

TEAS Working Group
Internet Draft
Intended status: Informational

Fabio Peruzzini
TIM
Jean-Francois Bouquier
Vodafone
Italo Busi
Huawei
Daniel King
Old Dog Consulting
Daniele Ceccarelli
Cisco

Expires: August 2024

February 22, 2024

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

draft-ietf-teas-actn-poi-applicability-11

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 defined by the IETF to support this deployment architecture and specific scenarios relevant to 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/lid-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 August 22, 2024.

Copyright Notice

Copyright (c) 2024 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
1.1. Terminology.....	5
2. Reference Network Architecture.....	7
2.1. Multi-domain Service Coordinator (MDSC) functions.....	9
2.1.1. Multi-domain L2/L3 VPN Network Services.....	11
2.1.2. Multi-domain and Multi-layer Path Computation.....	14
2.2. IP/MPLS Domain Controller and NE Functions.....	17
2.3. Optical Domain Controller and NE Functions.....	19
3. Interface Protocols and YANG Data Models for the MPIs.....	19
3.1. RESTCONF Protocol at the MPIs.....	19
3.2. YANG Data Models at the MPIs.....	20
3.2.1. Common YANG Data Models at the MPIs.....	20
3.2.2. YANG models at the Optical MPIs.....	21
3.2.3. YANG data models at the Packet MPIs.....	22
3.3. Path Computation Element Protocol (PCEP).....	23
4. Inventory, Service and Network Topology Discovery.....	24
4.1. Optical Topology Discovery.....	25
4.2. Optical Path Discovery.....	28
4.3. Packet Topology Discovery.....	29
4.4. TE Path Discovery.....	30
4.5. Inter-domain Link Discovery.....	31

4.5.1. Cross-layer Link Discovery.....	32
4.5.2. Inter-domain IP Link Discovery.....	34
4.6. Multi-layer IP Link Discovery.....	37
4.6.1. Single-layer Intra-domain IP Links.....	39
4.7. LAG Discovery.....	41
4.8. L2/L3 VPN Network Services Discovery.....	43
4.9. Inventory Discovery.....	44
5. Establishment of L2/L3 VPN Services with TE Requirements.....	44
5.1. Optical Path Computation.....	46
5.2. Multi-layer IP Link Setup.....	47
5.2.1. Multi-layer LAG Setup.....	49
5.2.2. Multi-layer LAG Update.....	50
5.2.3. Multi-layer TE path properties Configuration.....	50
5.3. TE Path Setup and Update.....	51
5.4. L2/L3 VPN Network Service Setup.....	52
6. Conclusions.....	53
7. Security Considerations.....	54
8. Operational Considerations.....	55
9. IANA Considerations.....	55
10. References.....	55
10.1. Normative References.....	55
10.2. Informative References.....	57
Appendix A. Additional Scenarios.....	60
A.1. OSS/Orchestration Layer.....	60
A.1.1. MDSC NBI.....	60
A.2. Multi-layer and Multi-domain Resiliency.....	62
A.2.1. Maintenance Window.....	62
A.2.2. Router Port Failure.....	62
A.3. Muxponders.....	63
Acknowledgments.....	65
Contributors.....	65
Authors' Addresses.....	67

1. 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 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 other vendors in both domains. The operation of these complex packet and optical networks is often siloed, as these technology domains require specific skill sets.

The packet/optical network deployment and operation separation are inefficient for many reasons. First, both capital expenditure (CAPEX) and operational expenditure (OPEX) could be significantly reduced by integrating the packet and the optical networks. Second, multi-layer online topology insight can speed up troubleshooting (e.g., alarm correlation) and network operation (e.g., coordination of maintenance events), and multi-layer offline topology inventory can improve service quality (e.g., detection of diversity constraint violations). Third, 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 across layers (e.g., to achieve bandwidth-on-demand or to perform activities during maintenance windows).

This document uses packet-based Traffic Engineered (TE) service examples. These are described as "TE-path" in this document. Unless otherwise stated, these TE services may be instantiated using RSVP-TE-based or SR-TE-based, forwarding plane mechanisms.

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

This document describes critical scenarios for POI from the packet service layer perspective and identifies the required coordination

between packet and optical layers to improve POI deployment and operation. These scenarios focus on multi-domain packet networks operated as a client of optical networks.

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

Multi-layer and multi-domain scenarios, based on the reference network described in section 2 and very relevant for Service Providers, are described in sections 4 and 5.

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

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

A summary of the gaps identified in this analysis is provided in section 6.

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.

1.1. 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 address, routing protocol etc.)

Port:

the physical entity that transmits and receives physical signals

Interface:

a physical or logical entity that transmits and receives traffic

Link:

an association between two interfaces that can exchange traffic directly

Ethernet link:

a link between two Ethernet interfaces

IP link:

a link between two IP interfaces

Cross-layer link:

an Ethernet link between an Ethernet interface on a router and an Ethernet interface on an optical NE

Intra-domain single-layer Ethernet link:

an Ethernet link between between two Ethernet interfaces on physically adjacent routers that belong to the same P-PNC domain

Intra-domain single-layer IP link:

an IP link supported by an intra-domain single-layer Ethernet link

Inter-domain single-layer Ethernet link:

an Ethernet link between between two Ethernet interfaces on physically adjacent routers which belong to different P-PNC domains

Inter-domain single-layer IP link:

an IP link supported by an inter-domain single-layer Ethernet link.

Intra-domain multi-layer Ethernet link:

an Ethernet link supported by two cross-layer links and an optical tunnel in between

Intra-domain multi-layer IP link:

an IP link supported an intra-domain multi-layer Ethernet link

2. 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:

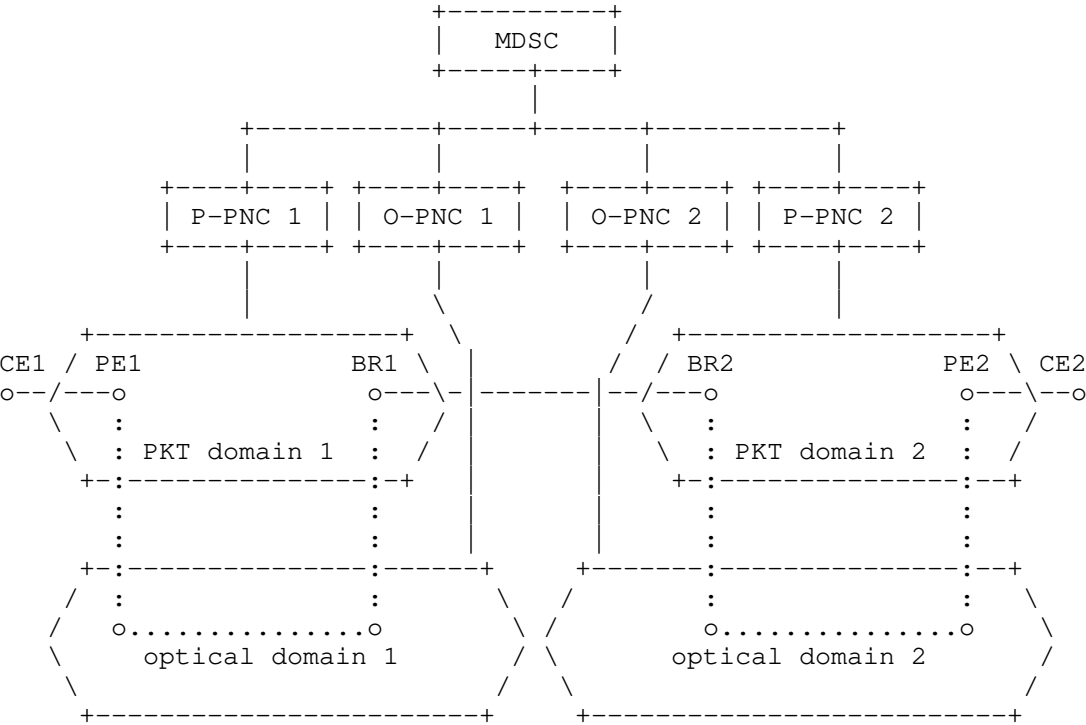


Figure 1 - Reference Network

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 a packet network).

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

- o 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;
- o There are no inter-domain physical links between optical domains. Inter-domain physical links exist only:
 - o 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;
 - o between packet and optical domains (i.e., between routers and optical NEs): these links are called cross-layer links within this document;
 - o between customer sites and the packet network (i.e., between CE devices and PE routers): these links are called access links within this document.

- o All the physical interfaces at inter-domain links are Ethernet physical interfaces.

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

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.

2.1. 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 and provides multi-layer/multi-domain L2/L3 VPN network services requested by an OSS/Orchestration layer.

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

1. 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.

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] described in point 2 above. Therefore, the MDSC deals 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 outside this document's scope.

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

- o Initiated by request from the OSS/Orchestration layer to setup L2/L3 VPN network services that require multi-layer/multi-domain coordination;
- o The MDSC initiates them 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 fulfilment, these workflows are not related to a network service provisioning request 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.

2.1.1. Multi-domain L2/L3 VPN Network Services

Figure 2 and Figure 3 provide an example of a 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.

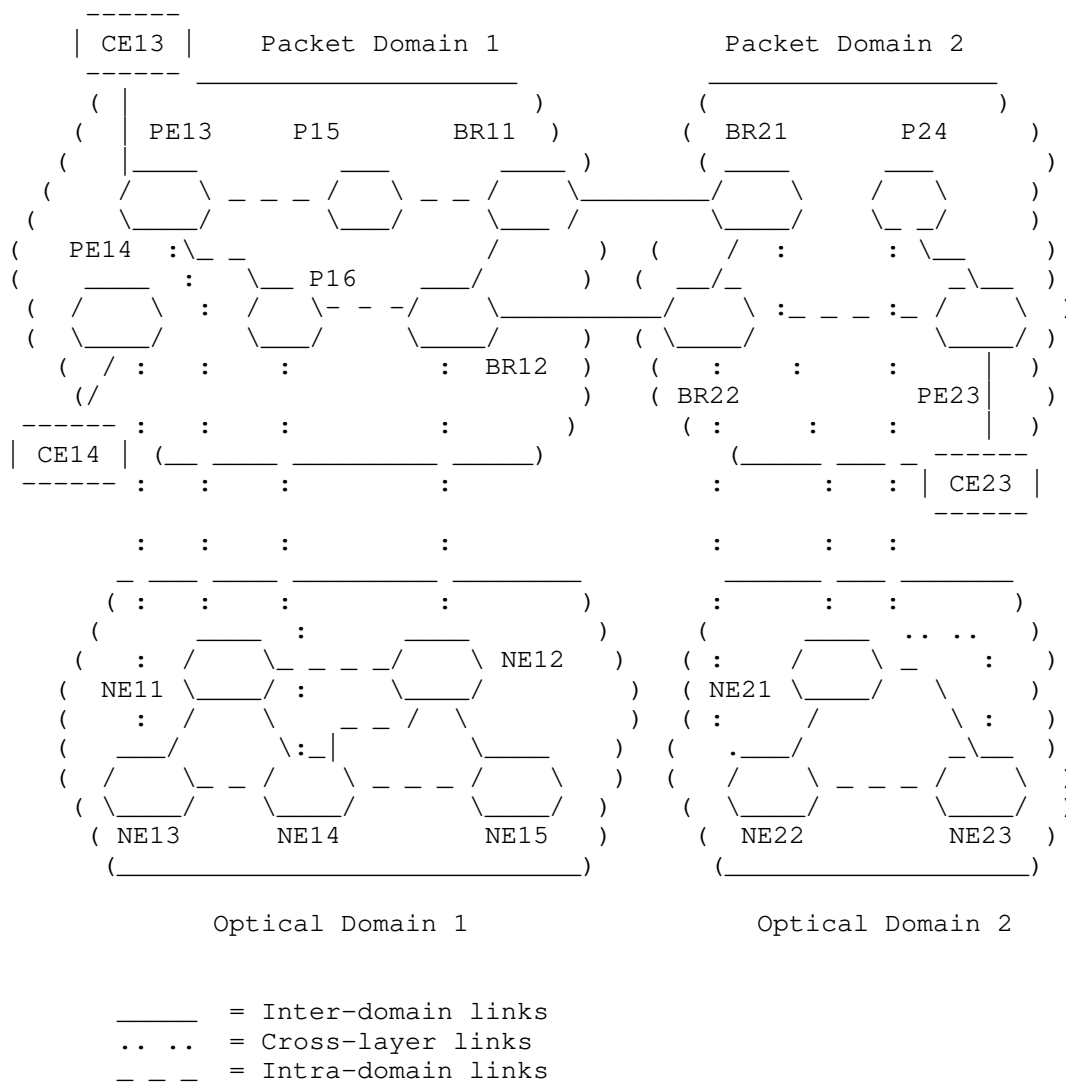


Figure 2 - Multi-domain VPN topology example

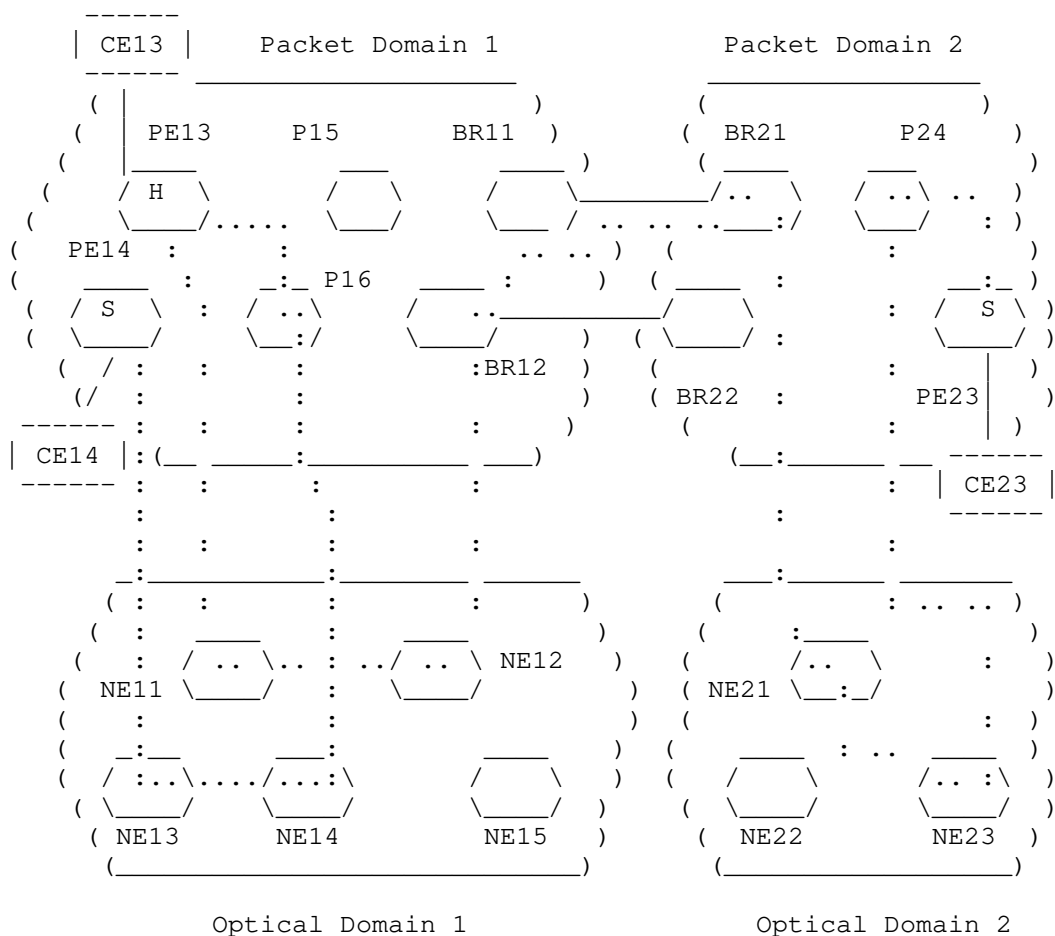


Figure 3 - Multi-domain VPN TE paths example

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

1. BGP-LU ([RFC8277])

2. Inter-domain RSVP-TE
3. Inter-domain SR-TE

This document analyses the inter-domain TE options for which the TE tunnel model, defined in [TE-TUNNEL], could be used at the MPI for intra-domain or inter-domain TE configuration. The analysis of other options is outside the scope of this draft.

It is also assumed that:

- o the bandwidth of each intra-domain TE path is managed by its respective P-PNC;
- o technology-specific mechanisms (in the case of inter-domain SR-TE, the binding SID) are used for the inter-domain TE path stitching;
- o each packet domain in Figure 2 uses technology-specific local protection mechanisms (in the case of SR-TE, TI-LFA), with the awareness of multi-layer TE path properties (e.g., SRLG).

In the case of inter-domain TE-paths, it is also assumed that each packet domain in Figure 2 and Figure 3 implements the same TE technology, and the stitching between two domains is done using inter-domain TE.

In this scenario, one of the key MDSC functions is to identify the multi-domain/multi-layer TE paths to be 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 TE paths.

The selection of the TE path should take into account the TE requirements and the binding requirements for the L2/L3 VPN network service.

In general, the binding requirements for a network service (e.g., L2/L3 VPN) can be summarized within three cases:

1. The customer is asking for VPN isolation to dynamically create and bind tunnels to the service so that they are not shared by other services (e.g. VPN).
The level of isolation can be different:
 - a) Hard isolation with deterministic latency means L2/L3 VPN requires a set of dedicated TE Tunnels (neither sharing

- with other services nor competing for bandwidth with other tunnels), providing deterministic latency performances
 - b) Hard isolation but without deterministic characteristics
 - c) Soft isolation means the tunnels associated with L2/L3 VPN are dedicated to that but can compete for bandwidth with other tunnels.
2. The customer does not ask for isolation and could request a VPN service where associated tunnels can be shared across multiple VPNs.

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

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

This document analyses scenarios where only one TE path is used to carry the VPN traffic between PEs. Scenarios, where multiple parallel TE paths are used in load-balancing to carry the VPN traffic between PEs, are possible but their analysis is outside the scope of this document.

2.1.2. Multi-domain and Multi-layer Path Computation

When a new TE path needs to be setup, the MDSC is also responsible for coordinating 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. Then, 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 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 abstract 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 complete 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. It uses the information received by the O-PNCs and 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 a 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 different vendors provide them).

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.

2.2. IP/MPLS Domain Controller and NE Functions

Each packet domain in Figure 1, corresponding to either an IGP area or an Autonomous System (AS) within the same operator network, is controlled by a packet domain controller (P-PNC).

P-PNCs are responsible for setting up the TE paths between any two PEs or BRs in their respective controlled domains, as requested by MDSC, and providing topology information to the MDSC.

For example, for inter-domain SR-TE, the setup bidirectional SR-TE path from PE13 in domain 1 to PE23 in domain 2, as shown in Figure 3, requires the MDSC to coordinate the actions of:

- o P-PNC1 to push a SID list to PE13 including the Binding SID associated to the SR-TE path in Domain 2 with PE23 as the target destination (forward direction);
- o P-PNC2 to push a SID list to PE23, including the Binding SID associated with the SR-TE path in Domain 1 with PE13 as the target destination (reverse direction).

With reference to Figure 4, P-PNCs are then responsible:

1. To expose to MDSC their respective detailed TE topology
2. To perform single-layer single-domain local TE path computation, when requested by MDSC between two PEs (for single-domain end-to-end TE path) or between PEs and BRs for an inter-domain TE path selected by MDSC;
3. To configure the routers in their respective domain to setup a TE path;
4. To configure the VRF and PE-CE interfaces (Service access points) of the intra-domain and inter-domain network services requested by the MDSC.

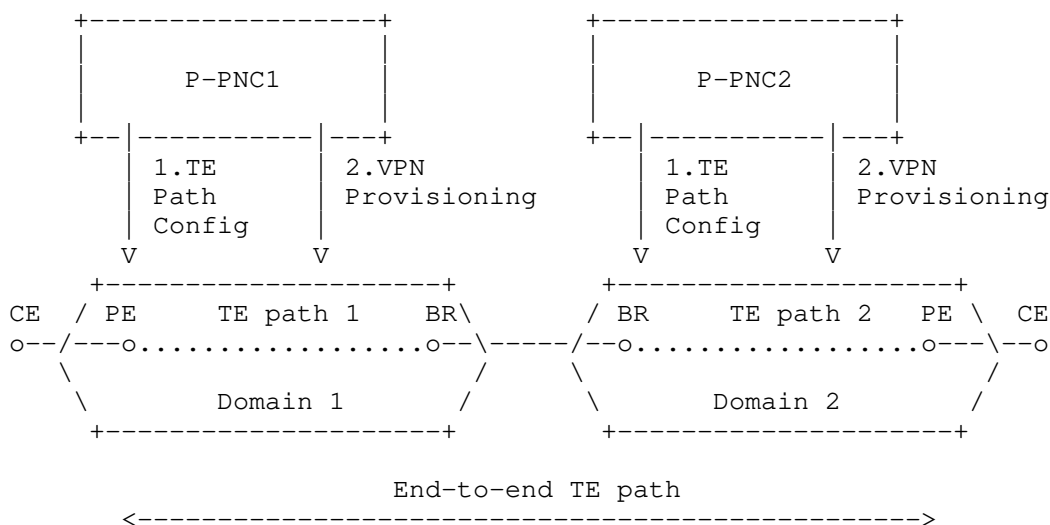


Figure 4 Domain Controller & NE Functions

When requesting the setup of a new TE path, the MDSC provides the P-PNCs with the explicit path to be created or modified. In other words, the MDSC can communicate to the P-PNCs the complete list of nodes involved in the path (strict mode). In this case, the P-PNC is just responsible to set up that explicit TE path. For example:

- o with SR-TE, the P-PNC pushes to headend PE or BR the list of SIDs to create the explicit SR-TE path, provided by the MDSC;
- o with RSVP-TE, the P-PNC requests the headend PE or BR to start signaling the explicit RSVP-TE path, provided by the MDSC.

To scale in large SR-TE packet domains, the MDSC can provide P-PNC a loose path, together with per-domain TE constraints. The P-PNC can then select the complete path within its domain.

In such a case, it is mandatory that P-PNC signals back to the MDSC which path it has chosen so that the MDSC keeps track of the relevant resources utilization.

From the Figure 3 example, the TE path requested by the MDSC touches PE13 - P16 - BR12 - BR21 - PE23. P-PNC2 is aware of two paths with the same topology metric, e.g. BR21 - P24 - PE23 and BR21 - BR22 -

PE23, but with different loads. It may prefer to steer the traffic on the latter because it is less loaded.

For the purposes of this document it is assumed that the MDSC always provides the explicit list of all the hops to the P-PNCs to setup or modify the TE path.

2.3. Optical Domain Controller and NE Functions

The optical network provides 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:

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

3. Interface Protocols and YANG Data Models for the MPIs

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

3.1. 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.

3.2. 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].

This section describes the YANG data models that are applicable to the Packet and Optical MPIs. Some of these YANG data models can be optional depending on the specific network configuration detailed in section 4 and section 5.

3.2.1. 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 can use the following common topology YANG data models at the MPI:

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

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

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

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

3.2.2. YANG models at the Optical MPIs

The Optical PNC can use the following technology-specific topology YANG data models, which augment the generic TE Topology Model:

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

The optical PNC can use the following technology-specific tunnel YANG data models, which augments the generic TE Tunnel Model:

- o The WDM Tunnel Model, defined in the "ietf-wdm-tunnel" YANG module of [WDM-TUNNEL];
- o the OTN Tunnel Model, defined in the "ietf-otn-tunnel" YANG module of [OTN-TUNNEL].

The optical PNC can 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 can use the following client signal YANG data models:

- o the CBR Client Signal Model, defined in the "ietf-trans-client-service" YANG module of [CLIENT-SIGNAL];
- o the Ethernet Client Signal Model, defined in the "ietf-eth-tran-service" YANG module of [CLIENT-SIGNAL].

3.2.3. YANG data models at the Packet MPIS

The Packet PNC can use the following technology-specific topology YANG data models:

- o The L3 Topology Model, defined in the "ietf-l3-unicast-topology" YANG module of [RFC8346], which augments the Base Network Topology Model;
- o the packet TE Topology Model, defined in the "ietf-te-topology-packet" YANG module of [L3-TE-TOPO], which augments the generic TE Topology Model;
- o The MPLS-TE Topology Model, defined in the "ietf-te-mpls-topology" YANG module of [MPLS-TE-TOPO], which augments the TE Packet Topology Model with or without the L3 TE Topology Model, defined in "ietf-l3-te-topology" YANG module of [L3-TE-TOPO];
- o the SR Topology Model, defined in the "ietf-sr-mpls-topology" YANG module of [SR-TE-TOPO].

The Packet PNC can use the following technology-specific tunnel YANG data models, which augments the generic TE Tunnel Model:

- o The MPLS-TE Tunnel Model, defined in the "ietf-te-mpls" YANG modules of [MPLS-TE-TUNNEL];
- o the SR-TE Tunnel Model which is to be defined as described in section 6.

The packet PNC can use the following network service YANG data models:

- o L3VPN Network Model (L3NM), defined in the "ietf-l3vpn-ntw" YANG module of [RFC9182];
- o L3NM TE Service Mapping, defined in the "ietf-l3nm-te-service-mapping" YANG module of [TSM];

- o L2VPN Network Model (L2NM), defined in the "ietf-l2vpn-ntw" YANG module of [RFC9291];
- o L2NM TE Service Mapping, defined in the "ietf-l2nm-te-service-mapping" YANG module of [TSM].

3.3. Path Computation Element Protocol (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 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, 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 the 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 perform full multi-layer path computation or delegate path computation to the underneath PNCs.

This approach is desirable for operators from a multi-vendor integration perspective as it is simple. We need only one type of interface (RESTCONF) and use the relevant YANG data models depending on the operator use case considered. The benefits of having only one protocol for the MPI between MDSC and PNC have already been highlighted in [PATH-COMPUTE].

4. 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 RESTCONF/YANG interface is added to an existing PCEP deployment (brownfield scenario).

Section 4 and section 5 of this draft analyse the case where a single RESTCONF/YANG interface is deployed at the MPI (i.e., option 1 above).

4. Inventory, Service and Network Topology Discovery

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

- o 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;
- o the optical tunnels supporting multi-layer intra-domain IP links;
- o both intra-domain and inter-domain L2/L3 VPN network services deployed within the network;
- o the TE paths supporting those L2/L3 VPN network services;
- o 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 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.

Reporting the entire inventory and detailed topology information of packet and optical networks to the MDSC may present scalability issues as a potential drawback. The analysis of the scalability of

this approach and mechanisms to address potential issues is outside the scope of this document.

Each PNC provides the MDSC the topology view of the domain it controls, as described in section 4.1 and 4.3. The MDSC uses this information to discover the complete topology view of the multi-layer multi-domain networks it controls.

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

It should also be possible to correlate information from IP and optical layers (e.g., which port, lambda/OTSi, and direction are 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 correlate optical and IP layers alarms to speed-up troubleshooting easily.

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.

4.1. Optical Topology Discovery

The WSON Topology Model and the Flexi-grid Topology model can be 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 or a mix of fixed-grid and flexible-grid.

It is worth noting that, as described in Appendix I of [ITU-T_G.694.1], a fixed-grid can also be described as a flexible grid with constraints: for example a 50GHz fixed-grid can be described as a flexible-grid which supports only $m=4$ and values of n which are only multiplier of 8.

As a consequence:

- o A flexible-grid DWDM network topology can only be reported using the Flexi-grid Topology model;
- o A fixed-grid DWDM network topology, can be reported using either the WSON Topology model or the Flexi-grid Topology model;
- o A mixed fixed and flexible grid DWDM network topology can be reported using either the Flexi-grid Topology model or both WSON and Flexi-grid topology models.

Clarifying how both WSON and Flexi-grid topology models could be used together (e.g., through multi-inheritance as described in [TE-TOPO-PF]) has been identified as a gap.

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.

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:

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

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

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

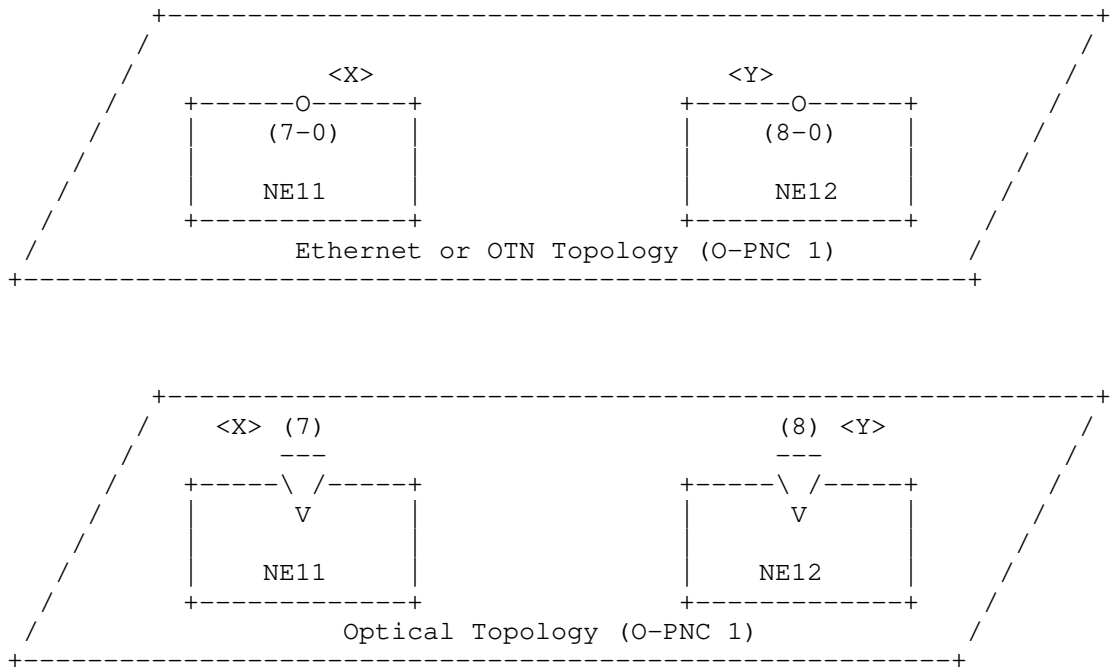
The OTN Topology Model also reports the CBR client LTPs that terminates the cross-layer links: once CBR client LTP is reported for each CBR or multi-function client interface on the optical NEs (see sections 4.4 and 5.1 of [TNBI] for the description of multi-function client interfaces).

The Ethernet Topology Model reports the Ethernet client LTPs that terminate the cross-layer links: one Ethernet client LTP is reported for each Ethernet or multi-function client interface on the optical NEs.

The optical transponders and, optionally, the OTN access cards, are abstracted at MPI by the O-PNC as Trail Termination Points (TTPs), defined in [RFC8795], within the optical network topology. This 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 in both cases the optical transponders and the WDM node are under the control of the same O-PNC.

The association between the Ethernet or CBR client LTPs terminating the Ethernet cross-layer links and the optical TTPs is reported using the Inter Layer Lock-id (ILL) identifiers, defined in [RFC8795].

For example, with a reference to Figure 5, the ILL values X and Y are used to associate the client LTPs (7-0) in NE11 and (8-0) in NE12 with the corresponding optical TTPs (7) in NE11 and (8) in NE12, respectively.



Legenda:

=====
O LTP

\ / TTP
V

< > Inter-Layer Lock-id reported by the PNC

Figure 5 - Multi-layer optical topology discovery

The intra-domain optical links 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 WDM tunnels: this relationship is reported by the mechanisms described in section 4.2.

4.2. Optical Path Discovery

The WDM Tunnel Model is used to report all the WDM 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 and the Transparent CBR client signal model are used to report all the connectivity services provided by the underlay optical tunnels between Ethernet or CBR client LTPs, depending on whether the connectivity service is frame-based or transparent. The underlay optical tunnels can be either WDM tunnels or, when the optional OTN switching layer is deployed, OTN tunnels.

The WDM tunnels can be used to support either Ethernet or CBR client signals or multi-layer intra-domain OTN links. In the latter case, the hierarchical-link container, defined in [TE-TUNNEL], associates the underlay WDM tunnel with the supported multi-layer intra-domain OTN link and it allows discovery of the multi-layer path supporting all the connectivity services provided by the optical network.

The O-PNCs report in their operational datastores all the Ethernet and CBR client connectivities and all the optical tunnels deployed within their optical domain regardless of the mechanisms being used to set them up, such as the mechanisms described in section 5.2, as well as other mechanism (e.g., static configuration), which are outside the scope of this document.

4.3. Packet Topology Discovery

The L3 Topology Model is used report the IP network topology.

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].

The TE Topology Model, TE Packet Topology Model and MPLS-TE Topology Model are used together to report the MPLS-TE network topology, as described in [MPLS-TE-TOPO].

As described in [L3-TE-TOPO], the relationship between the IP network topology and the MPLS-TE network topology depends on whether the two network topologies are congruent or not: in the latter case, the L3 TE Topology Model is used, together with the L3 Topology Model to provide the association between the two network topologies.

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 TE path computation, the P-PNCs report the full packet network, including all the information that the MDSC requires to perform TE path computation. In particular, one TE node is reported for each router and one TE link is reported for each intra-domain IP link. The packet topology also reports the IP LTPs terminating the inter-domain IP links.

The Ethernet Topology Model is used to report the intra-domain Ethernet links supporting the intra-domain IP links as well as the Ethernet LTPs that might terminate cross-layer links, inter-domain Ethernet links or access links, as described in detail in section 4.5 and in section 4.6.

All the intra-domain Ethernet and IP links are discovered by the P-PNCs, using mechanisms, such as LLDP [IEEE 802.1AB], which are outside the scope of this document, and reported at the MPIs within the Ethernet or the packet network topology.

4.4. TE Path Discovery

We assume that the discovery of existing TE paths, including their bandwidth, at the MPI is done using the generic TE tunnel YANG data model, defined in [TE-TUNNEL], with packet technology-specific (e.g., MPLS-TE or SR-TE) augmentations.

Note that technology-specific augmentations of the generic path TE tunnel model for SR-TE path setup and discovery is outlined in section 1 of [TE-TUNNEL] but are currently identified as a gap in section 6.

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

For example, considering the L3VPN in Figure 2, the TE path 1 in one direction (PE13-PE16-PE14) and the TE path in the reverse direction (between PE14 and PE13) should be reported by the P-PNC1 to the MDSC as TE primary and primary-reverse paths of the same TE tunnel instance. The bandwidth of these TE paths represents the bandwidth allocated by P-PNC1 to the two 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 TE paths established within their packet domain regardless of the

mechanism being used to set them up; i.e., independently on whether the mechanisms described in section 5.3 or other means, such as static configuration, which are outside the scope of this document, are used.

4.5. Inter-domain Link Discovery

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

- o Inter-domain Ethernet links supporting inter-domain IP links between two adjacent IP domains;
- o Cross-layer links between an IP domain and an adjacent optical domain;
- o Access links between a CE device and a PE router.

All the three types of links are Ethernet 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 Ethernet link or an access link. The TE Topology Model supports the discovery for all these types of links with no need for the P-PNC to know the type of inter-domain link.

There are two possible models to report the access links between CEs and PEs: the TE Topology Model, defined in [RFC8795], or the Service Attachment Points (SAP) Model, defined in [RFC9408].

Although the discovery of access links is outside the scope of this document, clarifying the relationship between these two models has been identified as a gap.

The inter-domain Ethernet 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].

The plug-id attribute can also be used to discover the access-links, but the analysis of the access-link discovery is outside the scope of this document.

This document considers the following two options for discovering inter-domain 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 specific LLDP information reported within the plug-id value, such as the Chassis ID and Port ID mandatory TLVs, is implementation specific and needs to be consistent across all the PNCs within the network.

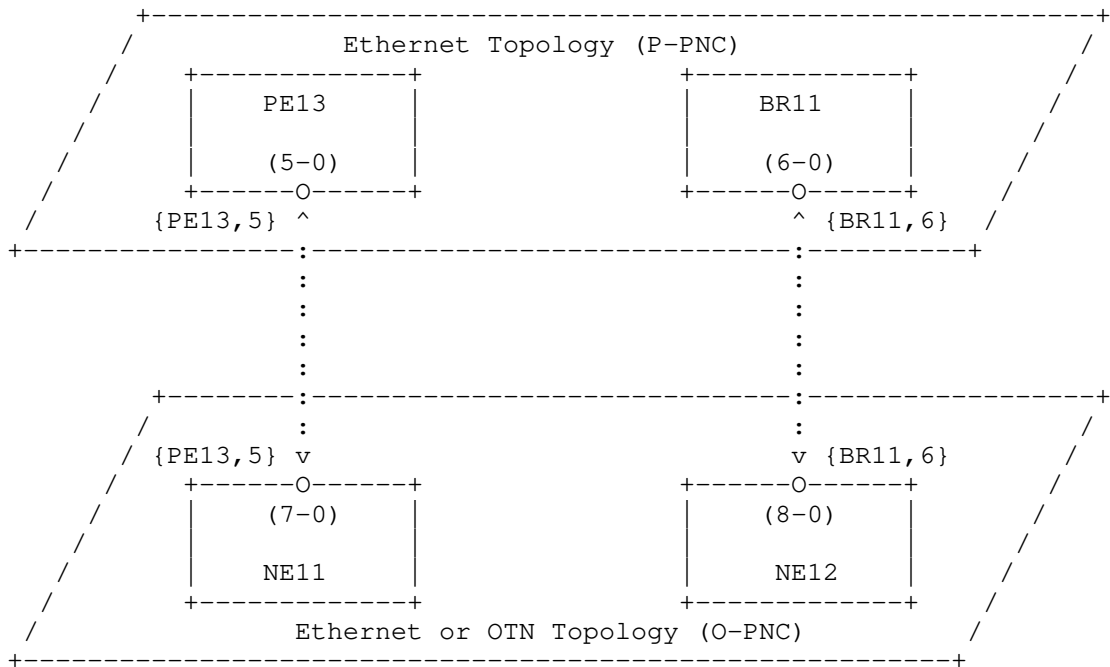
The static configuration requires an administrative burden to configure network-wide unique identifiers: it is therefore more viable for inter-domain Ethernet links. For the cross-layer links, the automatic discovery solution based on LLDP snooping is preferable when possible.

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 TLVs.

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

4.5.1. Cross-layer Link Discovery

The MDSC can discover a cross-layer link by matching the plug-id values of the two LTPs reported by two adjacent O-PNC and P-PNC: in case LLDP snooping is used, the P-PNC reports the LLDP information sent by the corresponding Ethernet interface on the router while the O-PNC reports the LLDP information received by the corresponding Ethernet interface on the optical NE, e.g., between LTP 5-0 on PE13 and LTP 7-0 on NE11, as shown in Figure 6.



Legenda:
=====

O	LTP
---	-----

<...> Link discovered by the MDSC
{ } LTP Plug-id reported by the PNC

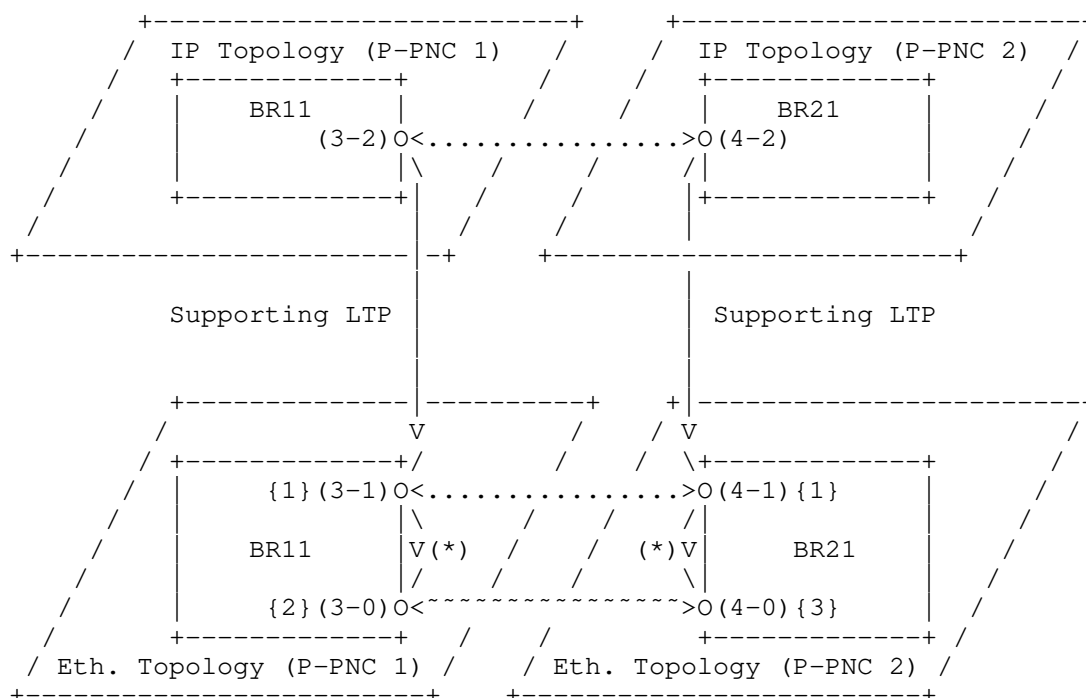
Figure 6 - Cross-layer link discovery

As described in section 4.1, the LTP terminating a cross-layer link is reported by an O-PNC in the Ethernet topology or in the OTN topology model or in both models, depending on the type of corresponding physical port on the optical NE.

It is worth noting that the discovery of cross-layer links is based only on the LLDP information sent by the Ethernet interfaces of the routers and received by the Ethernet interfaces of the optical NEs. Therefore the MDSC can discover these links also before optical paths, supporting overlay multi-layer IP links, are setup.

4.5.2. Inter-domain IP Link Discovery

The MDSC can discover an inter-domain Ethernet link which supports 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 the LLDP information being sent and being received from the corresponding Ethernet interfaces, e.g., between the Ethernet LTP 3-1 on BR11 and the Ethernet LTP 4-1 on BR21 shown in Figure 7.



Notes:

=====

(*) Supporting LTP
 {1} {BR11,3,BR21,4}
 {2} {BR11,3}
 {3} {BR21,4}

Legenda:

=====

O LTP
 ----> Supporting LTP
 <...> Link discovered by the MDSC
 <~~~> Link inferred by the MDSC
 { } LTP Plug-id reported by the PNC

Figure 7 - Inter-domain Ethernet and IP link discovery

Different information is required to be encoded by the P-PNC within the plug-id attribute of the Etherent LTPs to discover cross-layer links and inter-domain Ethernet links.

If the P-PNC does not know a priori whether an Ethernet interface on a router terminates a cross-layer link or an inter-domain Ethernet link, it has to report at the MPI two Ethernet LTPs representing the same Ethernet interface, e.g., both the Ethernet LTP 3-0 and the Ethernet LTP 3-1, supported by LTP 3-0, shown in Figure 7:

- o The physical Ethernet LTP (e.g., LTP 3-0 in BR11, as shown in Figure 7) is used to represent the physical adjacency between the router Ethernet interface and either the adjacent router Ethernet interface (in case of a single-layer Ethernet link) or the optical NE Ethernet interface (in case of a multi-layer Ethernet link). Therefore, as described in section 4.5.1, the P-PNC reports, within the plug-id attribute of this LTP, the LLDP information sent by the corresponding router Ethernet interface; such as the {BR11,3} and {BR21,4} plug-id values reported, respectively, by the Ethernet LTP 3-0 on BR11 and by the Ethernet LTP 4-0 on BR21, as shown in Figure 7;
- o The logical Ethernet LTP (e.g., LTP 3-1 in BR11, as shown in Figure 7), supported by a physical Ethernet LTP (e.g., LTP 3-0 in BR11, as shown in Figure 7), is used to discover the logical adjacency between router Ethernet interfaces, which can be either single-layer or multi-layer. Therefore, the P-PNC reports, within the plug-id attribute of this LTP, the LLDP information sent and received by the corresponding router Ethernet interface; such as the {BR11,3, BR21,4} plug-id values reported by the Ethernet LTP 3-1 on BR11 and by the Ethernet LTP 4-1 on BR21, as shown in Figure 7.

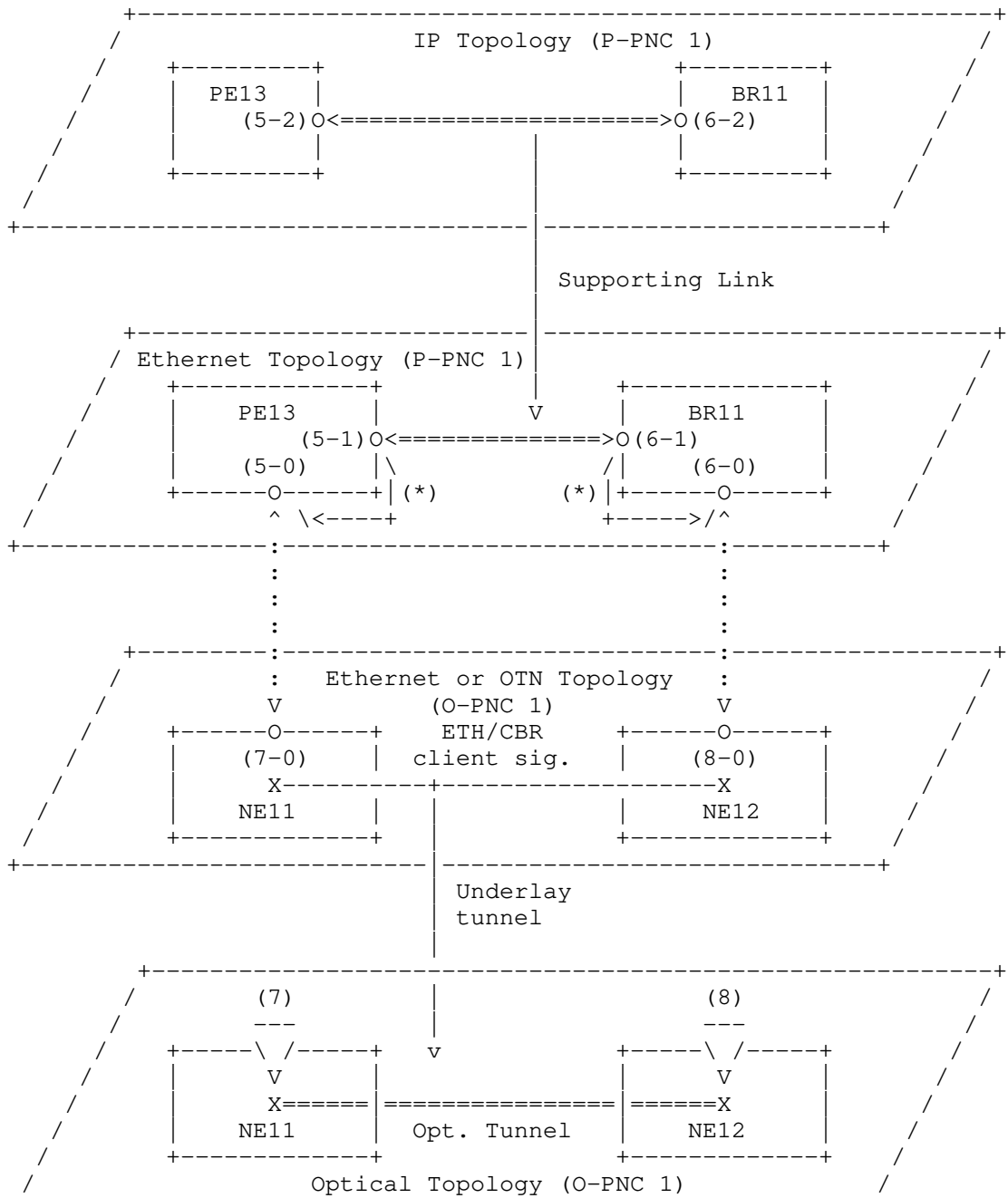
It is worth noting that in case of an inter-domain single-layer Ethernet links, the MDSC cannot discover, using the the LLDP information reported in the plug-id attributes, the physical adjacency between the two router Ethernet interfaces because these two plug-id values do not match, such as the plug-id values {BR11,3} and {BR21,4} shown in Figure 7. However, the MDSC may infer the physical intra-domain Ethernet links if it knows a priori, using mechanisms which are outside the scope of this document, that the Ethernet interfaces on the routers either terminates a cross-layer link or a single-layer (intra-domain or inter-domain) Ethernet link, e.g., as shown in Figure 7.

The P-PNC can omit reporting the physical Ethernet LTPs when it knows, by mechanisms which are outside the scope of this document, that the corresponding router Ethernet interfaces terminate single-layer inter-domain Ethernet links.

The MDSC can then discover an inter-domain IP link between the two IP LTPs that are supported by the two Ethernet LTPs terminating an inter-domain Ethernet link, discovered as described in section 4.5.2, e.g., between the IP LTP 3-2 on BR21 and the IP LTP 4-2 on BR22, supported respectively by the Ethernet LTP 3-1 on BR11 and by the Ethernet LTP 4-1 on BR21, as shown in Figure 7.

4.6. Multi-layer IP Link Discovery

A multi-layer intra-domain IP link and its supporting multi-layer intra-domain Ethernet link are discovered by the P-PNC like any other intra-domain IP and Ethernet links, as described in section 4.3, and reported at the MPI within the packet and the Ethernet network topologies, e.g., as shown in Figure 8.



+-----+

Notes:

=====

(*) Supporting LTP

Legenda:

=====

O LTP

\ / TTP

V

----> Supporting LTP or Supporting Link or Underlay tunnel

<==> Link discovered by the PNC and reported at the MPI

<...> Link discovered by the MDSC

x---x Ethernet/CBR client signal

X===X Optical tunnel

Figure 8 - Multi-layer intra-domain Ethernet and IP link discovery

The P-PNC does not report any plug-id information on the logical Ethernet LTPs terminating intra-domain Ethernet links, such as the LTP 5-1 on PE13 and LTP 6-1 in BR11 shown in Figure 8, since these links are discovered by the PNC.

In addition, the P-PNC also reports the physical Ethernet LTPs that terminate the cross-layer links supporting the multi-layer intra-domain Ethernet links, e.g., the Ethernet LTP 5-0 on PE13 and the Ethernet LTP 6-0 on BR11, shown in Figure 8.

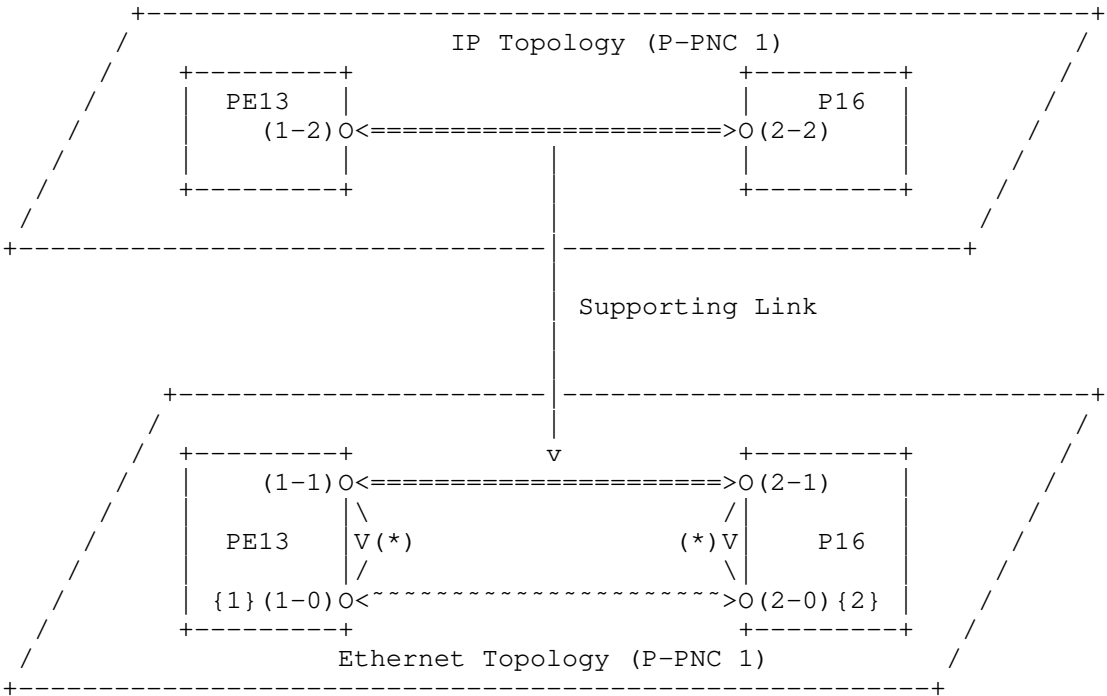
The MDSC discovers, using the mechanisms described in section 4.5, which Ethernet cross-layer links support the multi-layer intra-domain Ethernet links, e.g., the link between LTP 5-0 on PE13 and LTP 7-0 on NE11, shown in Figure 8.

The MDSC also discovers, from the information provided by the O-PNC and described in section 4.2, 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, e.g., as shown in Figure 8.

4.6.1. Single-layer Intra-domain IP Links

It is worth noting that the P-PNC may not be aware of whether an Ethernet interface on the router terminates a multi-layer or a single-layer intra-domain Ethernet link.

In this case, the P-PNC, always reports two Ethernet LTPs for each Ethernet interface on the router, e.g., the Ethernet LTP 1-0 and 1-1 on PE13, shown in Figure 9.



Notes:

=====

- (*) Supporting LTP
- {1} {PE13,1}
- {2} {P16,2}

Legenda:

=====

- O LTP
- > Supporting LTP
- <==> Link discovered by the PNC and reported at the MPI
- <~~~> Link inferred by the MDSC
- { } LTP Plug-id reported by the PNC

Figure 9 - Single-layer intra-domain Ethernet and IP link discovery

It is worth noting that in case of an intra-domain single-layer Ethernet links, the MDSC cannot discover, using the LLDP information reported in the plug-id attributes, the physical adjacency between the two router Ethernet interfaces because the two plug-id values do not match, such as the plug-id values {PE13,1} and {P16,2} shown in Figure 9. However, the MDSC may infer the physical intra-domain Ethernet links, e.g., between LTP 1-0 on PE13 and LTP 2-0 on P16, as shown in Figure 9, if it knows a priori, using mechanisms which are outside the scope of this document, that all the Ethernet interfaces on the routers either terminates a cross-layer link or a single-layer (intra-domain or inter-domain) Ethernet link, e.g., as shown in Figure 9.

The P-PNC can omit reporting the physical Ethernet LTP if it knows, by mechanisms which are outside the scope of this document, that the intra-domain Ethernet link is single-layer.

4.7. LAG Discovery

The P-PNCs can discover the configuration of the LAG groups within its domain and report each intra-domain LAG as an Ethernet bundle link, within the Ethernet topology exposed at the MPI.

This is done bundling multiple single-domain Ethernet links, as shown in Figure 10. For example, the Ethernet bundled link between the Ethernet LTP 5-1 on BR21 and the Ethernet LTP 6-1 on P24, is built from the Ethernet links setup respectively:

- o between the Ethernet LTP 1-1 on BR21 and the Ethernet LTP 2-1 on P24; and
- o between the Ethernet LTP 3-1 on BR21 and the Ethernet LTP 4-1 on P24.

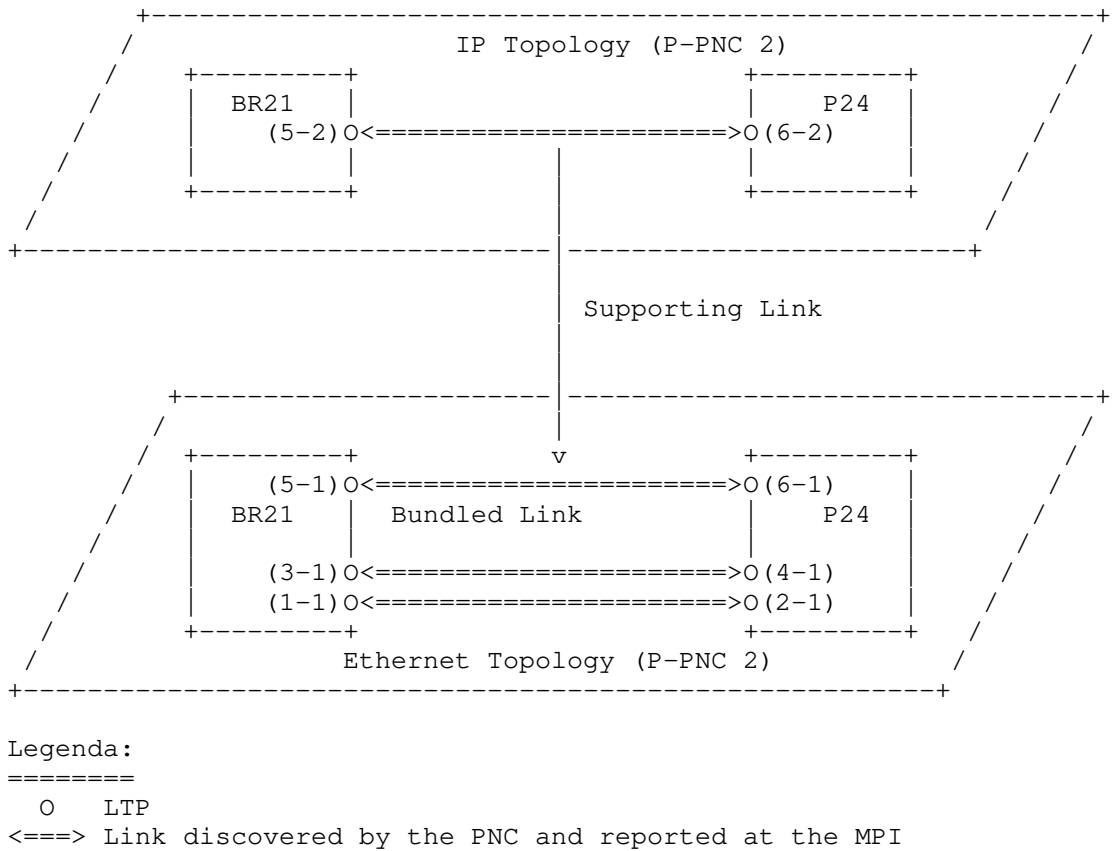


Figure 10 - LAG

The mechanisms used by the MDSC to discover single-layer and multi-layer intra-domain LAG link is the same (the only difference being whether the bundled links are single-layer or multi-layer).

Instead, the mechanisms used by the MDSC to discover single-layer inter-domain LAG links between two BRs are different and outside the scope of this document since they do not imply any cross-layer coordination between packet and optical domains.

As described in section 4.3, the mechanisms used by the P-PNC to discover the configuration of the LAG groups within its domain, such as LLDP [IEEE 802.1AB], are outside the scope of this document.

However, it is worth noting that according to [IEEE 802.1AB], LLDP can be configured on a LAG group (Aggregated Port) and/or on any number of its LAG members (Aggregation Ports).

If LLDP is enabled on both LAG members and groups, two types of LLDP packets are transmitted by the routers and received by the optical NEs on some cross-layer links: one sent for the LLDP session configured at LAG member (Aggregation Port) level and another one for the LLDP session configured at LAG group (Aggregated Port) level. This could cause some issues when LLDP snooping is used to discover the cross-layer links, as defined in section 4.5.1.

The cross-layer link discovery is based only on the LLDP session configured on the LAG members (Aggregation Ports) to allow discovery of these links independently from the configuration of the underlay optical tunnel or from the LAG group.

To avoid any ambiguity on how the optical NEs can identify which LLDP packets belong to which LLDP session, the P-PNC can disable the LLDP sessions on the LAG groups configured by the MDSC (e.g., the multi-layer single-domain LAG groups configured using the mechanisms described in section 5.2.1), keeping the LLDP sessions on the LAG members enabled.

Another option is to rely on other mechanisms (e.g., the Port type field in the Link Aggregation TLV defined in Annex F of [IEEE 802.1AX]) that allow the optical NE to identify which LLDP packets belong to which LLDP session: the O-PNC can then use only the LLDP information from the LLDP sessions configured on the LAG members to support the cross-layer link discovery mechanisms defined in section 4.5.1.

4.8. L2/L3 VPN Network Services Discovery

The P-PNC reports the L2/L3 VPN services configured within its domain, using the L2NM and L3NM network service models, and which packet TE tunnels (e.g., MPLS-TE or SR-TE) are used by each L2/L3 VPN service, using the L2NM and L3NM TE service mapping models.

The MDSC can use the information mentioned above together with the packet TE path, packet topology, multi-layer IP links, optical topology and optical path information discovered as described in the previous sections, to discover the multi-layer path used to carry the traffic for each L2/L3 VPN service.

4.9. Inventory Discovery

There 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.

5. Establishment of L2/L3 VPN 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 TE paths that meet these TE requirements (see section 2.1.1).

For example, considering the L3VPN in Figure 2 and Figure 3, the MDSC finds that:

- o PE13-P16-PE14 TE path already exists but have not enough bandwidth to support the new L3VPN, as described in section 4.4;, and that:
 - o the IP link(s) between PE13 and P16 has not enough bandwidth to support increasing the bandwidth of that TE path, as described in section 4.3;

- o a new underlay optical tunnel could be setup to increase the bandwidth of the IP link(s) between PE13 and P16 to support increasing the bandwidth of that overlay TE path, as described in section 5.1. The dimensioning of the underlay optical tunnel is decided by the MDSC based on the TE requirements (e.g., the bandwidth) requested by the TE path and on its multi-layer optimization policy, which is an internal MDSC implementation issue;
- o a new multi-domain TE path needs to be setup between PE13 and PE23, e.g., either because existing TE paths between PE13 and PE23 are not able to meet the TE and binding requirements of the L2/L3 VPN service or because there is no TE path 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 5.1, 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:

- o a new underlay optical tunnel between PE13 and BR11, supporting a new IP link, as described in section 5.2;
- o 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 5.2.

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:

- o the O-PNC1 to setup a new optical tunnel between the ROADMs connected to PE13 and P16, as described in section 5.2;
- o 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 PE13 and P16, as described in section 5.2;

- o the P-PNC1 to update the bandwidth of the selected TE path between PE13 and PE14, as described in section 5.3.

After that, the MDSC requests P-PNC2 to setup a TE path between BR21 and PE23, with an explicit path (BR21, P24, PE23) to constrain this new TE path to use the new underlay optical tunnel setup between BR21 and P24, as described in section 5.3. The P-PNC2 properly configures the routers within its domain to setup the requested path and returns to the MDSC the information which is needed for multi-domain TE path stitching. For example, in case of inter-domain SR-TE, 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 binding SID it has assigned to this policy in BR21.

Then the MDSC requests P-PNC1 to setup a TE path between PE13 and BR11, with an explicit path (PE13, BR11) to constrain this new TE path to use the new underlay optical tunnel setup between PE13 and BR11, specifying also which inter-domain link should be used to send traffic to BR21 and the information to be used for the multi-domain TE path stitching, as described in section 4.4 (e.g., in case of inter-domain SR-TE, the binding SID that has been assigned by P-PNC2 to the corresponding SR policy in BR21). The P-PNC1 properly configures the routers within its domain to setup the requested path and the multi-domain TE path stitching. For example, in case of inter-domain SR-TE, the P-PNC1, knowing also the node and the adjacency SIDs assigned within its domain and the EPE SID assigned by P-PNC1 to the inter-domain link between BR11 and BR21, and the binding SID assigned by P-PNC2, installs the proper policy, or policies, within PE13.

Once the 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 TE paths, as described in section 5.4.

5.1. 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 described in sections 4.1, 4.5 and 4.6, there is a one-to-one relationship between a multi-layer intra-domain IP link and its underlay optical tunnel. 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 associated multi-layer single-domain IP link.

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.

There 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.

5.2. Multi-layer IP Link Setup

As described in section 5.1, there is a one-to-one relationship between a multi-layer intra-domain IP link and its underlay optical tunnel.

Therefore, to setup a new multi-layer intra-domain IP link, the MDSC requires the O-PNC to setup the optical tunnel (using either the WDM Tunnel model or the OTN Tunnel model, if the optional OTN switching is supported) within the optical network and to steer the client traffic between the two cross-layer links over that optical tunnel, using either the Ethernet Client Signal Model (for frame-based transport) or the Transparent CBR Client Signal Model (for transparent transport).

For example, with a reference to Figure 11, the MDSC can request the O-PNC1 to setup an optical tunnel between the optical TTPs (7) on NE11 and (8) on NE12 and to steer over this tunnel the client traffic between LTP (7-0) on NE11 and LTP (8-0) on NE12.

Editors Note: Add a new Figure 11 which is an exact copy of Figure 8

Figure 11 - Multi-layer IP link setup

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

If LLDP [IEEE 802.1AB] or any other discovery mechanisms, which are 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 and report it to the P-PNC, as described in section 4.6.

Otherwise, if there are no automatic discovery mechanisms, the MDSC can configure 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, e.g., the LTP 5-1 on PE13 and 6-1 on BR11 shown in Figure 11.

After the multi-layer single-domain Ethernet link has been configured by the MDSC or discovered by the P-PNC, the corresponding multi-layer single-domain IP link can also be configured either by the MDSC or by the P-PNC.

This document assumes that this IP link is configured by the P-PNC.

It is worth noting that if LAG is not supported within the domain controlled by the P-PNC, the P-PNC can configure the multi-layer single-domain IP link as soon as the underlay multi-layer single-domain Ethernet link is either discovered by the P-PNC or configured by the MDSC at the MPI. However, if LAG is supported the P-PNC has not enough information to know whether the discovered/configured multi-layer single-domain Ethernet link would be:

1. Used to support a multi-layer single-domain IP link;
2. Used to create a new LAG group;
3. Added to an existing LAG group.

Therefore the P-PNC does not take any further action after a multi-layer single-domain Ethernet link is discovered or configured by the MDSC at the MPI.

The MDSC can request the P-PNC to configure a new multi-layer single-domain IP link, supported by the the just discovered or configured multi-layer single-domain Ethernet link, by creating an IP link within the running datastore of the P-PNC MPI. Only the IP link, IP LTPs and the reference to the supporting multi-layer single-domain

Ethernet link are configured by the MDSC. All the other configuration is provided by the P-PNC.

For example, with a reference to Figure 11, the MDSC can request the P-PNC1 to setup a multi-layer single-domain IP Link between IP LTP 5-2 on PE13 and IP LTP 6-2 on BR11 supported by the multi-layer single-domain Ethernet link between ETH LTP 5-1 on PE13 and ETH LTP 6-1 on BR11.

The P-PNC configures the requested multi-layer single-domain IP link and, once finished, reports it to the MDSC within the IP topology exposed at its MPI.

5.2.1. Multi-layer LAG Setup

The P-PNC configures a new LAG group between two routers when the MDSC creates at the MPI a new Ethernet bundled link (using the bundled-link container defined in [RFC8795]) bundling the multi-layer single-domain Ethernet link(s) being created, as described above.

When a new LAG link is created, it is also recommended to configure the minimum number of active member links required to consider the LAG link as being up. For example, a LAG link with three members can be considered up when only one member link fails and down when at least two member links fail.

The attribute required to configure the minimum number of active member links is missing in [CLIENT-TOPO] and this is identified as a gap in section 6.

It is worth noting that a new LAG group can be created to bundle one or more multi-layer single-domain Ethernet link(s).

For example, with a reference to Figure 10, the MDSC can request the P-PNC2 to setup an Ethernet bundled link between the Ethernet LTP 5-1 on BR21 and the Ethernet LTP 6-1 on P24, bundling the multi-layer single-domain Ethernet link between the Ethernet LTP 1-1 on BR21 and the Ethernet LTP 2-1 on P24.

It is worth noting that the MDSC needs to create also the Ethernet LTPs terminating the Ethernet bundled link.

The MDSC can request the P-PNC to configure a new multi-layer single-domain IP link, supported by the just configured Ethernet bundled link, following the same procedure described in section 5.2 above.

For example, with a reference to Figure 10, the MDSC can request the P-PNC2 to setup a multi-layer single-domain IP Link between IP LTP 5-2 on BR21 and IP LTP 6-2 on P24 supported by the Ethernet bundle link between ETH LTP 5-1 on BR21 and the Ethernet LTP 6-1 on P24.

5.2.2. Multi-layer LAG Update

The P-PNC adds new member(s) to an existing LAG group when the MDSC updates at the MPI the configuration of an existing Ethernet bundled link adding the multi-layer single-domain Ethernet link(s) being created, as described above.

When member links are added or removed from a LAG link, the minimum number of active member links required to consider the LAG link as being up may also need to be updated.

For example, with a reference to Figure 10, the MDSC can request the P-PNC2 to add the multi-layer single-domain Ethernet link setup between the Ethernet LTP 3-1 on BR21 and the Ethernet LTP 4-1 on P24 to the existing Ethernet bundle link setup between the Ethernet LTP 5-1 on node BR21 and the Ethernet LTP 6-1 on node P24.

After the LAG configuration has been updated, the P-PNC can also update the bandwidth information of the multi-layer single-domain IP link supported by the updated Ethernet bundled link.

5.2.3. Multi-layer TE path properties Configuration

The MDSC can discover the TE path properties (e.g., the list of SRLGs, the delay) of a multi-layer IP link from the TE properties of:

- o the IP LTPs terminating the multi-layer IP link (e.g., the list of SRLGs reported by the P-PNC using the packet TE topology model);
- o the optical path (e.g., the list of SRLGs reported by the O-PNC using the WDM or OTN tunnel model); and
- o the cross-domain links (e.g., the list of SRLGs reported by the O-PNC and P-PNC respectively, using the WSON and/or flexi-grid, the OTN and the packet TE topology models).

The MDSC can also report this information to the P-PNC by properly configuring the multi-layer IP link properties using the packet TE topology model at the packet PNC MPI.

This information is used by the P-PNC at least when computing the local protection path, as described in section 5.3, e.g., to ensure that the local protection path is SRLG disjoint with the primary path.

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.

5.3. TE Path Setup and Update

This version of the draft assumes that TE path setup and update at the MPI could be done using the generic TE tunnel YANG data model, defined in [TE-TUNNEL], with packet technology-specific augmentations, described in section 3.2.3.

When a new TE path needs to be setup, the MDSC can use the [TE-TUNNEL] model to request the P-PNC to set it up, properly specifying the path constraints, such as the explicit path, to force the P-PNC to setup an TE path that meets the end-to-end TE and binding constraints and uses the optical tunnels setup by the MDSC for the purpose of supporting this new TE path.

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

In the latter case, the technology-specific augmentations should allow the configuration of the information needed for multi-domain TE path stitching.

For example, the 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 multi-domain SR-TE path stitching and to allow the P-PNC to report the binding SID assigned to the segment TE paths. Note that 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.

[Editor's Note:] Add some text about the protection options (to further discuss whether to put this text here or in section 4.2.2).

The MDSC also request the P-PNC to configure local protection mechanisms. For example, in case of SR-TE domain, the TI-LFA local protection, as defined in [TI-LFA]: 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 requested local protection mechanisms within the P-PNC domain are 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 TE path properties (e.g., SRLG information), configured by the MDSC as described in section 5.2.3, when computing the protection configuration (e.g., in case of SR-TE domains, 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.

5.4. L2/L3 VPN Network Service Setup

The MDSC can use the L2NM and L3NM network service models to request the P-PNCs to setup L2/L3 VPN services and the L2NM and L3NM TE service mapping models to request the P-PNCs to configure the PE routers to steer the L2/L3 VPN traffic to the selected TE tunnels (e.g., MPLS-TE or SR-TE).

It is worth noting that the L2NM and L3NM TE service mapping models, defined in [TSM], provide a list of TE tunnel(s) that should be used to forward L2/L3 VPN traffic between the two PEs terminating the listed TE tunnel(s). If the list contains more than one TE tunnel for the same pair of PEs, these TE tunnels are used for load balancing the associated L2/L3 VPN traffic between the same set of two PEs.

The possibility to request splitting the traffic, between multiple TE tunnels for the same PEs pair, in a different way than load balancing is identified as a gap requiring further work and outside of the scope of this draft.

6. 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:

- o how both WSON and Flexi-grid topology models could be used together (through multi-inheritance): this gap has been identified in section 4.1;.
- o network inventory model: this gap has been identified in section 4.9 and the solution in [NETWORK-INVENTORY] has been proposed to resolve it;
- o technology-specific augmentations of the path computation RPC, defined in [PATH-COMPUTE] for optical networks: this gap has been identified in section 5.1 and the solution in [OPTICAL-PATH-COMPUTE] has been proposed to resolve it;
- o 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 [RFC9408]: this gap has been identified in section 4.3;
- o 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 5.3;
- o an attribute, which is used to configure the minimum number of active member links required to consider the LAG link as being up, is missing from the topology model defined in [CLIENT-TOPO]: this gap has been identified in section 5.2.1;
- o a mechanism to configure splitting the L2/L3 VPN traffic, between multiple TE tunnels for the same PEs pair, in a different way than load balancing: this gap has been identified in section 5.4;
- o a mechanism to report client connectivity constraints imposed by some muxponder design: this gap has been identified in appendix A.3.

Although not applicable to this document, it has been noted that being able to use WSON and Flexi-grid topology models together (through multi-inheritance) is not only useful in cases of mixed fixed-grid and flexible-grid DWDM network topology but also the only viable option in case of a mixed CWDM and DWDM network topology.

Although not applicable to this document, it has been noted that the WDM tunnel model would support also optical tunnel setup in case of a mixed CWDM and DWDM network topology.

Although not analysed in this document, it has been noted that the TE Tunnel model, defined in [TE-TUNNEL], needs also to be enhanced to support scenarios where multiple parallel TE paths are used in load-balancing to carry the traffic between two end-points (e.g., VPN traffic between two PEs).

7. Security Considerations

This document highlights how the ACTN architecture can deploy packet over optical infrastructure services. It highlights how existing IETF protocols and data models may be used for multi-layer services. It reuses several existing IETF protocols and data models for the MPI interfaces between each PNC (Optical or Packet) and the MDSC, including:

- o RESTCONF
- o NETCONF
- o PCEP
- o YANG

Several existing authentication and encryption practices and techniques may be used to help secure these MPI interfaces. These mechanisms include using Transport Layer Security (TLS) to provide secure transport for RESTCONF, NETCONF and PCEP. Furthermore, access control techniques can also provide additional security. NETCONF supports an Access Control Model (NACM), and RESTCONF supports Role Based Access Control (RBAC), which should also ensure that MDSC to PNC communication is based on authorised use and granular control of connectivity and resource requests.

8. 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.

9. IANA Considerations

This document requires no IANA actions.

10. References

10.1. 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 (WSOs)", RFC 9094, August 2021.
- [ITU-T_G.694.1] International Telecommunication Union, "Spectral grids for WDM applications: DWDM frequency grid", ITU-T Recommendation G.694.1, February 2012.
- [IEEE 802.1AB] IEEE 802.1AB-2016, "IEEE Standard for Local and metropolitan area networks - Station and Media Access Control Connectivity Discovery", March 2016.
- [IEEE 802.1AX] IEEE 802.1AB-2014, "IEEE Standard for Local and metropolitan area networks - Link Aggregation", December 2014.
- [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.

- [MPLS-TE-TOPO] Busi, I. et al., "A YANG Data Model for MPLS-TE Topology", draft-busizheng-teas-yang-te-mpls-topology, 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.
- [WDM-TUNNEL] Guo, A. et al., "A Yang Data Model for WDM Tunnels", draft-ietf-ccamp-wdm-tunnel-yang, work in progress.
- [OTN-TUNNEL] Zheng, H. et al., "OTN Tunnel YANG Model", draft-ietf-ccamp-otn-tunnel-model, work in progress.
- [MPLS-TE-TUNNEL] Saad, T. et al., "A YANG Data Model for MPLS Traffic Engineering Tunnels", draft-ietf-teas-yang-te-mpls, 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.

10.2. 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.
- [RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, October 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.
- [RFC9182] S. Barguil, et al., "A YANG Network Data Model for Layer 3 VPNs", RFC 9182, February 2022.
- [RFC9291] Boucadair, M. et al., "A YANG Network Data Model for Layer 2 VPNs", RFC 9291, September 2022.
- [RFC9408] Boucadair, M. et al., "A YANG Network Data Model for Service Attachment Points (SAPs)", RFC 9408, June 2023.
- [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.
- [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.
- [TI-LFA] Litkowski, S. et al., "Topology Independent Fast Reroute using Segment Routing", draft-ietf-rtgwg-segment-routing-ti-lfa, work in progress.
- [TE-TOPO-PF] Busi I. et al., " Profiles for Traffic Engineering (TE) Topology Data Model and Applicability to non-TE Use Cases", draft-busi-teas-te-topology-profiles, work in progress

Appendix A. Additional Scenarios

A.1. OSS/Orchestration Layer

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

- o to abstract (through MDSC and PNCs) the underlying transport network complexity to the Business Systems Support layer;
- o to coordinate NFV, Transport (e.g. IP, optical and microwave networks), Fixed Access, Core and Radio domains enabling full automation of end-to-end services to the end customers;
- o 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.

A.1.1. 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 12 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], [RFC9291], [RFC9182] and [TSM].

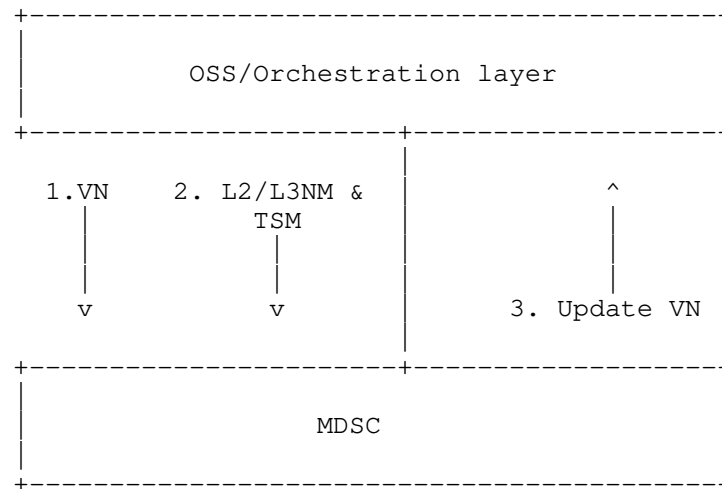


Figure 12 Service Request Process

- o 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.
 - o 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).
 - o The model also provides VN-level preference information (e.g., VN member diversity) and VN-level admin-status and operational-status.
- o The L2NM and L3NM YANG data models, defined in [RFC9291] and [RFC9182], 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.
- o The TE & Service Mapping YANG data model [TSM] provides TE-service mapping.

- o TE-service mapping provides the mapping between a L2/L3 VPN instance and the corresponding VN instances.
- o The TE-service mapping also provides the binding requirements 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).
- o Site mapping provides the site reference information across L2/L3 VPN Site ID, VN Access Point ID, and the LTP of the access link.

A.2. Multi-layer and Multi-domain Resiliency

A.2.1. 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.

A.2.2. 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:

- a) 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.

- b) 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.

A.3. Muxponders

The setup of a client connectivity service between two transponders is relatively clear and its implementation simple.

There is a one to one relationship between the transponder's client and trunk (or DWDM) port. The client port bitrate determines the trunk port bit rate which will also determine the Baud-rate, the modulation format, the FEC etc.

The controller, when asked to set up a client connectivity service, needs to find a WDM tunnel suitable to comply the DWDM port parameters.

The setup of a client connectivity service between two muxponders is different since there is a one to many relationship between the muxponder's trunk (or DWDM) port and client ports. For example, there might be a 100Gb/s trunk port shared by ten 10GE client ports.

The controller, when asked to set a 10GE client connectivity service between two muxponder's client ports, needs first to check whether there is already an existing WDM tunnel between the two muxponders and then take different actions:

1. if the WDM tunnel already exists, the controller needs only to enable the 10GE client ports to establish the 10GE client connectivity service;
2. if the WDM tunnel does not exist, the controller has to first establish the WDM tunnel, finding a proper optical path matching the optical parameters of the two muxponders' trunk ports (e.g., an OTSi carrying an OTU4), and then enable the 10GE client ports to establish the 10GE client connectivity service.

Since multiple client connectivity services are sharing the same WDM tunnel, a multiplexing label shall be assigned to each client connectivity service. The multiplexing label can either be a standard label (e.g., an OTN timeslot) or a vendor-specific label. The multiplexing label can be either configurable (flexible configuration) or assigned by design to each muxponder's client port (fixed configuration). In the former case, any muxponder client port can be connected with any other client port of the peer muxponder (for example client port 1 on one muxponder can be connected with client port 5 on the peer muxponder) while in the latter case only client ports with the same port number can be connected (for example client port 2 on one muxponder can be connected only with client port 2 on the peer muxponder and not with any other client port).

In case of flexible configuration, since the two muxponders are under the control of the same O-PNC, the configuration of the multiplexing label, regardless of whether it is a standard or vendor-specific label, can be done by the O-PNC using mechanisms which are vendor-specific and outside the scope of this document. The MDSC can just request the O-PNC to setup a client connectivity service over a WDM tunnel.

In case of fixed configuration, the multiplexing label is assigned by the muxponder but the O-PNC and MDSC needs to be aware of the connectivity constraints to avoid try and fail.

It is worth noting that the current WSON and Flexi-grid topology models in [RFC9094] and [Flexi-TOPO] do not provide sufficient information to the MDSC about this connectivity constraint and this is identified as a gap.

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

Email: ggalimbe56@gmail.com

Zheng Yanlei
China Unicom

Email: zhengyanlei@chinaunicom.cn

Anton Snitser
Cisco

Email: asnizar@cisco.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

Oscar Gonzalez de Dios
Telefonica

Email: oscar.gonzalezdedios@telefonica.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
Cisco

Email: daniele.ietf@gmail.com

