniquely identified by an ALLOC-ID, which is assigned by the OLT. Each T-CONT supports dynamic bandwidth allocation (DBA), which allows for real-time adjustment of bandwidth allocation based on network traffic and user demand, optimizing performance and user experience. This ensures that different types of traffic receive the necessary bandwidth and priority, improving the overall QoS.

In addition to XGS-PON native QoS, Ciena's host switch offers hierarchical egress QoS (HE-QoS), which can be used for per-ONU-port egress shaping. This allows operators to use deep packet buffers on the host switch to precisely control traffic flow, enabling administrators to tailor QoS policies to meet specific network requirements.

Hyperscale DC management

In the last three decades, there has been little evolution in DC management. Console access servers, serving as out-of-band management solutions, remain a common choice for ensuring continuous uptime. This enables secure remote console management of any device with a serial, USB, or Ethernet console management port. Out-of-band management is physically separate from the "in-band" network connection. Console (access/terminal) servers are used to configure devices and are responsible for managing and monitoring the network's operations. Hyperscalers continue to deploy many Ethernet switches in their server racks as part of their DC growth, and many are using white box switches with packet software for these switches.

In DCs, management switches offer out-of-band management access for network equipment, which is crucial for handling both the actual data traffic moving east-west and north-south. In the scenario illustrated in Figure 1, each rack features 20GbE ports for conventional management purposes alongside 20 console ports for backup in case of issues like failed boot-ups during software upgrades or reconfigurations. The out-of-band connectivity provided by console access servers also offers crucial disaster recovery capabilities, ensuring hyperscalers maintain visibility and redundancy in their DC infrastructure. All DC operators require robust management tools, with hyperscalers necessitating advanced solutions.

In the depicted use case, 921 server racks are used, requiring cabling of 36,000 individual Ethernet ports (18,000 for management, 18,000 for console) to compact 1 rack unit (RU) aggregation switches in adjacent racks. These aggregation switches then link to gateway routers in the DC using n x 10G ports. To accommodate these 36,000 management ports, the management plane necessitates 1,152 RUs of Ethernet switches (equivalent to 27 racks) for consolidation, highlighting the space and power requirements associated with managing the DC infrastructure.

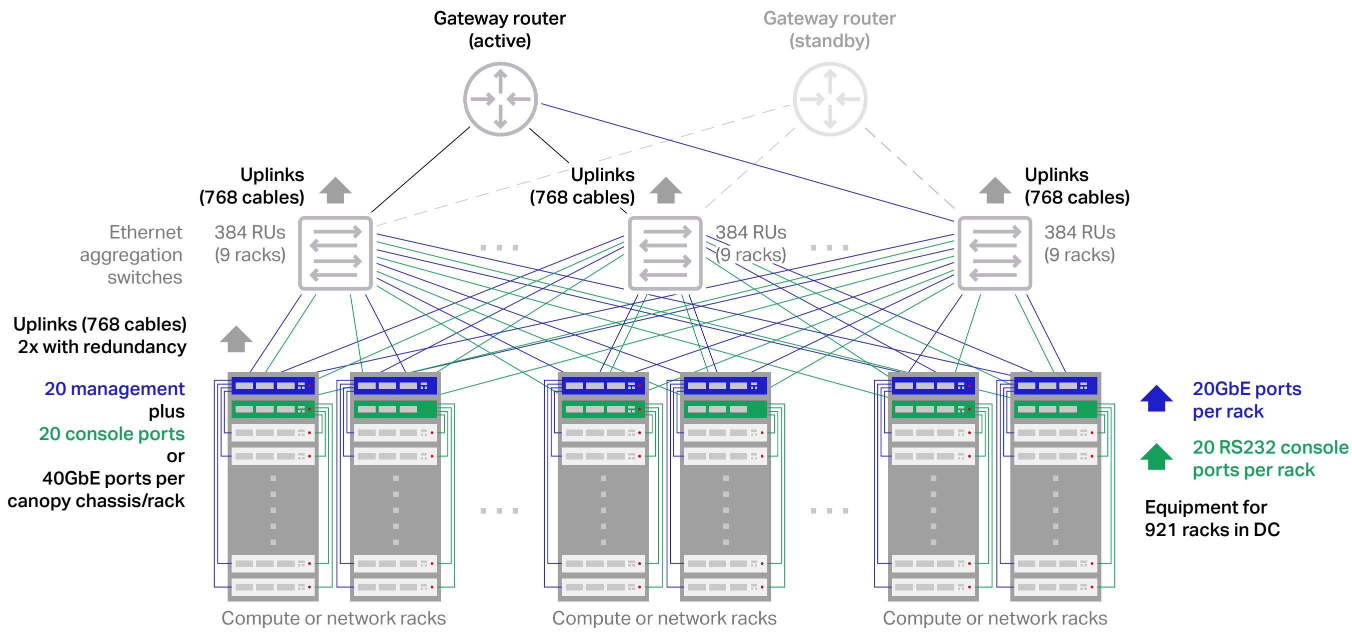

Figure 1. Legacy DC management model

To connect to these 36,000 management ports, the management plane needs 1,152 RUs of Ethernet switches (27 racks) to consolidate all these ports. Consider the space and power of all this aggregation equipment—just for the DC's management plane.

This formula does not support sustainability. The legacy approach of n x 10G ports to the gateway router uses many unnecessary ports where 100G ports would be far more efficient. Conversely, the management plane footprint is excessive in terms of space and power when compared to the minimally loaded management traffic.

PON as an out-of-band management infrastructure

PON uses only passive equipment in the optical distribution link (ODL) between the OLT and ONU. PON was originally targeted as a low-cost point-to-multi-point shared medium for residential broadband markets. This was enabled by using cost-effective ONU laser technology and TDM / time-division multiple access (TDMA) media access control (MAC) on a single fiber. In recent years, symmetrical XGS-PON has been favored in most of the world, transforming fiber access and capturing a significant portion of the residential broadband market.

With the focus on keeping ONT/ONU equipment as low cost as possible, XGS-PON upstream lasers transmit between 1,260 nm and 1,280 nm, and the OLT downstream lasers transmit between 1,575 nm and 1,581 nm on a single fiber. Depending on the ODL, a single OLT can be used to serve up to 128 ONUs. The OLT class of laser fiber type determines the reach and number of ONUs that the OLT can serve.

A micro OLT (uOLT) 10G PON transceiver, when plugged into a router, turns that router port into an OLT, enabling XGS-PON service on a single fiber. This unique uOLT plug capability enables both 10G OLT and routing. Even more importantly, as management traffic increases, those 10G router ports are 25GS-PON ready, enabling 2.5 times faster speeds by simply plugging in a 25GS-PON uOLT.

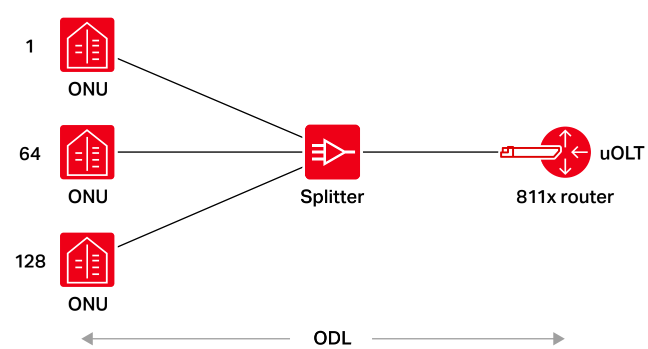

Figure 2. PON ODL

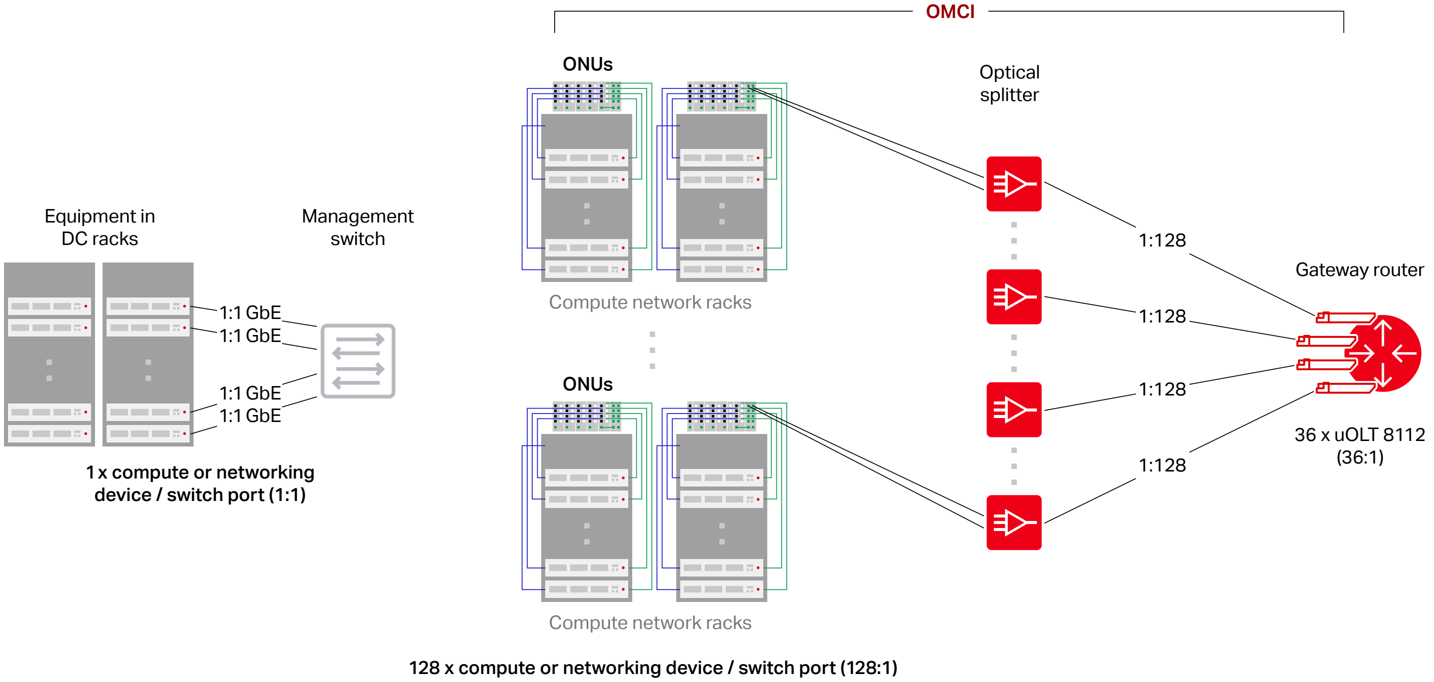

Figure 3. PON's OMCI: a refined solution to management plane density growth

OMCI

OMCI is a protocol used in fiber optic networks to remotely manage and control ONUs on a PON. It is an ITU-T defined standard interface, which simplifies network management, enables remote access and monitoring, and supports various functions like configuration, monitoring, and troubleshooting. Its technical value lies in streamlining operations, improving efficiency, and enhancing network performance and reliability.

OMCI eliminates the need for separate management solutions in DCs by merging the management switch and console, creating a uniform management layer. This streamlines operations, reduces complexity, and provides a standardized approach for managing network elements.

Each ONU endpoint consists of up to four Ethernet management ports and four console ports. Instead of each port having an individual physical cable into an aggregation switch, up to 128 ONUs at eight ports

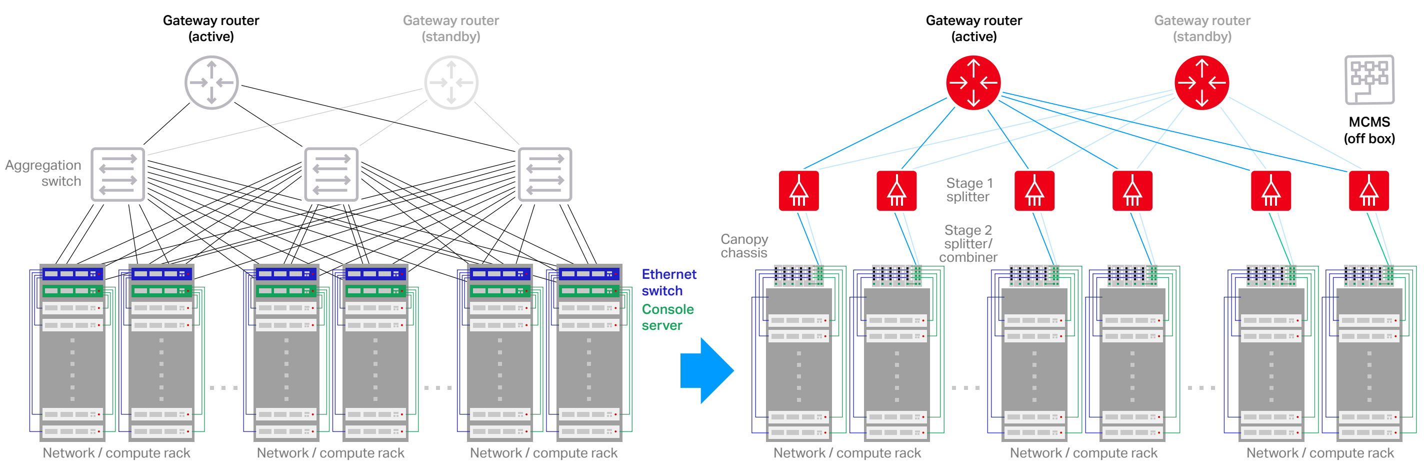

Figure 4. Legacy switch versus XGS-PON approach

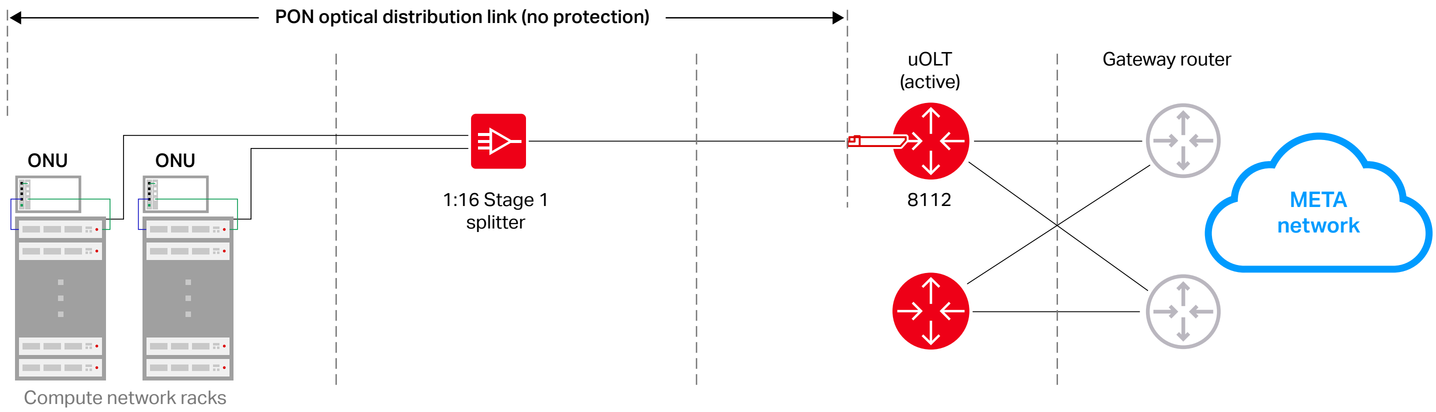

Figure 5. DC PON infrastructure without protection

each (or 1,024GbE ports) can be managed by a single fiber pair. This significantly reduces the space and power over the legacy management plane approach.

PON resiliency

Figure 5 illustrates an unprotected PON topology specifically designed for DC management plane applications. While management traffic is supported, any disruptive event or failure would have a direct impact on traffic flow. To ensure PON and management plane reliability and availability, it is crucial to implement PON protection mechanisms. These mechanisms play a vital role in safeguarding the network against potential failures and disruptions, enhancing the network's overall performance and resilience.

Additionally, PON topology can support PON infrastructure and uOLT protection. In Figure 6, Type F protection dual parenting in PONs is used, which simultaneously connects each ONU to two OLTs.

One OLT acts as the primary parent, responsible for handling traffic, while the other serves as the secondary parent for failover purposes. This redundancy ensures uninterrupted service, mitigates the risks associated with single points of failure, and enhances fault tolerance. Dual parenting is particularly beneficial for critical applications where any form of downtime is unacceptable.

Many Ciena routers support uOLT XGS-PON transceivers. For example, consider Ciena's 8112 Coherent Aggregation Router (36 ports), where up to 36 uOLTs can terminate 18,000 management and console ports (36 uOLT x 128 ONUs x 4 ports per ONU = 18,432 ports). Only a pair of 8112 routers is needed to support 36,000 management and console ports over 921 DC racks. To add router resiliency, these ONUs can be connected to another 8112 to offer 1:16 redundancy at the ONU level. The 8112s now aggregate this traffic into n x 100GbE—as opposed

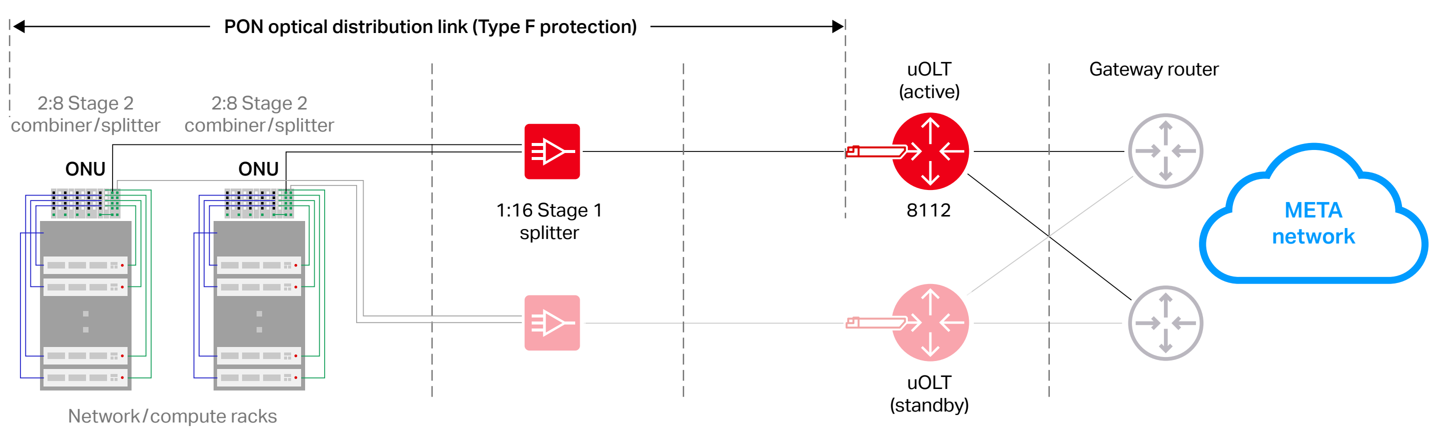

Figure 6. uOLT with dual parenting (PON protection)

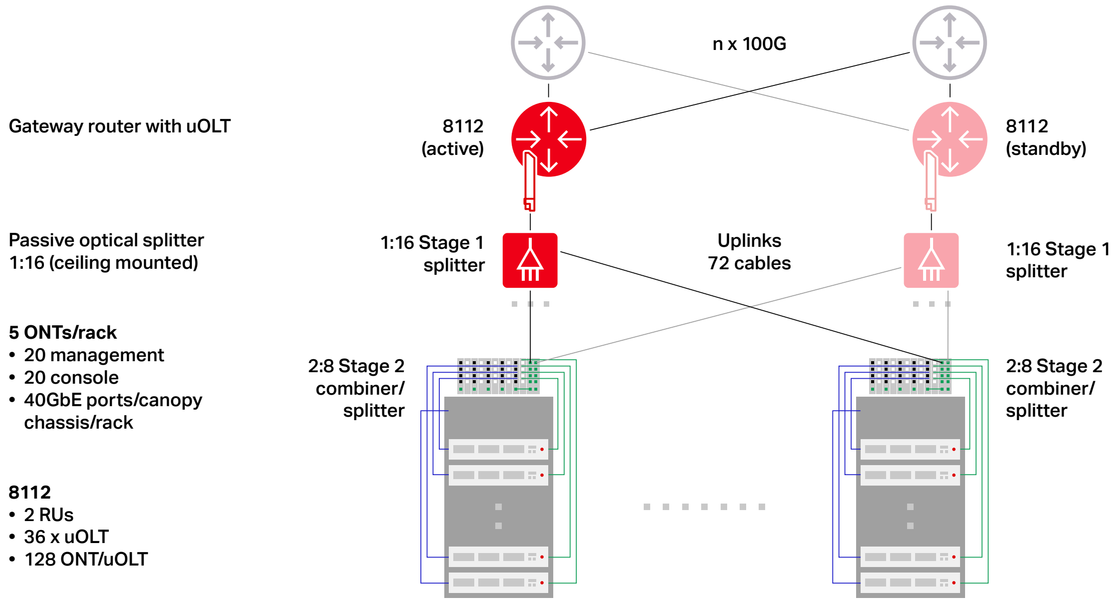

Figure 7. DC PON management model

to the 10GbE used by legacy aggregation switches—freeing up ports on the DC gateway routers. Now a typical hyperscale DC can house up to 10,000 racks or 195 management and console ports, resulting in 10 times the savings. This results in significant space and a reduction in power savings, shown in Figure 8.

When considering a DC—and hyperscalers are no exception—the three key considerations are operational efficiency (DC layout for ease of use), cooling efficiency (minimizing overall cooling costs), and availability (power, cooling, cable management, and so on).

Passively cooled ONUs offer efficient and silent operation, eliminating the need for fans and reducing energy consumption while delivering high-performance connectivity for network access.

Additionally, Ciena's console servers provide secure and reliable remote network device management, enabling efficient troubleshooting and configuration to ensure uninterrupted network operations.

In the legacy model, each Ethernet switch takes up 1 RU or 768 RUs in total. Ciena's DC out-of-band management approach is purpose-built for the restricted canopy space above the rack, meaning the only rack space needed would be for the 8112 routers. Each 8112 router takes up 1 RU of rack space, so only 2 x RU is needed versus the legacy 768 RUs. The sleek chassis is designed to support multiple pairs of ONUs and console servers on the top of the rack, with DC powering options, without taking up useable rack space.

As the volume of traffic in hyperscale DCs escalates alongside the emergence of GenAI, the existing

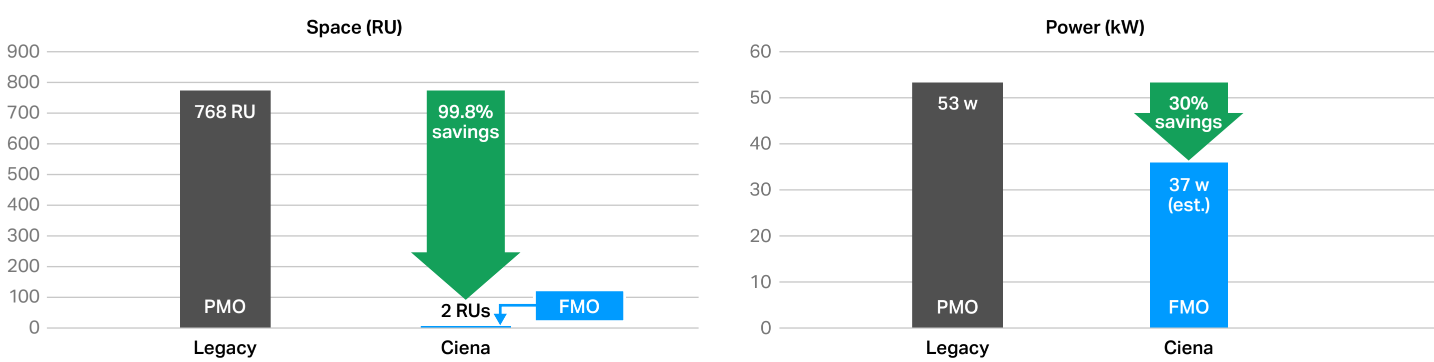

Figure 8. PON versus legacy DC space and power savings

management plane architecture of hyperscalers proves to be unsustainable due to its resource-intensive nature. In this scenario, the integration of PON technology offers a promising solution by streamlining management, ensuring uniformity, and providing simplified access to both compute and network racks or console and Ethernet connections.

By harnessing PON technology, DCs can achieve substantial reductions in space, power consumption, and cabling requirements, thereby bolstering their overall sustainability. Ciena's innovative DC PON technology presents an opportunity to realize space savings exceeding 99%—optimizing rack utilization, reducing power overheads, and supporting hyperscalers in advancing their sustainability goals. Moreover, the potential for significant rack savings opens avenues for DC providers to accommodate a greater number of compute elements compared to traditional legacy approaches.

Contributor: Mike Joseph

Mike Joseph is currently the tech lead / manager of the Infrastructure Network Engineering team at Meta. In this role, he leads the team that provides management, out-of-band, facilities, and disaster recovery networks for all their global DCs and points of presence (PoPs), as part of the backbone networking organization inside their production infrastructure department.

Mike previously oversaw technical operations for Google Fiber, a startup GPON ISP within Alphabet. He has also held roles as Director of Operations for Oracle's public cloud, Director of Network Operations for Yahoo!, and VP of Technical Operations for an SD-WAN and edge computing startup.

Contributor: Furqan Haq

Furqan Haq is a Principal Product Line Manager at Ciena and is a key member of the team defining the SAOS 10.x network operating system roadmap and strategy. His primary areas of responsibility include IP routing (IGP, BGP, routing policies), MPLS transport (BGP-LU, BGP-SR, RSVP-TE), and security. Furqan is also the lead technical architect of Ciena's DC PON solution. With extensive experience in the networking industry, Furqan provides trusted guidance to Ciena's broad IP customer base and helps shape Ciena's roadmap strategy across key market segments, such as mobile backhaul, business and residential services, and IP/MPLS core networks and coherent routing.

With over 20 years of experience in the telecommunications industry, Furqan previously served as a Consulting Engineer in Nokia's IP Routing & Transport unit, specializing in core IP routing, MPLS VPNs, BNG, Carrier Ethernet, NFV, and SDN. His career also includes roles in R&D engineering, mobile backhaul support, and network design across IP and fixed-access products.

Contributor: Wayne Hickey

Wayne Hickey is a Senior Advisor to Ciena Systems, Inc. In this capacity, he is tasked with the responsibility for Ciena's routing and switching go-to-market and portfolio strategy.

With more than 30 years of experience in the telecommunications industry, Wayne has been building networks for decades. He has held numerous positions of increasing responsibility as Senior Systems Engineer Manager, Senior Product Manager, and Solutions Marketing leadership, delivering comprehensive networking solutions to service providers, DC operators, and mission-critical utilities.

Contributor: Prabhakar Nagral

Prabhakar Nagral is a Quality Architect at Ciena, focusing on IP routing protocols, MPLS transport, SRv6, and PON within the routing and switching portfolio. He serves as the test architect for Ciena's DC PON solution, overseeing the testbed design and leading system testing. Recently, he significantly contributed to integrating AI into the RSP R&D quality team to enhance efficiency.

Was this content useful?

◯ Yes ◯ No