

100-Gigabit Ethernet and super-Lambda services: The next frontier for Carrier Ethernet

Carriers and cable multiple-system operators are widely deploying 10-Gigabit Ethernet (10-GbE) to interconnect IP routers and carrier-class Ethernet switches. The proliferation of 10-GbE services necessitates the next factor of 10: 100 GbE.

By Drew Perkins and Rick Dodd, Infinera

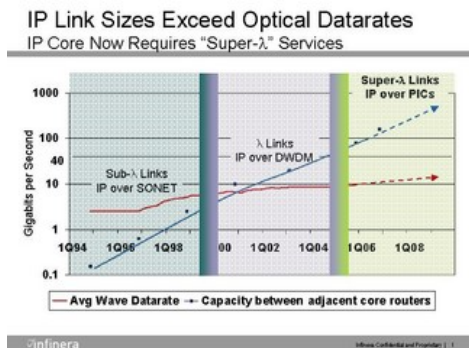
Dust off a 1999-era marketing presentation from any number of optical vendors and you likely will find the following story among the first few slides: "IP traffic is growing exponentially and will dominate voice in terms of network bandwidth. As such, TDM and IP networks need to converge, and 10G provides that convergence. OC-192 and next-generation 10-Gigabit Ethernet (10-GbE) will share technology, saving money all around."

Those presentations were right; 10G was indeed a convergence point for SONET/SDH and Ethernet, and the continued demand for IP bandwidth has led to widescale deployment of 10G circuits both across the WAN (in the form of ITU G.709-wrapped 10-Gbit wavelengths) and between IP routers (as either 10-Gbit/sec Packet-over-SONET (PoS) or, increasingly, 10-GbE LAN PHY).

By mid-2000, 10-Gbit/sec long haul transport capacity outpaced 2.5 Gbits/sec. Today, 91% of long haul DWDM capacity added worldwide each quarter is 10G. 40G is slowly beginning to grow and likely will accelerate once the price of 40G falls significantly below the cost of four x 10G. Meanwhile, IP router-to-router demand continues to grow exponentially. Many carriers quote annual IP network growth between 70% and 100%. Some Internet network access points (NAPs), like the Amsterdam NAP, are seeing 120%+ annual growth. As a result, many large IP backbones have tens of gigabits per second of traffic demand between adjacent routers, nearing--if not exceeding--100 Gbits/sec. In other words, IP routers have the ability to consume 100-GbE WAN ports, even though the capacity provided by the transport layer supports only 10-Gbit/sec and, in some cases, 40-Gbit/sec increments.

"Super-Lambda" circuits

As a result, we've entered into a new phase of the IP network evolution. In the early to mid 1990s, when large data backbones ran DS-3s or OC-3s between routers or switches, data transport occupied timeslots on SONET/SDH or even asynchronous fiber-optic transmission systems (FOTS), and IP connections were "sub-lambda." Around 2000, thanks to exponential growth in IP and the availability of 10G ports on both core routers and DWDM systems, IP service providers could choose to bypass SONET/SDH equipment and carry their 10G PoS traffic directly on DWDM or "lambda." However, router-to-router traffic demand in large IP networks today exceeds the capacity provided by a DWDM wavelength, necessitating the virtual combination of multiple lambdas into single IP links. This phase of the network can be thought of as "super-lambda."



Given the economics of DWDM networks--namely the advantages of multiple wavelengths compared to 100+G data rates per wavelength--the network has entered a phase in which "super-lambda" capabilities will be as important as lambda and sub-lambda services.

Making the reasonable assumption that IP traffic demand will continue to increase faster than the (economically practical) carrying capacity of a single wavelength of light, we need a scalable, reliable, cost-effective mechanism for aggregating multiple optical "lambdas" into one virtual port. Luckily, the industry already has implementations of such technology in today's network--albeit in other forms. Examples include Virtual Concatenation (VCAT), link aggregation, multi-link bonding, parallel, and CWDM optical interfaces like the OIF's VSR-4 and VSR-5 and the popular 10GBASE-LX4, as well as electrical specifications like IEEE's XAUI and the OIF's SFI-4 and SFI-5. And various companies are working to extend this technology to enable 100 GbE in numerous networking applications.

Lessons from 10 GbE

Understanding the path to 100 GbE can be enhanced by reviewing the state of the industry around 10 GbE. Today, 10 GbE is available in many forms, depending on the application and the transmission distance of the Ethernet frame--thousands of kilometers across a WAN, tens or hundreds of meters across a data center or central office (CO), a meter across a backplane, or centimeters between chips on a printed circuit board.

In the WAN case, 10 GbE is generally carried as one serial data stream. Infinera's DWDM product, for example, transports 10-GbE LAN PHY as a transparent 10.3125-Gbit/sec data stream inside a digital wrapper. In this case, serial transmission at 10 Gbits/sec (or 11 Gbits/sec once it's wrapped) and 15xx nm provides the best economics across the wide area.

To connect, say, an IP router to a transport layer product, most service providers would use optical fiber connected to XFP modules using the 10GBASE-LR interface. This is, again, a serial 10-Gbit/sec bit stream but operating at 1310 nm. Shorter distance data center applications may use the 10GBASE-LX4 specification on multimode fiber. 10GBASE-LX4 is widely deployed, with tens of thousands of modules shipped and numerous vendors capable of supplying quality parts. LX4 has improved the economics of multimode fiber largely because of its use of multiple lower speed lasers to create one high-speed signal. In 10GBASE-LX4, four 3.125-Gbit/sec optical signals are generated, each by a different colored laser. This allows LX4 modules to take advantage of the lower dispersion and improved economics of lower speed channels on multimode fiber. The four signals are logically combined by the module into one virtual 10-Gbit/sec connection.

A similar approach is used to send 10-Gbit/sec signals between line cards in IP routers, Ethernet switches, optical systems, and the like. Just as sending a 100-Gbit/sec signal serially across a 500-km WAN route would be cost prohibitive (due to modulator cost, significant polarization mode dispersion (PMD) compensation, and frequent regeneration,) so too is sending 10-Gbit/sec electrical signals across copper connectors and electrical backplanes cost prohibitive. Today, most high-capacity networking equipment uses 2.5G or 3.125G signals between line cards. 10-Gbit/sec signals are demultiplexed using TFI-5 or XAUI into lower speed bit streams to reduce noise and attenuation and generally save cost. The same is true between chipsets that might even be directly adjacent on a printed circuit board. Examine a datasheet for, say, a 10G SONET framer, a key component inside every 10G SONET ADM on the planet, and you'll find that the signals in and out of that framer do not travel at 10 Gbits/sec. Instead, they are translated into multiple "lanes" of lower speed signals, typically four x 2.488 Gbits/sec.

In all cases, the technology decision is driven by economics. What is the most economical way to carry my signal from point A to point B, considering both capital costs (e.g. How expensive is my laser?) and operational cost (e.g. Do I need a new fiber type in my data center? or Must I limit the distance between these line cards in my chassis?) In some cases, the economics favor a serial

approach. In others, the economics favor development of parallel approaches and standards like SFI-4, VSR-4, XAUI, 10GBASE-LX4, and VCAT. All have been widely deployed in today's networks.

100-GbE efforts

The issues are similar in the standardization, development, and deployment of 100 GbE. Across boards and backplanes, RF engineers continue to push the speed at which signals can travel reliably across a copper trace. Even so, parallel approaches will continue to dominate this segment. The OIF today is working on common electrical interface (CEI) standards, which will combine parallel electrical signals at six, ten, and perhaps even 25 Gbits/sec. These efforts should extend even beyond 100G to allow efficient scaling of systems and networks.

Within the point-of-presence (PoP), central office (CO), or data center, the parallels continue. Several vendors are working on proposals to standardize 100-GbE formats around both ten lambdas x 10 Gbits/sec per lambda as well as four lambdas x 25 Gbits/sec per lambda. Other researchers have described results for serial transmission at 100G in industry forums like the Optical Fiber Communications (OFC) Conference. We should expect that all solutions could exist, just as 10 GbE supports SR, LR, LX4, ER, and soon LRM and even 10GBASE-T. Deployments will be driven, as always, by the economics in a given application.

Over the near to medium term, the economics of optical networks will strongly favor "super-lambda" parallel optical approaches to 100 GbE. Capital and operational costs of dealing with 100-Gbit/sec serial streams in the network are simply too high compared with the economics of multiple 10G or 40G streams combined into one 100G signal--in part because 100G data rates on optical fiber make compensation of non-linear effects very costly and because the parallel optical alternative is so cost effective, thanks to recent advances in photonic integration.

Taken together it is possible and even likely that future 100-Gbit/sec links will traverse the network and network equipment in many different parallel forms. As such, a signal traveling from one core router switch fabric to the next may start off as a 20 x 6.25G electrical signal to get to a 100-GbE MAC chip, continue electrically as a 10 x 10 Gbit/sec signal to reach an optical transmitter, where it is optically carried as a 4 x 25G from router to transport system. The signal would then traverse the network as a 10 x 10G DWDM wave and return to the next switch fabric in reverse order. All the while, as far as the routing protocol and network operator is concerned, this is a 100-Gbit/sec signal, end-to-end.

Conclusions

With adjacent router capacities already in excess of 40G in some large core networks, the industry needs to devote significant resources to developing 100-GbE technologies within systems, between systems, and across the WAN. Given the economics of DWDM networks, namely the advantages of multiple wavelengths compared to 100+G data rates per wavelength, the network has entered a phase in which "super-lambda" capabilities will be as important as lambda and sub-lambda services. Fortunately, the techniques and standards required to implement super-lambda services already are well understood and deployed with standards like 10GBASE-LX4, VCAT, and CEI. Extension of these standards will allow carriers to continue to decouple their IP services offerings and optical infrastructure decisions, leading to maximum flexibility, responsiveness, and scale.

Drew Perkins is the CTO and Rick Dodd is the senior director of product marketing at Infinera (Sunnyvale, CA). They may be reached via the company's Web site at www.infinera.com.