Realizing Packet/Optical Network Convergence and Efficient Router Interconnect through Packet-aware Optical Transport

Chris Liou
Infinera Corp., 140 Caspian Ct, Sunnyvale, CA 94089, USA.

Abstract

In conventional network architectures, where IP/MPLS operates as a client layer over an optical transport layer, routers are employed throughout the WAN to provide a multitude of packet-centric bandwidth management functions, including packet services termination, aggregation, switching, QoS and transport.

New emerging high-speed packet management functions at 100GbE rates are emerging within next-generation optical transport systems that combine WDM, OTN and packet functionality in a way that provides cable operators with more cost-effective tools for optimizing packet traffic, increasing network resource efficiency, and reducing the total cost of the combined network.

This paper compares and contrasts the current PMO with an architectural approach based on packet-aware OTN/optical transport, discusses typical use cases, and presents the benefits this provides operators.

INTRODUCTION

The growth in packet traffic resulting from IP-based services, various cloud initiatives, and the shear growth in data/video-centric services is creating new strains on the operator's infrastructure as the challenge of growing or maintaining ARPU continues. As new competitive forces gain more traction in the marketplace, additional pressure is placed on network operators to not only differentiate their product offerings to gain and retain customers, but also ensure that the network infrastructure employs the best technologies and solutions for enabling the right combination of scalability, reachability, survivability and economics.

This presents a challenge for the network architect who must not only understand the relevant networking technologies applicable to the MSO market, but also consider new advances in technologies when exploring the best way to achieve optimal networking economics. One important technology advance that is driving network upgrades is the introduction of 100Gb/s coherent optics and the resulting fiber capacity that it enables. The upgrade from 10Gb/s to 100Gb/s transmission is vital for supporting the growth in network traffic and is being broadly deployed today.

While the use of 100Gb/s Layer 0 WDM technologies is a broadly accepted strategy for fiber capacity upgrades, there are many options for building the next layers of networking. Two methods used today include building a router-based L2/L3 layer directly over 100Gb wavelengths and building a router-based layer over a converged OTN/WDM layer. In both of these models, topology amongst routers is generally defined using IP link sizes commensurate with the underlying "wavelength" service, which is either 100GbE in the case of IP over WDM or a mix of 10/40/100GbE in the case of IP over OTN/WDM. All packet-level functions (statistical multiplexing, aggregation, QoS traffic management, switching) is relegated to the router layer, and the optical transport layer is providing router interconnect via transparent pipes.

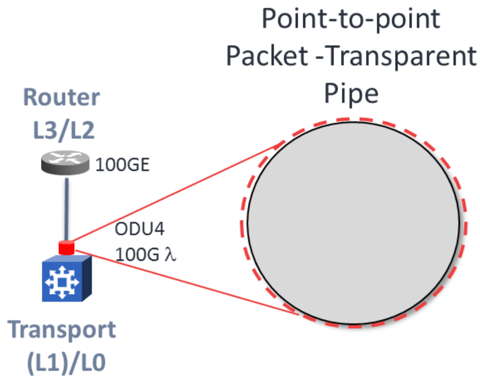

Fig 1 – Traditional approaches leverage optical transport layer for L0/L1 functions only, with all packet traffic processing performed solely by routers. The transport layer is oblivious to packets.

With new innovations emerging within the transport layer that integrates high-speed packet capabilities along with traditional L1/L0 functionality, however, new architectural options are available for building cost-efficient, scalable networks. The convergence of packet and optical transport capabilities into an integrated system, along with Transport SDN, is giving rise to a new set of tools that can help improve resource utilization across both layers, eliminating over-provisioning of router links and allowing the transport layer to cost-efficiently offload certain transport functions currently performed by routers.

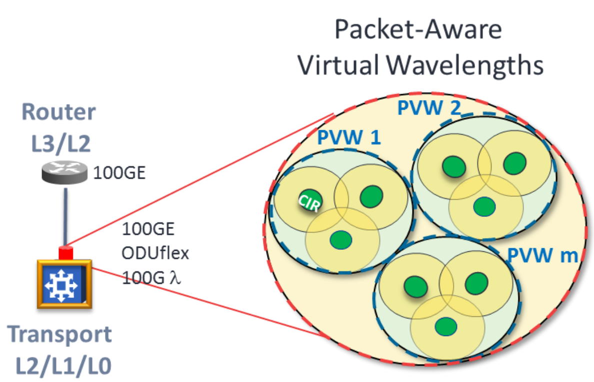

Fig 2 – Packet-aware transport employs PVWs to add packet processing, enabling more flexible packet-centric transport functionality that improves overall network cost and offloads routers from basic packet transport functions.

This paper discusses the architectural capabilities facilitated by a packet-aware transport layer that combines L2/L1/L0 functionality and compares this against current architectures. It introduces the concept of a Packet-aware Virtual Wavelength (PVW) and discusses how it plays a role in evolving the transport network, as well as the flexibility and savings it can bring to the combined IP/MPLS and optical transport network layers.

PMO CHALLENGES

The current approach for building MSO core networks generally involves the construction of a router layer that runs over an optical transport layer that is either based on Fixed or Reconfigurable OADM technologies for steering analog wavelengths between routers or on converged OTN/WDM technologies that digitally switch and groom router port traffic on to optical wavelengths.

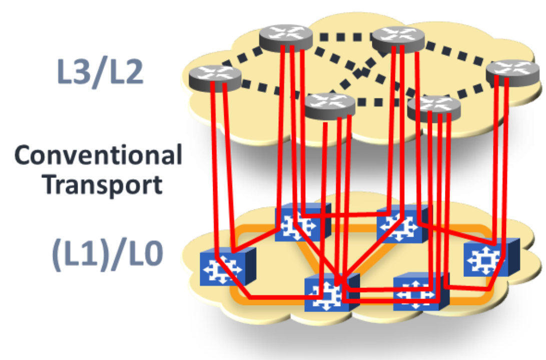

Fig 3 - Conventional architecture relies on routers to perform all packet functions. IP links and topology are rigidly coupled to wavelengths. This presents scale challenges when attempting to bypass transit routers and achieve fuller mesh connectivity between edge routers

The choice in optical transport technology is influenced by a number of factors beyond cost and fiber capacity – other considerations such as protection and restoration capabilities, operational complexity, reliability, and speed of service turn-up play an important role.

The cost associated with delivering traffic across the network is predominantly influenced by the amount of bandwidth required at both the transport layer as well as the IP/MPLS layer. At scale, these costs are generally dominated by the total number of router ports required along with the total number of wavelengths required. The choice of network architecture to support a network's traffic demands involves determining what type of traffic to handle at which layer, identifying the appropriate service recovery architecture to protect against failures, and understanding where opportunities for "arbitrage" exist, i.e., determining the most appropriate layer to transport bits with an end objective of minimizing the cost/bit metric, while still meeting service SLA objectives.

Lack of Cross-Layer Coordination

Traditionally, networks are often planned, engineered, and deployed layer by layer. One organization generally determines the best solutions for the router layer to support all the L2/L3 packet-based services, and a separate organization generally determines the transport solution, focusing on L0/L1, using inputs from other organizations to determine overall traffic forecasts. Additionally, each organization may have strategies for protecting against network failures. This often gives rise to inefficiencies, such as in the form of over-provisioning, non-optimal placement of traffic flows, or non-optimized use of available resources at each of the network layers.

Inefficient Network Resource Utilization

One of the reasons for networking inefficiency can be attributed to where packets are managed and where certain networking features such as protection are accessed. Traditionally, all packet functions from service termination and adaptation to traffic management and aggregation and even fast protection against network failures is performed within the router. While this may provide operational convenience, it does not leverage the lower-cost capabilities of the optical transport layer and can lead to inefficient usage of router resources.

One common practice for reducing the costs at this layer consists of leveraging the transport layer to enable more router bypass and minimize intermediate packet hops, thereby reducing the transit router tax [1]. This approach, however, runs into challenges as the network migrates to 100G transmission, as direct connectivity to destination routers requires a dedicated port and dedicated wavelength service. Full mesh connectivity between edge or service routers using dedicated ports would eliminate the need for transit routers but is often impractical and economically non-viable beyond small networks. In large networks, optical bypass opportunities generally focus on network hot-spots where there is sufficient amount of persistent traffic to warrant express, while other sites interconnect via a set of transit core routers.

Over-provisioning due to FRR Protection

The reliance on routers to perform protection is another source of inefficiencies, as spare bandwidth in the router layer required for potential network failures also translates into spare bandwidth in the transport layer. Over-provisioning of the router layer by a factor of 2x to support router-based protection schemes like Fast Reroute (FRR) is not uncommon. As an alternative or supplementary approach, some networks employ 1+1 protection at the optical layer, which can achieve sub-50ms protection transparent to the IP/MPLS layer, but which requires 100% dedicated protection bandwidth at the optical layer. Other multi-layer recovery strategies to help increase the utilization of router links include combining optical wavelength restoration along with MPLS FRR [2], but with sub-50ms protection still resident within the router layer, only so much reduction in overprovisioning can be achieved, as some spare router bandwidth must be available for failover scenarios, even if just for high priority traffic, such as Expedited Forwarding (EF) or Assured Forwarding (AF) DiffServ traffic.

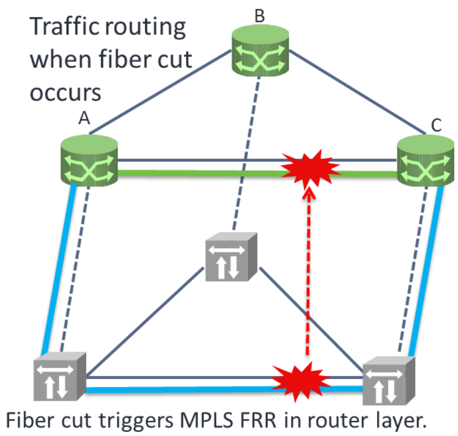

Fig 4a – with MPLS FRR, a fiber cut triggers protection at higher router layer, consuming spare resources at both the router and optical layers.

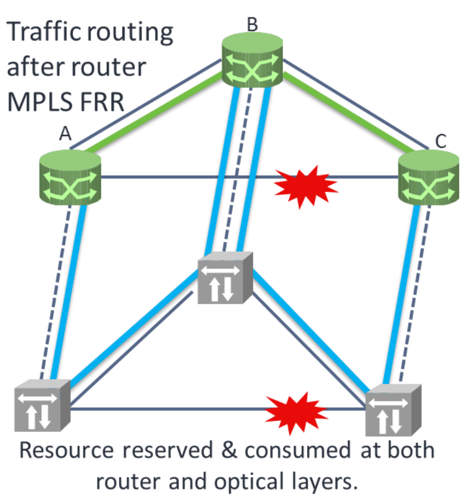

Fig 4b – Traffic from A to C is restored via router B in sub-50ms, but requires over-dimensioning of IP links, which also requires corresponding spare optical capacity.

With advances in the optical transport layer, however, new alternatives to relying solely upon routers to host basic packet transport, aggregation, multicast, and protection schemes are now emerging. These new tools are enabling architects to more efficiently manage traffic internally within the transport layer, thereby reducing the amount of traffic that needs to be processed by routers.

PACKET AWARE TRANSPORT

State of the art converged L1/L0 transport systems today combine sub-wavelength granularity circuit switching fabrics (typically OTN) along with WDM transmission, enabling not just multi-service and multi-rate transport services over the optical infrastructure, but also efficient switching and grooming of services into wavelength containers. These platforms can also support new advanced, deterministic sub-50ms shared mesh protection (SMP) schemes [3,4,5]. Such systems provide operators with the ability to decouple bandwidth services from the analog transmission layer, and facilitate a much more flexible operational model for dynamically provisioning transport bandwidth services, when compared to traditional transponder/muxponder and ROADM based approaches.

The next evolutionary phase of the optical transport network involves adding packet-awareness – the ability to perform Ethernet and MPLS packet processing (e.g., classification, metering, traffic management, QoS) at 100GbE speeds (and lower), mapping packet flows based on configuration rules into multiple flexibly sized ODUflex circuits, and expressing these individually encapsulated packet flows within the transport network to their end destinations.

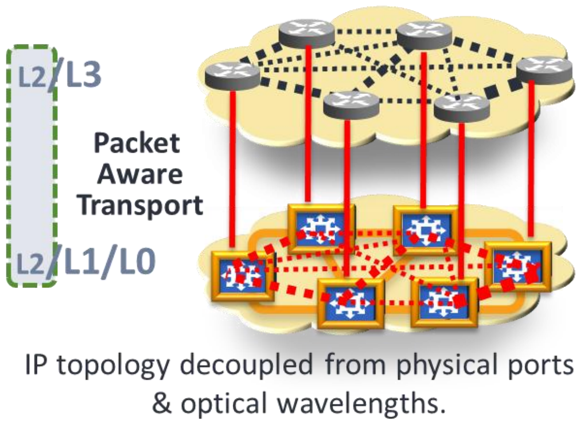

Fig 5 – packet-awareness in the transport layer converges L2/L1/L0 functionality, enabling more flexible and bandwidth-efficient router interconnect, as well as natively supporting Ethernet E-LINE/E-LAN/E-TREE services. L2 technology in both the router and transport layer enables better cross-layer resource coordination.

This convergence of L2/L1/L0 functionality transforms the transport network from one that delivers only 10/40/100G transparent wavelength services between routers to one that can natively support flexible packet-transport interconnect solutions without necessarily having to pass traffic back up to the router, thereby offloading routers and router ports from having to handle L2 transport functions and the associated traffic, and instead focus on higher-level L2/L3 service delivery. The new fundamental networking building block this enables is a **Packet-aware Virtual Wavelength** (PVW).

Packet-aware Virtual Wavelengths Overview

PVWs enable the transport layer to now become fully aware of the packet traffic coming from the routers while providing the performance and predictability of switched sub-wavelength granularity circuits. With Ethernet packets serving as a common technology denominator between the router and optical layers, greater synergies can be realized between both network layers. In addition to packet processing and traffic management fnctions typically found in routers and switches, PVWs can map individual packet flows into individual ODUflex circuits at the optical layer and apply integrated L1 and L2 QoS mechanisms to concurrently manage packet and circuit characteristics. Flexible packet mapping can be performed based on address fields such as VLAN IDs or MPLS labels, and can be further enhanced through the use of ACLs to perform policy-based forwarding.

As in packet systems, PVWs can apply different CoS models, such as flow-based and class-based queuing, and apply them on a per-client and per network-facing interface. Internally, MPLS-TP Pseudowires (PWEs) can provide a scalable encapsulation mechanism for the packet flows, before being mapped into ODUflex container circuits for optical transport. With both L1 and L2 functions localized, the amount of packet oversubscsription can be well coordinated with the size of the ODUflex circuit in increments of 1.25Gb/s, creating the cross-layer glue between packets and circuits and ensuring a proper over-subscription ratio is maintained for optimal balance between performance and economics. As more or less dedicated bandwidth is needed for the mapped packet flows, the ODUflex circuits can be hitlessly resized.

The seamless integration between L2 and L1 also provides a vehicle to differentiate the transport of packet services, even between the same source/destination routers. Routers that are adjacent from the L3 control plane perspective can use PVWs to distinguish single-hop traffic, based on some packet attributes, such as priority bits. High and low latency transport paths, for example, can be used to distinguish different application flows coming from a single router port. This provides operators with simpler way to control how transport network assets are utilized for supporting specific packet traffic flows.

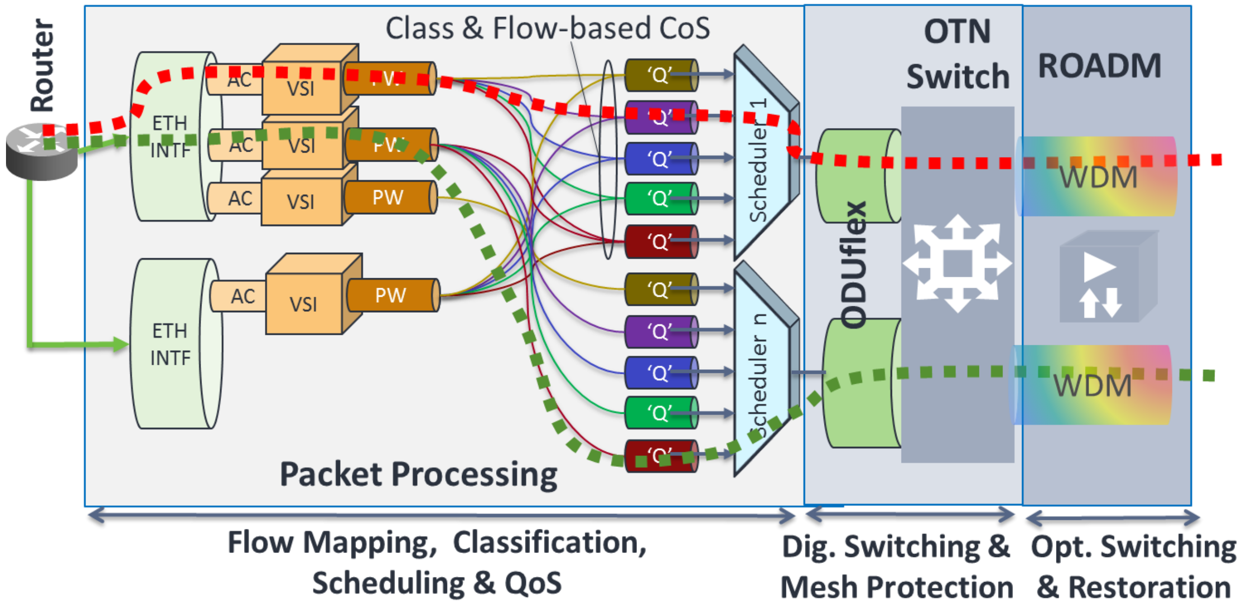

Fig 6 - Integration of L2 packet processing functions along with L1 OTN switching/grooming and L0 WDM ROADM switching creates a highly flexible packet-transport and service delivery platform. Highly differentiated transport capabilities can be offered for packet flows from routers, as well as native MEF Ethernet services.

<u>Flexible Packet Transport Services</u>

The flexible mapping of packet flows from the router into PWEs and then into ODUflex circuits also gives rise to new point-to-point (P2P), point-to-multipoint (P2MP) and multipoint-to-multipoint (MP2MP) packet transport services (including Carrier Ethernet services), delivered natively by the transport network. In addition to E-LINE type services (EPL, EVPL), PVWs can enable E-TREE and E-LAN based services without relying on an external system. Both Ethernet UNI and ENNI based services are supported, with full support for untagged traffic as well as C-VLAN and S-VLAN tagged traffic.

Standards-based L2 traffic management capabilities within the PVW facilitate the support for MEF 2.0 compliant services. Packet metering, queuing and scheduling along with standard CoS support and configurable ingress and egress bandwidth profiles provide a full complement of packet QoS capabilities for hosting native high-speed Ethernet services directly on the transport platform, up to 100GbE.

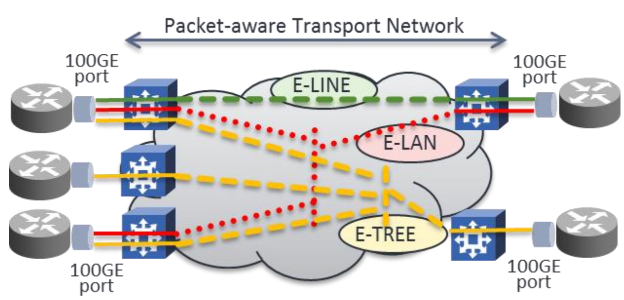

Fig 7 – The transport network can cost-effectively offload routers for high-speed Ethernet services, including E-LINE, E-TREE, and E-LAN.

Unlike conventional packet-switched networks for delivering packets to their destination, however, PVW-based networks intrinsically provide fine-granular, guaranteed dedicated bandwidth directly between the service termination sites, as the transport network fabric is based on switched ODUflex circuits. This ensures low-latency transport of traffic and guaranteed jitter for packet services, compared to packet-switched networks, since packet flows within PVWs are not subject to congestion or contention from any flows other than the ones mapped into these transport tunnels at the source location.

Since separate PVWs can be used for different flows originating from a common router interface, operators can leverage this as a new tool for engineering differentiated, bandwidth efficient, SLA-enforceable packet transport services. PVWs provide operators with controls over typical L2 QoS parameters as well as L1/L0 circuit-centric bandwidth metrics (size, latency, transport layer path diversity, etc.). Oversubscription can be controlled per PVW, as well as the protection or restoration policy as part of a multi-layer protection strategy.

For P2MP and MP2MP services such as E-TREE and E-LAN, PVWs facilitate the creation of isolated transport overlay meshes, interconnecting multiple service termination sites. Multi-point EVCs can then be established that can perform the MAC-based learning and forwarding needed for ETREE and ELAN services. This packet transport capability can be leveraged not just for hosting high-speed MEF 2.0 multi-point services within the transport layer, but also for simplifying and reducing the cost of implementing IP multicast based services, as described later.

Transport-layer Protection of Packet Services

The convergence of L1 and L2 technologies into a packet-aware transport network also provides a rich set of flexible protection schemes at both the packet and OTN layers, and can significantly reduce the protection bandwidth requirements at the router layer. State-of-the-art packet-aware transport systems can offer ITU G.808.3 and G.ODUSMP compliant Shared Mesh Protection (SMP) schemes that provide resiliency against network failures in under 50ms for PVWs. Fast SMP schemes provide the performance guarantees typically associated with 1+1 protection schemes, but with much better bandwidth efficiency, by utilizing shared protection bandwidth and sophisticated algorithms for ensuring sub-50ms performance times for all impacted SMP-protected PVWs. Priority and preemption capabilities further enhance the differentiated transport protection schemes available.

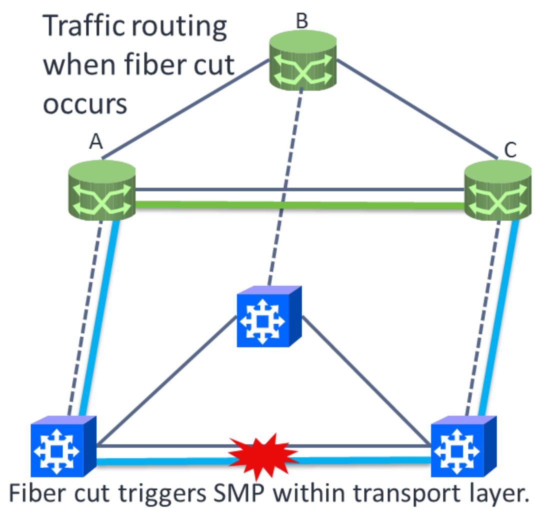

Fig 8a – with transport-layer SMP, the fiber cut is reacted to locally within 50ms.

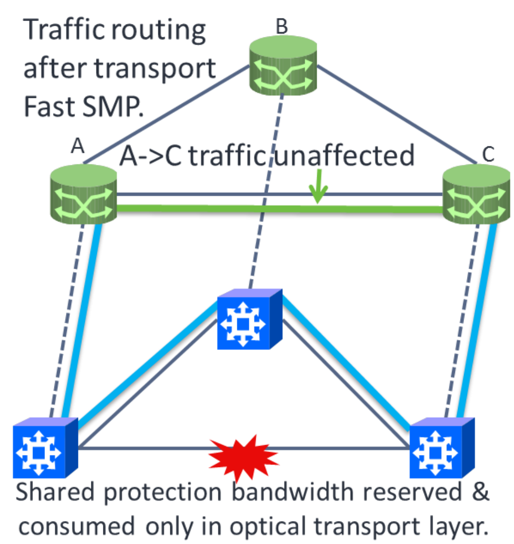

Fig 8b – since the SMP protection restores traffic within 50ms, the A->C traffic is restored without any impact to the LSP path. This reduces LSP churn and reduces over-dimensioning of IP links for protection bandwidth.

Fast SMP gives network operators a new alternative to MPLS FRR for protecting packet services, thus reducing overprovisioning of bandwidth at the router layer [5]. A proper mix of schemes at both the router and transport layers can be utilized, depending on use-case scenarios. For example, in a long-haul core backbone, larger LSPs or flows aggregated into PVWs can be mesh protected at the transport layer, closer to where network failures occur, while smaller flows can still be protected at the router layer, based on the operator's own traffic mix and service protection requirements. Operators that use PVWs to distinguish different levels of business services by mapping them to individual SLA-managed PVWs can apply different protection strategies for further differentiation.

Additionally, optical layer restoration leveraging ROADMs can provide operators with yet another layer of network resiliency by restoring optical capacity at the wavelengths and super-channels granularity. Such analog restoration schemes generally take several seconds, and can complement the sub-50ms performance of digital protection schemes when additional resiliency is required.

CABLE NETWORK APPLICATIONS & USE-CASES LEVERAGING PACKET-AWARE TRANSPORT

With the features, flexibility and economic benefits packet-aware transport networking can offer, cable operators can leverage this new set of capabilities across multiple applications. Some key use-case applications are described below.

Transport Layer Master Head-End Packet Aggregation

In metro and regional cable networks, it is common to find hub-and-spoke packet traffic patterns as traffic is aggregated upstream from multiple head-ends or hubs to a master head-end. With a conventional optical transport network, the packet-layer aggregation function is performed in routers, sometimes before passing traffic to the optical transport layer. In other scenarios, dedicated wavelengths are used to hub the routers from multiple sites along a ring back to the hub site, where the wavelengths terminate into a larger router.

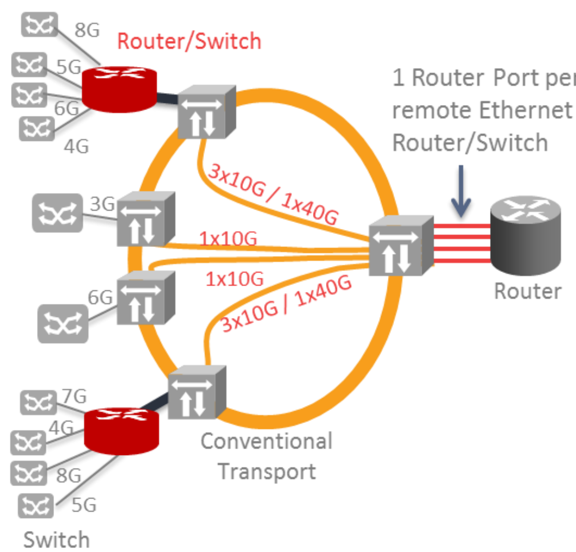

Fig 9a – in a conventional architecture, router ports aggregate packet traffic and transport the packets over dedicated wavelengths back to the hub location.

Packet-aware transport offloads the packet aggregation function from the router and handles it natively within the transport layer, thereby reducing router ports or potentially eliminating the aggregation router altogether. Multiple 10Gb/s ports can be aggregated locally at edge locations, while at the master head-end, packet traffic from multiple sites can be aggregated and handed off to a core router. Additionally, with the PVWs being configurable in 1.25Gb/s increments, transport bandwidth can be efficiently allocated and resized, based on actual packet traffic load, instead of burning an entire optical wavelength.

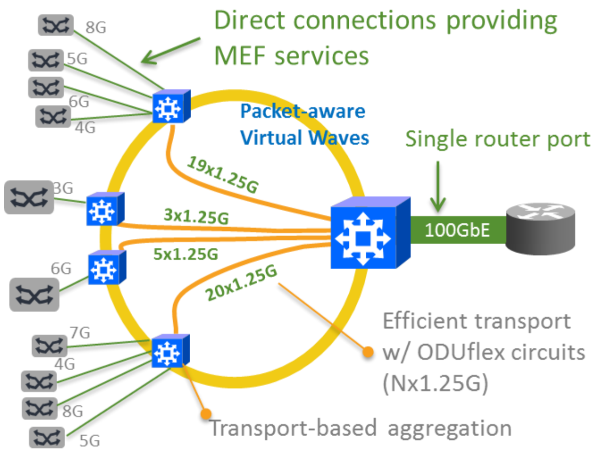

Fig 9b – in the packet-aware transport architecture, aggregation is handled natively in transport layer, and an appropriate amount of optical bandwidth is allocated, eliminating stranded capacity. In this example, the aggregation router is no longer needed.

Cost-efficient, Scalable LSP Protection

In multi-layer long-haul core networks that overlay IP/MPLS over an optical transport layer, the combination of packet-aware transport and Fast SMP can enable a much more cost-efficient traffic-engineered backbone through various router offload and bypass tools.

With MPLS-TP packet aware transport, LSPs can be mapped and aggregated into appropriately dimensioned ODUflex circuits at the optical layer, creating express lanes through the multi-layer network and bypassing intermediate LSRs. Transport of LSPs between the edge routers can further be differentiated through the use of separate ODUflex circuits for different groups of LSPs, each of which may have different latency characteristics and which may employ different protection schemes. New protection schemes such as Fast SMP can be used in lieu of MPLS FRR to provide sub-50ms protection capabilities, but with better bandwidth efficiency, resulting from reduced over-dimensioning at the router layer. By aggregating multiple LSPs into Fast SMP protected ODUflex circuits, a scalable end-to-end recovery scheme is enabled while simultaneously reducing LSP churn.

High-speed Pay-as-you-Grow Ethernet Service Delivery

A packet-awareness transport system can natively support MEF 2.0 services for high-speed interfaces, and offload higher-layer routers/switches. High-bandwidth Ethernet services that are hosted on routers/switches consume resources at both network layers, and incur interconnect costs associated with transporting the service from the router port to the optical transport layer.

By hosting services such as fractional 100GbE EVPL or EPL services directly off the transport platform, the cost of utilizing router/switch resources and the interconnect cost for those services are eliminated. Furthermore, an optimal amount of optical capacity can be allocated in increments of 1.25Gb/s to transport MEF services efficiently without stranding capacity, as can be the case when PVWs are not utilized, and an entire wavelength is allocated to transport an Ethernet service (e.g., 50Gb/s) that is less than the optical transport wavelength (e.g., 100Gb/s). PVWs also provide coordinated cross-layer configuration of L2 QoS parameters along with L0/L1 bandwidth characteristics for Ethernet services. Customer Ethernet services can be individually classified into ODUflex circuits, each of which can be switched to different destinations, and which can offer different latency, diversity, and protection/restoration characteristics. Hitless resizability of the PVW ensures the proper amount of capacity is provided for the supported Ethernet services, facilitating bandwidth-efficient pay-as-you-grow services for customers.

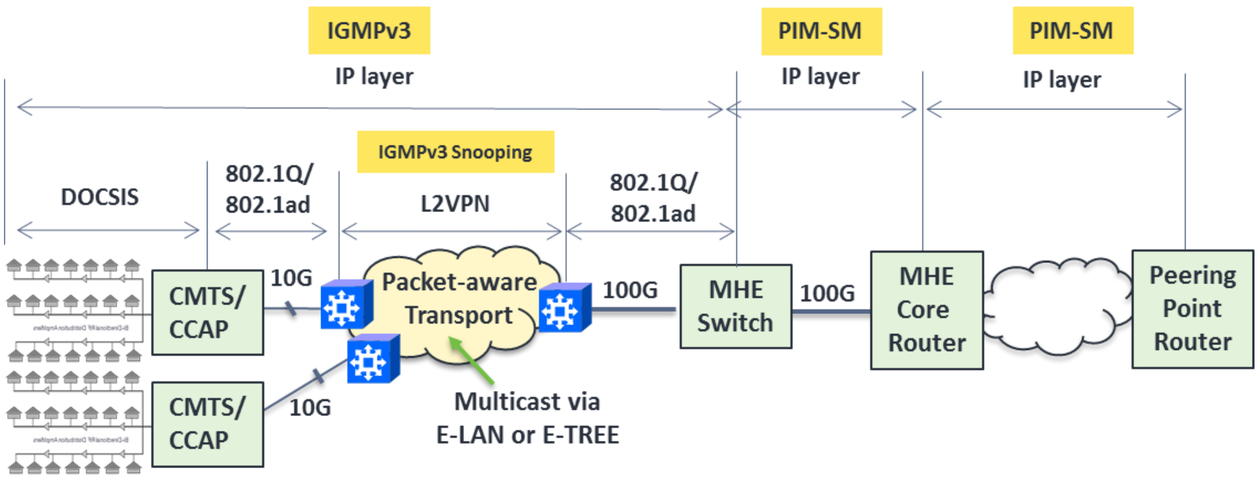

Fig 10 – IPTV multicast based on embedded E-LAN/E-TREE functionality and IGMP snooping within the transport layer can simplify the network by eliminating the aggregation router at CMTS/CCAP sites.

Transport Layer IPTV Multicast

Multi-cast IPTV is an increasingly strategic broadcast digital video delivery method for many cable operators as they make the transition to all-IP based service delivery. Conventional IPTV architectures leverage the router infrastructure to perform all the multi-cast functions – the combination of one or more routers at the master head-end and at each of the regional head-ends interconnected with static wavelengths generally provides support for IGMP/PIM and handles the multi-cast control plane and data plane functions.

With configurable E-LAN and E-TREE services natively available within the packet-aware transport network, however, packet flows can be multi-cast within the transport layer instead of within the router. This can eliminate the need for routers at the regional head-end performing packet aggregation from multiple local CMTS/CCAPs, and additionally can help enhance the utilization of the router port at the master head-end by relieving it from performing the multi-cast function. IGMP snooping enabled in the transport layer for each ELAN/ETREE instance ensures proper establishment of multi-cast forwarding entries downstream of the master head-end router to the CMTS/CCAP.

SUMMARY

The growth in bandwidth requirements on MSO networks necessitates exploration of new technologies and networking architectures in order to scale networks economically. The integration of packet-awareness into the converged L0/L1 network enables an evolution of the optical transport network from one that delivers point-to-point packet-agnostic bandwidth pipes to one that can deliver more dynamic, flexible packet transport services. This not only enables the transport network to natively host packet services such as E-LINE, E-LAN, and E-TREE services, but also facilitates more flexible and bandwidth-efficient interconnect solutions for routers, reducing the total cost of the network. The features enabled by packet-aware transport networks create a set of new networking tools available to MSOs, and include:

- Flexible mapping of packets into PVWs based on packet header fields, including labels & IDs, or additional fields such as address or packet classification information

- Coordinated cross-layer configuration of packet & transport QoS parameters
- Per-PVW protection and restoration options, including sub-50ms SMP, serving as an alternative to MPLS FRR
- Integrated P2MP and MP2MP multicast forwarding for E-TREE and E-LAN support
- Hitless resizability of the optical bandwidth associated with PVWs to accommodate elastic scale-up and scale-down packet traffic

With these tools, many cable network applications can be more efficiently and cost-effectively implemented as compared to today's conventional approaches.

ACKNOWLEDGEMENTS

The author would like to acknowledge Gaylord Hart for reviewing this paper and providing insightful comments.

REFERENCES

[1] Network Cost Savings from Router Bypass in IP over WDM Core Networks, S. Melle et alia, OFC/NFOEC 2008.

[2] Multi-layer Restoration – The Impact on the Optical Layer, M. Gunkel, OFC 2014.

[3] Draft Recommendation ITU-T G.808.3., "Generic protection switching - Shared Mesh Protection"

[4] Draft Recommendation ITU-T G.ODUSMP, "OTN protection switching - Shared Mesh Protection"

[5] Novel Design of G.ODUSMP to Achieve Sub-50 ms Performance with Shared Mesh Protection in Carrier Networks, W. Wauford et alia, ECOC 2013.