Applicability of Abstraction and Control of Traffic Engineered Networks (ACTN) to Packet Optical Integration (POI)

draft-ietf-teas-actn-poi-applicability-05

Abstract

This document considers the applicability of Abstraction and Control of TE Networks (ACTN) architecture to Packet Optical Integration (POI)in the context of IP/MPLS and optical internetworking. It identifies the YANG data models being defined by the IETF to support this deployment architecture and specific scenarios relevant for Service Providers.

Existing IETF protocols and data models are identified for each multi-layer (packet over optical) scenario with a specific focus on the MPI (Multi-Domain Service Coordinator to Provisioning Network Controllers Interface)in the ACTN architecture.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on April 9, 2021.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

[1. Introduction 3](#_Toc96368394)

[1.1. Terminology 5](#_Toc96368395)

[2. Reference network architecture 5](#_Toc96368396)

[2.1. Multi‑domain Service Coordinator (MDSC) functions 8](#_Toc96368397)

[2.1.1. Multi-domain L2/L3 VPN network services 9](#_Toc96368398)

[2.1.2. Multi-domain and multi-layer path computation 13](#_Toc96368399)

[2.2. IP/MPLS Domain Controller and NE Functions 16](#_Toc96368402)

[2.3. Optical Domain Controller and NE Functions 17](#_Toc96368403)

[3. Interface protocols and YANG data models for the MPIs 18](#_Toc96368404)

[3.1. RESTCONF protocol at the MPIs 18](#_Toc96368405)

[3.2. YANG data models at the MPIs 18](#_Toc96368406)

[3.2.1. Common YANG data models at the MPIs 18](#_Toc96368407)

[3.2.2. YANG models at the Optical MPIs 19](#_Toc96368408)

[3.2.3. YANG data models at the Packet MPIs 20](#_Toc96368409)

[3.3. PCEP 21](#_Toc96368410)

[4. Multi-layer and multi-domain services scenarios 22](#_Toc96368411)

[4.1. Scenario 1: inventory, service and network topology discovery 22](#_Toc96368412)

[4.1.1. Optical topology discovery 23](#_Toc96368413)

[4.1.2. Optical path discovery 25](#_Toc96368414)

[4.1.3. Packet topology discovery 25](#_Toc96368415)

[4.1.4. SR-TE paths discovery 26](#_Toc96368416)

[4.1.5. Inter-domain links discovery 27](#_Toc96368417)

[4.1.6. Multi-layer IP links discovery 29](#_Toc96368418)

[4.1.7. LAG discovery 30](#_Toc96368419)

[4.1.8. L2/L3 VPN network services discovery 30](#_Toc96368420)

[4.1.9. Inventory discovery 30](#_Toc96368421)

[4.2. Establishment of L2/L3 VPN network services with TE requirements 30](#_Toc96368422)

[4.2.1. Optical Path Computation 33](#_Toc96368423)

[4.2.2. Multi-layer IP link Setup and Update 33](#_Toc96368424)

[4.2.3. SR-TE Path Setup and Update 34](#_Toc96368425)

[5. Conclusions 35](#_Toc96368426)

[6. Security Considerations 36](#_Toc96368427)

[7. Operational Considerations 36](#_Toc96368428)

[8. IANA Considerations 36](#_Toc96368429)

[9. References 36](#_Toc96368430)

[9.1. Normative References 36](#_Toc96368431)

[9.2. Informative References 38](#_Toc96368432)

[Appendix A. OSS/Orchestration Layer 41](#_Toc96368433)

[A.1. MDSC NBI 41](#_Toc96368434)

[Appendix B. Multi-layer and multi-domain resiliency 44](#_Toc96368435)

[B.1. Maintenance Window 44](#_Toc96368436)

[B.2. Router port failure 44](#_Toc96368437)

[Acknowledgments 45](#_Toc96368438)

[Contributors 45](#_Toc96368439)

[Authors’ Addresses 47](#_Toc96368440)

# Introduction

The complete automation of the management and control of Service Providers transport networks (IP/MPLS, optical, and microwave transport networks) is vital for meeting emerging demand for high-bandwidth use cases, including 5G and fiber connectivity services. The Abstraction and Control of TE Networks (ACTN) architecture and interfaces facilitate the automation and operation of complex optical and IP/MPLS networks through standard interfaces and data models. This allows a wide range of of network services that can be requested by the upper layers fulfilling almost any kind of service level requirements from a network perspective (e.g. physical diversity, latency, bandwidth, topology, etc.)

Packet Optical Integration (POI) is an advanced use case of traffic engineering. In wide-area networks, a packet network based on the Internet Protocol (IP), and often Multiprotocol Label Switching (MPLS) or Segment Routing (SR), is typically realized on top of an optical transport network that uses Dense Wavelength Division Multiplexing (DWDM)(and optionally an Optical Transport Network (OTN)layer).

In many existing network deployments, the packet and the optical networks are engineered and operated independently. As a result, there are technical differences between the technologies (e.g., routers compared to optical switches) and the corresponding network engineering and planning methods (e.g., inter-domain peering optimization in IP, versus dealing with physical impairments in DWDM, or very different time scales). In addition, customers needs can be different between a packet and an optical network, and it is not uncommon to use different vendors in both domains. The operation of these complex packet and optical networks is often siloed, as these technology domains require specific skills sets.

The packet/optical network deployment and operation separation are inefficient for many reasons. Both capital expenditure (CAPEX) and operational expenditure (OPEX) could be significantly reduced by integrating the packet and the optical networks. Multi-layer online topology insight can speed up troubleshooting (e.g., alarm correlation) and network operation (e.g., coordination of maintenance events), multi-layer offline topology inventory can improve service quality (e.g., detection of diversity constraint violations) and multi-layer traffic engineering can use the available network capacity more efficiently (e.g., coordination of restoration). In addition, provisioning workflows can be simplified or automated as needed across layers (e.g., to achieve bandwidth-on-demand or to perform maintenance events).

ACTN framework enables this complete multi-layer and multi-vendor integration of packet and optical networks through Multi-Domain Service Coordinator (MDSC) and packet and optical Provisioning Network Controllers (PNCs).

In this document, critical scenarios for POI are described from the packet service layer perspective and identified the required coordination between packet and optical layers to improve POI deployment and operation. Precise definitions of scenarios can help with achieving a common understanding across different disciplines. The focus of the scenarios are multi‑domain packet networks operated as a client of optical networks.

This document analyses the case where the packet networks support multi‑domain SR‑TE paths and the optical networks could be either a DWDM network or an OTN network (without DWDM layer) or multi‑layer OTN/DWDM network. DWDM networks could be either fixed-grid or flexible-grid.

For each multi-layer scenario, the document analyzes how to use the interfaces and data models of the ACTN architecture.

Understanding the level of standardization and the possible gaps will help assess the feasibility of integration between packet and optical DWDM domains (and optionally OTN layer) in an end-to-end multi-vendor service provisioning perspective.

## Terminology

This document uses the ACTN terminology defined in [RFC8453]

In addition this document uses the following terminology.

Customer service: the end-to-end service from CE to CE

Network service: the PE to PE configuration including both the network service layer (VRFs, RT import/export policies configuration) and the network transport layer (e.g. RSVP-TE LSPs). This includes the configuration (on the PE side) of the interface towards the CE (e.g. VLAN, IP adress, routing protocol etc.)

# Reference network architecture

This document analyses several deployment scenarios for Packet and Optical Integration (POI) in which ACTN hierarchy is deployed to control a multi‑layer and multi-domain network, with two optical domains and two packet domains, as shown in Figure 1:

+----------+

| MDSC |

+-----+----+

|

+-----------+-----+------+-----------+

| | | |

+----+----+ +----+----+ +----+----+ +----+----+

| P-PNC 1 | | O-PNC 1 | | O-PNC 2 | | P-PNC 2 |

+----+----+ +----+----+ +----+----+ +----+----+

| | | |

| \ / |

+-------------------+ \ / +-------------------+

CE1 / PE1 BR1 \ | / / BR2 PE2 \ CE2

o--/---o o---\-|-------|--/---o o---\--o

\ : : / | | \ : : /

\ : PKT domain 1 : / | | \ : PKT domain 2 : /

+-:---------------:-+ | | +-:---------------:--+

: : | | : :

: : | | : :

+-:---------------:------+ +-------:---------------:--+

/ : : \ / : : \

/ o...............o \ / o...............o \

\ optical domain 1 / \ optical domain 2 /

\ / \ /

+------------------------+ +--------------------------+

1. – Reference Network

The ACTN architecture, defined in [RFC8453], is used to control this multi-layer and multi-domain network where each Packet PNC (P-PNC) is responsible for controlling its packet domain and where each Optical PNC (O-PNC) in the above topology is responsible for controlling its optical domain. The packet domains controlled by the P-PNCs can be Autonomous Systems (ASes), defined in [RFC1930], or IGP areas, within the same operator network.

The routers between the packet domains can be either AS Boundary Routers (ASBR) or Area Border Router (ABR): in this document, the generic term Border Router (BR) is used to represent either an ASBR or an ABR.

The MDSC is responsible for coordinating the whole multi-domain multi-layer (packet and optical) network. A specific standard interface (MPI) permits MDSC to interact with the different Provisioning Network Controller (O/P-PNCs).

The MPI interface presents an abstracted topology to MDSC hiding technology-specific aspects of the network and hiding topology details depending on the policy chosen regarding the level of abstraction supported. The level of abstraction can be obtained based on P-PNC and O-PNC configuration parameters (e.g., provide the potential connectivity between any PE and any BR in an SR-TE network).

In the reference network of Figure 1, it is assumed that:

* The domain boundaries between the packet and optical domains are congruent. In other words, one optical domain supports connectivity between routers in one and only one packet domain;
* There are no inter-domain physical links between optical domains. Inter-domain physical links exist only:
  + between packet domains (i.e., between BRs belonging to different packet domains): these links are called inter-domain Ethernet or IP links within this document;
  + between packet and optical domains (i.e., between routers and optical NEs): these links are called cross-layer links within this document;
  + between customer sites and the packet network (i.e., between CE devices and PE routers): these links are called access links within this document.
* All the physical interfaces at inter-domain links are Ethernet physical interfaces.

Although the new optical technologies (e.g., QSFP-DD ZR 400G) allows providing DWDM pluggable interfaces on the routers, the deployment of those pluggable optics is not yet widely adopted by the operators. The reason is that most operators are not yet ready to manage packet and optical networks in a single unified domain. The analysis of the unified use case is outside the scope of this draft.

This document analyses scenarios where all the multi-layer IP links, supported by the optical network, are intra-domain (intra-AS/intra-area), such as PE-BR, PE-P, BR-P, P-P IP links. Therefore the inter-domain IP links are always single-layer links supported by Ethernet physical links.

The analysis of scenarios with multi-layer inter-domain IP links is outside the scope of this document.

Therefore, if inter-domain links between the optical domains exist, they would be used to support multi-domain optical services, which are outside the scope of this document.

The optical network elements (NEs) within the optical domains can be ROADMs or OTN switches, with or without an integrated ROADM function.

## Multi‑domain Service Coordinator (MDSC) functions

The MDSC in Figure 1 is responsible for multi-domain and multi-layer coordination across multiple packet and optical domains, as well as to provide multi-layer/multi-domain L2/L3 VPN network services requested by an OSS/Orchestration layer.

From an implementation perspective, the functions associated with MDSC and described in [RFC8453] may be grouped in different ways.

1. Both the service- and network-related functions are collapsed into a single, monolithic implementation, dealing with the end customer service requests received from the CMI (Customer MDSC Interface) and adapting the relevant network models. An example is represented in Figure 2 of [RFC8453].
2. An implementation can choose to split the service-related and the network-related functions into different functional entities, as described in [RFC8309] and in section 4.2 of [RFC8453]. In this case, MDSC is decomposed into a top-level Service Orchestrator, interfacing the customer via the CMI, and into a Network Orchestrator interfacing at the southbound with the PNCs. The interface between the Service Orchestrator and the Network Orchestrator is not specified in [RFC8453].
3. Another implementation can choose to split the MDSC functions between an "higher-level MDSC" (MDSC-H) responsible for packet and optical multi-layer coordination, interfacing with one Optical "lower-level MDSC" (MDSC-L), providing multi-domain coordination between the O-PNCs and one Packet MDSC-L, providing multi‑domain coordination between the P-PNCs (see for example Figure 9 of [RFC8453]).
4. Another implementation can also choose to combine the MDSC and the P-PNC functions together.

In the current service provider’s network deployments, at the North Bound of the MDSC, instead of a CNC, typically there is an OSS/Orchestration layer. In this case, the MDSC would implement only the Network Orchestration functions, as in [RFC8309] and described in point 2 above. Therefore, the MDSC is dealing with the network services requests received from the OSS/Orchestration layer.

The functionality of the OSS/Orchestration layer and the interface toward the MDSC are usually operator-specific and outside the scope of this draft. Therefore, this document assumes that the OSS/Orchestrator requests the MDSC to set up L2/L3 VPN network services through mechanisms that are outside the scope of this document.

There are two prominent workflow cases when the MDSC multi-layer coordination is initiated:

* Initiated by a request from the OSS/Orchestration layer to setup L2/L3 VPN network services that requires multi-layer/multi-domain coordination;
* Initiated by the MDSC itself to perform multi-layer/multi-domain optimizations and/or maintenance activities (e.g. rerouting LSPs with their associated services when putting a resource, like a fibre, in maintenance mode during a maintenance window). Unlike service fulfillment, these workflows are not related to a network service provisioning request being received from the OSS/Orchestration layer.

The latter workflow cases are outside the scope of this document.

This document analyses the use cases where multi-layer coordination is triggered by a network service request received from the OSS/Orchestration layer.

### Multi-domain L2/L3 VPN network services

Figure 2 provides an example of an hub & spoke multi-domain L2/L3 VPN with three PEs where the hub PE (PE13) and one spoke PE (PE14) are within the same packet domain and the other spoke PE (PE23) is within a different packet domain.

------

| CE13 |\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

------ ) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

( | ) ( )

( | PE13 P15 BR11 ) ( BR21 P24 )

( \_\_\_\_ \_\_\_ \_\_\_\_ ) ( \_\_\_\_ \_\_\_ )

( / H \ \_ \_ \_ / \ \_ \_ / \ \_)\_ \_ \_(\_ / \ \_ \_ \_ / \ )

( \\_\_\_\_/... \\_\_\_/ \\_\_\_\_/ ) ( \\_\_\_\_/ \\_\_\_/ )

( :..... ) ( | )

( \_\_\_\_ :\_\_ \_\_\_\_ ) ( \_\_\_\_ \_|\_\_ )

( / S \...../ \.\_.\_./ \\_\_\_\_\_\_\_\_\_\_/ \.\_.\_.\_.\_./ S \ )

( \\_\_\_\_/ \\_\_\_/ \\_\_\_\_/ ) ( \\_\_\_\_/ \\_\_\_\_/ )

( | ) ( | )

( | PE14 P16 BR12 ) ( BR22 PE23 | )

( | ) ( | )

------ ) ( ------

| CE14 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_) (\_\_\_\_\_\_\_\_\_\_\_\_\_| CE23 |

------ ------

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

( ) ( )

( \_\_\_\_ \_\_\_\_ ) ( \_\_\_\_ )

( /NE11\ \_\_ \_ \_ \_ \_ /NE12\ ) ( /NE21\ \_ \_ )

( \\_\_\_\_/.. \\_\_\_\_/ ) ( \\_\_\_\_/ \ )

( | :..... ...: \ ) ( / \ )

( \_|\_\_ :\_\_: \\_\_\_\_ ) ( \_\_\_/ \_\_\\_ )

( /NE13\\_ \_ /NE14\ \_ \_ \_ /NE15\ ) ( /NE22\ \_ \_ \_ /NE23\ )

( \\_\_\_\_/ \\_\_\_\_/ \\_\_\_\_/ ) ( \\_\_\_\_/ \\_\_\_\_/ )

( ) ( )

(\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_) (\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

optical domain 1 optical domain 2

H / S = Hub VRF / Spoke VRF

\_\_\_\_ = Inter-domain interconnections

..... = SR policy Path 1

\_ \_ \_ = SR policy Path 2

1. - Multi-domain L3VPN example

There are many options to implement multi‑domain L2/L3 VPNs, including:

1. BGP-LU (seamless MPLS)
2. Inter-domain RSVP-TE
3. Inter-domain SR-TE

This document provides an analysis of the inter‑domain SR-TE option. The analysis of other options is outside the scope of this draft.

It is also assumed that:

* each packet domain in Figure 2 is implementing SR-TE and the stitching between two domains is done using end-to-end/multi-domain SR-TE;
* the bandwidth of each intra‑domain SR-TE path is managed by its respective P-PNC;
* binding SID is used for the end-to-end SR-TE path stitching;
* each packet domain in Figure 2 is using TI-LFA, with SRLG awareness, for local protection within each domain.

In this scenario, one of the key MDSC functions is to identify the underlay multi‑domain/multi-layer SR-TE paths used to carry the L2/L3 VPN traffic between PEs belonging to different packet domains and to relay this information to the P-PNCs, to ensure that the PEs’ forwarding tables (e.g., VRF) are properly configured to steer the L2/L3 VPN traffic over the intended multi‑domain/multi-layer SR-TE paths.

For each underlay SR-TE path required to support the L2/L3 VPN network service, it is possible that:

1. A SR-TE path that meets the TE requirements already exist in the network.
2. An existing SR-TE path could be modified (e.g., through bandwidth increase) to meet the TE requirements:
   1. The SR-TE path characteristics can be modified only in the packet layer.
   2. One or more new underlay optical tunnels need to be setup to support the requested changes of the overlay SR-TE paths (multi-layer coordination is required).
3. A new SR-TE path needs to be setup:
   1. The new SR-TE path reuses existing underlay optical tunnels;
   2. One or more new underlay optical tunnels need to be setup to support the setup of the new SR-TE path (multi-layer coordination is required).

In general the binding requirements for a service (e.g L2/L3 VPN) to the related TE tunnels supporting it, can be summarized within three cases:

1. The customer is asking for VPN isolation dynamically creating and binding tunnels to the service such that they are not shared by others services (e.g. VPN).

The level of isolation can be different:

1. Hard isolation with deterministic latency that means L2/L3 VPN requiring a set of dedicated TE Tunnels (neither sharing with other services nor competing for bandwidth with other tunnels) providing deterministic latency performances
2. Hard isolation but without deterministic characteristics
3. Soft isolation that means the tunnels associated with L2/L3 VPN are dedicated to that but can compete for bandwidth with other tunnels.
4. The customer does not ask isolation,and could request a VPN service where associated tunnels can be shared across multiple VPNs.

TE binding requirement types [TSM] are:

1. Hard Isolation with deterministic latency: The L2/L3VPN network service requires a set of dedicated TE Tunnels providing deterministic latency performances and that cannot be not shared with other services, nor compete for bandwidth with other Tunnels.
2. Hard Isolation: This is similar to the above case without deterministic latency requirements.
3. Soft Isolation: The L2/L3VPN network service requires a set of dedicated MPLS-TE tunnels that cannot be shared with other services, but which could compete for bandwidth with other Tunnels.
4. Sharing: The L2/L3VPN network service allows sharing the MPLS-TE Tunnels supporting it with other services.

There could be additional TE binding requirements for the first three types with respect to different VN members of the same VN (on how different VN members, belonging to the same VN, can share or not network resources). For the first two cases, VN members can be hard‑isolated, soft-isolated, or shared. For the third case, VN members can be soft-isolated or shared.

### Multi-domain and multi-layer path computation

When a new SR-TE path needs to be setup, the MDSC is also responsible to coordinate the multi-layer/multi-domain path computation.

Depending on the knowledge that MDSC has of the topology and configuration of the underlying network domains, three approaches for performing multi-layer/multi-domain path computation are possible:

1. Full Summarization: In this approach, the MDSC has an abstracted TE topology view of all of its, packet and optical, underlying domains.  
     
   In this case, the MDSC does not have enough TE topology information to perform multi-layer/multi-domain path computation. Therefore the MDSC delegates the P-PNCs and O-PNCs to perform local path computation within their respective controlled domains and it uses the information returned by the P-PNCs and O-PNCs to compute the optimal multi-domain/multi-layer path.  
     
   This approach presents an issue to P-PNC, which does not have the capability of performing a single-domain/multi-layer path computation, since it can not retrieve the topology information from the O‑PNCs nor delegate the O‑PNC to perform optical path computation.  
     
   A possible solution could be to include a CNC function within the P-PNC to request the MDSC multi-domain optical path computation, as shown in Figure 10 of [RFC8453].  
     
   Another solution could be to rely on the MDSC recursive hierarchy, as defined in section 4.1 of [RFC8453], where, for each IP and optical domain pair, a "lower-level MDSC" (MDSC-L) provides the essential multi-layer correlation and the "higher-level MDSC" (MDSC-H) provides the multi-domain coordination.  
   In this case, the MDSC-H can get an abstact view of the underlying multi-layer domain topologies from its underlying MDSC-L. Each MDSC-L gets the full view of the IP domain topology from P-PNC and can get an abstracted view of the optical domain topology from its underlying O-PNC. In other words, topology abstraction is possible at the MPIs between MDSC-L and O-PNC and between MDSC-L and MDSC-H.
2. Partial summarization: In this approach, the MDSC has full visibility of the TE topology of the packet network domains and an abstracted view of the TE topology of the optical network domains.   
     
   The MDSC then has only the capability of performing multi-domain/single-layer path computation for the packet layer (the path can be computed optimally for the two packet domains).  
      
   Therefore, the MDSC still needs to delegate the O-PNCs to perform local path computation within their respective domains and it uses the information received by the O-PNCs, together with its TE topology view of the multi-domain packet layer, to perform multi-layer/multi-domain path computation.
3. Full knowledge: In this approach, the MDSC has the complete and enough detailed view of the TE topology of all the network domains (both optical and packet).  
     
   In such case MDSC has all the information needed to perform multi-domain/multi-layer path computation, without relying on PNCs.  
     
   This approach may present, as a potential drawback, scalability issues and, as discussed in section 2.2. of [PATH-COMPUTE], performing path computation for optical networks in the MDSC is quite challenging because the optimal paths depend also on vendor‑specific optical attributes (which may be different in the two domains if they are provided by different vendors).

This document analyses scenarios where the MDSC uses the partial summarization approach to coordinate multi‑domain/multi‑layer path computation.

Typically, the O-PNCs are responsible for the optical path computation of services across their respective single domains. Therefore, when setting up the network service, they must consider the connection requirements such as bandwidth, amplification, wavelength continuity, and non-linear impairments that may affect the network service path.

The methods and types of path requirements and impairments, such as those detailed in [OIA-TOPO], used by the O-PNC for optical path computation are not exposed at the MPI and therefore out of scope for this document.

## IP/MPLS Domain Controller and NE Functions

*[****Editors’ Note****:] Can we remove this section?*

IP/MPLS networks are assumed to have multiple domains. Each domain, corresponding to either an IGP area or an Autonomous System (AS) within the same operator network, is controlled by an IP/MPLS domain controller (P‑PNC).

Among the functions of the P‑PNC, there are the setup or modification of the intra‑domain SR‑TE paths, between PEs and BRs, and the configuration of the L2/L3 VPN network services, such as the VRF, in the PE nodes, as shown in Figure 3:

+------------------+ +------------------+

| | | |

| P‑PNC1 | | P‑PNC2 |

| | | |

+--|-----------|---+ +--|-----------|---+

| 1.Tunnel | 2.VPN | 1.Tunnel | 2.VPN

| Config | Provisioning | Config | Provisioning

V V V V

+---------------------+ +---------------------+

CE / PE tunnel 1 BR\ / BR tunnel 2 PE \ CE

o--/---o..................o--\-----/--o..................o---\--o

\ / \ /

\ Domain 1 / \ Domain 2 /

+---------------------+ +---------------------+

End-to-end tunnel

<------------------------------------------------->

1. IP/MPLS Domain Controller & NE Functions

It is assumed that BGP is running in the inter-domain link(s). The P‑PNC controller is also responsible for configuring the BGP speakers within its control domain, if necessary.

The BGP would be responsible for the end-to-end tunnel label distribution on PE and BR nodes. The MDSC is responsible for selecting the BRs and the intra-domain MPLS‑TE tunnels between PE/BR nodes.

If new MPLS‑TE tunnels are needed or modifications (e.g., bandwidth increase) to existing MPLS\_TE tunnels are needed, as outlined in section 2.2, the MDSC would request their setup or modifications to the P‑PNCs (step 1 in Figure 3). Then the MDSC would request the P‑PNC to configure the VPN, including selecting the intra‑domain TE tunnel (step 2 in Figure 3).

The P‑PNC should configure, using mechanisms outside the scope of this document, the ingress PE forwarding table, e.g., the VRF, to forward the VPN traffic, received from the CE, with the following three labels:

* VPN label: assigned by the egress PE and distributed by BGP;
* end‑to‑end LSP label: assigned by the egress BR, selected by the MDSC, and distributed by BGP;
* MPLS-TE tunnel label, assigned by the next hop P node of the tunnel selected by the MDSC and distributed by mechanism internal to the IP/MPLS domain (e.g., RSVP-TE).

## Optical Domain Controller and NE Functions

*[****Editors’ Note****:] Can we remove this section?*

The optical network provides the underlay connectivity services to IP/MPLS networks. The packet and optical multi-layer coordination is done by the MDSC, as shown in Figure 1.

The O‑PNC is responsible to:

* provide to the MDSC an abstract TE topology view of its underlying optical network resources;
* perform single‑domain local path computation, when requested by the MDSC;
* perform optical tunnel setup, when requested by the MDSC.

The mechanisms used by O‑PNC to perform intra‑domain topology discovery and path setup are usually vendor‑specific and outside the scope of this document.

Depending on the type of optical network, TE topology abstraction, path computation and path setup can be single‑layer (either OTN or WDM) or multi-layer OTN/WDM. In the latter case, the multi-layer coordination between the OTN and WDM layers is performed by the O‑PNC.

# Interface protocols and YANG data models for the MPIs

This section describes general assumptions applicable at all the MPI interfaces, between each PNC (Optical or Packet) and the MDSC, to support the scenarios discussed in this document.

## RESTCONF protocol at the MPIs

The RESTCONF protocol, as defined in [RFC8040], using the JSON representation defined in [RFC7951], is assumed to be used at these interfaces. In addition, extensions to RESTCONF, as defined in [RFC8527], to be compliant with Network Management Datastore Architecture (NMDA) defined in [RFC8342], are assumed to be used as well at these MPI interfaces and also at MDSC NBI interfaces.

## YANG data models at the MPIs

The data models used on these interfaces are assumed to use the YANG 1.1 Data Modeling Language, as defined in [RFC7950].

### Common YANG data models at the MPIs

As required in [RFC8040], the "ietf-yang-library" YANG module defined in [RFC8525] is used to allow the MDSC to discover the set of YANG modules supported by each PNC at its MPI.

Both Optical and Packet PNCs use the following common topology YANG data models at the MPI:

* The Base Network Model, defined in the “ietf-network” YANG module of [RFC8345];
* The Base Network Topology Model, defined in the “ietf-network-topology” YANG module of [RFC8345], which augments the Base Network Model;
* The TE Topology Model, defined in the “ietf-te-topology” YANG module of [RFC8795], which augments the Base Network Topology Model.

Both Optical and Packet PNCs use the common TE Tunnel Model, defined in the “ietf‑te” YANG module of [TE‑TUNNEL], at the MPI.

All the common YANG data models are generic and augmented by technology-specific YANG modules, as described in the following sections.

Both Optical and Packet PNCs also use the Ethernet Topology Model, defined in the “ietf-eth-te-topology” YANG module of [CLIENT-TOPO], which augments the TE Topology Model with Ethernet technology-specific information.

Both Optical and Packet PNCs use the following common notifications YANG data models at the MPI:

* Dynamic Subscription to YANG Events and Datastores over RESTCONF as defined in [RFC8650];
* Subscription to YANG Notifications for Datastores updates as defined in [RFC8641].

PNCs and MDSCs are compliant with subscription requirements as stated in [RFC7923].

### YANG models at the Optical MPIs

The Optical PNC uses at least one of the following technology-specific topology YANG data models, which augment the generic TE Topology Model:

* The WSON Topology Model, defined in the “ietf-wson-topology” YANG module of [RFC9094];
* the Flexi‑grid Topology Model, defined in the “ietf-flexi-grid-topology” YANG module of [Flexi‑TOPO];
* the OTN Topology Model, as defined in the “ietf-otn-topology” YANG module of [OTN-TOPO].

The optical PNC uses at least one of the following technology-specific tunnel YANG data models, which augments the generic TE Tunnel Model:

* The WSON Tunnel Model, defined in the “ietf-wson-tunnel” YANG modules of [WSON-TUNNEL];
* the Flexi‑grid Tunnel Model, defined in the “ietf-flexi-grid-tunnel” YANG module of [Flexi‑TUNNEL];
* the OTN Tunnel Model, defined in the “ietf-otn-tunnel” YANG module of [OTN-TUNNEL].

The optical PNC can optionally use the generic Path Computation YANG RPC, defined in the “ietf-te-path-computation” YANG module of [PATH‑COMPUTE].

Note that technology-specific augmentations of the generic path computation RPC for WSON, Flexi-grid and OTN path computation RPCs have been identified as a gap.

The optical PNC uses the Ethernet Client Signal Model, defined in the “ietf-eth-tran-service” YANG module of [CLIENT-SIGNAL].

### YANG data models at the Packet MPIs

The Packet PNC also uses at least the following technology-specific topology YANG data models:

* The L3 Topology Model, defined in the “ietf‑l3‑unicast‑topology” YANG module of [RFC8346], which augments the Base Network Topology Model;
* the L3 specific data model including extended TE attributes (e.g. performance derived metrics like latency), defined in “ietf-l3-te-topology” and in “ietf-te-topology-packet” YANG modules of [L3-TE-TOPO];
* the SR Topology Model, defined in the “ietf-sr-mpls-topology” YANG module of [SR‑TE‑TOPO].

*Need to check the need/applicability of the “ietf-l3-te-topology” in this scenario since it is not described in [SR-TE-TOPO].*

The packet PNC uses at least the following YANG data models:

* L3VPN Network Model (L3NM), defined in the “ietf-l3vpn-ntw” YANG module of [L3NM];
* L3NM TE Service Mapping, defined in the “ietf-l3nm-te-service-mapping” YANG module of [TSM];
* L2VPN Network Model (L2NM), defined in the “ietf-l2vpn-ntw” YANG module of [L2NM];
* L2NM TE Service Mapping, defined in the “ietf-l2nm-te-service-mapping” YANG module of [TSM].

## PCEP

[RFC8637] examines the applicability of a Path Computation Element (PCE) [RFC5440] and PCE Communication Protocol (PCEP) to the ACTN framework. It further describes how the PCE architecture applies to ACTN and lists the PCEP extensions that are needed to use PCEP as an ACTN interface. The stateful PCE [RFC8231], PCE-Initiation [RFC8281], stateful Hierarchical PCE (H-PCE) [RFC8751], and PCE as a central controller (PCECC) [RFC8283] are some of the key extensions that enable the use of PCE/PCEP for ACTN.

Since the PCEP supports path computation in the packet and optical networks, PCEP is well suited for inter-layer path computation. [RFC5623] describes a framework for applying the PCE-based architecture to interlayer (G)MPLS traffic engineering. Furthermore, the section 6.1 of [RFC8751] states the H-PCE applicability for inter-layer or POI.

[RFC8637] lists various PCEP extensions that apply to ACTN. It also lists the PCEP extension for optical network and POI.

Note that the PCEP can be used in conjunction with the YANG data models described in the rest of this document. Depending on whether ACTN is deployed in a greenfield or brownfield, two options are possible:

1. The MDSC uses a single RESTCONF/YANG interface towards each PNC to discover all the TE information and request TE tunnels. It may either perform full multi-layer path computation or delegate path computation to the underneath PNCs.  
     
   This approach is desirable for operators from an multi-vendor integration perspective as it is simple, and we need only one type of interface (RESTCONF) and use the relevant YANG data models depending on the operator use case considered. Benefits of having only one protocol for the MPI between MDSC and PNC have been already highlighted in [PATH-COMPUTE].
2. The MDSC uses the RESTCONF/YANG interface towards each PNC to discover all the TE information and requests the creation of TE tunnels. However, it uses PCEP for hierarchical path computation.  
     
   As mentioned in Option 1, from an operator perspective, this option can add integration complexity to have two protocols instead of one, unless the RESTOCONF/YANG interface is added to an existing PCEP deployment (brownfield scenario).

Section 4 of this draft analyses the case where a single RESTCONF/YANG interface is deployed at the MPI (i.e., option 1 above).

# Multi-layer and multi-domain services scenarios

Multi-layer and multi-domain scenarios, based on reference network described in section 2, and very relevant for Service Providers, are described in the next sections.

For each scenario, existing IETF protocols and data models, identified in section 3.1 and section 3.2, are analysed with particular focus on the MPI in the ACTN architecture.

## Scenario 1: inventory, service and network topology discovery

In this scenario, the MSDC needs to discover through the underlying PNCs:

* the network topology, at both optical and IP layers, in terms of nodes and links, including the access links, inter-domain IP links as well as cross-layer links;
* the optical tunnels supporting multi-layer intra-domain IP links;
* both intra-domain and inter-domain L2/L3 VPN network services deployed within the network;
* the SR-TE paths supporting those L2/L3 VPN network services;
* the hardware inventory information of IP and optical equipment.

The O-PNC and P-PNC could discover and report the hardware network inventory information of their equipment that is used by the different management layers. In the context of POI, the inventory information of IP and optical equipment can complement the topology views and facilitate the packet/optical multi-layer view, e.g., by providing a mapping between the lowest level LTPs in the topology view and corresponding physical port in the network inventory view.

The MDSC could also discover the entire network inventory information of both IP and optical equipment and correlate this information with the links reported in the network topology.

Each PNC provides to the MDSC the topology view of the domain it controls, as described in section 4.1.1 and 4.1.3. The MDSC uses this information to discover the complete topology view of the multi-layer multi-domain network it controls.

The MDSC should also maintain up-to-date inventory, service and network topology databases of both IP and optical layers through the use of IETF notifications through MPI with the PNCs when any network inventory/topology/service change occurs.

It should be possible also to correlate information coming from IP and optical layers (e.g., which port, lambda/OTSi, and direction, is used by a specific IP service on the WDM equipment).

In particular, for the cross-layer links, it is key for MDSC to automatically correlate the information from the PNC network databases about the physical ports from the routers (single link or bundle links for LAG) to client ports in the ROADM.

The analysis of multi-layer fault management is outside the scope of this document. However, the discovered information should be sufficient for the MDSC to easily correlate optical and IP layers alarms to speed-up troubleshooting.

Alarms and event notifications are required between MDSC and PNCs so that any network changes are reported almost in real-time to the MDSC (e.g., NE or link failure). As specified in [RFC7923], MDSC must subscribe to specific objects from PNC YANG datastores for notifications.

### Optical topology discovery

The WSON Topology Model or, alternatively, the Flexi‑grid Topology model is used to report the DWDM network topology (e.g., ROADM nodes and links), depending on whether the DWDM optical network is based on fixed grid or flexible-grid.

The OTN Topology Model is used to report the OTN network topology (e.g., OTN switching nodes and links), when the OTN switching layer is deployed within the optical domain.

In order to allow the MDSC to discover the complete multi-layer and multi-domain network topology and to correlate it with the hardware inventory information, the O-PNCs report an abstract optical network topology where:

* one TE node is reported for each optical NE deployed within the optical network domain; and
* one TE link is reported for each OMS link and, optionally, for each OTN link.

The Ethernet Topology Model is used to report the Ethernet client LTPs that terminate the cross-layer links: one Ethernet client LTP is reported for each Ethernet client interface on the optical NEs.

Since the MDSC delegates optical path computation to its underlay O-PNCs, the following information can be abstracted and not reported at the MPI:

* the optical parameters required for optical path computation, such as those detailed in [OIA-TOPO];
* the underlay OTS links and ILAs of OMS links;
* the physical connectivity between the optical transponders and the ROADMs.

otThis abstraction is valid independently of the fact that optical transponders are physically integrated within the same WDM node or are physically located on a device external to the WDM node since it both cases the optical transponders and the WDM node are under the control of the same O-PNC.

o

All the optical links are intra-domain and they are discovered by O-PNCs, using mechanisms which are outside the scope of this document, and reported at the MPIs within the optical network topology.

In case of a multi-layer DWDM/OTN network domain, multi-layer intra-domain OTN links are supported by underlay DWDM tunnels, which can be either WSON tunnels or, alternatively, Flexi-grid tunnels, depending on whether the DWDM optical network is based on fixed grid or flexible-grid. This relationship is reported by the mechanisms described in section 4.1.2.

### Optical path discovery

The WSON Tunnel Model or, alternatively, the Flexi‑grid Tunnel model, depending on whether the DWDM optical network is based on fixed grid or flexible-grid, is used to report all the DWDM tunnels established within the optical network.

When the OTN switching layer is deployed within the optical domain, the OTN Tunnel Model is used to report all the OTN tunnels established within the optical network.

The Ethernet client signal Model is used to report all the Ethernet connectivity provided by the underlay optical tunnels between Ethernet client LTPs. The underlay optical tunnels can be either DWDM tunnels or, when the optional OTN switching layer is deployed, OTN tunnels.

The DWDM tunnels can be used as underlay tunnels to support either Ethernet client connectivity or multi-layer intra-domain OTN links. In the latter case, the hierarchical-link container, defined in [TE-TUNNEL], is used to reference which multi-layer intra-domain OTN links are supported by the underlay DWDM tunnels.

The O-PNCs report in their operational datastores all the Ethernet client connectivities and all the optical tunnels deployed within their optical domain regarless of the mechanisms being used to set them up, such as the mechanisms described in section 4.2.2, as well as other mechanism (e.g., static configuration), which are outside the scope of this document.

### Packet topology discovery

The L3 Topology Model, SR Topology Model, TE Topology Model and the TE Packet Topology Model are used together to report the SR‑TE network topology, as described in figure 2 of [SR‑TE‑TOPO].

In order to allow the MDSC to discover the complete multi-layer and multi-domain network topology and to correlate it with the hardware inventory information as well as to perform multi-domain SR-TE path computation, the P-PNCs report the full SR-TE network, including all the information that is required by the MDSC to perform SR-TE path computation. In particular, one TE node is reported for each IP router and one TE link is reported for each intra-domain IP link.

All the intra-domain IP links are discovered by the P‑PNCs, using either LLDP [IEEE 802.1AB] or any other mechanisms, which are outside the scope of this document, and reported at the MPIs within the SR-TE network topology.

The Ethernet Topology Model is used to report the underlay intra-domain Ethernet links supporting the intra-domain IP links as well as the Ethernet LTPs that can terminate cross-layer links, inter-domain Ethernet links or access links, as described in detail in section 4.1.5 and 4.1.6.

### SR-TE paths discovery

This version of the draft assumes that discovery of existing SR­‑TE paths, including their bandwidth, at the MPI is done using the generic TE tunnel YANG data model, defined in [TE‑TUNNEL], with SR‑TE specific augmentations, as also outlined in section 1 of [TE‑TUNNEL].

Note that technology-specific augmentations of the generic path TE tunnel model for SR-TE path setup and discovery have been identified as a gap.

To enable MDSC to discover the full end-to-end SR-TE path configuration, the SR-TE specific augmentation of the [TE-TUNNEL] should allow the P-PNC to report the SID list assigned to an SR-TE path within its domain.

For example, considering the L3VPN in Figure 2, the PE13-P16-PE14 SR-TE path and the SR-TE path in the reverse direction (between PE14 and PE13) could be reported by the P-PNC1 to the MDSC as TE paths of the same TE tunnel instance. The bandwidth of these TE paths represents the bandwidth allocated by P‑PNC1 to the two SR-TE paths,which can be symmetric or asymmetric in the two directions.

The P-PNCs use the TE tunnel model to report, at the MPI, all the SR-TE paths established within their packet domain regardless of the mechanism being used to set them up. In other words, the TE tunnel data model reports within the operational datastore both the SR-TE paths being setup by the MDSC at the MPI, using the mechanisms described in section 4.2.3, as well as the SR-TE paths being setup by other means, such as static configuration, which are outside the scope of this document.

### Inter-domain links discovery

In the reference network of Figure 1, there are three types of inter‑domain links:

* Inter-domain IP links between two adjancent IP domains;
* Cross-layer links between an an IP domain and an adjacent optical domain;
* Access links between a CE device and a PE router.

All the three types of links are Ethernet physical links.

It is worth noting that the P‑PNC may not be aware whether an Ethernet interface terminates a cross-layer link, an inter-domain IP link or an access link.

It is not yet clarified which model can be used to report the access links between CEs and PEs (e.g., by using the Ethernet Topology Model defined in [CLIENT-TOPO] or by using the UNI Topology Model defined in [UNI-TOPO]). This has been identified as a gap.

The inter‑domain IP links and cross-layer links are discovered by the MDSC using the plug‑id attribute, as described in section 4.3 of [RFC8795].

More detailed description of how the plug-id can be used to discover inter-domain links is also provided in section 5.1.4 of [TNBI].

This document considers the following two options for discovering inter‑domain Ethernet links:

1. Static configuration
2. LLDP [IEEE 802.1AB] automatic discovery

Other options are possible but not described in this document.

As outlined in [TNBI], the encoding of the plug-id namespace and the LLDP information within the plug-id value is implementation specific and needs to be consistent across all the PNCs within the network.

The static configuration requires an administrative burden to configure network-wide unique identifiers: it is therefore more viable for inter‑domain IP links. For the cross-layer links, the automatic discovery solution based on LLDP snooping is preferable when possible.

The routers exchange standard LLDP packets as defined in [IEEE 802.1AB] and the optical NEs snoop the LLDP packets received from the local Ethernet interface and report to the O-PNCs the extracted information, such as the Chassis ID, the Port ID, System Name.

Note that the optical NEs do not actively participate in the LLDP packet exchange and does not send any LLDP packets.

The MDSC can discover a cross-layer link by matching the plug-id values of the two Ethernet LTPs reported by the adjacent O-PNC and P-PNC: the P-PNC reports the LLDP information sent by the corresponding router port while the O-PNC reports the LLDP information received by the corresponding optical NE local Ethernet interface.

It is worth noting that the discovery of cross-layer links is based only on the LLDP information sent by the IP router ports and received by the optical NE local Ethernet interfaces and therefore the MDSC can discover these links also before overlay multi-layer IP links are setup.

The MDSC can discover an inter-domain Ethernet link supporting an inter-domain IP link, by matching the plug-id values of the two Ethernet LTPs reported by the two adjacent P-PNCs: the two P-PNCs report both the LLDP information being sent and received from the corresponding router ports. The MDSC can then discover the inter-domain IP link between the two IP LTPs that are supported by the two Ethernet LTPs terminating an inter-domain Ethernet link.

The LLDP information required to be encoded in the plug-id attribute for inter-domain link discovery are the Chassis ID, or optionally the System Name, and the Port ID.

Different information is required to be encoded within the plug-id attribute of Etherent LTPs to discover cross-layer links and inter-domain IP links. If the P-PNC does not know a priori whether a router port terminates a cross-layer link or an inter-domain Ethernet link, it has to report two two Ethernet LTPs:

* An underlay Ethernet LTP reporting, within the plug-id attribute, the LLDP information sent by the corresponding router port;
* An overlay Ethernet LTP, supported by the other Ethernet LTP, reporting, within the plug-id attribute, both the LLDP information sent and received by the corresponding router port.

It is worth noting that in case of inter-domain IP link, no inter-domain Ethernet link can be discovered by the MDSC between the underlay Ethernet LTPs using the plug-id attribute. However, the MDSC may infer these inter-domain Ethernet link if it knows a priori, by mechanisms which are outside the scope of this document, that no other underlay Ethernet connectivity is possible.

For example, in the reference netwok of Figure 1, the MDSC can discover an underlay inter-domain Ethernet link between the two underlay Ethernet LTPs that supports two overlay Ethernet LTPs terminating an inter-domain Ethernet link discovered using the plug-id attribute.

also a multi-layer or a single-layer

, as described above, but no plug-id information is reported on the overlay Ethernet LTP since it is connected via an intra-domain Ethernet link to another Ethernet LTP

### Multi-layer IP links discovery

A multi-layer intra-domain IP link is discovered by the P-PNC like any other intra-domain IP link as described in section 4.1.3 and reported at the MPI within the Ethernet and SR-TE network topology.

Since in this case the LLDP information exchanged by the two adjacent routers’ ports are used by the P-PNC to discover the intra-domain Ethernet link, no LLDP information is provided in the plug-id attribute of the Ethenet LTPs that terminate intra-domain Ethernet links.

In addition, the P‑PNC also reports the two underlay Ethernet LTPs that supports the two Ethernet LTPs terminating the intra-domain Ethernet link supporting that multi-layer intra-domain IP link.

The MDSC therefore discovers which Ethernet cross-layer links support the multi-layer intra-domain IP links, as described in section 4.1.5.

The MDSC also discovers, from the information provided by the O-PNC and described in section 4.1.2, which optical tunnels support the multi-layer intra-domain IP links and therefore the path within the optical network that supports a multi-layer intra-domain IP link.

### LAG discovery

TBA

### L2/L3 VPN network services discovery

TBA

### Inventory discovery

The are no YANG data models in IETF that could be used to report at the MPI the whole inventory information discovered by a PNC.

[RFC8345] had foreseen some work for inventory as an augmentation of the network model, but no YANG data model has been developed so far.

There are also no YANG data models in IETF that could be used to correlate topology information, e.g., a link termination point (LTP), with inventory information, e.g., the physical port supporting an LTP, if any.

Inventory information through MPI and correlation with topology information is identified as a gap requiring further work and outside of the scope of this draft.

## Establishment of L2/L3 VPN network services with TE requirements

In this scenario the MDSC needs to setup a multi‑domain L2VPN or a multi-domain L3VPN with some SLA requirements.

The MDSC receives the request to setup a L2/L3 VPN network service from the OSS/Orchestration layer (see Appendix A).

The MDSC translates the L2/L3 VPN SLA requirements into TE requirements (e.g., bandwidth, TE metric bounds, SRLG disjointness, nodes/links/domains inclusion/exclusion) and find the SR-TE paths that meet these TE requirements (see section 2.1.1).

For example, considering the L3VPN in Figure 2, the MDSC finds that:

* a PE13-P16-PE14 SR-TE path already exists but have not enough bandwidth to support the new L3VPN, as described in section 4.1.4;
* the IP link(s) between P16 and PE14 has not enough bandwidth to support increasing the bandwidth of that SR-TE path, as described in section 4.1.3;
* a new underlay optical tunnel could be setup to increase the bandwidth IP link(s) between P16 and PE14 to support increasing the bandwidth of that overlay SR-TE path, as described in section 4.2.2. The dimensioning of the underlay optical tunnel is decided by the MDSC based on the bandwidth requested by the SR-TE path and on its multi-layer optimization policy, which is an internal MDSC implementation issue.

Considering for example the L3VPN in Figure 2, the MDSC can also decide that a new multi‑domain SR‑TE path needs to be setup between PE13 and PE23.

As described in section 2.1.2, with partial summarization, the MDSC will use the TE topology information provided by the P-PNCs and the results of the path computation requests sent to the O-PNCs, as described in section 4.2.1, to compute the multi‑layer/multi-domain path between PE13 and PE23.

For example, the multi-layer/multi-domain performed by the MDSC could require the setup of:

* a new underlay optical tunnel between PE13 and BR11, supporting a new IP link, as described in section 4.2.2;
* a new underlay optical tunnel between BR21 and P24 to increase the bandwidth of the IP link(s) between BR21 and P24, as described in section 4.2.2.

When the setup of the L2/L3 VPN network service requires multi-domain and multi-layer coordination, the MDSC is also responsible for coordinating the network configuration required to realize the request network service across the appropriate optical and packet domains.

The MDSC would therefore request:

* the O-PNC1 to setup a new optical tunnel between the ROADMs connected to P16 and PE14, as described in section 4.2.2;
* the P-PNC1 to update the configuration of the existing IP link, in case of LAG, or configure a new IP link, in case of ECMP, between P16 and PE14, as described in section 4.2.2;
* the P-PNC1 to update the bandwidth of the selected SR‑TE path between PE13 and PE14, as described in section 4.2.3.

After that, the MDSC requests P-PNC2 to setup an SR-TE path between BR21 and PE23, with an explicit path (BR21, P24, PE23) as described in section 4.2.3. The P‑PNC2, knowing the node and the adjacency SIDs assigned within its domain, can install the proper SR policy, or hierarchical policies, within BR21 and returns to the MDSC the assigned binding SID.

***[Editor’s Note]*** *Further investigation is needed for the SR specific extensions to the TE tunnel model.*

MDSC request P-PNC1 to setup an SR-TE path between PE13 and BR11, with an explicit path (PE13, BR11), specifying the inter‑domain link toward BR21 and the binding SID to be used for the end‑to-end SR-TE path stitching, as described in section 4.2.3. The P‑PNC1, knowing also the node and the adjacency SIDs assigned within its domain and the EPE SID assigned by BR11 to the inter‑domain link toward BR21, installs the proper policy, or policies, within PE13.

Once the SR-TE paths have been selected and, if needed, setup/modified, the MDSC can request to both P-PNCs to configure the L3VPN and its binding with the selected SR-TE paths using the [L3NM] and [TSM] YANG data models.

***[Editor’s Note]*** *Further investigation is needed to understand how the binding between a L3VPN and this new end‑to‑end SR-TE path can be configured.*

### Optical Path Computation

As described in section 2.1.2, the optical path computation is usually performed by the O-PNCs.

When performing multi-layer/multi-domain path computation, the MDSC can delegate the O-PNC for single-domain optical path computation.

As discussed in [PATH‑COMPUTE], there are two options to request an O-PNC to perform optical path computation: either via a "compute-only" TE tunnel path, using the generic TE tunnel YANG data model defined in [TE‑TUNNEL] or via the path computation RPC defined in [PATH‑COMPUTE].

This draft assumes that the path computation RPC is used.

As described in sections 4.1.1 and 4.1.5, there is a one-to-one relationship between the router ports, the cross-layer links and the optical TTPs. Therefore, the properties of an optical path between two optical TTPs, as computed by the O-PNC, can be used by the MDSC to infer the properties of the multi-layer single-domain IP link between the router ports associated with the two optical TTPs.

The are no YANG data models in IETF that could be used to augment the generic path computation RPC with technology‑specific attributes.

Optical technology-specific augmentation for the path computation RPC is identified as a gap requiring further work outside of this draft's scope.

### Multi-layer IP link Setup and Update

To setup a new multi-layer IP link between two router ports, the MDSC requires the O‑PNC to setup an optical tunnel (either a WSON Tunnel or a Flexi‑grid Tunnel or an OTN Tunnel) within the optical network between the two TTPs associated, as described in section 4.2.1, with these two router ports.

The MDSC also requires the O‑PNC to steer the Ethernet client traffic between the two cross-layer links over the optical tunnel using the Ethernet Client Signal Model.

After the optical tunnel has been setup and the client traffic steering configured, the two IP routers can exchange Ethernet packets between themselves, including LLDP messages.

If LLDP [IEEE 802.1AB] or any other discovery mechanism, which is outside the scope of this document, is used between the adjacency between the two routers’ ports, the P-PNC can automatically discover the underlay multi-layer single-domain Ethernet link being set up by the MDSC. Otherwise, the MDSC configures this multi-layer single-domain Ethernet link at the MPI of the P-PNC.

The two Ethernet LTPs terminating this multi-layer single-domain Ethernet link are supported by the two underlay Ethernet LTPs terminating the two cross-layer links.

In both cases, the MDSC configures the multi-layer single-domain IP link between the two IP LTPs supported by the two Ethernet LTPs terminating its underlay multi-layer single-domain Ethernet link.

The MDSC also configures the multi-layer single-domain IP link between the two IP LTPs supported by the two Ethernet LTPs terminating the multi-layer single-domain Ethernet link which has been either discovered by the P-PNC or configured by the MDSC.

***[Editor’s Note]*** *Add text for IP link update and clarify that the IP link bandwidth increase can be done either by LAG or by ECMP. Both options are valid and widely deployed and more or less the same from POI perspective.*

***[Editor’s Note]*** *Add text about the configuration of multi-layer SRLG information (issue #45).*

It is worth noting that the list of SRLGs for a multi-layer IP link can be quite long. Implementation-specific mechanisms can be implemented by the MDSC or by the O-PNC to summarize the SRLGs of an optical tunnel. These mechanisms are implementation-specific and have no impact on the YANG models nor on the interoperability at the MPI, but cares have to be taken to avoid missing information.

### SR-TE Path Setup and Update

This version of the draft assumes that SR-TE path setup and update at the MPI could be done using the generic TE tunnel YANG data model, defined in [TE-TUNNEL], with SR TE specific augmentations, as also outlined in section 1 of [TE-TUNNEL].

The MDSC can use the [TE-TUNNEL] model to request the P‑PNC to setup TE paths specifying the explicit path to force the P‑PNC to setup the actual path being computed by MDSC.

The [TE-TUNNEL] model supports requesting the setup of both end-to‑end as well as segment TE paths (within one domain).

In the latter case, SR‑TE specific augmentations of the [TE‑TUNNEL] model should be defined to allow the MDSC to configure the binding SIDs to be used for the end to-end SR-TE path stitching and to allow the P‑PNC to report the binding SID assigned to the segment TE paths.

The assigned binding SID should be persistent in case router or P-PNC rebooting.

The MDSC can also use the [TE-TUNNEL] model to request the P‑PNC to increase the bandwidth allocated to an existing TE path, and, if needed, also on its reverse TE path. The [TE-TUNNEL] model supports both symmetric and asymmetric bandwidth configuration in the two directions.

The MDSC also request the P-PNC to configure TI-LFA local protection: the mechanisms to request the configuration TI-LFA local protection for SR-TE paths using the [TE-TUNNEL] are a gap in the current YANG models.

The TI-LFA local protection within the P-PNC domain is configured by the P-PNC through implementation specific mechanisms which are outside the scope of this document. The P-PNC takes into account the multi-layer SRLG information, configured by the MDSC as described in section 4.2.2, when computing the TI-LFA post-convergence path for multi-layer single-domain IP links.

SR-TE path setup and update (e.g., bandwidth increase) through MPI is identified as a gap requiring further work, which is outside of the scope of this draft.

# Conclusions

The analysis provided in this document has shown that the IETF YANG models described in 3.2 provides useful support for Packet Optical Integration (POI) scenarios for resource discovery (network topology, service, tunnels and network inventory discovery) as well as for supporting multi-layer/multi-domain L2/L3 VPN network services.

Few gaps have been identified to be addressed by the relevant IETF Working Groups:

* network inventory model: this gap has been identified in section 4.1.9 and the solution in [NETWORK-INVENTORY] has been proposed to resolve it;
* technology-specific augmentations of the path computation RPC, defined in [PATH-COMPUTE] for optical networks: this gap has been identified in section 4.2.1 and the solution in [OPTICAL-PATH-COMPUTE] has been proposed to resolve it;
* relationship between a common discovery mechanisms applicable to access links, inter-domain IP links and cross-layer links and the UNI topology discover mechanism defined in [UNI-TOPO]: this gap has been identified in section 4.1.3;
* a mechanism applicable to the P-PNC NBI to configure the SR-TE paths. Technology-specific augmentations of TE Tunnel model, defined in [TE-TUNNEL], are foreseen in section 1 of [TE-TUNNEL] but not yet defined: this gap has been identified in section 4.2.3.

# Security Considerations

Several security considerations have been identified and will be discussed in future versions of this document.

# Operational Considerations

Telemetry data, such as collecting lower-layer networking health and consideration of network and service performance from POI domain controllers, may be required. These requirements and capabilities will be discussed in future versions of this document.

# IANA Considerations

This document requires no IANA actions.

# References

## Normative References

[RFC7923] Voit, E. et al., "Requirements for Subscription to YANG Datastores", RFC 7923, June 2016.

[RFC7950] Bjorklund, M. et al., "The YANG 1.1 Data Modeling Language", RFC 7950, August 2016.

[RFC7951] Lhotka, L., "JSON Encoding of Data Modeled with YANG", RFC 7951, August 2016.

[RFC8040] Bierman, A. et al., "RESTCONF Protocol", RFC 8040, January 2017.

[RFC8342] Bjorklund, M. et al., "Network Management Datastore Architecture (NMDA)", RFC 8342, March 2018.

[RFC8345] Clemm, A., Medved, J. et al., “A Yang Data Model for Network Topologies”, RFC8345, March 2018.

[RFC8346] Clemm, A. et al., “A YANG Data Model for Layer 3 Topologies”, RFC8346, March 2018.

[RFC8453] Ceccarelli, D., Lee, Y. et al., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC8453, August 2018.

[RFC8525] Bierman, A. et al., "YANG Library", RFC 8525, March 2019.

[RFC8527] Bjorklund, M. et al., "RESTCONF Extensions to Support the Network Management Datastore Architecture", RFC 8527, March 2019.

[RFC8641] Clemm, A. and E. Voit, "Subscription to YANG Notifications for Datastore Updates", RFC 8641, September 2019.

[RFC8650] Voit, E. et al., "Dynamic Subscription to YANG Events and Datastores over RESTCONF", RFC 8650, November 2019.

[RFC8795] Liu, X. et al., "YANG Data Model for Traffic Engineering (TE) Topologies", RFC8795, August 2020.

[RFC9094] Zheng H., Lee, Y. et al., "A YANG Data Model for Wavelength Switched Optical Networks (WSONs)", RFC 9094, August 2021.

[IEEE 802.1AB] IEEE 802.1AB-2016, "IEEE Standard for Local and metropolitan area networks - Station and Media Access Control Connectivity Discovery", March 2016.

[Flexi‑TOPO] Lopez de Vergara, J. E. et al., "YANG data model for Flexi-Grid Optical Networks", draft-ietf-ccamp-flexigrid-yang, work in progress.

[OTN-TOPO] Zheng, H. et al., "A YANG Data Model for Optical Transport Network Topology", draft-ietf-ccamp-otn-topo-yang, work in progress.

[CLIENT-TOPO] Zheng, H. et al., "A YANG Data Model for Client-layer Topology", draft-zheng-ccamp-client-topo-yang, work in progress.

[L3-TE-TOPO] Liu, X. et al., "YANG Data Model for Layer 3 TE Topologies", draft-ietf-teas-yang-l3-te-topo, work in progress.

[SR‑TE‑TOPO] Liu, X. et al., "YANG Data Model for SR and SR TE Topologies on MPLS Data Plane", draft-ietf-teas-yang-sr-te-topo, work in progress.

[TE-TUNNEL] Saad, T. et al., "A YANG Data Model for Traffic Engineering Tunnels and Interfaces", draft-ietf-teas-yang-te, work in progress.

[WSON‑TUNNEL] Lee, Y. et al., "A Yang Data Model for WSON Tunnel", draft-ietf-ccamp-wson-tunnel-model, work in progress.

[Flexi‑TUNNEL] Lopez de Vergara, J. E. et al., "A YANG Data Model for Flexi-Grid Tunnels ", draft-ietf-ccamp-flexigrid-tunnel-yang, work in progress.

[OTN-TUNNEL] Zheng, H. et al., "OTN Tunnel YANG Model", draft-ietf-ccamp-otn-tunnel-model, work in progress.

[PATH-COMPUTE] Busi, I., Belotti, S. et al, "Yang model for requesting Path Computation", draft-ietf-teas-yang-path-computation, work in progress.

[CLIENT-SIGNAL] Zheng, H. et al., "A YANG Data Model for Transport Network Client Signals", draft-ietf-ccamp-client-signal-yang, work in progress.

## Informative References

[RFC1930] J. Hawkinson, T. Bates, “Guideline for creation, selection, and registration of an Autonomous System (AS)”, RFC 1930, March 1996.

[RFC5440] Vasseur, JP. et al., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[RFC5623] Oki, E. et al., "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 5623, September 2009.

[RFC8231] Crabbe, E. et al., "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, September 2017.

[RFC8281] Crabbe, E. et al., "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, December 2017.

[RFC8283] Farrel, A. et al., "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, December 2017.

[RFC8309] Q. Wu, W. Liu, and A. Farrel, “Service Model Explained”, RFC 8309, January 2018.

[RFC8637] Dhody, D. et al., "Applicability of the Path Computation Element (PCE) to the Abstraction and Control of TE Networks (ACTN)", RFC 8637, July 2019.

[RFC8751] Dhody, D. et al., "Hierarchical Stateful Path Computation Element (PCE)", RFC 8751, March 2020.

[L2NM] S. Barguil, et al., “A Layer 2 VPN Network YANG Model”, draft-ietf-opsawg-l2nm, work in progress.

[L3NM] S. Barguil, et al., “A Layer 3 VPN Network YANG Model”, draft-ietf-opsawg-l3sm-l3nm, work in progress.

[TSM] Y. Lee, et al., “Traffic Engineering and Service Mapping Yang Model”, draft-ietf-teas-te-service-mapping-yang, work in progress.

[TNBI] Busi, I., Daniel, K. et al., "Transport Northbound Interface Applicability Statement", draft-ietf-ccamp-transport-nbi-app-statement, work in progress.

[VN] Y. Lee, et al., “A Yang Data Model for ACTN VN Operation”, draft-ietf-teas-actn-vn-yang, work in progress.

[OIA-TOPO] Lee Y. et al., "A YANG Data Model for Optical Impairment-aware Topology", draft-ietf-ccamp-optical-impairment-topology-yang, work in progress.

[UNI-TOPO] Gonzalez de Dios O. et al., "A Network YANG Model for Service Attachment Points", draft-dbwb-opsawg-sap, work in progress.

[NETWORK-INVENTORY] Yu C. et al., "A YANG Data Model for Optical Network Inventory", draft-yg3bp-ccamp-optical-inventory-yang, work in progress.

[OPTICAL-PATH-COMPUTE] Busi I. et al., "YANG Data Models for requesting Path Computation in Optical Networks", draft-gbb-ccamp-optical-path-computation-yang, work in progress.

1. OSS/Orchestration Layer

The OSS/Orchestration layer is a vital part of the architecture framework for a service provider:

* to abstract (through MDSC and PNCs) the underlying transport network complexity to the Business Systems Support layer;
* to coordinate NFV, Transport (e.g. IP, optical and microwave networks), Fixed Acess, Core and Radio domains enabling full automation of end-to-end services to the end customers;
* to enable catalogue-driven service provisioning from external applications (e.g. Customer Portal for Enterprise Business services), orchestrating the design and lifecycle management of these end-to-end transport connectivity services, consuming IP and/or optical transport connectivity services upon request.

As discussed in section 2.1, in this document, the MDSC interfaces with the OSS/Orchestration layer and, therefore, it performs the functions of the Network Orchestrator, defined in [RFC8309].

The OSS/Orchestration layer requests the creation of a network service to the MDSC specifying its end-points (PEs and the interfaces towards the CEs) as well as the network service SLA and then proceeds to configuring accordingly the end-to-end customer service between the CEs in the case of an operator managed service.

* 1. MDSC NBI

As explained in section 2, the OSS/Orchestration layer can request the MDSC to setup L2/L3VPN network services (with or without TE requirements).

Although the OSS/Orchestration layer interface is usually operator-specific, typically it would be using a RESTCONF/YANG interface with a more abstracted version of the MPI YANG data models used for network configuration (e.g. L3NM, L2NM).

Figure 4 shows an example of possible control flow between the OSS/Orchestration layer and the MDSC to instantiate L2/L3 VPN network services, using the YANG data models under the definition in [VN], [L2NM], [L3NM] and [TSM].

+-------------------------------------------+

| |

| OSS/Orchestration layer |

| |

+-----------------------+-------------------+

|

1.VN 2. L2/L3NM & | ^

| TSM | |

| | | |

| | | |

v v | 3. Update VN

|

+-----------------------+-------------------+

| |

| MDSC |

| |

+-------------------------------------------+

1. Service Request Process

* The VN YANG data model, defined in [VN], whose primary focus is the CMI, can also provide VN Service configuration from an orchestrated network service point of view when the L2/L3 VPN network service has TE requirements. However, this model is not used to setup L2/L3 VPN service with no TE requirements.
  + It provides the profile of VN in terms of VN members, each of which corresponds to an edge-to-edge link between customer end-points (VNAPs). It also provides the mappings between the VNAPs with the LTPs and the connectivity matrix with the VN member. The associated traffic matrix (e.g., bandwidth, latency, protection level, etc.) of VN member is expressed (i.e., via the TE-topology’s connectivity matrix).
  + The model also provides VN-level preference information (e.g., VN member diversity) and VN-level admin-status and operational-status.
* The L2NM and L3NM YANG data models, defined in [L2NM] and [L3NM], whose primary focus is the MPI, can also be used to provide L2VPN and L3VPN network service configuration from a orchestrated connectivity service point of view.
* The TE & Service Mapping YANG data model [TSM] provides TE-service mapping.
  + TE-service mapping provides the mapping between a L2/L3VPN instance and the corresponding VN instances.
  + The TE-service mapping also provides the service mapping requirement type as to how each L2/L3 VPN/VN instance is created concerning the underlay TE tunnels (e.g., whether they require a new and isolated set of TE underlay tunnels or not.
  + Site mapping provides the site reference information across L2/L3VPN Site ID, VN Access Point ID, and the LTP of the access link.

1. Multi-layer and multi-domain resiliency
   1. Maintenance Window

Before planned maintenance operation on DWDM network takes place, IP traffic should be moved hitless to another link.

MDSC must reroute IP traffic before the events takes place. It should be possible to lock IP traffic to the protection route until the maintenance event is finished, unless a fault occurs on such path.

* 1. Router port failure

The focus is on client-side protection scheme between IP router and reconfigurable ROADM. Scenario here is to define only one port in the routers and in the ROADM muxponder board at both ends as back-up ports to recover any other port failure on client-side of the ROADM (either on router port side or on muxponder side or on the link between them). When client-side port failure occurs, alarms are raised to MDSC by IP-PNC and O-PNC (port status down, LOS etc.). MDSC checks with OP-PNC(s) that there is no optical failure in the optical layer.

There can be two cases here:

1. LAG was defined between the two end routers. MDSC, after checking that optical layer is fine between the two end ROADMs, triggers the ROADM configuration so that the router back-up port with its associated muxponder port can reuse the OCh that was already in use previously by the failed router port and adds the new link to the LAG on the failure side.  
     
   While the ROADM reconfiguration takes place, IP/MPLS traffic is using the reduced bandwidth of the IP link bundle, discarding lower priority traffic if required. Once back-up port has been reconfigured to reuse the existing OCh and new link has been added to the LAG then original Bandwidth is recovered between the end routers.  
     
   Note: in this LAG scenario let assume that BFD is running at LAG level so that there is nothing triggered at MPLS level when one of the link member of the LAG fails.
2. If there is no LAG then the scenario is not clear since a router port failure would automatically trigger (through BFD failure) first a sub-50ms protection at MPLS level :FRR (MPLS RSVP-TE case) or TI-LFA (MPLS based SR-TE case) through a protection port. At the same time MDSC, after checking that optical network connection is still fine, would trigger the reconfiguration of the back-up port of the router and of the ROADM muxponder to re-use the same OCh as the one used originally for the failed router port. Once everything has been correctly configured, MDSC Global PCE could suggest to the operator to trigger a possible re-optimization of the back-up MPLS path to go back to the MPLS primary path through the back-up port of the router and the original OCh if overall cost, latency etc. is improved. However, in this scenario, there is a need for protection port PLUS back-up port in the router which does not lead to clear port savings.

Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

Some of this analysis work was supported in part by the European Commission funded H2020-ICT-2016-2 METRO-HAUL project (G.A. 761727).

Contributors

Sergio Belotti  
Nokia

Email: sergio.belotti@nokia.com

Gabriele Galimberti  
Cisco

Email: ggalimbe@cisco.com

Zheng Yanlei  
China Unicom

Email: zhengyanlei@chinaunicom.cn

Anton Snitser  
Sedona

Email: antons@sedonasys.com

Washington Costa Pereira Correia  
TIM Brasil

Email: wcorreia@timbrasil.com.br

Michael Scharf  
Hochschule Esslingen - University of Applied Sciences

Email: michael.scharf@hs-esslingen.de

Young Lee  
Sung Kyun Kwan University

Email: younglee.tx@gmail.com

Jeff Tantsura  
Apstra

Email: jefftant.ietf@gmail.com

Paolo Volpato  
Huawei

Email: paolo.volpato@huawei.com

Brent Foster  
Cisco

Email: [brfoster@cisco.com](mailto:brfoster@cisco.com)

Authors’ Addresses

Fabio Peruzzini

TIM

Email: fabio.peruzzini@telecomitalia.it

Jean-Francois Bouquier  
Vodafone

Email: jeff.bouquier@vodafone.com

Italo Busi  
Huawei

Email: Italo.busi@huawei.com

Daniel King  
Old Dog Consulting

Email: daniel@olddog.co.uk

Daniele Ceccarelli  
Ericsson

Email: daniele.ceccarelli@ericsson.com