www.pipelinepub.com
Volume 20, Issue 10

The Network Economics of AI are Driving Bandwidth to the Metro Edge

By: Marco Bertolini

The explosion of interest in AI is expected to result in a market size of $2.5 trillion by 2032. Network and cloud interconnection architectures, especially in the metro area, are also scaling rapidly to align with the demand for AI-based services. Similarly, the size of the edge computing market in 2023 reached $16.5 billion, and edge analytics will reach $41 billion globally within five years.

The focus on the edge reflects the economics of AI-driven use cases in everything from smart cities and transportation to logistics and advanced manufacturing. Data is being processed at the edge because localized exchanges reduce transportation costs, provide the lowest latencies, and are closest to dense populations with their increasingly automated service infrastructures.

Even IoT data, which is individually negligible, begins to add up as it scales. It is estimated that 40 billion IoT devices will be available in 2025 and that these devices will be capable of generating 79.4 zettabytes of data. Along with the $1 trillion 5G services market by 2028 and the $35 billion PON access market by 2030, it's expected that in just two years, 33 petabytes of data will require interconnection bandwidth, and the demand is growing fast.

Delivery of massive volumes of data to the edge is being facilitated by the ever-increasing deployment of optical fiber deeper into metro and access networks, as it provides nearly unlimited bandwidth over traditional copper or radio access technologies.

All this data will be serviced by a wide range of service providers including colocation companies, cloud service providers and hyperscalers, internet exchange points, content delivery networks, research and education networks (RENs), and telecom communications service providers (CSPs), all of whom need to build out their metro edge and interconnect networks to meet the surging demand for bandwidth to the network edge.

Introducing the Metro Edge

The networks that deliver our communications services are not one-size-fits-all. Tailored solutions are essential and are dictated by the scale and distances data must traverse, the end-users served, and the specific needs of each network operator. This isn't inherently different from moving people and freight: small, light vehicles, even bicycles and drones, are used for local deliveries, while massive container ships and freight trains carry goods across oceans and continents.

In the realm of optical networks, there are three distinct segments. Subsea and long-haul networks bridge oceans and continents. They are the backbone of global data transport. Meanwhile, regional networks carry data across metropolitan areas in a region or an entire small country, serving as the intermediaries between long-haul and metro networks. Metro networks are positioned at the network's edge and connect platforms and customers, ensuring seamless local data delivery (See Figure 1, below).

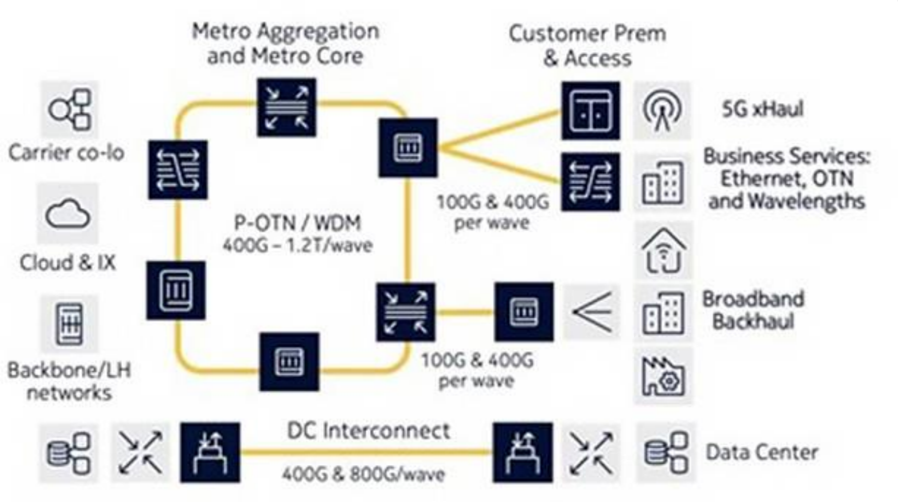

Figure 1: Metro edge networks expand connectivity past the core to a wide array of use cases and applications.

Modern regional and long-haul optical networks boast impressive scalability, often utilizing 1.2 Tb/s coherent optics and conventional or long-band (C+L) wavelength division multiplexing (WDM) line systems to handle traffic growth rates of 30-40 percent annually. However, the network edge has quite different requirements. This includes supporting a diverse array of services with varying speeds and protocols, providing support for a wide range of applications.

Metro-focused equipment is being implemented in ever smaller and more compact platforms with reduced power and carbon footprints. Additionally, the capability to be deployed beyond traditional telecom or data center environments typical of both regional and long-haul deployments is essential, requiring support for AC power in enterprise settings and extended temperature ranges for outdoor cabinet installations.

Scaling network bandwidth at the metro edge must address and accommodate a variety of use cases. For instance, business services for enterprise and institutional users often involve a range of high- speed services, enabling end users to connect to offices around the world, in addition to private and public peering points, the internet, and edge cloud services.

These business services include fixed bit rate connections, such as 10 Gb/s, 16 Gb/s, 32 Gb/s, 100 Gb/s or greater. Service providers may even define adjustable and fractional bit rate services with either hard or soft isolation. They also span a range of protocols, including fiber channel, optical transport network (OTN), and Ethernet, which includes various carrier Ethernet services defined by the Metro Ethernet Forum (MEF).

In the realm of residential or enterprise services, end users are often provided “last mile” broadband backhaul technologies, such as fiber-to-the-home (FTTH) over PON technology, broadband services over hybrid fiber–coax (HFC) networks, xDSL over legacy copper access lines, or wireless connectivity using 5G or fixed wireless access (FWA) technology. In these scenarios, the broadband access platform aggregates traffic to and from multiple end users and backhauls it from the metro edge to the core over some form of fiber-based aggregation link or metro backhaul network.

Metro data center interconnection involves web scale companies, carriers, or enterprises connecting data centers to each other or to internet exchanges (IX) or peering points across short distances within a metropolitan area. These simple point-to-point connections are often spanned by pluggable coherent optics equipped directly in router ports. They provide scalable, energy-efficient connectivity between sites by using WDM technology to scale capacity across multiple channels over a single fiber pair.

Determining the Ideal Optical Network for the Metro Edge

As we’ve seen, with the growth of AI, autonomy, edge computing, and edge analytics, network operators face the challenge of scaling optical transport link capacity to meet the growing service demands at the metro edge. Unlike long-haul core networks, where fiber scarcity and vast distances necessitate optimization for maximum scale and reach, the metro edge requires a focus on low power consumption and compact form factors characteristic of pluggable transceivers. These include pluggable digital coherent optics (DCOs) in quad small form factor pluggable (QSFP) or C-form factors (CFPs), scaling transport link capacity up to 400 Gb/s or 800 Gb/s per wavelength. This provides the necessary scale to aggregate more business services and support metro data center interconnect (DCI) applications.

In addition, operators can opt for pluggable DCOs operating at 100 Gb/s per wavelength to enhance low-capacity edge nodes or backhaul links, enabling direct 100G uplinks over a common WDM infrastructure. By employing pluggable DCOs, network operators can right-size metro connection speeds while exploiting the higher capacity and longer reach of WDM as compared to gray optics.

Another critical requirement for optical network operators is selecting platforms and solutions optimized for metro edge applications. These platforms need to be small, compact, scalable, and energy-efficient, avoiding being burdened by the requirements of regional and long-haul network solutions. They must support both AC and DC power feeds for use in telco and customer premises applications and feature an extended operating temperature range and outside-plant compatibility for deployment in uncooled outdoor cabinets.

Edge-optimized line cards and features are essential elements to meeting the diverse requirements of the metro edge. They include compact multi-service aggregation cards that support packet and OTN aggregation of various protocols and port speeds. The aggregated traffic is then fed into high-speed uplinks using 100 Gb/s, 400 Gb/s and 800 Gb/s pluggable DCOs.

Photonic management of wavelengths can now be extended to the metro edge, using compact reconfigurable optical add/drop multiplexers (ROADMs) that support four- or nine-degree connections. These ROADMs are ideal for adding or dropping a few wavelengths at a node, functioning as a two-degree node on an aggregation ring, or as a four-degree node to couple rings as traffic is gathered across the core and edge. The use of ROADMs in metro scenarios allows for the drastic reduction of the cost of deploying new services and provides an additional layer of reliability, which is becoming increasingly relevant as the capacity in the metro edge grows.

An optimized metro edge solution should seamlessly integrate into a network operator’s end-to-end optical network, ensuring maximum operational efficiency and enabling end-to-end services. This includes supporting synchronization distribution from primary clocks to the edge.

Despite the differences between core, regional and metro networks, the ability to share standardized elements, such as common line cards, will minimize sparing costs and training requirements. Consistent operating and management systems across the network are also crucial, allowing the service provider to minimize training, standardize operational processes, and support optical network automation capabilities from end to end.

The Focus is on the Edge

Although the hype around AI sometimes seems overblown, it would be imprudent to ignore the seismic shift these technologies represent. Some are even arguing that the limiting factor to AI adoption may not be processing capabilities but sufficient power. This highlights the stress that will be put on some of our most fundamental infrastructure.

Along with power grids, networks serving dense, highly-connected and automated population centers will be under extreme stress as AI-based use cases explode over the next decade, and the metro edge is where most of this action will occur. There will be a wide diversity of applications, players and rapidly shifting loads to meet. This represents an enormous opportunity for network providers but a daunting challenge too — one that will only be met with a metro edge that is designed to be highly scalable, efficient, resilient, and adaptable.