MPLS-TP and OTN

IP is the predominant traffic type for packet services in transport networks; any expansion strategy for transport should also consider the nature of IP demands. This paper shares insights on how to plan the expansion of IP and transport networks in a holistic way and how IP can co-exist with lower layers of Transmission (DWDM, OTN, SDH, and L2 (MPLS-TP)). The paper also discusses how such an approach of carrying traffic on the appropriate layer will also help reduce the TCO (Total Cost of Ownership) of the core network.
White Paper

Introduction

The explosive growth in data traffic driven by the emergence of high-speed internet access technologies (4G/5G, FTTx) and the rising usage of powerful smartphones with high-quality multimedia capabilities is challenging telecom infrastructure providers to evolve new strategies for network build-outs and expansion of existing TDM-based networks. Packet traffic puts intense cost and capacity pressure on telecom networks since the nature of this traffic is such that the peak data rates are often an order of magnitude higher than the average data rate. The consequent demands on network scalability and stringent quality of service requirements cannot be economically handled by existing TDM transport equipment that is primarily designed to handle deterministic circuit and voice traffic.

As the amount of traffic increases in today’s networks, there is a need to handle traffic at larger granularities. Besides, as the ratio of TDM to data is constantly shifting in favor of data, there is a need to handle data traffic more efficiently in transport networks. As the transport networks grow larger, they have to be partitioned into several regions or subnets. Since many operators synergize their networks to provide end-to-end services, an easy management of operator-to-operator handoffs is needed. To solve these problems ITU-T laid down the specifications of the Optical Transport Network (OTN). ITU-T provides the general requirements in the G.872 standard and the interface definitions in the G.709 standard.

OTN transport has been in use for a decade, beginning with its use in Submarine Line Terminal Equipment (SLTE) applications in the 1990s. G.709 OTN was originally defined as a point-to-point protocol, designed to provide a protocol-independent wrapper of client data. The objective was to use a single homogenous protocol to wrap various clients, providing 100% transparent transport, something SONET/SDH was incapable of for Gigabit-speed services such as Ethernet and Fibre Channel, and wavelength services.

MPLS-TP is a flexible WAN technology that cost effectively scales services from 10 Mbps to 10Gbps. This is done by avoiding two key limitations of traditional TDM WAN technologies:

- Per service dedicated bandwidth reservations through the network
- Large step functions in the reservation of bandwidth

MPLS-TP also avoids the following limitations in traditional Ethernet and IP/MPLS technologies:

- Over-subscription of routing/switching resources in the WAN
- Complex OAM implementation of connectionless based networks

Packet transport using Layer 2 technologies like MPLS-TP and scalable Layer 1 technologies like OTN have emerged as effective options to reduce operational costs and complexity, contain capital expenditures and deliver premium real-time services with optimized performance on carrier networks.

Reducing CAPEX, OPEX AND network complexity

In the traditional approach, IP traffic uses transmission as a transport pipe. Service aggregation, protection and intelligence functions are relegated completely to the IP layer. In this approach, Routers Grow Fast (GRF) and Lower Layers Grow Fast (GLF) are both built in parallel, thus implying higher capex and opex for the telecom operator. In the GRF-GLF approach, IP layer manages all the network operations such as aggregation, protection or routing. Since all bits have to pass through the transit routers as well, the operator ends up building interfaces for three layers: originating core router, terminating router and all intermediate/transit routers. Scaling up router capacity implies scaling the transmission resources as well.
An alternate and better approach is referred to as "router offload". In its simplest definition, router offload means that any transit traffic at any intermediate core router should be offloaded to a lower layer like SDH/OTN/DWDM. Internet data flows to fixed destinations in a network i.e., towards internet gateways from originating routers. This is different from other packet services such as layer 3 VPN services that follow any-to-any/mesh connectivity and have to be processed at every core router. This point-to-point nature of internet traffic makes it more conducive for transport over lower layers of transmission. This is because the cost to carry data on the lower layers is less than carrying it over higher IP layer. Thus by intelligently designing core networks, routers can use lower layers such as L2 (MPLS-TP), SDH, OTN or DWDM to carry internet traffic more cost-effectively to the internet gateway site. In fact, core routers do not need to grow as fast as lower layers need to, thus bringing significant cost savings for telecom operators. This strategy results in “Grow-Routers-Slowly, Grow-Lower-Layers-Fast” (GRS-GLF) thereby reducing the total cost of ownership (TCO) of the network as a whole.

**SDN/NFV AND CPO (Converged Packet Optical Transport)**

The emergence of cloud services and network function virtualization (NFV) has led to fundamental changes in how telecom networks are built. As network services are increasingly embedded in the cloud, it becomes more cost-effective to adopt a network architecture where multiple layers of service delivery aggregation routers are connected to each other by simple transport equipment subsumed within converged packet optical transport equipment (CPO) which then backhauls it to a central router. CPOs will do basic packet services as Multicast, Ethernet and MPLS (MPLS-TP).

Modern CPO equipment is characterized by three types of technology sub-layers, namely:
- Connectionless packet switching service provided by IP or Ethernet,
- Connection-oriented packet-switched service provided by MPLS or MPLS-TP
- Connection-oriented circuit-switched service provided by OTN.

For connectionless packet switching, ethernet is typically preferred since unlike IP, ethernet is agnostic to the type of service being transported. Moreover, not all fast growing services such as data center traffic are IP based, hence limiting service variety.

---

Comparing MPLS-TP with IP/MPLS

MPLS-TP is an enhanced profile of IP/MPLS designed to meet transport network operational requirements. It borrows critical elements from IP/MPLS such as its forwarding mechanisms, while including additional functionality such as performance monitoring, OAM, Tandem Connection Monitoring (TCM) and protection switching. MPLS-TP feature set is implemented as per RFC 5654. These are divided into general, layering, control plane, and protection and recovery.

Three key characteristics of MPLS-TP:

- Reduced dependency on "routable" IP protocols thus lowering vulnerability to network layer cyber-attacks
- Provides superior OAM capabilities with pro-active and reactive fault management and performance monitoring, similar to those provided by SONET/SDH and OTN.
- Uses LSPs (Label Switched Paths) and PW (Pseudo-wires) to deliver connection-oriented services. Network provisioning via centralised Network Management System (NMS) is possible without using a control plane.

Drawbacks of IP/MPLS technology

**Non-Routability in Data Plane:** The need to build a flexible IP communication network has to consider some of its security implications. When IP protocols were being deployed in smart grids, tough security regulations were developed to protect the communication network. For example, the North American Electric Reliability Corporation (NERC) has outlined cyber security requirements in eight Critical Infrastructure Protection (CIP) standards that will

---


5 Hulseboch, T., Belmont, D., Manske, Mike.: Smart Grid Network. MPLS Design Approach, Technical Paper by West Monroe Partners (2010)
MPLS VPWS and VPLS services are seen as potential alternatives that meet "non-routable" standards compared to standard IP. However, MPLS continues to make extensive use of IP based protocols such as BGP (I-BGP, E-BGP), OSPF, MOSPF, IGMP, MMRP, RSVP-TE and LDP. By eliminating the use of these protocols, MPLS-TP enables a true "non-routable" service.

Operational costs and simplicity: Packet transport approaches, unlike vanilla IP or IP/MPLS Routers, adopt a Connection-oriented Ethernet (CoE) paradigm that provides for traffic-engineered, "pinned-down" packet flows. The operator has full knowledge and visibility into the service topology, network resource availability and an ability to control provisioning from a central management system. Although the initial focus of this standard was on creating "point-to-point" circuits, it has been further extended to support traffic engineered multicast (E-Tree) and broadcast (E-LAN) flows. The provisioned model of operations enables the use of operationally proven concepts such as protection switching of services. Backup connections for critical service flows may be configured from the management system but the actual protection switching is delegated to data plane mechanisms.

Also operators are concerned with rising power bills; hence the use of energy efficient network architectures and equipment is important. IP routers have been widely recognized as the highest power-per-bit consumers in the network and highly inefficient at low utilization levels. Just turning on an IP Router with no load can require 60-80% of peak power. A recent energy modeling study by Ovum suggests that an "overall network power reduction of the order of 50% is possible through straightforward but challenging changes such as the elimination of IP transit traffic and how IP traffic is optimally routed and switched at various layers". Thus the use of enhanced Packet Transport equipment serves as a concrete initiative in "greening" of the access network and significantly reduces the long term OPEX of the network.

Enhanced OAM Features: Ethernet uses bi-directional flow symmetry, i.e., the forward and reverse paths of all flows, whether unicast or multicast in nature, follow the same path through the network. This allows the creation of advanced OAM mechanisms on Carrier Ethernet switches that provide network operators a powerful set of tools to ensure carrier-grade predictability, reliability and measurability for SLA based services. These OAM standards are able to manage Layer 2 based MEF services (E-LINE, E-LAN, E-TREE) without having to overlay an expensive IP infrastructure such as MPLS. Carrier Ethernet OAM includes both fault management and performance management functions and is defined in standards IEEE 802.1ag and ITU Y.1731. Fault management permits detection and localization of network defects using CCM (Continuity Check Message), LB (Loopback) and LT (Link Trace) messages. Carrier Ethernet OAM borrows the notion of TCM-like (Tandem Connection Monitoring) maintenance domains as defined in the SDH, SONET and OTN standards. Up to eight hierarchical levels are possible and OAM checks may be performed independently on each of these layers with the lower layers transparently passing the OAM messages transmitted by the higher layers. Performance monitoring allows measurement of different service parameters such as loss, throughput, latency and jitter using message frames such as LM (loss measurement) and DMM (Delay Measurement Message). MPLS-TP products leverage carrier Ethernet standards to deliver a comprehensive OAM solution that augments the basic fault management suite provided by MPLS Bi-directional Forwarding Detection (BFD), VCCV and LSP Ping/Traceroute with extensive performance monitoring (delay, frame loss, delay variation measurements) features defined in ITU Y.1731. By supporting MPLS-based OAM messages, state-of-the-art MPLS-TP implementations can inter-operate with popular IP/MPLS routers also. For end-to-end OAM of each pseudo-wire, one of the three options – Y.1731, VCCV and BFD can be chosen for network performance monitoring and fault management. MPLS-TP platforms also make use of AIS (Alarm Indication Signal), RDI (Remote Defect Indication) and CSF (Client Signal Failure) signals to improve alarm correlation and single-ended fault management at the Tunnel and Pseudo-wire levels.

Centralized Policy-based Operations: MPLS-TP provides support for static traffic-engineered "pinned-down" service from a centralized network management system (NMS). This is perfectly compatible with traditional transport-style operations that assure enhanced reliability, predictability and determinism for utility networks. Both service and protection paths can be pre-configured with the operator having full knowledge and visibility into service topology, network resource availability and provisioning from a hierarchical management system. Another benefit of this model, when compared to control plane based approaches like MPLS, is the ability to isolate and manage sub-networks as virtually independent regions similar to utility operational models. "Troubled" regions can be
isolated by the NMS with critical traffic re-routed away from these parts of the network to maximize network uptime and productivity, while minimizing threat impact. Also, MPLS-TP makes it easier to evolve networks towards an SDN WAN as it employs a centralized control plane. Moreover, unlike IP/MPLS, which has a tight integration of data and control planes, MPLS-TP provides a clean separation between the two thus making it easier to introduce SDN in any layer in an independent manner. This permits technology-agnostic control of connectivity across multiple network layers (L1, L2 or L3) and enables superior cost and performance optimization across circuit and packet layers\(^6\). Migration of the current multi-layer NMS into SDN based architecture allows for the introduction of orchestration and new applications. The network elements that provide MPLS-TP functionality have appropriate interfaces to allow the introduction of 3rd party controllers without architecture change.

Migration of the current multi-layer NMS into SDN based architecture allows for the introduction of orchestration and new applications. The network elements that provide MPLS-TP functionality have appropriate interfaces to allow the introduction of 3rd party controllers without architecture change.

**Figure 2: Evolving the Layered Architecture with SDN**

**Transport-style Network Protection:** Existing transport networks based on SDH/SONET are known for a variety of fast protection switching mechanisms like 1+1 Linear MSP, 1:1, SNCP ring protection that delivers network recovery from failures in less than 50 msecs. However, the concept of "rapid" protection switching is mostly alien to data networks. IP/MPLS networks support FRR (Fast ReRoute) mechanisms that work for link and node failures but not at the LSP or Pseudo-wire (PW) granularity. Also, FRR is very complex compared to SDH protection schemes and requires significant expertise and knowledge of IP and MPLS protocols. MPLS-TP has bridged this gap considerably by defining 1+1 and 1:1 LSP and PW protection in uni- and bi-directional modes. Also, PWs provisioned in different operator or administrative domains can be stitched together into a single multi-segment PW (MS-PW) with end-to-end protection switching. New packet ring protection standards based on ERPS (ITU G.8032) are now capable of sub-50ms failure recovery in single ring and multi-ring topologies without explicit bandwidth reservation on the protect path.

Failure statistics suggest that transport networks average approximately 4x the fiber cuts witnessed in core networks within a year. Moreover, most service providers seem to know a-priori which fiber segments are more prone to failures. Also, transport networks often provide anywhere between two and five alternate paths for several traffic demands that can be utilized by the network operators to assure higher availability for SLA driven services. New MPLS-TP implementations incorporate novel multi-segment "stitching" method by combining MS-PW stitched protection with 1:1 linear LSP and PW protection for supporting arbitrary meshed topologies. This advanced feature is especially useful for backhaul networks with dispersed clusters and possibilities of more than two alternate fiber paths for services within/between clusters.

**Efficient Multicast for Video-based Services:** Emerging services such as digital video simulcasts, multi-site conferencing and high-definition tele-presence require support for point-to-multipoint mode of transmission on a transport network that efficiently supports multicast. Though the ability to offer multicast and "any-to-any" services has been a challenge for most technologies including IP/MPLS, it is a key strength of Ethernet. Current implementations based on IP/MPLS primarily rely on fully meshed topologies with point-to-point LSPs and Pseudo-wires to realize multipoint and multicast services. MPLS technology for multicast delivery is being actively discussed in IETF but there is limited convergence in the vendor and service provider community. IP multicast protocols such as PIM are also an option for distributing multicast traffic but would be far more complex to implement and are more prone to jitter. Recent studies have shown that IP Routers have on an average, a wider

spread of packet delays when compared to Carrier Ethernet switches even for low bandwidth utilization in the network. However, if the bandwidth utilization in the backbone increases above 5%, the packet delay spread for an IP Routed network grows significantly, thus adversely impacting service performance. Hence, advanced Layer 2 MPLS-TP platforms that are typically supplemented by Carrier Ethernet capabilities have a significant advantage over traditional IP/MPLS Routers in the multicast domain.

**“IP Aware” vs “IP Routed” Transport:** A common misconception is the necessity of using IP Routers to transport IP packets. Although networks may carry IP traffic in the form of various services such as VoIP, IPTV, Web video from a transport perspective, the only operator requirement is the ability to faithfully preserve a range of L2, L3 and L4 service markings as appropriate so that end-to-end QoS consistency is retained. Support for “IP Awareness” to perform differentiated classification and aggregation of up to eight classes is a valuable feature available in new MPLS-TP platforms. It eliminates the need to aggregate service types with different performance criteria to a single CoS class due to equipment limitations. For example, real-time and synchronization traffic classes can be treated differently, since the latter service typically has more stringent performance requirements. Another useful concept is the connection-oriented mapping of services. Connection-oriented mapping mandates the use of a range of optional traffic handles on a packet in addition to the limited packet header parameters listed below. A packet belonging to such a service has to be suitably “shaped” by a hierarchical scheduler and QoS engine to comply to pre-defined service bounds on admission - such as CIR/EIR (Committed/Excess Information Rate), CBS/EBS (Committed/Excess Burst Rate), tolerated Jitter/Latency/Packet Loss. Most of these sophisticated QoS features are currently not supported on traditional Access Routers but are available on MPLS-TP based Packet Transport equipment.

**Benefits of MPLS-TP with OTN**

In today’s next generation transport networks the optical layer, typically referred to as “Layer 0” or the “DWDM layer”, is the most cost effective layer for maximizing fiber network utilization and managing point to point traffic that is 10Gbps and higher. Layer 1 switching is the most cost-effective layer for transporting point-to-point traffic between 2Mbps and 10Gbps, while also offering robust traffic protection mechanisms and simple provisioning methodologies. While SDH grooming provides cost savings for point to point traffic below 1Gbps, OTN extends cost savings for traffic between 1Gbps and 10Gbps.

In addition, OTN provides several other advantages as listed below:

**Tandem Connection Monitoring (TCM):** OTN supports six levels of TCM. This is a significant improvement over the one level provided by SDH. With six levels of TCM, the network can be partitioned into six level of hierarchy from a management perspective. Some of these levels might be defined by a single operator to manage his large network, or for managing multi-vendor subnets within his own networks or for hand-offs between multiple operators.

**Efficient Sub-Lambda Grooming:** DWDM layer is capable of handling traffic at a wavelength layer. This provides either a 10Gbps or a 40Gbps granularity in the DWDM network depending on the interfaces used. Since DWDM by itself does not offer any way to consolidate partially filled wavelengths, this leads to a lot of bandwidth wastage. OTN solves this problem by providing sub lambda grooming through an OTN cross-connect.

**Reduced Build and Operational Costs:** On a cost per Gbps of switching capacity, routers tend to be much more expensive than OTN fabrics. By substituting OTN for a significant part of the switching capacity (the transit traffic), we are reducing the overall costs significantly. Also, OTN fabrics consume less power compared to routing fabrics. They also consume less power than SDH fabrics since they groom traffic at 1Gbps granularity instead of 2Mbps.

**Transport-style Reliability:** OTN’s reliability functions are modeled on the lines of SONET/SDH technology with the same level of versatility. OTN incorporates the following reliability features:

- Comprehensive OAMP includes messages related to Loss of Signal/Frame (LOS/F), alignment errors (TIM), frame errors (BIP-8, BEI), alarm indications (AIS, BDI) that can enable fault diagnostics at the OTN layer
• Support for a range of protection schemes is available: OCh 1+1, OCh-SPRing, ODU-1+1 linear, ODU-SNC/I, ODUSNC/N, ODU-SNC/S, and ODU-SPRing are analogous to popular linear and ring protection protocols in SDH/SONET
• Support for mesh-based automatic restoration with a GMPLS control-plane that offers additional flexibility by reducing the need for pre-allocated protection bandwidth on a backup path and for manual intervention by a network operator

**Forward Error Correction (FEC):** OTN supports FEC based on Reed-Solomon (255/239) with 16-byte parity. While SDH too has defined an in-band FEC, it is fairly rudimentary. OTN FEC, on the other hand, can correct eight bytes of error per sub-row or detect up to 16 byte errors (without any correction) resulting in a 6.2 dB improvement in SNR. The availability of an enhanced FEC function can result in tangible cost savings in the access by reducing the need for external amplifier elements (often unmanaged) or more expensive long-reach interfaces.

**Transparent service transport:** Unlike SDH, OTN can achieve truly transparent service delivery because it uses asynchronous mapping. For example, if OTN is being used to transport an SDH client, it will neither modify any of the SDH overhead bits nor will the timing information be touched. This is especially important in the context of defence networks and other security-sensitive transport requirements.

**Packet switching bypass:** There are many emerging time-critical applications (e.g., video on demand, video conferencing) that require simple and effective transport of client traffic with the lowest possible latency. In such cases, even if the client happens to be a packet interface, it is particularly advantageous to use OTN as the transport protocol in the access network. This is because OTN does not have a store-and-forward architecture like in the case of packet switching thus avoiding unnecessary delays, jitter and service impairments. This approach also frees up capacity in packet switching fabrics (e.g., Ethernet switch) for applications that actually require intermediate processing thus lowering the overall cost of the access network.

**ODUFlex for Bandwidth Reclamation:** ODUFlex is particularly suitable to accommodate a range of new and custom client signal rates that are emerging in Access networks. Examples are EPON, GPON, CPRI that are fast gaining popularity as the access is increasingly fiberized to deliver gigabits of bandwidth to user locations, whether offices or residences. With ODUFlex, the container can be the exact size of its client, leaving the remaining space for other client signals. ODUFlex can carry both Constant Bit Rate (CBR) and Variable Bit Rate (VBR) packet-based clients. CBR clients are mapped using Bit-Synchronous Mapping Procedure (BMP) and packet-based client signals are accommodated by using Frame-mapped Generic Framing Procedure (GFP-F). ODUFlex is then mapped into a number of time slots in a High-Order ODU (HO-ODU) by using Generic Mapping Procedure (GMP). The clear advantage of ODUFlex is that unused bandwidth on existing fiber can be reclaimed and the operator can avoid burning an additional fiber pair to meet new traffic requirements. This is critical in access networks with fiber exhaust issues.

**Beyond-100G OTN Transport:** The 5G era demands support for new data oriented client signals such as 25GE, 40GE, 100GE, 200GE, 400GE and FlexE. OTN3.0 which supports nx100G interfaces, in addition to a flexible choice of FEC has poised to be the technology of choice to provide efficient transport capabilities for explosive bandwidth requirements.

**Tejas Networks POTP platform with integrated OTN**

Tejas Networks’ TJ1600 platform increases bandwidth efficiency by enabling deployment of fewer wavelengths leading to significantly lower network costs in large deployments. By converging multiple technologies such as viz., SDH, DWDM, PTN and OTN in a single shelf, TJ1600 reduces the number of nodes required to be co-located at each site. Thus the overall capital expenditure is lowered. TJ1600 is coupled with a multilayered management system that simplifies end-to-end service provisioning and network configuration to reduce the operational expenditure.
**Benefits of a Tejas’ Converged WDM+OTN DXC platform**

**Savings in Client Interfaces:** In a pure-play approach, the OTN DXC is an independent switching shelf that is co-located with a stand-alone optical WDM platform. While the service add-drop requirements are typically driven by customer needs, a pure-play approach calls for additional client interfaces to realize back-to-back connectivity between the two shelves. Depending on the quantum of traffic handled and grooming desired at a given node location, the savings in client interfaces could be significant with a converged approach. Moreover, in order to meet intermediate regenerator requirements on a network, a “pure-play” approach will require additional electrical systems with associated power and real-estate to do electrical processing of wavelengths. In the case of converged platforms, regeneration can be done through the electrical backplane hence no separate platform is required.

**Greater Bandwidth Efficiency:** A pure-play approach will always result in the use of a greater number of high-speed wavelengths in the network. In converged platforms, presence of integrated OTN switching enables superior fill ratios of existing 10G/40G/100G+ wavelengths resulting in minimizing the number of wavelengths required and eliminating stranded bandwidth. Bandwidth efficiency improves with the capacity of the wavelengths and traffic volume in the network. Optical layer restoration being at a coarser wavelength level requires longer backup paths and more regenerators, thus potentially resulting in higher cost. When converged platforms are further combined with advanced CDC-ROADM capabilities up to 18% savings in OTN line cards can be achieved.

**Bandwidth Reclamation:** ODUFlex is particularly suitable to accommodate a range of custom client signals with varying rates such as EPON, GPON, CPRI. The clear advantage of ODUflex is that unused bandwidth on a wavelength can be reclaimed and the operator can avoid burning an additional wavelength to meet new traffic requirements. State-of-the-art converged platforms with integrated OTN DXC such as TJ1600 can deliver this additional benefit not possible in WDM-only platforms.

**Savings in Opex:** Historically, WDM transport and electrical switching functions were separated into two different platforms because of density mismatch between client optics and long-haul WDM optics. However, in recent years, this mismatch between the two has been bridged and it is now possible to efficiently integrate WDM transport and DXC capabilities (SDH, Ethernet or OTN) without incurring a capacity penalty. Converged platforms eliminate duplication of hardware, additional cost of short-reach optical inter-connections between multiple co-located systems, besides yielding considerable savings in space and power consumption.

**CONCLUSION**

Telecom infrastructure providers are actively exploring various alternatives to realize a cost-effective yet scalable transport network using packet based technologies. Layer 3 technologies such as IP routing or MPLS traffic engineering that are widely employed in core networks could prove to be very expensive for traffic backhaul when extended to the access regions of the network. Routed networks are more complex to control, less energy-efficient and have poor and/or unpredictable latency and jitter performance for advanced real-time services. IP/MPLS was not designed to implement large metro-scale networks with thousands of network elements. Advanced MPLS-TP implementations when combined with OTN are the best choice to meet the stringent performance demands and compelling economics requirements of a next-generation, packet-optimized optical transport infrastructure.