GÉANT SELECTS INFINERA PACKET-AWARE OPTICAL TRANSPORT NETWORK TO SOLVE "ELEPHANT FLOW" PROBLEM

∫infinera®

CUSTOMER
GÉANT

CHALLENGE
Overcome Ethernet LAG limitations to better support "elephant flows" above 10 Gb/s

Facilitate the upgrade path from 10 Gb/s to 100 Gb/s link speeds

Reduce the cost of packet forwarding on backbone routers

Flexibly upgrade GÉANT IP trunk capacity to keep up with traffic growth

SOLUTION
Use Infinera PXM in the existing DTN-X network

Enable traffic routing across super-channel backbone based on VLAN or MPLS tags

RESULTS
Almost a 50 percent reduction in expensive router ports compared to non-packet-aware solutions

Ability to evolve to dedicated 100 Gb/s backbone trunks as demand increases

Forms the foundation of a packet-aware, software-defined network-capable Intelligent Transport Network

GÉANT

GÉANT is Europe's leading collaboration on network and related e-infrastructure and services for the benefit of research and education. The knowledge-sharing among members leads to innovation that contributes to Europe's economic growth and competitiveness. GÉANT operates a pan-European super-channel core network with nodes in 19 countries that includes high-speed connections to the Large Hadron Collider in Switzerland and supercomputing centers across Europe.

The GÉANT project is a European success story. For over 15 years, through the joined forces of national research and education network (NREN) organizations, the project has been a vital element of Europe's e-infrastructure strategy. GÉANT provides the high-speed connectivity needed to share, access and process massive volumes of data generated by, and essential to, diverse research and education communities working in areas such as particle physics, bioinformatics, earth observation, drug discovery, and arts and culture.

GÉANT's Challenge

The rapid growth in network traffic has challenged GÉANT's ability to grow the network beyond 10 gigabit per second (Gb/s) flows in a way that ensures network performance is both responsive and cost-effective. While Ethernet link aggregation group (LAG) works well for conventional flow sizes, there are well-documented issues with "elephant flows," which are flows with bandwidth demand that exceeds the individual link data rate of the LAG group.

Figure 2 shows the problem. Traffic flows from A, B and C to the remote location X (shown as the green dashed lines) consist of conventional, mixed flow sizes, with none of the
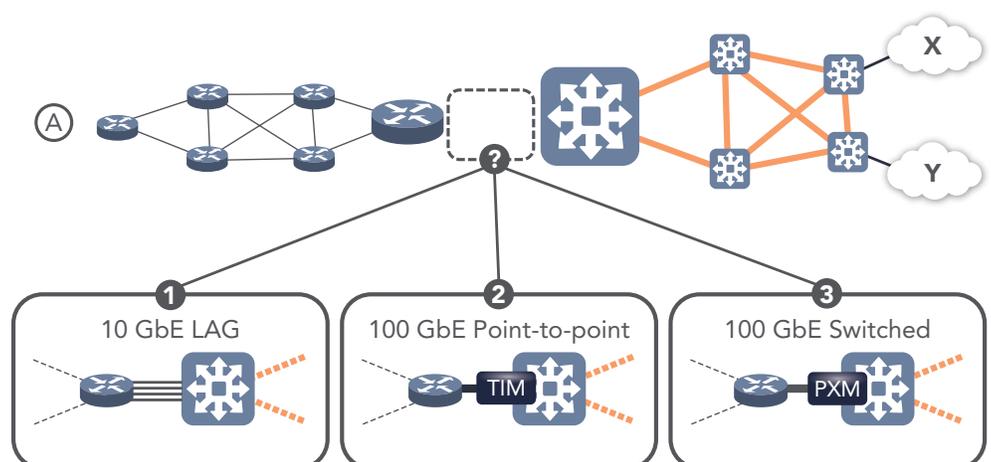

Figure 1: Three Options for Scalable Evolution from 10 Gb/s to 100 Gb/s Service Rates
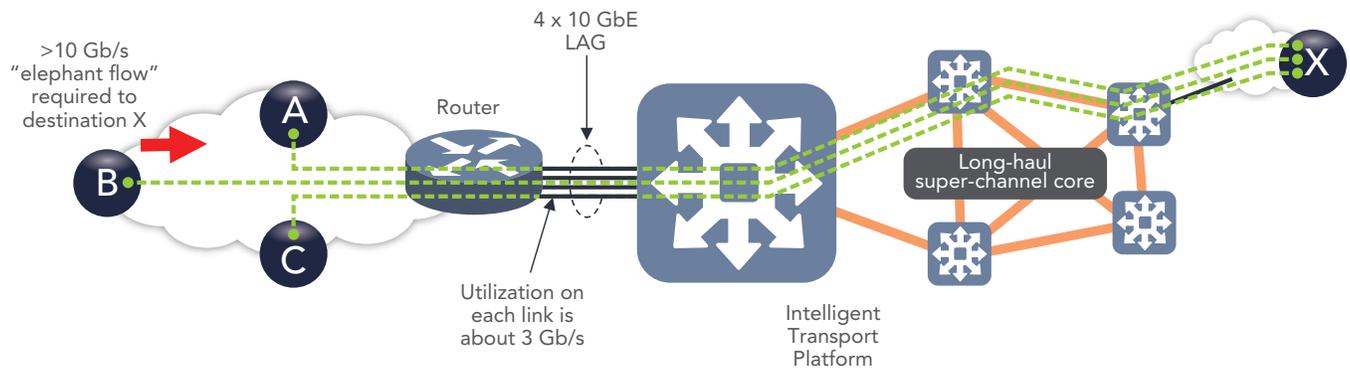
Figure 2: Using Ethernet Link Aggregation Groups with N x 10 GbE Links

flows exceeding about 3 Gb/s. Ethernet LAG will effectively distribute the traffic from these, and many other sub-10 Gb/s flows, across the four 10 gigabit Ethernet (GbE) links between the router and the DTN-X Intelligent Transport Network. This enables GÉANT to cost-effectively support aggregate traffic using an N x 10 GbE configuration on the router because 10 GbE router ports are relatively inexpensive.

The challenge occurs when the customer at location B, for example, generates a flow greater than 5 Gb/s (indicated by the red arrow in Figure 3). These flows are usually referred to as "elephant flows" and, because of the limitations of the Transmission Control Protocol (TCP), Ethernet LAG will try to direct this flow over a single link within the LAG. But because the flow size exceeds the capacity of individual LAG links, the only options are either to buffer the flow over a single link, or to fragment the flow over multiple links. In the case of sustained elephant flows, buffering will inevitably lead to packet loss, while fragmenting the flow may result in out-of-sequence packets arriving at the destination. Both of these conditions will cause retransmission of entire TCP windows, resulting in the sustained overloading of one or more network links.[1]

The problem with the configuration shown in Figure 3 is the need to support both N x 10 GbE for conventional

"Infinera's packet optical solution offers an effective approach to optimize the GÉANT network and efficiently and optimally handle large traffic flows. We look forward to realizing the expected efficiencies and operational benefits of this solution, and continuing our excellent collaborative work with Infinera on future applications."

Mark Johnston
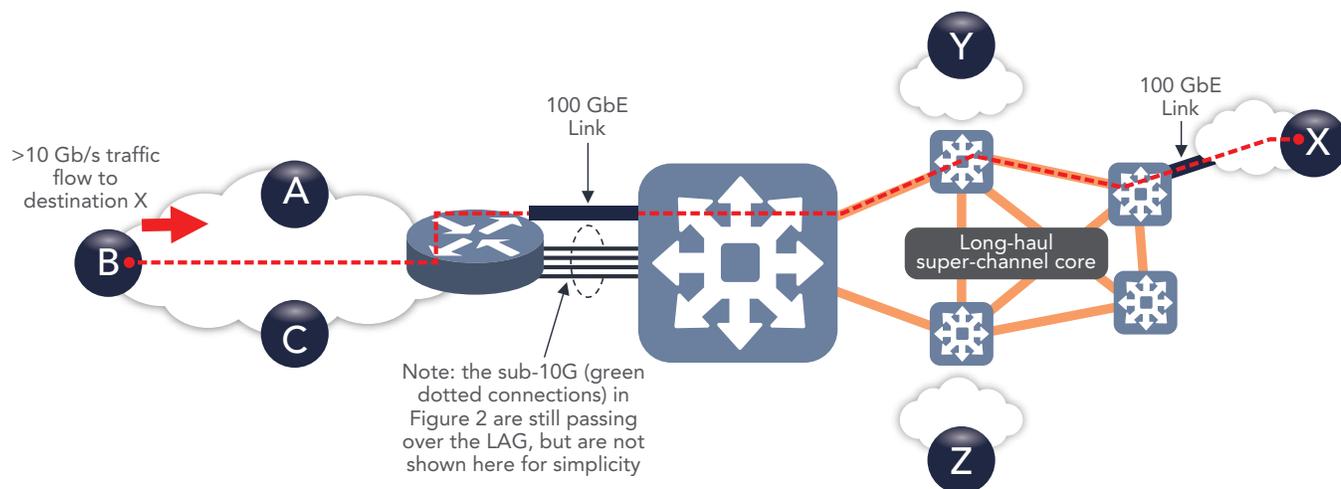*Chief Network Operations Officer*
GÉANT

Figure 3: Adding a 100 GbE Point-to-point Link to Support Elephant Flows

traffic patterns and a 100 GbE link dedicated to the elephant flows between locations B and X. GÉANT and Infinera worked together on a solution to maximize the utilization of their existing DTN-X Intelligent Transport Network to resolve these challenges.

The Solution: Enhanced Packet Switching

The Infinera DTN-X platform was chosen by GÉANT in 2011after a rigorous competitive bidding process. GÉANT particularly valued the ability of Infinera's multi-layer Intelligent Transport Network to provide terabit-class scalability, operational simplicity and efficiency, and high levels of flexibility and programmability. In 2012 GÉANT upgraded the transmission layer of their network using Infinera to enable multi-terabit capacity to be delivered to research and academic institutions in 19 European countries. Infinera recently enhanced the DTN-X Family by adding a Packet Switching Module (PXM) that provides advanced packet processing and quality of service (QoS) capabilities for Ethernet and multi-protocol label switching (MPLS) traffic flows.

Figure 4 shows how Infinera's DTN-X Platform with the PXM solves GÉANT's elephant flow problem. The PXM is a packet-aware Optical Transport Network (OTN) interface that identifies standard Ethernet

"One of the aspects of the Infinera solution that we really appreciate is the extensibility of the DTN-X. Being able to add new capabilities with the Packet Switching Module enhances the services we offer our users and extends the life of the network infrastructure. It's what makes the DTN-X a genuine platform, and not just a product."

Mark Johnston
Chief Network Operations Officer
GÉANT

>10 Gb/s traffic flow to destination X

100 GbE Link

100 GbE Link

Note: The N x 10 GbE LAG is no longer required
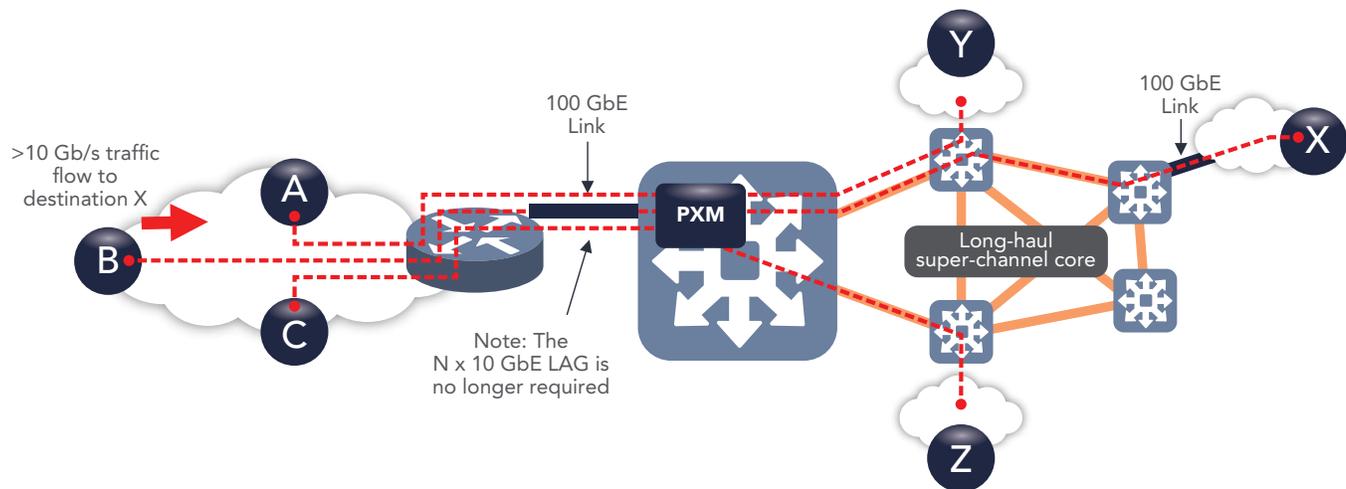
PXM

Long-haul super-channel core

Figure 4: Using the Packet-aware PXM to Cost-effectively Support Elephant Flows

virtual local area network (VLAN) tags and MPLS labels to direct traffic via the DTN-X backplane, and on to the destination across the super-channel backbone. This is achieved by the existing routers assigning aggregated flows to specific VLAN-tagged channels, upon which the PXM's role is to forward these flows onto IP trunks that have been dimensioned to suit.

The connections from A to Y, B to X and C to Z shown in Figure 4 can be individually sized using the ITU-T ODUflex protocol. Note that this single, 100 GbE link not only replaces the 100 GbE link shown in Figure 3, but it can also replace the N x 10

GbE links used for conventional traffic flows. This approach effectively "right-sizes" the OTN capacity available in the core, allowing resources to be utilized more efficiently. ODUflex resizing enables changes to be made dynamically in response to changing demands, and classic packet shaping and policing in the PXM can be matched to Metro Ethernet Forum Carrier Ethernet 2.0 bandwidth profiles. Furthermore, right-sizing in this way allows a smooth and cost-effective evolution to dedicated 100 Gb/s links as traffic flows increase by simply migrating high-demand flows onto dedicated point-to-point 100 Gb/s connections, and thereby freeing up the PXM capacity

to support new sub-100 Gb/s elephant flows.

The Results: Cost-effectiveness Through Packet Aggregation

GÉANT has completed the initial deployment of PXM interfaces in the portion of the backbone that carries traffic to Eastern Europe. This traffic travels over fiber rings in Austria, Croatia, Czech Republic, France, Germany, Italy, Hungary, Slovakia, Slovenia and Switzerland.

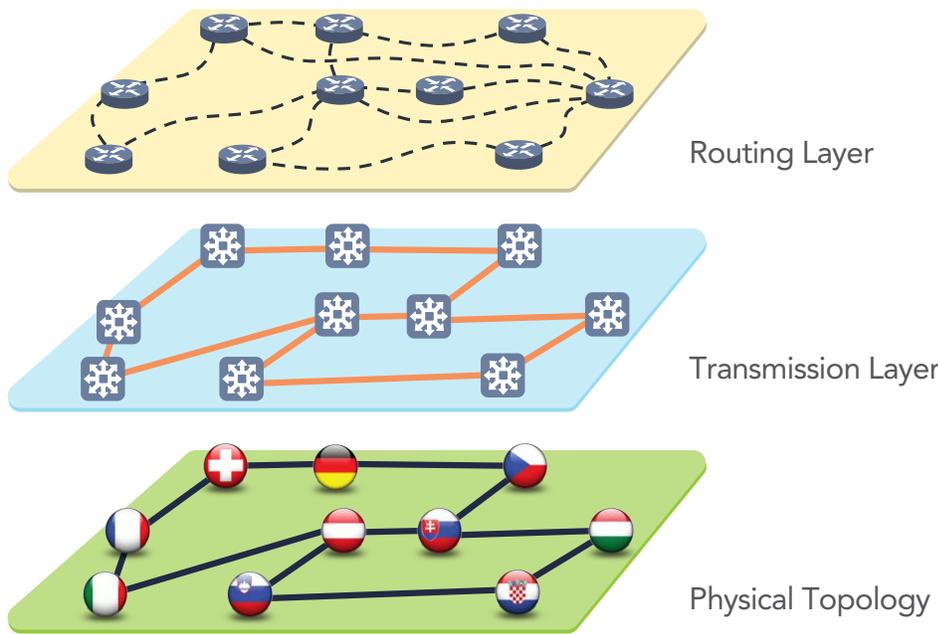As a result of deploying Infinera's Packet Switching Module, GÉANT has the

Routing Layer

Transmission Layer

Physical Topology

Figure 5: Physical, Transmission and Routing Topologies can be Common or Distinct

option to reduce the number of expensive router ports needed by almost 50 percent compared to non-packet-aware solutions, and as the initial pilot deployment is scaled up. In one of the design options, GÉANT would be able to

reduce the number of 100 GbE router ports from eight to five by replacing eight Infinera 100 Gb/s ports with five PXM cards—a reduction factor that would continue to accumulate as GÉANT extends the roll-out to create a more meshed packet-aware topology. In addition, GÉANT has the ability

to transition higher-demand connections to dedicated backbone trunks as demand increases. The deployment of PXMs also forms the foundation of a packet-aware, Intelligent Transport Network that can be controlled at a more granular level using a software-defined networking solution such as Infinera's Xceed Software Suite.

For more information on *Infinera Intelligent Transport Networks, click https://www. infinera.com/company/contact-us/ to contact us.*

References

1 Hacker, T.J. University of Michigan. "The Effects of Systemic Packet Loss on Aggregate TCP Flows". SC '02: Proceedings of the 2002 ACM/IEEE Conference on Supercomputing.

Global Headquarters
140 Caspian Court
Sunnyvale, CA 94089
USA
Tel: 1 408 572 5200
Fax: 1 408 572 5454
www.infinera.com

Asia and Pacific Rim
Infinera Asia Limited
8th floor
Samsung Hub
3 Church Street
Singapore 049483
Tel: +65 6408 3320

Europe, Middle East, Africa
Infinera Limited
125 Finsbury Pavement
London EC2A 1NQ,
United Kingdom
Tel: +44 207 065 1340

Customer Service and Technical Support
North America
Tel: 877 INF 5288
Outside North America
Tel: 1 408 572 5288

For more information
Contact Us
infinera.com/contact-us

∫infinera®