TEAS Working Group   Fabio Peruzzini
Internet Draft   TIM
Intended status: Informational   Jean-Francois Bouquier
                                                   Vodafone
                                                 Italo Busi
                                                     Huawei
                                                Daniel King
                                         Old Dog Consulting
                                          Daniele Ceccarelli
                                                    Ericsson

Expires: May 2021   November 2, 2020

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

draft-ietf-teas-actn-poi-applicability-01

Status of this Memo

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,
   identifying the YANG data models being defined by the IETF to
   support this deployment architecture as well as specific scenarios
   relevant for Service Providers.

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

Table of Contents

1. Introduction

The full automation of the management and control of Service
Providers transport networks (IP/MPLS, Optical and also Microwave)
is key for achieving the new challenges coming now with 5G as well
as with the increased demand in terms of business agility and
mobility in a digital world. ACTN architecture, by abstracting the
network complexity from Optical and IP/MPLS networks towards MDSC
and then from MDSC towards OSS/BSS or Orchestration layer through
the use of standard interfaces and data models, is allowing a wide
range of transport connectivity 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 possibly Multiprotocol Label Switching
(MPLS) 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 of each other. There are
technical differences between the technologies (e.g., routers vs.
optical switches) and the corresponding network engineering and
planning methods (e.g., inter-domain peering optimization in IP vs.
dealing with physical impairments in DWDM, or very different time
scales). In addition, customers needs can be different between a
packet and an optical network, and it is not uncommon to use
different vendors in both domains. Last but not least, state-of-the-
art packet and optical networks use sophisticated but complex
technologies, and for a network engineer it may not be trivial to be

a full expert in both areas. As a result, packet and optical
networks are often operated in technical and organizational silos.

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

ACTN framework enables this complete multi-layer and multi-vendor
integration of packet and optical networks through MDSC and packet
and optical PNCs.

In this document, key scenarios for Packet Optical Integration (POI)
are described from the packet service layer perspective. The
objective is to explain the benefit and the impact for both the
packet and the optical layer, and to identify the required
coordination between both layers. Precise definitions of scenarios
can help with achieving a common understanding across different
disciplines. The focus of the scenarios are IP/MPLS networks
operated as client of optical DWDM networks. The scenarios are
ordered by increasing level of integration and complexity. For each
multi-layer scenario, the document analyzes how to use the
interfaces and data models of the ACTN architecture.

Understanding the level of standardization and the possible gaps
will help to better assess the feasibility of integration between IP
and Optical DWDM domain (and optionally OTN layer), in an end-to-end
multi-vendor service provisioning perspective.

2. Reference architecture and network scenario

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

```
                          +----------+
                          |   MDSC   |
                          +-----+----+
                                |
             +-----------+-----+------+-----------+
             |           |            |           |
         +----+----+ +----+----+  +----+----+ +----+----+
         | P-PNC 1 | | O-PNC 1 |  | O-PNC 2 | | P-PNC 2 |
         +----+----+ +----+----+  +----+----+ +----+----+
             |           |            |           |
             |           \            /           |
      +------------------+  \        /  +------------------+
   CE / PE               BR \ |     / / BR              PE \  CE
   o--/---o                 o---\-|-------|--/---o              o---\--o
     \   :                  :  / |      | \   :              :  /
      \  : PKT Domain 1  :  /  |      |  \  : PKT Domain 2  :  /
       +-:---------------:-+   |      |   +-:---------------:--+
         :               :     |      |     :               :
         :               :     |      |     :               :
       +-:---------------:------+     +-------:---------------:--+
      / :               :      \     /      :               :  \
     /   o..............o       \   /       o..............o    \
     \      Optical Domain 1     / \         Optical Domain 2     /
      \                         /   \                           /
       +----------------------+     +------------------------+
```
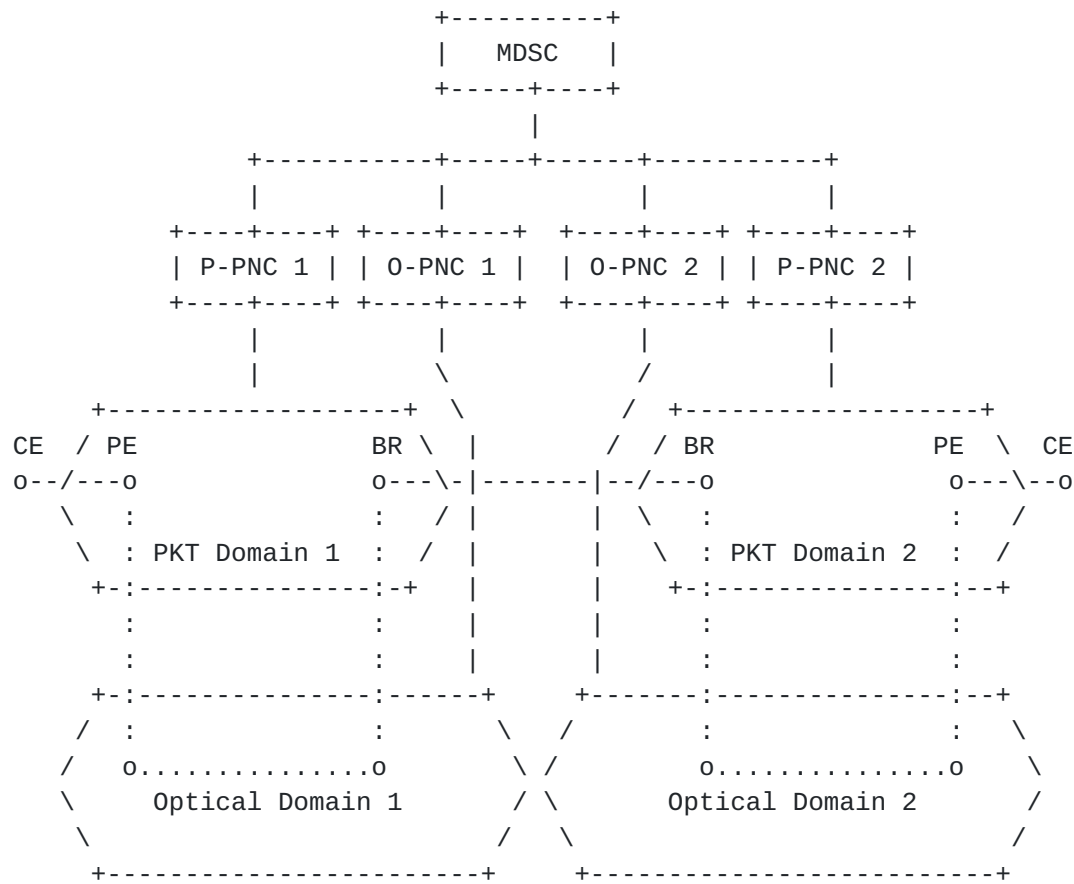
                    Figure 1 - Reference Scenario

   The ACTN architecture, defined in [RFC8453], is used to control this
   multi-domain network where each Packet PNC (P-PNC) is responsible
   for controlling its IP domain, which can be either an Autonomous
   System (AS), [RFC1930], or an IGP area within the same operator
   network, and each Optical PNC (O-PNC) is responsible for controlling
   its Optical Domain.

   The routers between IP 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 a
   ABR.

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

   The MPI interface presents an abstracted topology to MDSC hiding
   technology-specific aspects of the network and hiding topology

   details depending on the policy chosen regarding the level of
   abstraction supported. The level of abstraction can be obtained
   based on P-PNC and O-PNC configuration parameters (e.g. provide the
   potential connectivity between any PE and any BR in an MPLS-TE
   network).

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

   o The domain boundaries between the IP and Optical domains are
      congruent. In other words, one Optical domain supports
      connectivity between Routers in one and only one Packet Domain;

   o Inter-domain links exist only between Packet domains (i.e.,
      between BR routers) and between Packet and Optical domains (i.e.,
      between routers and Optical NEs). In other words, there are no
      inter-domain links between Optical domains;

   o The interfaces between the Routers and the Optical NEs are
      "Ethernet" physical interfaces;

   o The interfaces between the Border Routers (BRs) are "Ethernet"
      physical interfaces.

   This version of the document assumes that the IP Link supported by
   the Optical network are always intra-AS (PE-BR, intra-domain BR-BR,
   PE-P, BR-P, or P-P) and that the BRs are co-located and connected by
   an IP Link supported by an Ethernet physical link.

   The possibility to setup inter-AS/inter-area IP Links (e.g.,
   inter-domain BR-BR or PE-PE), supported by Optical network, is for
   further study.

   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 NEs within the optical domains can be ROADMs or OTN
   switches, with or without a ROADM.

   The MDSC in Figure 1 is responsible for multi-domain and multi-layer
   coordination across multiple Packet and Optical domains, as well as
   to provide L2/L3VPN services.

   Although the new technologies (e.g. QSFP-DD ZR 400G) are making
   convenient to fit the DWDM pluggable interfaces on the Routers, the
   deployment of those pluggable is not yet widely adopted by the
   operators. The reason is that most of operators are not yet ready to
   manage Packet and Transport networks in a unified single domain. As

a consequence, this draft is not addressing the unified scenario.
This matter will be described in a different draft.

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

1. Both the service- and network-related functions are collapsed into
   a single, monolithic implementation, dealing with the end customer
   service requests, received from the CMI (Customer MDSC Interface),
   and the adaptation to the relevant network models. Such case is
   represented in Figure 2 of [RFC8453]
2. An implementation can choose to split the service-related and the
   network-related functions in 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 H-MDSC responsible for packet-optical multi-layer
   coordination, interfacing with one Optical L-MDSC, providing
   multi-domain coordination between the O-PNCs and one Packet
   L-MDSC, providing multi-domain coordination betweeh the P-PNCs
   (see for example Figure 9 of [RFC8453]).
4. Another implementation can also choose to combine the MDSC and the
   P-PNC functions together.

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

[Editors'note:] Check for a better term to define the network
services. It may be worthwhile defining what are the customer and
network services.

The OSS/Orchestration layer is a key 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 Acess, 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.

The functionality of the OSS/Orchestration layer as well as the
interface toward the MDSC are usually operator-specific and outside
the scope of this draft. This document assumes that the
OSS/Orchestrator requests MDSC to setup L2VPN/L3VPN services through
mechanisms which are outside the scope of the draft.

There are two main cases when MDSC coordination of underlying PNCs
in POI context is initiated:

o Initiated by a request from the OSS/Orchestration layer to setup
   L2VPN/L3VPN services that requires multi-layer/multi-domain
   coordination.

o Initiated by the MDSC itself to perform multi-layer/multi-domain
   optimizations and/or maintenance works, beyond discovery (e.g.
   rerouting LSPs with their associated services when putting a
   resource, like a fibre, in maintenance mode during a maintenance
   window). Different to service fulfillment, the workflows then are
   not related at all to a service provisioning request being
   received from the OSS/Orchestration layer.

Above two MDSC workflow cases are in the scope of this draft or in
future versions.

2.1. L2/L3VPN Service Request in North Bound of MDSC

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

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

Figure 2 shows an example of a possible control flow between the
OSS/Orchestration layer and the MDSC to instantiate L2/L3VPN

services, using the YANG models under definition in [VN], [L2NM],
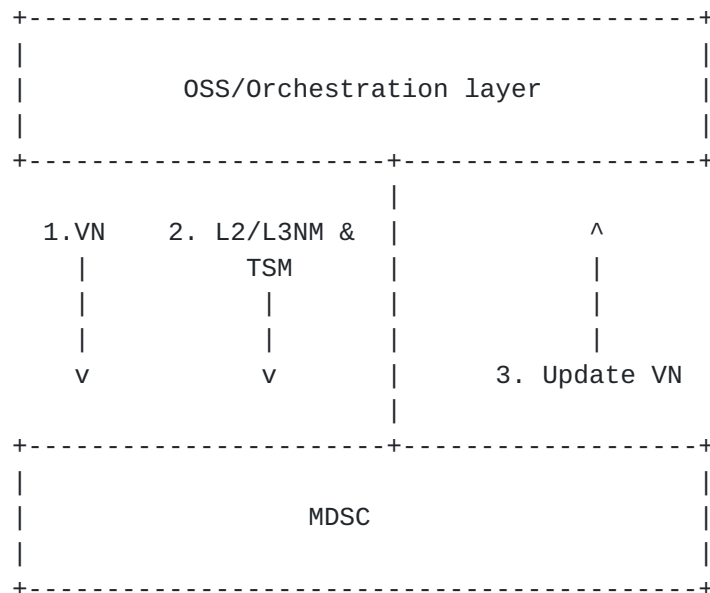[L3NM] and [TSM].

```
            +--------------------------------------------+
            |                                            |
            |          OSS/Orchestration layer           |
            |                                            |
            +----------------------+---------------------+
                                   |
        1.VN    2. L2/L3NM &  |                ^
           |           TSM         |                |
           |            |          |                |
           |            |          |                |
           v            v          |       3. Update VN
                                   |
            +----------------------+---------------------+
            |                                            |
            |                   MDSC                      |
            |                                            |
            +--------------------------------------------+
```

                    Figure 2 Service Request Process

o The VN YANG model [VN], whose primary focus is the CMI, can also
   be used to provide VN Service configuration from a orchestrated
   connectivity service point of view, when the L2/L3VPN service has
   TE requirements. This model is not used to setup L2/L3VPN 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 between the connectivity matrix with
       the VN member from which 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 YANG model [L2NM], whose primary focus is the MPI, can
   also be used to provide L2VPN service configuration and site
   information, from a orchestrated connectivity service point of
   view.

   o The L3NM YANG model [L3NM], whose primary focus is the MPI, can
      also be used to provide all L3VPN service configuration and site
      information, from a orchestrated connectivity service point of
      view.

   o The TE & Service Mapping YANG model [TSM] provides TE-service
      mapping as well as site mapping.

     o TE-service mapping provides the mapping between a L2/L3VPN
        instance and the corresponding VN instances.

     o The TE-service mapping also provides the service mapping
        requirement type as to how each L2/L3VPN/VN instance is
        created with respect to the underlay TE tunnels (e.g.,
        whether they require a new and isolated set of TE underlay
        tunnels or not). See Section 2.2 for detailed discussion on
        the mapping requirement types.

     o Site mapping provides the site reference information across
        L2/L3VPN Site ID, VN Access Point ID, and the LTP of the
        access link.

2.2. Service and Network Orchestration

   From a functional standpoint, MDSC represented in Figure 2
   interfaces with the OSS/Orchestration layer and decouples L2/L3VPN
   service configuration functions from network configuration
   functions. Therefore in this document the MDSC performs the
   functions of the Network Orchestrator, as defined in [RFC 8309].

   One of the important MDSC functions is to identify which TE Tunnels
   should carry the L2/L3VPN traffic (e.g., from TE & Service Mapping
   configuration) and to relay this information to the P-PNCs, to
   ensure the PEs' forwarding tables (e.g., VRF) are properly
   populated, according to the TE binding requirement for the L2/L3VPN.

   TE binding requirement types [TSM] are:

   1. Hard Isolation with deterministic latency: The L2/L3VPN service
      requires a set of dedicated TE Tunnels providing deterministic
      latency performances and that cannot be not shared with other
      services, nor compete for bandwidth with other Tunnels.

   2. Hard Isolation: This is similar to the above case without
      deterministic latency requirements.

   3. Soft Isolation: The L2/L3VPN service requires a set of dedicated
      MPLS-TE tunnels which cannot be shared with other services, but
      which could compete for bandwidth with other Tunnels.

   4. Sharing: The L2/L3VPN service allows sharing the MPLS-TE Tunnels
      supporting it with other services.

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

   In order to fulfill the the L2/L3VPN end-to-end TE requirements,
   including the TE binding requirements, the MDSC needs to perform
   multi-layer/multi-domain path computation to select the BRs, the
   intra-domain MPLS-TE Tunnels and the intra-domain Optical Tunnels.

   Depending on the knowledge that MDSC has of the topology and
   configuration of the underlying network domains, three models for
   performing path computation are possible:

   1. Summarization: MDSC has an abstracted TE topology view of all of
      the underlying domains, both packet and optical. MDSC does not
      have enough TE topology information to perform
      multi-layer/multi-domain path computation. Therefore MDSC
      delegates the P-PNCs and O-PNCs to perform a local path
      computation within their controlled domains and it uses the
      information returned by the P-PNCs and O-PNCs to compute the
      optimal multi-domain/multi-layer path.
      This model presents an issue to P-PNC, which does not have the
      capability of performing a single-domain/multi-layer path
      computation (that is, P-PNC does not have any possibility to
      retrieve the topology/configuration information from the Optical
      controller). A possible solution could be to include a CNC
      function in the P-PNC to request the MDSC multi-domain Optical
      path computation, as shown in Figure 10 of [RFC8453].

2. Partial summarization: MDSC has full visibility of the TE
   topology of the packet network domains and an abstracted view of
   the TE topology of the optical network domains.
   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 MDSC still needs to delegate the O-PNCs to perform
   local path computation within their respective domains and it
   uses the information received by the O-PNCs, together with its TE
   topology view of the multi-domain packet layer, to perform
   multi-layer/multi-domain path computation.
   The role of P-PNC is minimized, i.e. is limited to management.

3. Full knowledge: MDSC has the complete and enough detailed view of
   the TE topology of all the network domains (both optical and
   packet). In such case MDSC has all the information needed to
   perform multi-domain/multi-layer path computation, without
   relying on PNCs.
   This model may present, as a potential drawback, scalability
   issues and, as discussed in section 2.2. of [PATH-COMPUTE],
   performing path computation for optical networks in the MDSC is
   quite challenging because the optimal paths depend also on
   vendor-specific optical attributes (which may be different in the
   two domains if they are provided by different vendors).

The current version of this draft assumes that MDSC supports at
least model #2 (Partial summarization).

[Note: check with opeerators for some references on real deployment]

2.2.1. Hard Isolation

For example, when "Hard Isolation with or w/o deterministic latency"
TE binding requirement is applied for a L2/L3VPN, new Optical
Tunnels need to be setup to support dedicated IP Links between PEs
and BRs.

The MDSC needs to identify the set of IP/MPLS domains and their BRs.
This requires the MDSC to request each O-PNC to compute the
intra-domain optical paths between each PEs/BRs pairs.

When requesting optical path computation to the O-PNC, the MDSC
needs to take into account the inter-layer peering points, such as
the interconnections between the PE/BR nodes and the edge Optical
nodes (e.g., using the inter-layer lock or the transitional link
information, defined in [RFC8795]).

When the optimal multi-layer/multi-domain path has been computed, the MDSC requests each O-PNC to setup the selected Optical Tunnels and P-PNC to setup the intra-domain MPLS-TE Tunnels, over the selected Optical Tunnels. MDSC also properly configures its BGP speakers and PE/BR forwarding tables to ensure that the VPN traffic is properly forwarded.

2.2.2. Shared Tunnel Selection

In case of shared tunnel selection, the MDSC needs to check if there is multi-domain path which can support the L2/L3VPN end-to-end TE service requirements (e.g., bandwidth, latency, etc.) using existing intra-domain MPLS-TE tunnels.

If such a path is found, the MDSC selects the optimal path from the candidate pool and request each P-PNC to setup the L2/L3VPN service using the selected intra-domain MPLS-TE tunnel, between PE/BR nodes.

Otherwise, the MDSC should detect if the multi-domain path can be setup using existing intra-domain MPLS-TE tunnels with modifications (e.g., increasing the tunnel bandwidth) or setting up new intra-domain MPLS-TE tunnel(s).

The modification of an existing MPLS-TE Tunnel as well as the setup of a new MPLS-TE Tunnel may also require multi-layer coordination e.g., in case the available bandwidth of underlying Optical Tunnels is not sufficient. Based on multi-domain/multi-layer path computation, the MDSC can decide for example to modify the bandwidth of an existing Optical Tunnel (e.g., ODUflex bandwidth increase) or to setup new Optical Tunnels to be used as additional LAG members of an existing IP Link or as new IP Links to re-route the MPLS-TE Tunnel.

In all the cases, the labels used by the end-to-end tunnel are distributed in the PE and BR nodes by BGP. The MDSC is responsible to configure the BGP speakeers in each P-PNC, if needed.

2.3. IP/MPLS Domain Controller and NE Functions

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

Among the functions of the P-PNC, there are the setup or modification of the intra-domain MPLS-TE Tunnels, between PEs and BRs, and the configuration of the VPN services, such as the VRF in the PE nodes, as shown in Figure 3:

```
       +------------------+              +------------------+
       |                  |              |                  |
       |     P-PNC1       |              |     P-PNC2       |
       |                  |              |                  |
       +--|-----------|---+              +--|-----------|---+
          | 1.Tunnel  | 2.VPN              | 1.Tunnel  | 2.VPN
          | Config    | Provisioning       | Config    | Provisioning
          V           V                    V           V
        +---------------------+          +---------------------+
     CE  / PE     tunnel 1     BR\        / BR    tunnel 2    PE \  CE
     o--/---o.................o--\-----/--o.................o---\--o
       \                         /     \                         /
        \        Domain 1       /       \        Domain 2       /
        +---------------------+          +---------------------+


                        End-to-end tunnel
           <---------------------------------------------------->
```
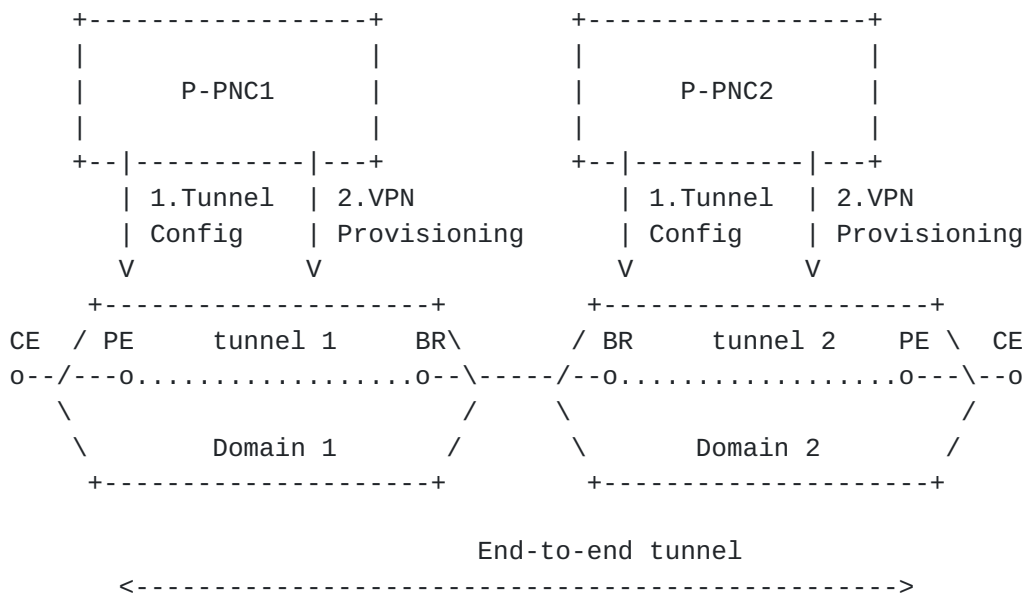
           Figure 3 IP/MPLS Domain Controller & NE Functions

   It is assumed that BGP is running in the inter-domain IP/MPLS
   networks for L2/L3VPN and that the P-PNC controller is also
   responsible for configuring the BGP speakers within its control
   domain, if necessary.

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

   If new MPLS-TE Tunnels are needed or mofications (e.g., bandwidth
   ingrease) to existing MPLS_TE Tunnels are needed, as outlined in
   section 2.2, the MDSC would request their setup or modifications to
   the P-PNCs (step 1 in Figure 3). Then the MDSC would request the
   P-PNC to configure the VPN, including the selection of the
   intra-domain TE Tunnel (step 2 in Figure 3).

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

   o VPN label: assigned by the egress PE and distributed by BGP;

   o end-to-end LSP label: assigned by the egress BR, selected by the
      MDSC, and distributed by BGP;

o MPLS-TE tunnel label, assigned by the next hop P node of the
   tunnel selected by the MDSC and distributed by mechanism internal
   to the IP/MPLS domain (e.g., RSVP-TE).

2.4. Optical Domain Controller and NE Functions

Optical network provides the underlay connectivity services to
IP/MPLS networks. The coordination of Packet/Optical multi-layer 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-speicific and outside
the scope of this document.

Depending on the type of optical network, TE topology abstraction,
path compution 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 which are applicable at
all the MPI interfaces, between each PNC (Optical or Packet) and the
MDSC, and also to all 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. 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 CMI 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].

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 use the following common topology YANG
models at the MPI to report their abstract topologies:

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 with TE specific information.

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

Both Optical and Packet PNCs must use the following common
notifications YANG models at the MPI so that any network changes can
be reported almost in real-time to MDSC by the PNCs:

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 must be compliant with subscription requirements as
stated in [RFC7923].

3.2.2. YANG models at the Optical MPIs

The Optical PNC also uses at least the following technology-specific
topology YANG models, providing WDM and Ethernet technology-specific
augmentations of the generic TE Topology Model:

o The WSON Topology Model, defined in the "ietf-wson-topology" YANG
   modules of [WSON-TOPO], or the Flexi-grid Topology Model, defined
   in the "ietf-flexi-grid-topology" YANG module of [Flexi-TOPO].

o Optionally, when the OTN layer is used, the OTN Topology Model,
   as defined in the "ietf-otn-topology" YANG module of [OTN-TOPO].

o The Ethernet Topology Model, defined in the "ietf-eth-te-
   topology" YANG module of [CLIENT-TOPO].

o Optionally, when the OTN layer is used, the network data model
   for L1 OTN services (e.g. an Ethernet transparent service) as
   defined in "ietf-trans-client-service" YANG module of draft-ietf-
   ccamp-client-signal-yang [CLIENT-SIGNAL].

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

The Ethernet Topology is used to report the access links between the
IP routers and the edge ROADMs.

The optical PNC uses at least the following YANG models:

o The TE Tunnel Model, defined in the "ietf-te" YANG module of
   [TE-TUNNEL]

o The WSON Tunnel Model, defined in the "ietf-wson-tunnel" YANG
   modules of [WSON-TUNNEL], or the Flexi-grid Media Channel Model,
   defined in the "ietf-flexi-grid-media-channel" YANG module of
   [Flexi-MC]

o Optionally, when the OTN layer is used, the OTN Tunnel Model,
   defined in the "ietf-otn-tunnel" YANG module of [OTN-TUNNEL].

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

The TE Tunnel model is generic and augmented by technology-specific
models such as the WSON Tunnel Model and the Flexi-grid Media
Channel Model.

The WSON Tunnel Model or, alternatively, the Flexi-grid Media
Channel Model are used to setup connectivity within the DWDM network
depending on whether the DWDM optical network is based on fixed grid
or flexible-grid.

The Ethernet Client Signal Model is used to configure the steering
of the Ethernet client traffic between Ethernet access links and TE
Tunnels, which in this case could be either WSON Tunnels or
Flexi-Grid Media Channels. This model is generic and applies to any
technology-specific TE Tunnel: technology-specific attributes are
provided by the technology-specific models which augment the generic
TE-Tunnel Model.

3.2.3. YANG data models at the Packet MPIs

The Packet PNC also uses at least the following technology-specific
topology YANG models, providing IP and Ethernet technology-specific
augmentations of the generic Topology Models described in section
3.2.1:

o The L3 Topology Model, defined in the "ietf-l3-unicast-topology"
   YANG modules of [RFC8346], which augments the Base Network
   Topology Model

o The L3 specific data model including extended TE attributes (e.g.
   performance derived metrics like latency), defined in "ietf-l3-
   te-topology" and in "ietf-te-topology-packet" in draft-ietf-teas-
   l3-te-topo [L3-TE-TOPO]

o The Ethernet Topology Model, defined in the "ietf-eth-te-
   topology" YANG module of [CLIENT-TOPO], which augments the TE
   Topology Model

The Ethernet Topology Model is used to report the access links
between the IP routers and the edge ROADMs as well as the
inter-domain links between ASBRs, while the L3 Topology Model is
used to report the IP network topology (e.g., IP routers and links).

o The User Network Interface (UNI) Topology Model, being defined in
   the "ietf-uni-topology" module of the draft-ogondio-opsawg-uni-
   topology [UNI-TOPO] which augment "ietf-network" module defined
   in [RFC8345] adding service attachment points to the nodes to
   which L2VPN/L3VPN IP/MPLS services can be attached.

o L3VPN network data model defined in "ietf-l3vpn-ntw" module of
   draft-ietf-opsawg-l3sm-l3nm [L3NM] used for non-ACTN MPI for
   L3VPN service provisioning

   o L2VPN network data model defined in "ietf-l2vpn-ntw" module of
      draft-ietf-barguil-opsawg-l2sm-l2nm [L2NM] used for non-ACTN MPI
      for L2VPN service provisioning

   [Editor's note:] Add YANG models used for tunnel and service
   configuration.

4. Multi-layer and multi-domain services scenarios

   Multi-layer and multi-domain scenarios, based on reference network
   described in section 2, and very relevant for Service Providers, are
   described in the next sections. For each scenario existing IETF
   protocols and data models are identified with particular focus on
   the MPI in the ACTN architecture. Non ACTN IETF data models required
   for L2/L3VPN service provisioning between MDSC and IP PNCs are also
   identified.

4.1. Scenario 1: network and service topology discovery

   In this scenario, the MSDC needs to discover through the underlying
   PNCs, the network topology, at both WDM and IP layers, in terms of
   nodes (NEs) and links, including inter AS domain links as well as
   cross-layer links but also in terms of tunnels (MPLS or SR paths in
   IP layer and OCh and optionally ODUk tunnels in optical layer).MDSC
   discovers also the IP/MPLS transport services (L2VPN/L3VPN)
   deployed, both intra-domain and inter-domain wise.

   Each PNC provides to the MDSC an abstracted or full topology view of
   the WDM or the IP topology of the domain it controls. This topology
   can be abstracted in the sense that some detailed NE information is
   hidden at the MPI, and all or some of the NEs and related physical
   links are exposed as abstract nodes and logical (virtual) links,
   depending on the level of abstraction the user requires. This
   information is key to understand both the inter-AS domain links
   (seen by each controller as UNI interfaces but as I-NNI interfaces
   by the MDSC) as well as the cross-layer mapping between IP and WDM
   layer.

   The MDSC also maintains an up-to-date network database of both IP
   and WDM layers (and optionally OTN layer) through the use of IETF
   notifications through MPI with the PNCs when any topology change
   occurs. It should be possible also to correlate information coming
   from IP and WDM layers (e.g.: which port, lambda/OTSi, direction is
   used by a specific IP service on the WDM equipment)

   In particular, For the cross-layer links it is key for MDSC to be
   able to correlate automatically 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.

   It should be possible at MDSC level to easily correlate WDM and IP
   layers alarms to speed-up troubleshooting

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

4.1.1. Inter-domain link discovery

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

   o Links between two IP domains (ASes)

   o Links between an IP router and a ROADM

   Both types of links are Ethernet physical links.

   The inter-domain link information is reported to the MDSC by the two
   adjacent PNCs, controlling the two ends of the inter-domain link.
   The MDSC needs to understa how to merge the these inter-domain
   Ethernet links together.

   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.

   The MDSC can understand how to merge these inter-domain links
   together using the plug-id attribute defined in the TE Topology
   Model [RFC8795], as described in as described in section 4.3 of
   [RFC8795].

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

Both types of inter-domain links are discovered using the plug-id
attributes reported in the Ethernet Topologies exposed by the two
adjacent PNCs. The MDSC can also discover an inter-domain IP
link/adjacency between the two IP LTPs, reported in the IP
Topologies exposed by the two adjacent P-PNCs, supported by the two
ETH LTPs of an Ethernet Link discovered between these two P-PNCs.

The static configuration requires an administrative burden to
configure network-wide unique identifiers: it is therefore more
viable for inter-AS links. For the links between the IP routers and
the Optical NEs, the automatic discovery solution based on LLDP
snooping is preferable when LLDP snooping is supported by the
Optical NEs.

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

4.1.2. IP Link Setup Procedure

The MDSC requires the O-PNC to setup a WDM Tunnel (either a WSON
Tunnel or a Flexi grid Tunnel) within the DWDM network between the
two Optical Transponders (OTs) associated with the two access links.

The Optical Transponders are reported by the O-PNC as Trail
Termination Points (TTPs), defined in [TE TOPO], within the WDM
Topology. The association between the Ethernet access link and the
WDM TTP is reported by the Inter Layer Lock (ILL) identifiers,
defined in [TE TOPO], reported by the O PNC within the Ethernet
Topology and WDM Topology.

The MDSC also requires the O-PNC to steer the Ethernet client
traffic between the two access Ethernet Links over the WDM Tunnel.

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

If LLDP [IEEE 802.1AB] is used between the two routers, the P PNC
can automatically discover the IP Link being set up by the MDSC. The
IP LTPs terminating this IP Link are supported by the ETH LTPs
terminating the two access links.

Otherwise, the MDSC needs to require the P PNC to configure an IP
Link between the two routers: the MDSC also configures the two ETH
LTPs which support the two IP LTPs terminating this IP Link.

[4.2](#). L2VPN/L3VPN establishment

To be added

[Editor's Note] What mechanism would convey on the interface to the
IP/MPLS domain controllers as well as on the SBI (between IP/MPLS
domain controllers and IP/MPLS PE routers) the TE binding policy
dynamically for the L3VPN? Typically, VRF is the function of the
device that participate MP-BGP in MPLS VPN. With current MP-BGP
implementation in MPLS VPN, the VRF's BGP next hop is the
destination PE and the mapping to a tunnel (either an LDP or a BGP
tunnel) toward the destination PE is done by automatically without
any configuration. It is to be determined the impact on the PE VRF
operation when the tunnel is an optical bypass tunnel which does not
participate either LDP or BGP.

New text to answer the yellow part:

The MDSC Network-related function will then coordinate with the PNCs
involved in the process to provide the provisioning information
through ACTN MDSC to PNC (MPI) interface. The relevant data models
used at the MPI may be in the form of L3NM, L2NM or others and are
exchanged through MPI API calls. Through this process MDSC Network-
related functions provide the configuration information to realize a
VPN service to PNCs. For example, this process will inform PNCs on
what PE routers compose a L3VPN, the topology requested, the VPN
attributes, etc.

At the end of the process PNCs will deliver the actual configuration
to the devices (either physical or virtual), through the ACTN
Southbound Interface (SBI). In this case the configuration policies
may be exchanged using a Netconf session delivering configuration
commands associated to device-specific data models (e.g. BGP[], QOS
[], etc.).

Having the topology information of the network domains under their
control, PNCs will deliver all the information necessary to create,
update, optimize or delete the tunnels connecting the PE nodes as
requested by the VPN instantiation.

[5](#). Security Considerations

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

6. Operational Considerations

Telemetry data, such as the collection of 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.

7. IANA Considerations

This document requires no IANA actions.

8. References

8.1. Normative References

[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.

[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.

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

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

[WSON-TOPO] Lee, Y. et al., " A YANG Data Model for WSON (Wavelength
          Switched Optical Networks)", draft-ietf-ccamp-wson-yang,
          work in progress.

   [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.

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

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

   [Flexi-MC] Lopez de Vergara, J. E. et al., "YANG data model for
              Flexi-Grid media-channels", draft-ietf-ccamp-flexigrid-
              media-channel-yang, work in progress.

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

8.2. Informative References

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

   [RFC4364] E. Rosen and Y. Rekhter, "BGP/MPLS IP Virtual Private
              Networks (VPNs)", RFC 4364, February 2006.

   [RFC4761] K. Kompella, Ed., Y. Rekhter, Ed., "Virtual Private LAN
              Service (VPLS) Using BGP for Auto-Discovery and
              Signaling", RFC 4761, January 2007.

   [RFC6074] E. Rosen, B. Davie, V. Radoaca, and W. Luo, "Provisioning,
              Auto-Discovery, and Signaling in Layer 2 Virtual Private
              Networks (L2VPNs)", RFC 6074, January 2011.

   [RFC6624] K. Kompella, B. Kothari, and R. Cherukuri, "Layer 2
              Virtual Private Networks Using BGP for Auto-Discovery and
              Signaling", RFC 6624, May 2012.

   [RFC7209] A. Sajassi, R. Aggarwal, J. Uttaro, N. Bitar, W.
              Henderickx, and A. Isaac, "Requirements for Ethernet VPN
              (EVPN)", RFC 7209, May 2014.

   [RFC7432] A. Sajassi, Ed., et al., "BGP MPLS-Based Ethernet VPN",
              RFC 7432, February 2015.

   [RFC7436] H. Shah, E. Rosen, F. Le Faucheur, and G. Heron, "IP-Only
              LAN Service (IPLS)", RFC 7436, January 2015.

   [RFC8214] S. Boutros, A. Sajassi, S. Salam, J. Drake, and J.
              Rabadan, "Virtual Private Wire Service Support in Ethernet
              VPN", RFC 8214, August 2017.

   [RFC8299] Q. Wu, S. Litkowski, L. Tomotaki, and K. Ogaki, "YANG Data
              Model for L3VPN Service Delivery", RFC 8299, January 2018.

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

   [RFC8466] G. Fioccola, ed., "A YANG Data Model for Layer 2 Virtual
              Private Network (L2VPN) Service Delivery", RFC8466,
              October 2018.

   [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.

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

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

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

   [ACTN-PM] Y. Lee, et al., "YANG models for VN & TE Performance
             Monitoring Telemetry and Scaling Intent Autonomics",
             draft-lee-teas-actn-pm-telemetry-autonomics, work in
             progress.

   [BGP-L3VPN] D. Jain, et al. "Yang Data Model for BGP/MPLS L3 VPNs",
             draft-ietf-bess-l3vpn-yang, work in progress.

Appendix A. Multi-layer and multi-domain resiliency

A.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. 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 backup 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-
   optimisation of the back-up MPLS path to go back to the MPLS
   primary path through the back-up port of the router and the
   original OCh if overall cost, latency etc. is improved. However,
   in this scenario, there is a need for protection port PLUS back-
   up port in the router which does not lead to clear port savings.

Acknowledgments

Contributors

Sergio Belotti
Nokia

Email: sergio.belotti@nokia.com

Gabriele Galimberti
Cisco

Email: ggalimbe@cisco.com

Zheng Yanlei
China Unicom

Email: zhengyanlei@chinaunicom.cn

Anton Snitser
Sedona

      Email: antons@sedonasys.com

      Washington Costa Pereira Correia
      TIM Brasil

      Email: wcorreia@timbrasil.com.br

      Michael Scharf
      Hochschule Esslingen - University of Applied Sciences

      Email: michael.scharf@hs-esslingen.de

      Young Lee
      Sung Kyun Kwan University

      Email: younglee.tx@gmail.com

      Jeff Tantsura
      Apstra

      Email: jefftant.ietf@gmail.com

      Paolo Volpato
      Huawei

      Email: paolo.volpato@huawei.com

Authors' Addresses

      Fabio Peruzzini
      TIM

      Email: fabio.peruzzini@telecomitalia.it

      Jean-Francois Bouquier
      Vodafone

      Email: jeff.bouquier@vodafone.com

      Italo Busi
      Huawei

      Email: Italo.busi@huawei.com

      Daniel King
      Old Dog Consulting

      Email: daniel@olddog.co.uk

      Daniele Ceccarelli
      Ericsson

      Email: daniele.ceccarelli@ericsson.com