

The AI middle mile and federated wavelengths: Connecting enterprises to AI training centers

Publication date:

18 September 2024

Author:

Ian Redpath, Research Director, Transport Networks and Components

Omdia view

Summary

Omdia had the opportunity to speak with Dave Ward, the new chief technology and product officer (CTPO) of Lumen Technologies (Lumen). Dave opened with the following:

AI and the cloud are driving the largest expansion of networks in our lifetime.

Dave's new role includes updating the AI middle mile and building a more agile and responsive digital platform to support the growing demand for temporal wavelength services. This is because an automated and federated environment is necessary, as no single service provider can operate independently.

The MEF Lifecycle Service Orchestration (LSO) framework has been implemented by over 39 communications service providers (CSPs) for carrier Ethernet services across various use cases. The industry's next steps include broadening the ecosystem to include more CSPs (already in progress) and extending the LSO-supported payloads to include 100G, 400G, and 800G wavelengths.

Dave summarized with the following:

The demand for ultra-high capacity is huge, and the MEF APIs are critical. The 400G market momentum is massive; we exceeded our entire 2024 400G service forecast by mid-year, driven by AI.

What is the wavelength market?

A wavelength is a high-capacity service sold directly from an optical platform. The wavelengths under discussion today are 100Gbps and 400Gbps. A wavelength seller can offer wavelengths from A to B using their optical network. Typical performance parameters include primary and backup path latency.

A wavelength buyer has complete access to all the uncontended bandwidth with guaranteed latency performance. In the past, typical wavelength buyers included communications or cloud service providers who needed to complete backbone portions of their networks where they did not have their own fiber assets. A select group of enterprises with very high bandwidth-intensive needs have also purchased wavelengths.

A wavelength service is a high bandwidth option that fills the gap between carrier Ethernet and IP/MPLS services, and a leased/owned fiber network. What is changing?

Enterprises re-architecting networks for AI and the cloud

Enterprises have embraced the cloud service model, resulting in over \$250bn in cloud service revenue for just three major cloud service providers that reported their revenue publicly. At the same time, enterprises desire flexibility and choice and prefer not to be locked into a single cloud supplier. What they seek is a multi-cloud model.

Moving on to AI, enterprises are interested in a flexible AI environment. They want to be able to use the large language model (LLM) of their choice for their specific needs. They prefer to work with multiple LLM providers and not to be restricted to a one-supplier ecosystem.

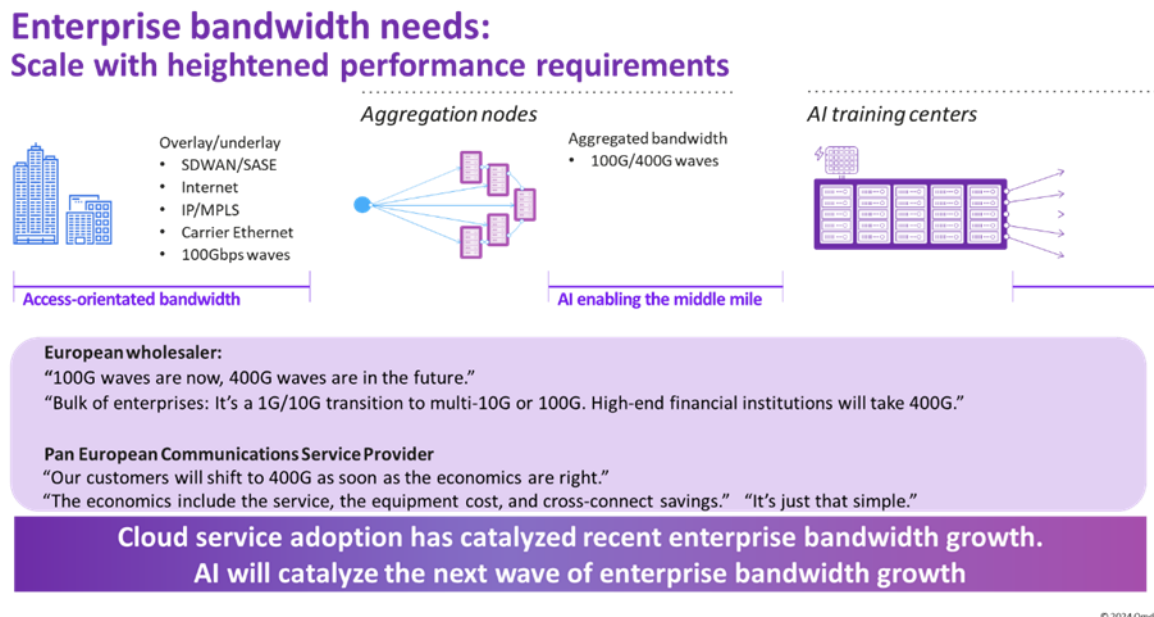
Enterprises have lakes and oceans of data. Often, an enterprise's data lake and the AI training center of choice are physically located in different locations (e.g., across town or across the country). This means a lot of data needs to be moved, including all the data ever created by a major company.

Dave noted the following:

Moving the data into the AI training data centers is a bottleneck. As the enterprises work through their AI and cloud strategy, their data strategy, the need for a refreshed network strategy becomes apparent. When the enterprises re-architect their networks, they go with 100G access. The middle mile CSPs will aggregate the client's 100G wavelengths to 400G for transport across their multi-terabit optical networks.

As enterprises consider upgrading their key sites to 100Gbps access, questions arise regarding the state of the middle mile and the core. If access is 100G, how many terabits are required in the middle mile and core portions of the network?

Performance expectations are also increasing. Enterprises have become much more sensitive to latency and want the network latency statistics on primary and secondary routes under consideration. Are the traditional optical network paths, often originally deployed along pipelines and rail lines, the optimal low-latency path between today's enterprise data lake and the LLM training center?

Figure 1: Enterprises consuming waves for the cloud and AI


Source: Omdia

Cloud AI provider perspective

Cloud and AI providers have massive network requirements. Many of these providers have their own optical cores, which include metro, long haul, and international subsea segments. The emergence of the new AI ecosystem is creating new and enhanced network needs for these providers. They are typically constrained in terms of network capillarity and have limited access to end customers, particularly enterprises.

For the cloud and AI provider communities, their immediate needs start at the physical fiber level:

- Expanded connectivity into all new and expanding cloud and AI properties
- Primary and secondary latency-optimized routes for today’s traffic patterns

The next set of requirements focuses on the ingestion of client traffic from access through the middle mile with the attributes of speed, flexibility, agility, and reliability. The traditional wavelength paradigm was quite static, but today’s high-capacity bandwidth needs are much more dynamic. AI training projects may be occasional but massive endeavors, and in such cases, the win-win solution for all parties would be large-scale but temporary bandwidth. Industry scale would involve numerous enterprises, verticals, and AI models, with growth and evolution over time.

How would this network be built? The traditional paradigm of static wavelength, with “set-and-forget” provisioning, would not be agile and responsive enough for the new environment. The new environment requires the following:

- A robust, federated ecosystem, including last- and middle-mile network operators
- An automated, agile, silicon-speed business process environment

- An automated and responsive operational environment

New expectations for business processes would include visibility on route inventory with key performance parameters, and pricing.

On the operations side, cloud and AI providers and more enterprises are becoming sophisticated optical network clients. They desire real-time photonic operational data, including:

- Photonic performance of the reconfigurable optical add-drop multiplexers (ROADMs), optical line amplifiers, and optical transceivers
- Performance statistics, such as latency, optical loss, and optical pluggable temperature

Access to the photonic performance datasets would enable a more predictive and proactive operational stance. Policies could be set to enable preemptive actions based on typical degradation scenarios to maintain optimal network performance.

Dave captured the current state of affairs as follows:

Non-standard APIs do exist, but the industry needs to evolve to a fully standardized environment to achieve the scale needed.

The MEF LSO framework

Over the course of decades, the MEF has established a framework that would meet the industry's needs—a multi-carrier Ethernet automated and federated network standard known as the LSO framework.

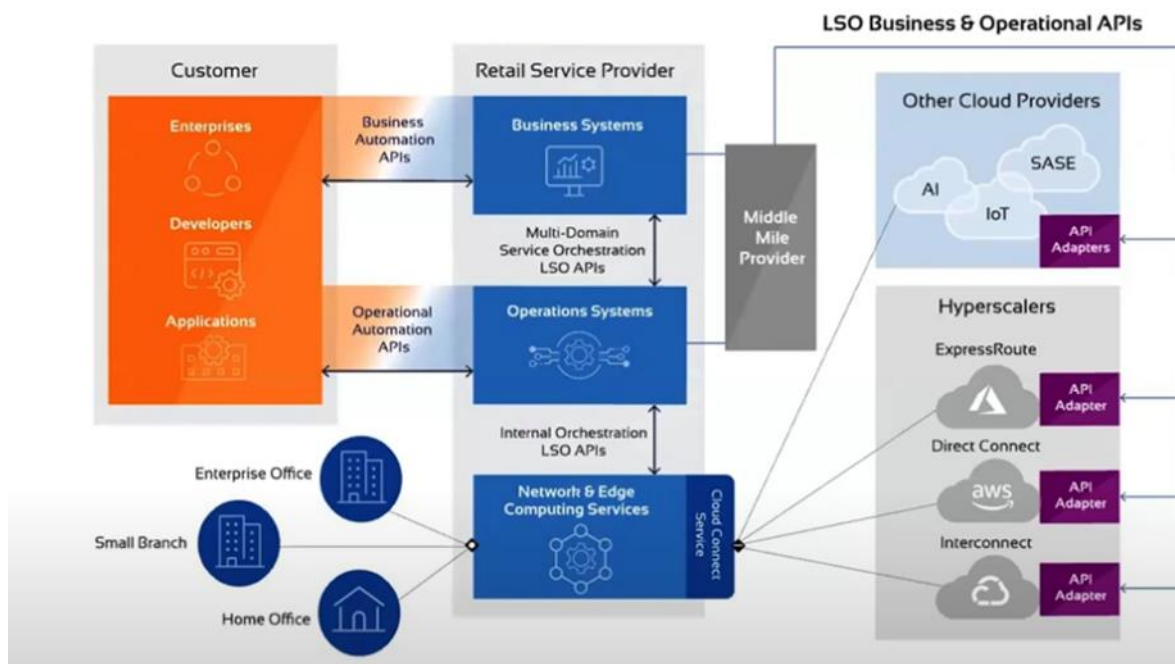
The LSO framework consists of MEF-defined specifications that enable standardized service orchestration based on the standardized lifecycles of end-to-end connectivity services across one or more network service domains. It is an agile approach to streamlining and automating the service lifecycle in a sustainable fashion for coordinated management and control across all network domains responsible for delivering an end-to-end connectivity service (e.g., carrier Ethernet, IP VPN, and MPLS).

The MEF 55.1 LSO standard describes a reference architecture and framework for orchestrating the service lifecycle. It includes a set of functional management entities that enable cooperative service lifecycle orchestration for the MEF 3.0 Connectivity Services standard. The framework also provides high-level functional requirements and outlines high-level operational threads describing orchestrated connectivity service behavior as well as interactions among management and control entities.

The management interface reference points that characterize interactions between LSO functional management entities are identified in the reference architecture and are described in such a manner that they can be realized by interface profiles (and further by APIs), which can be used to enable automated and orchestrated connectivity services.

The MEF has also focused industry efforts on specific high-value use cases. Multi-cloud connectivity has emerged as a key use case.

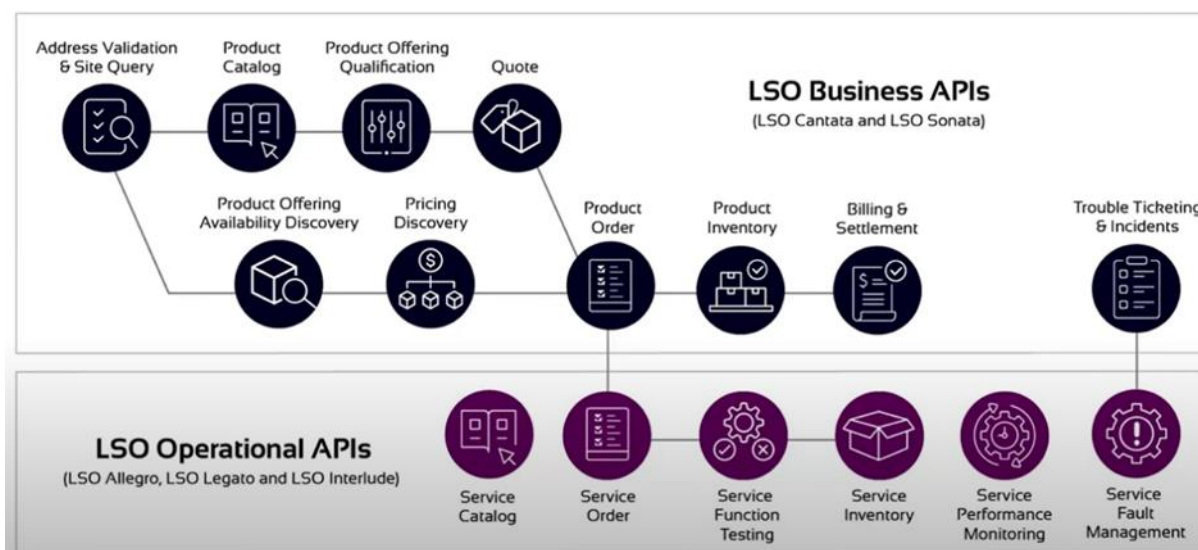
Figure 2: Network as a Service (NaaS) use case: Multi-cloud connectivity



Source: MEF 2024 Q3 Members Meeting – Opening Plenary

The MEF has over 39 CSPs with LSO networks in production and at least another 36 working towards production. The MEF has developed both business and operational process frameworks.

Figure 3: MEF LSO business and operational APIs support



Source: MEF 2024 Q3 Members Meeting – Opening Plenary

Commissioned Research

The next step is to build upon the Wavelength-as-a-Service definitions in the MEF 63 and MEF 64 standards to create standardized wavelength product models that can be used in the LSO business APIs as product payloads. The first pre-standard version of the LSO wavelength product payload for use with LSO APIs should be available on the LSO Marketplace by the end of 2024. This will be alongside carrier Ethernet, internet access, multi-cloud, and edge compute product payloads, which can already be integrated with the LSO business APIs using the LSO Blending Tool.

The ability to implement and use the same LSO business APIs with a wide range of LSO product payloads—including wavelength product payloads—enables a high ROI on LSO API implementation. It also provides NaaS and communication service providers to quickly support the latest connectivity and other products needed to support the emerging AI middle mile and federated wavelength use cases comprehensively.

Appendix

Further reading

[*Event Recap: OFC 2024 – AI in the Data Center Driving Datacom*](#) (May 2024)

[*Lumen, Microsoft, and Corning are building the next-generation global AI network*](#) (August 2024)

MEF

[“MEF Standards,”](#) MEF Wiki (retrieved September 14, 2024)

[“MEF 55.1 Lifecycle Service Orchestration \(LSO\): Reference Architecture and Framework,”](#) MEF (retrieved September 14, 2024)

[“MEF Specifications: MEF 63,”](#) MEF (retrieved September 14, 2024)

[“MEF 64 Operator Layer 1 Service Attributes and Services,”](#) MEF (retrieved September 14, 2024)

[*PTC 2024: Pairs AI, the Cloud, and the Sea Like No Other*](#) (February 2024)

Author(s)

Ian Redpath, Research Director, Transport Networks and Components

askananalyst@omdia.com



Omdia Commissioned Research

This piece of research was commissioned by MEF.

Citation policy

Request external citation and usage of Omdia research and data via citations@omdia.com.

Omdia consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help you. For more information about Omdia's consulting capabilities, please contact us directly at consulting@omdia.com.

Copyright notice and disclaimer

The Omdia research, data and information referenced herein (the "Omdia Materials") are the copyrighted property of Informa Tech and its subsidiaries or affiliates (together "Informa Tech") or its third party data providers and represent data, research, opinions, or viewpoints published by Informa Tech, and are not representations of fact.

The Omdia Materials reflect information and opinions from the original publication date and not from the date of this document. The information and opinions expressed in the Omdia Materials are subject to change without notice and Informa Tech does not have any duty or responsibility to update the Omdia Materials or this publication as a result.

Omdia Materials are delivered on an "as-is" and "as-available" basis. No representation or warranty, express or implied, is made as to the fairness, accuracy, completeness, or correctness of the information, opinions, and conclusions contained in Omdia Materials.

To the maximum extent permitted by law, Informa Tech and its affiliates, officers, directors, employees, agents, and third party data providers disclaim any liability (including, without limitation, any liability arising from fault or negligence) as to the accuracy or completeness or use of the Omdia Materials. Informa Tech will not, under any circumstance whatsoever, be liable for any trading, investment, commercial, or other decisions based on or made in reliance of the Omdia Materials.

CONTACT US

[omdia.com](https://www.omdia.com)

customersuccess@omdia.com