TEAS Working Group
Internet Draft

Category: Informational

Haomian Zheng
Xianlong Luo
Yi Lin
Huawei Technologies
Yang Zhao
China Mobile
Yunbin Xu
CAICT
Sergio Belotti
Dieter Beller
Nokia

Expires: July 16, 2022

January 12, 2022

Interworking of GMPLS Control and Centralized Controller System

draft-ietf-teas-gmpls-controller-inter-work-07

Abstract

Generalized Multi-Protocol Label Switching (GMPLS) control allows each network element (NE) to perform local resource discovery, routing and signaling in a distributed manner.

On the other hand, with the development of software-defined transport networking technology, a set of NEs can be controlled via centralized controller hierarchies to address the issue from multi-domain, multi-vendor and multi-technology. An example of such centralized architecture is ACTN controller hierarchy described in RFC 8453.

Instead of competing with each other, both the distributed and the centralized control plane have their own advantages, and should be complementary in the system. This document describes how the GMPLS distributed control plane can interwork with a centralized controller system in a transport network.

at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

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

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

This Internet-Draft will expire on July 16, 2022.

Table of Contents

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] extends MPLS to support different classes of interfaces and switching capabilities such as Time-Division Multiplex Capable (TDM), Lambda Switch Capable (LSC), and Fiber-Switch Capable (FSC). Each network element (NE) running a GMPLS control plane collects network information from other NEs and supports service provisioning through signaling in a distributed manner. More generic description for Traffic-engineering networking information exchange can be found in [RFC7926].

On the other hand, Software-Defined Networking (SDN) technologies have been introduced to control the transport network in a centralized manner. Central controllers can collect network information from each node and provision services to corresponding nodes. One of the examples is the Abstraction and Control of Traffic Engineered Networks (ACTN) [RFC8453], which defines a hierarchical architecture with Provisioning Network Controller (PNC), Multi-domain Service Coordinator (MDSC) and Customer Network Controller (CNC) as central controllers for different network abstraction levels. A Path Computation Element (PCE) based approach has been proposed as Application-Based Network Operations (ABNO) in [RFC7491].

In such centralized controller architectures, GMPLS can be applied for the NE-level control. A central controller may support GMPLS enabled domains and may interact with a GMPLS enabled domain where the GMPLS control plane does the service provisioning from ingress

to egress. In this case the centralized controller sends the request

to the ingress node and does not have to configure all NEs along the path through the domain from ingress to egress thus leveraging the GMPLS control plane. This document describes how GMPLS control interworks with centralized controller system in transport network.

2. Overview

In this section, overviews of GMPLS control plane and centralized controller system are discussed as well as the interactions between the GMPLS control plane and centralized controllers.

2.1. Overview of GMPLS Control Plane

GMPLS separates the control plane and the data plane to support time-division, wavelength, and spatial switching, which are significant in transport networks. For the NE level control in GMPLS, each node runs a GMPLS control plane instance. Functionalities such as service provisioning, protection, and restoration can be performed via GMPLS communication among multiple NEs. At the same time, the controller can also collect node and link resources in the network to construct the network topology and compute routing paths for serving service requests.

Several protocols have been designed for GMPLS control [RFC3945] including link management [RFC4204], signaling [RFC3471], and routing [RFC4202] protocols. The controllers applying these protocols communicate with each other to exchange resource information and establish Label Switched Paths (LSPs). In this way, controllers in different nodes in the network have the same view of the network topology and provision services based on local policies.

2.2. Overview of Centralized Controller System

With the development of SDN technologies, a centralized controller architecture has been introduced to transport networks. One example architecture can be found in ACTN [RFC8453]. In such systems, a controller is aware of the network topology and is responsible for provisioning incoming service requests.

Multiple hierarchies of controllers are designed at different levels implementing different functions. This kind of architecture enables multi-vendor, multi-domain, and multi-technology control. For example, a higher-level controller coordinates several lower-level controllers controlling different domains, for topology collection and service provisioning. Vendor-specific features can be abstracted between controllers, and standard API (e.g., generated from RESTconf/YANG) is used.

2.3. GMPLS Control Interwork with Centralized Controller System

Besides the GMPLS and the interactions among the controller hierarchies, it is also necessary for the controllers to communicate with the network elements. Within each domain, GMPLS control can be applied to each NE. The bottom-level central controller can act as a NE to collect network information and initiate LSP. Figure 1 shows an example of GMPLS interworking with centralized controllers (ACTN terminologies are used in the figure).

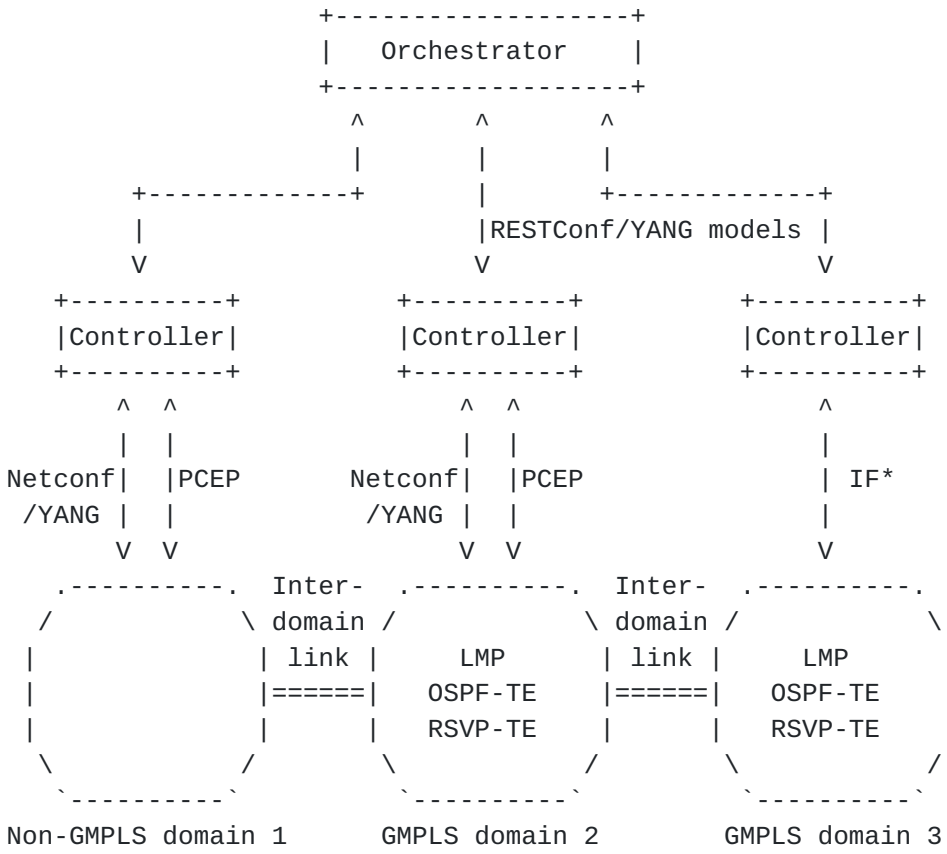


Figure 1: Example of GMPLS/non-GMPLS interworks with Controllers

Figure 1 shows the scenario with two GMPLS domains and one non-GMPLS domain. This system supports the interworking among non-GMPLS domain, GMPLS domain and the controller hierarchies. For domain 1, the network element were not enabled with GMPLS so the control can be purely from the controller, via Netconf/YANG and/or PCEP. For domain 2 and 3, each domain has the GMPLS control plane enabled at the physical network level. The PNC can exploit GMPLS capability implemented in the domain to listen to the IGP routing protocol

messages (OSPF LSAs for example) that the GMPLS control plane instances are disseminating into the network and thus learn the network topology. For path computation in the domain with PNC implementing a PCE, PCCs (e.g. NEs, other controller/PCE) use PCEP to ask the PNC for a path and get replies. The MDSC communicates with PNCs using for example REST/RESTConf based on YANG data models. As a PNC has learned its domain topology, it can report the topology to the MDSC. When a service arrives, the MDSC computes the path and coordinates PNCs to establish the corresponding LSP segment.

Alternatively, the NETCONF protocol can be used to retrieve topology information utilizing the e.g. [RFC8795] Yang model and the technology-specific YANG model augmentations required for the specific network technology. The PNC can retrieve topology information from any NE (the GMPLS control plane instance of each NE in the domain has the same topological view), construct the topology of the domain and export an abstracted view to the MDSC. Based on the topology retrieved from multiple PNCs, the MDSC can create topology graph of the multi-domain network, and can use it for path computation. To setup a service, the MDSC can exploit e.g. [TE-Tunnel] Yang model together with the technology-specific YANG model augmentations.

3. Discovery Options

In GMPLS control, the link connectivity need to be verified between each pair of nodes. In this way, link resources, which are fundamental resources in the network, are discovered by both ends of the link.

3.1. LMP

Link management protocol (LMP) [RFC4204] runs between a pair of nodes and is used to manage TE links. In addition to the setup and maintenance of control channels, LMP can be used to verify the data link connectivity and correlate the link property.

4. Routing Options

In GMPLS control, link state information is flooded within the network as defined in [RFC4202]. Each node in the network can build the network topology according to the flooded link state information. Routing protocols such as OSPF-TE [RFC4203] and ISIS-TE [RFC5307] have been extended to support different interfaces in GMPLS.

In centralized controller system, central controller can be placed at the GMPLS network and passively receive the information flooded in the network. In this way, the central controller can construct

and update the network topology.

4.1. OSPF-TE

OSPF-TE is introduced for TE networks in [RFC3630]. OSPF extensions have been defined in [RFC4203] to enable the capability of link state information for GMPLS network. Based on this work, OSPF protocol has been extended to support technology-specific routing. The routing protocol for OTN, WSON and optical flexi-grid network are defined in [RFC7138], [RFC7688] and [RFC8363], respectively.

4.2. ISIS-TE

ISIS-TE is introduced for TE networks in [RFC5305] and is extended to support GMPLS routing functions [RFC5307], and has been updated to [RFC7074] to support the latest GMPLS switching capability and Types fields.

4.3. Netconf/RESTconf

Netconf [RFC6241] and RESTconf [RFC8040] protocols are originally used for network configuration. Besides, these protocols can also be used for topology retrieval by using topology-related YANG models, such as [RFC8345] and [RFC8795]. These protocols provide a powerful mechanism for notification that permits to notify the client about topology changes.

5. Path Computation

Once a controller learns the network topology, it can utilize the available resources to serve service requests by performing path computation. Due to abstraction, the controllers may not have sufficient information to compute the optimal path. In this case, the controller can interact with other controllers by sending Yang Path Computation requests [PAT-COMP] to compute a set of potential optimal paths and then, based on its own constraints, policy and specific knowledge (e.g. cost of access link) can choose the more feasible path for service e2e path setup.

Path computation is one of the key objectives in various types of controllers. In the given architecture, it is possible for different components that have the capability to compute the path.

5.1. Constraint-based Path Computing in GMPLS Control

In GMPLS control, a routing path is computed by the ingress node [RFC3473] and is based on the ingress node TED. Constraint-based path computation is performed according to the local policy of the ingress node.

5.2. Path Computation Element (PCE)

PCE has been introduced in [RFC4655] as a functional component that provides services to compute path in a network. In [RFC5440], the path computation is accomplished by using the Traffic Engineering Database (TED), which maintains the link resources in the network. The emergence of PCE efficiently improve the quality of network planning and offline computation, but there is a risk that the computed path may be infeasible if there is a diversity requirement, because stateless PCE has no knowledge about the former computed paths.

To address this issue, stateful PCE has been proposed in [RFC8231]. Besides the TED, an additional LSP Database (LSP-DB) is introduced to archive each LSP computed by the PCE. In this way, PCE can easily figure out the relationship between the computing path and former computed paths. In this approach, PCE provides computed paths to PCC, and then PCC decides which path is deployed and when to be established.

In PCE Initiation [RFC8281], PCE is allowed to trigger the PCC to setup, maintenance, and teardown of the PCE-initiated LSP under the stateful PCE model. This would allow a dynamic network that is centrally controlled and deployed.

In centralized controller system, the PCE can be implemented in a central controller, and the central controller performs path computation according to its local policies. On the other hand, the PCE can also be placed outside of the central controller. In this case, the central controller acts as a PCC to request path computation to the PCE through PCEP. One of the reference architecture can be found at [RFC7491].

6. Signaling Options

Signaling mechanisms are used to setup LSPs in GMPLS control. Messages are sent hop by hop between the ingress node and the egress node of the LSP to allocate labels. Once the labels are allocated along the path, the LSP setup is accomplished. Signaling protocols such as RSVP-TE [RFC3473] have been extended to support different interfaces in GMPLS.

6.1. RSVP-TE

RSVP-TE is introduced in [RFC3209] and extended to support GMPLS signaling in [RFC3473]. Several label formats are defined for a generalized label request, a generalized label, suggested label and label sets. Based on [RFC3473], RSVP-TE has been extended to support technology-specific signaling. The RSVP-TE extensions for OTN, WSON,

optical flexi-grid network are defined in [RFC7139], [RFC7689], and [RFC7792], respectively.

7. Interworking Scenarios

7.1. Topology Collection & Synchronization

Topology information is necessary on both network elements and controllers. The topology on network element is usually raw information, while the topology on the controller can be either raw or abstracted. Three different abstraction methods have been described in [RFC8453], and different controllers can select the corresponding method depending on application.

When there are changes in the network topology, the impacted network element(s) need to report changes to all the other network elements, together with the controller, to sync up the topology information. The inter-NE synchronization can be achieved via protocols mentioned in section 3 and 4. The topology synchronization between NEs and controllers can either be achieved by routing protocols OSPF-TE/PCEP-LS in [PCEP-LS] or Netconf protocol notifications with YANG model.

7.2. Multi-domain Service Provisioning

Based on the topology information on controllers and network elements, service provisioning can be deployed. Plenty of methods have been specified for single domain service provisioning, such as using PCEP and RSVP-TE.

Multi-domain service provisioning would request coordination among the controller hierarchies. Given the service request, the end-to-end delivery procedure may include interactions at any level (i.e. interface) in the hierachy of the controllers (e.g. MPI and SBI for ACTN). The computation for a cross-domain path is usually completed by controllers who have a global view of the topologies. Then the configuration is decomposed into lower layer controllers, to configure the network elements to set up the path.

A combination of the centralized and distributed protocols may be necessary for the interaction between network elements and controller. Several methods can be used to create the inter-domain path:

1) With end-to-end RSVP-TE session:

In this method, the SDN controller of the source domain triggers the source node to create the end-to-end RSVP-TE session, and the assignment and distribution of the labels on the inter-domain links

are done by the boarder nodes of each domain, using RSVP-TE

protocol. Therefore, this method requires the interworking of RSVP-TE protocols between different domains.

There are two possible methods:

1.1) One single end-to-end RSVP-TE session

In this method, an end-to-end RSVP-TE session from the source NE to the destination NE will be used to create the inter-domain path. A typical example would be the PCE Initiation scenario, in which a PCE message (PCInitiate) is sent from the controller to the first-end node, and then trigger a RSVP procedure along the path. Similarly, the interaction between the controller and the ingress node of a domain can be achieved by Netconf protocol with corresponding YANG models, and then completed by running RSVP among the network elements.

1.2) LSP Stitching

The LSP stitching method defined in [RFC5150] can also be used to create the end-to-end LSP. I.e., when the source node receives an end-to-end path creation request (e.g., using PCEP or Netconf protocol), the source node starts an end-to-end RSVP-TE session along the end points of each LSP segment (refers to S-LSP in [RFC5150]) of each domain, to assign the labels on the inter-domain links between each pair of neighbor S-LSPs, and stitch the end-to-end LSP to each S-LSP. See Figure 2 as an example. Note that the S-LSP in each domain can be either created by each domain controller in advance, or created dynamically triggered by the end-to-end RSVP-TE session.

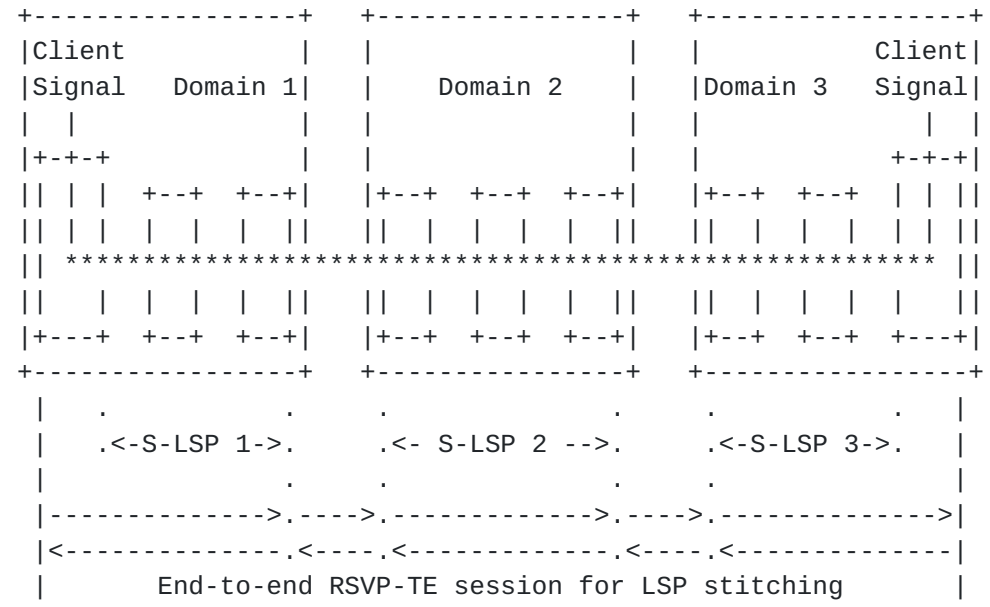


Figure 2: LSP stitching

2) Without end-to-end RSVP-TE session:

In this method, each SDN controller is responsible to create the path segment within its domain. The boarder node does not need to communicate with other boarder nodes in other domains for the distribution of labels on inter-domain links, so end-to-end RSVP-TE session through multiple domains is not required, and the interworking of RSVP-TE protocol between different domains is not needed.

Note that path segments in the source domain and the destination domain are "asymmetrical" segments, because the configuration of client signal mapping into server layer tunnel is needed at only one end of the segment, while configuration of server layer cross-connect is needed at the other end of the segment. For example, the path segment 1 and 3 in Figure 3 are asymmetrical segments, because one end of the segment requires mapping GE into ODU0, while the other end of the segment requires setting up ODU0 cross-connect.

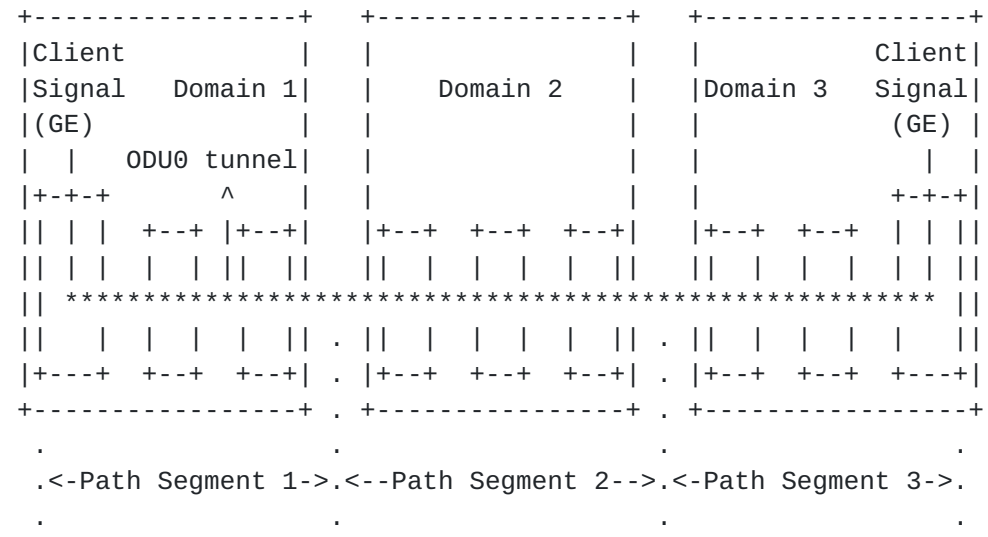


Figure 3: Example of asymmetrical path segment

The PCEP / GMPLS protocols should support creation of such asymmetrical segment.

Note also that mechanisms to assign the labels in the inter-domain links are also needed to be considered. There are two possible methods:

2.1) Inter-domain labels assigned by NEs:

The concept of Stitching Label that allows stitching local path segments was introduced in [RFC5150] and [sPCE-ID], in order to form the inter-domain path crossing several different domains. It also

describes the BRPC and H-PCE PCInitiate procedure, i.e., the ingress

boarder node of each downstream domain assigns the stitching label for the inter-domain link between the downstream domain and its upstream neighbor domain, and this stitching label will be passed to the upstream neighbor domain by PCE protocol, which will be used for the path segment creation in the upstream neighbor domain.

2.2) Inter-domain labels assigned by SDN controller:

If the resource of inter-domain links are managed by the multi-domain SDN controller, each single-domain SDN controller can provide to the multi-domain SDN controller the list of available labels (e.g. timeslots if OTN is the scenario) using IETF Topology model and related technology specific extension. Once that multi-domain SDN controller has computed e2e path RSVP-TE or PCEP can be used in the different domains to setup related segment tunnel consisting with label inter-domain information, e.g. for PCEP the label ERO can be included in the PCInitiate message to indicate the inter-domain labels, so that each boarder node of each domain can configure the correct cross-connect within itself.

7.3. Multi-layer Service Provisioning

GMPLS can interwork with centralized controller system in multi-layer networks.

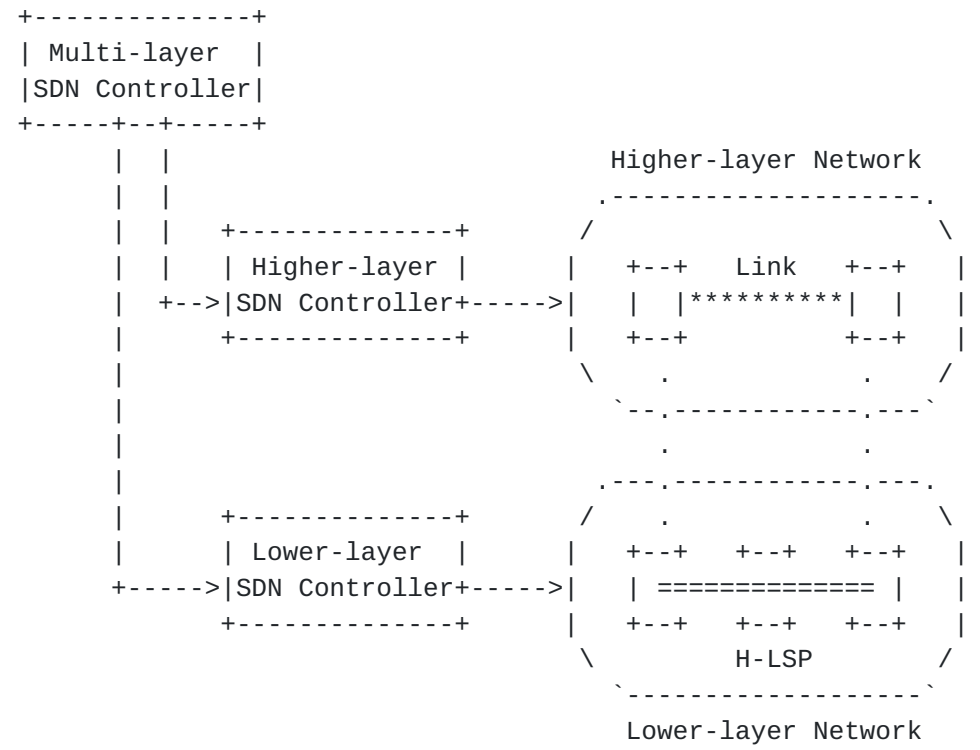


Figure 4: Example of GMPLS-SDN interworking in multi-layer network

An example with two layers of network is shown in Figure 4. In this example, the GMPLS control plane is enabled in each layer network, and interworks with the SDN controller of its domain (higher-layer SDN controller and lower-layer SDN controller, respectively). The multi-layer SDN controller, which acts as the Orchestrator, is used to coordinate the control of the multi-layer network.

7.3.1. Multi-layer Path Computation

[RFC5623] describes three inter-layer path computation models and four inter-layer path control models:

- 3 Path computation:
  - o Single PCE path computation model
  - o Multiple PCE path computation with inter-PCE communication model
  - o Multiple PCE path computation without inter-PCE communication model
- 4 Path control:
  - o PCE-VNTM cooperation model
  - o Higher-layer signaling trigger model
  - o NMS-VNTM cooperation model (integrated flavor)
  - o NMS-VNTM cooperation model (separate flavor)

Section 4.2.4 of [RFC5623] also provides all the possible combinations of inter-layer path computation and inter-layer path control models.

To apply [RFC5623] in multi-layer network with GMPLS-SDN interworking, the higher-layer SDN controller and the lower-layer SDN controller can act as the PCE Hi and PCE Lo respectively, and typically, the multi-layer SDN controller can act as a VNTM because it has the abstracted view of both the higher-layer and lower-layer networks.

Table 1 shows all possible combinations of path computation and path control models in multi-layer network with GMPLS-SDN interworking:

Table 1: Combinations of path computation and path control models

Path computation \ Path control	Single PCE (Not applicable)	Multiple PCE with inter-PCE	Multiple PCE w/o inter-PCE
PCE-VNTM cooperation	.....	Yes	Yes
Higher-layer signaling trigger	.....	Yes	Yes
NMS-VNTM cooperation (integrated flavor)	.....	.....	.....
NMS-VNTM cooperation (separate flavor)	.....	.....	.....

Not applicable because there are multiple PCEs

Typical models to be used

Note that:

- Since there is one PCE in each layer network, the path computation model "Single PCE path computation" is not applicable.

- For the other two path computation models "Multiple PCE with inter-PCE" and "Multiple PCE w/o inter-PCE", the possible combinations are the same as defined in [RFC5623]. More specifically:

  o The path control models "NMS-VNTM cooperation (integrated flavor)" and "NMS-VNTM cooperation (separate flavor)" are the typical models to be used in multi-layer network with GMPLS-SDN interworking. This is because in these two models, the path computation is triggered by the NMS or VNTM. And in SDN centralized controller system, the path computation requests are typically from the multi-layer SDN controller (acts as VNTM).

  o For the other two path control models "PCE-VNTM cooperation" and "Higher-layer signaling trigger", the path computation is

triggered by the NEs, i.e., NE performs PCC functions. These

two models are still possible to be used, although they are not the main methods.

7.3.2. Cross-layer Path Creation

In a multi-layer network, a lower-layer LSP in the lower-layer network can be created, which will construct a new link in the higher-layer network. Such lower-layer LSP is called Hierarchical LSP, or H-LSP for short, see [RFC6107].

The new link constructed by the H-LSP then can be used by the higher-layer network to create new LSPs.

As described in [RFC5212], two methods are introduced to create the H-LSP: the static (pre-provisioned) method and the dynamic (triggered) method.

1) Static (pre-provisioned) method

In this method, the H-LSP in the lower layer network is created in advance. After that, the higher layer network can create LSPs using the resource of the link constructed by the H-LSP.

The multi-layer SDN controller is responsible to decide the creation of H-LSP in the lower layer network if it acts as a VNTM. It then requests the lower-layer SDN controller to create the H-LSP via, for example, MPI interface under the ACTN architecture. See Section 3.3.2 of [TE-Tunnel].

The lower-layer SDN controller can trigger the GMPLS control plane to create the H-LSP. As a typical example, the PCInitiate message can be used for the communication between the lower-layer SDN controller and the source node of the H-LSP.

And the source node of the H-LSP can trigger the RSVP-TE signaling procedure to create the H-LSP, as described in [RFC6107].

2) Dynamic (triggered) method

In this method, the signaling of LSP creation in the higher layer network will trigger the creation of H-LSP in the lower layer network dynamically, if it is necessary.

In this case, after the cross-layer path is computed, the multi-layer SDN controller requests the higher-layer SDN controller for the cross-layer LSP creation. As a typical example, the MPI interface under the ACTN architecture could be used.

The higher-layer SDN controller can trigger the GMPLS control plane

to create the LSP in the higher-layer network. As a typical example,

the PCInitiate message can be used for the communication between the higher-layer SDN controller and the source node of the Higher-layer LSP, as described in Section 4.3 of [RFC8282]. At least two sets of ERO information should be included to indicate the routes of higher-layer LSP and lower-layer H-LSP.

The source node of the Higher-layer LSP follows the procedure defined in Section 4 of [RFC6001], to trigger the GMPLS control plane in both higher-layer network and lower-layer network to create the higher-layer LSP and the lower-layer H-LSP.

On success, the source node of the H-LSP should report the information of the H-LSP to the lower-layer SDN controller via, for example, PCRpt message.

7.3.3. Link Discovery

If the higher-layer network and the lower-layer network are under the same GMPLS control plane instance, the H-LSP can be an FA-LSP. Then the information of the link constructed by this FA-LSP, called FA, can be advertised in the routing instance, so that the higher-layer SDN controller can be aware of this new FA. [RFC4206] and the following updates to it (including [RFC6001] and [RFC6107]) describe the detail extensions to support advertisement of an FA.

If the higher-layer network and the lower-layer network are under separated GMPLS control plane instances, after an H-LSP is created in the lower-layer network, the link discovery procedure defined in LMP protocol ([RFC4204]) will be triggered in the higher-layer network to discover the information of the link constructed by the H-LSP. The information of this new link will be advertised to the higher-layer SDN controller.

7.4. Recovery

The GMPLS recovery functions are described in [RFC4426]. Span protection, end-to-end protection and restoration, are discussed with different protection schemes and message exchange requirements. Related RSVP-TE extensions to support end-to-end recovery is described in [RFC4872]. The extensions in [RFC4872] include protection, restoration, preemption, and rerouting mechanisms for an end-to-end LSP. Besides end-to-end recovery, a GMPLS segment recovery mechanism is defined in [RFC4873], which also intends to be compatible with Fast Reroute (FRR) (see [RFC4090] which defines RSVP-TE extensions for the FRR mechanism, and [RFC8271] which described the updates of GMPLS RSVP-TE protocol for FRR of GMPLS TE-LSPs).

7.4.1. Span Protection

Span protection refers to the protection of the link between two neighboring switches. The main protocol requirements include:

- Link management: Link property correlation on the link protection type;

- Routing: announcement of the link protection type;

- Signaling: indication of link protection requirement for that LSP.

GMPLS already supports the above requirements, and there are no new requirements in the scenario of interworking between GMPLS and centralized controller system.

7.4.2. LSP Protection

The LSP protection includes end-to-end and segment LSP protection. For both cases:

- In the provisioning phase:

  In both single-domain and multi-domain scenarios, the disjoint path computation can be done by the centralized controller system, as it has the global topology and resource view. And the path creation can be done by the procedure described in Section 7.2.

- In the protection switchover phase:

  In both single-domain and multi-domain scenarios, the existing standards provide the distributed way to trigger the protection switchover. For example, data plane Automatic Protection Switching (APS) mechanism described in [G.808.1], or GMPLS Notify mechanism described in [RFC4872] and [RFC4873]. In the scenario of interworking between GMPLS and centralized controller system, it is recommended to still use these distributed mechanisms rather than centralized mechanism (i.e., the controller triggers the protection switchover). This can significantly shorten the protection switching time.

7.4.3. Single-domain LSP Restoration

- Pre-planned LSP rerouting (including shared-mesh restoration):

  In pre-planned protecting, the protecting LSP is established only in the control plane in the provisioning phase, and will be activated in the data plane once failure occurs.

In the scenario of interworking between GMPLS and centralized controller system, the route of protecting LSP can be computed by the centralized controller system. This takes the advantage of making better use of network resource, especially for the resource sharing in shared-mesh restoration.

- Full LSP rerouting:

  In full LSP rerouting, the normal traffic will be switched to an alternate LSP that is fully established only after failure occurrence.

  As described in [RFC4872] and [RFC4873], the alternate route can be computed on demand when failure occurrence, or pre-computed and stored before failure occurrence.

  In a fully distributed scenario, the pre-computation method offers faster restoration time, but has the risk that the pre-computed alternate route may become out of date due to the changes of the network.

  In the scenario of interworking between GMPLS and centralized controller system, the pre-computation of the alternate route could be taken place in the centralized controller (and may be stored in the controller or the head-end node of the LSP). In this way, any changes in the network can trigger the refreshment of the alternate route by the centralized controller. This makes sure that the alternate route will not become out of date.

7.4.4. Multi-domain LSP Restoration

A working LSP may traverse multiple domains, each of which may or may not support GMPLS distributed control plane.

In the case that all the domains support GMPLS, both the end-to-end rerouting method and the domain segment rerouting method could be used.

In the case that only some of the domains support GMPLS, the domain segment rerouting method could be used in those GMPLS domains. For other domains which do not support GMPLS, other mechanisms may be used to protect the LSP segments, which are out of scope of this document.

1) End-to-end rerouting:

In this case, failure occurs on the working LSP inside any domain or on the inter-domain links will trigger the end-to-end restoration.

In both pre-planned and full LSP rerouting, the end-to-end protecting LSP could be computed by the centralized controller system, and could be created by the procedure described in Section 7.2. Note that the end-to-end protecting LSP may traverse different domains from the working LSP, depending on the result of multi-domain path computation for the protecting LSP.

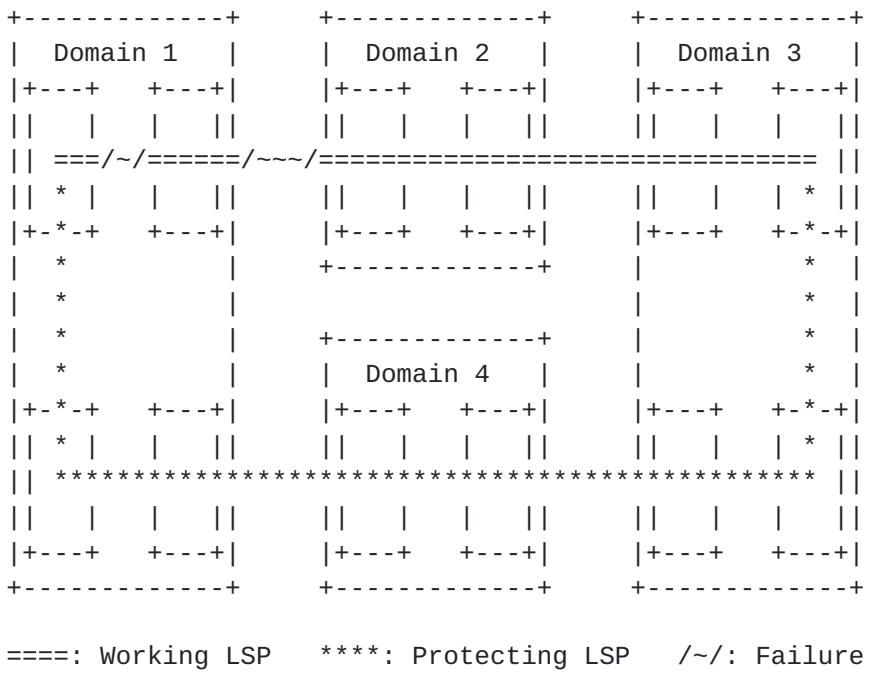
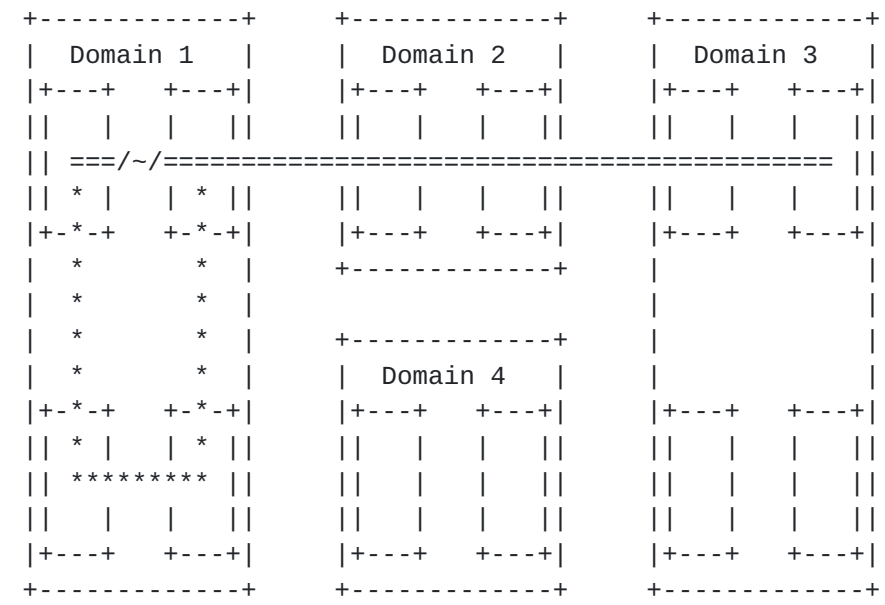


Figure 5: End-to-end restoration

2) Domain segment rerouting:

2.1) Intra-domain rerouting:

If failure occurs on the working LSP segment in a GMPLS domain, the segment rerouting ([RFC4873]) could be used for the working LSP segment in that GMPLS domain. Figure 6 shows an example of intra-domain rerouting.



====: Working LSP ****: Rerouting LSP segment /~/: Failure

Figure 6: Intra-domain segment rerouting

2.2) Inter-domain rerouting:

If intra-domain segment rerouting failed (e.g., due to lack of resource in that domain), or if failure occurs on the working LSP on an inter-domain link, the centralized controller system may coordinate with other domain(s), to find an alternative path or path segment to bypass the failure, and then trigger the inter-domain rerouting procedure. Note that the rerouting path or path segment may traverse different domains from the working LSP.

For inter-domain rerouting, the interaction between GMPLS and centralized controller system is needed:

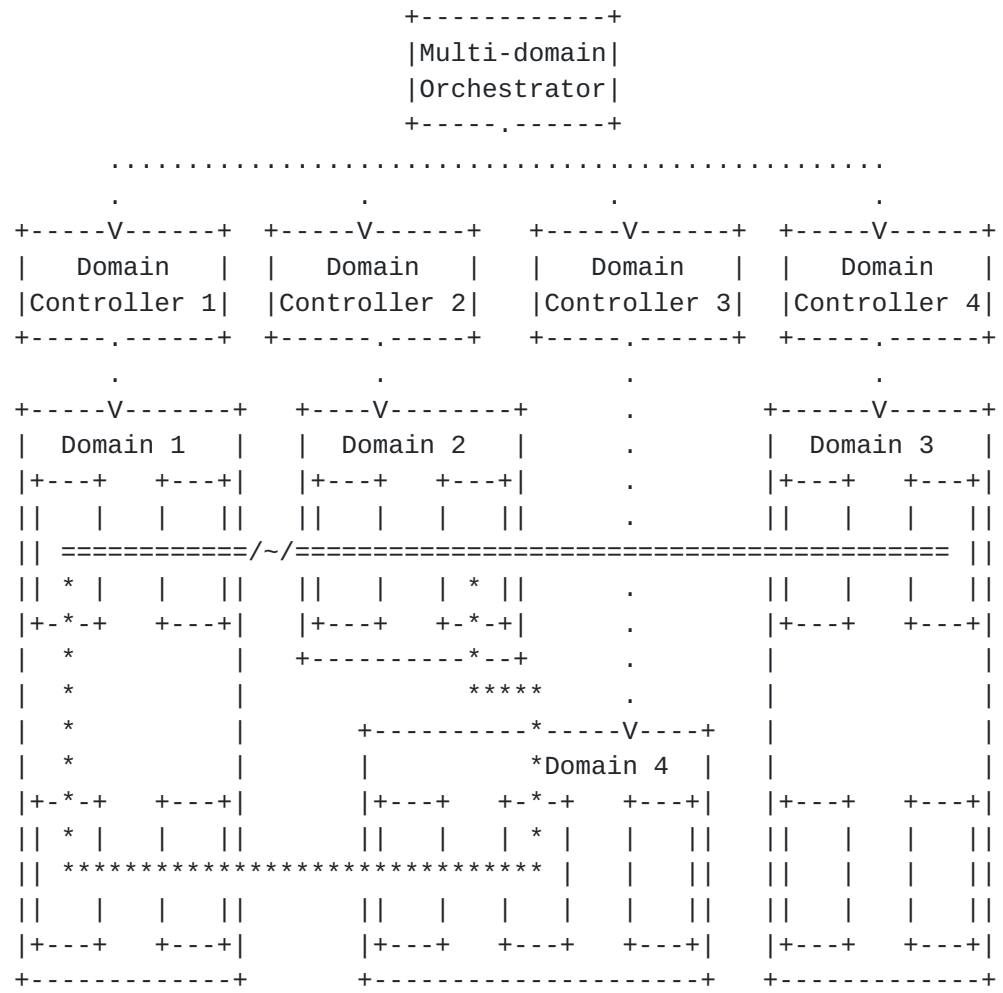
- Report of the result of intra-domain segment rerouting to its domain SDN controller, and then to the multi-domain orchestrator. The former one could be supported by the PCRpt message in [RFC8231], while the latter one could be supported by the MPI interface of ACTN.

- Report of inter-domain link failure to the two domain SDN controllers (by which the two ends of the inter-domain link are controlled respectively), and then to the multi-domain orchestrator. The former one could be done as described in Section 7.1 of this document, while the latter one could be supported by the MPI interface of ACTN.

- Computation of rerouting path or path segment crossing multi-

domains by the centralized controller system (see [PAT-COMP]);

- Creation of rerouting path segment in each related domain. The multi-domain orchestrator can send the path segment rerouting request to each related domain SDN controller via MPI interface, and then each domain SDN controller can trigger the creation of rerouting path segment in its domain. Note that the ingress and/or egress node of the rerouting path segment may be different from the working LSP segment in each related domain (e.g., Domain 1 and Domain 2 in Figure 7). Note also that the rerouting path segment may traverse a new domain which the working LSP does not traverse (e.g., Domain 4 in Figure 7).



====: Working LSP ****: Rerouting LSP segment /~/: Failure

Figure 7: Inter-domain segment rerouting

7.4.5. Fast Reroute

[RFC4090] defines two methods of fast reroute, the one-to-one backup method and the facility backup method. For both methods:

1) Path computation of protecting LSP:

In Section 6.2 of [RFC4090], the protecting LSP (detour LSP in one-to-one backup, or bypass tunnel in facility backup) could be computed by the Point of Local Repair (PLR) using, for example, Constraint-based Shortest Path First (CSPF) computation. In the scenario of interworking between GMPLS and centralized controller system, the protecting LSP could also be computed by the centralized controller system, as it has the global view of the network topology, resource and information of LSPs.

2) Protecting LSP creation:

In the scenario of interworking between GMPLS and centralized controller system, the Protecting LSP could still be created by the RSVP-TE signaling protocol as described in [RFC4090] and [RFC8271].

In addition, if the protecting LSP is computed by the centralized controller system, the Secondary Explicit Route Object defined in [RFC4873] could be used to explicitly indicate the route of the protecting LSP.

3) Failure detection and traffic switchover:

If a PLR detects that failure occurs, it is recommended to still use the distributed mechanisms described in [RFC4090] to switch the traffic to the related detour LSP or bypass tunnel, rather than in a centralized way. This can significantly shorten the protection switching time.

7.5. Controller Reliability

Given the important role in the network, the reliability of controller is critical. Once a controller is shut down, the network should operate as well. It can be either achieved by controller back up or functionality back up. There are several of controller backup or federation mechanisms in the literature. It is also more reliable to have some function back up in the network element, to guarantee the performance in the network.

8. Manageability Considerations

Each entity in the network, including both controllers and network elements, should be managed properly as it will interact with other entities. The manageability considerations in controller hierarchies and network elements still apply respectively. For the protocols applied in the network, manageability is also requested.

The responsibility of each entity should be clarified. The control of function and policy among different controllers should be consistent via proper negotiation process.

9. Security Considerations

This document provides the interwork between the GMPLS and controller hierarchies. The security requirements in both system still applies respectively. Protocols referenced in this document also have various security considerations, which is also expected to be satisfied.

Other considerations on the interface between the controller and the network element are also important. Such security includes the functions to authenticate and authorize the control access to the controller from multiple network elements. Security mechanisms on the controller are also required to safeguard the underlying network elements against attacks on the control plane and/or unauthorized usage of data transport resources.

10. IANA Considerations

This document requires no IANA actions.

11. References

11.1. Normative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC4206] Kompella, K. and Rekhter Y., "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

[RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

[RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, October 2008.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[RFC6001] Papadimitriou D., Vigoureux M., Shiomoto K., Brungard D. and Le Roux JL., "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, October 2010.

[RFC6107] Shiomoto K. and Farrel A., "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC 6107, February 2011.

[RFC6241] Enns, R., Bjorklund, M., Schoenwaelder J., Bierman A., "Network Configuration Protocol (NETCONF)", RFC 6241, June 2011.

[RFC7074] Berger, L. and J. Meuric, "Revised Definition of the GMPLS Switching Capability and Type Fields", RFC 7074, November 2013.

[RFC7491] King, D., Farrel, A., "A PCE-Based Architecture for Application-Based Network Operations", RFC7491, March 2015.

[RFC7926] Farrel, A., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D. and Zhang, X., "Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks", RFC7926, July 2016.

[RFC8040] Bierman, A., Bjorklund, M., Watsen, K., "RESTCONF Protocol", RFC 8040, January 2017.

[RFC8271] Taillon M., Saad T., Gandhi R., Ali Z. and Bhatia M., "Updates to the Resource Reservation Protocol for Fast Reroute of Traffic Engineering GMPLS Label Switched Paths", RFC 8271, October 2017.

[RFC8282] Oki E., Takeda T., Farrel A. and Zhang F., "Extensions to the Path Computation Element Communication Protocol (PCEP) for Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 8282, December 2017.

[RFC8453] Ceccarelli, D. and Y. Lee, "Framework for Abstraction and Control of Traffic Engineered Networks", RFC 8453, August 2018.

[RFC8795] Liu, X., Bryskin, I., Beeram, V., Saad, T., Shah, H., Gonzalez De Dios, O., "YANG Data Model for Traffic Engineering (TE) Topologies", RFC8795, August 2020.

11.2. Informative References

[RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC4202] Kompella, K., Ed. and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.

[RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.

[RFC4426] Lang, J., Ed., Rajagopalan, B., Ed., and D. Papadimitriou, Ed., "Generalized Multi-Protocol Label witching (GMPLS) Recovery Functional Specification", RFC 4426, March 2006.

[RFC5150] Ayyangar, A., Kompella, K., Vasseur, J.P., Farrel, A., "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February, 2008.

[RFC5212] Shiomoto K., Papadimitriou D., Le Roux JL., Vigoureux M. and Brungard D., "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, July 2008.

[RFC5623] Oki E., Takeda T., Le Roux JL. and Farrel A., "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 5623, September 2009.

[RFC7138] Ceccarelli, D., Ed., Zhang, F., Belotti, S., Rao, R., and J. Drake, "Traffic Engineering Extensions to OSPF for GMPLS Control of Evolving G.709 Optical Transport Networks", RFC 7138, March 2014.

[RFC7139] Zhang, F., Ed., Zhang, G., Belotti, S., Ceccarelli, D., and K. Pithewan, "GMPLS Signaling Extensions for Control of Evolving G.709 Optical Transport Networks", RFC 7139, March 2014.

[RFC7688] Lee, Y., Ed. and G. Bernstein, Ed., "GMPLS OSPF Enhancement for Signal and Network Element Compatibility for Wavelength Switched Optical Networks", RFC 7688, November 2015.

[RFC7689] Bernstein, G., Ed., Xu, S., Lee, Y., Ed., Martinelli, G., and H. Harai, "Signaling Extensions for Wavelength Switched Optical Networks", RFC 7689, November 2015.

[RFC7792] Zhang, F., Zhang, X., Farrel, A., Gonzalez de Dios, O., and D. Ceccarelli, "RSVP-TE Signaling Extensions in Support of Flexi-Grid Dense Wavelength Division Multiplexing (DWDM) Networks", RFC 7792, March 2016.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, September 2017.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", RFC 8281, October 2017.

[RFC8345] Clemm, A., Medved, J., Varga, R., Bahadur, N., Ananthakrishnan, H., Liu, X., "A YANG Data Model for Network Topologies", RFC 8345, March 2018.

[RFC8363] Zhang, X., Zheng, H., Casellas, R., Dios, O., and D. Ceccarelli, "GMPLS OSPF-TE Extensions in support of Flexi-grid DWDM networks", RFC8363, February 2017.

[PAT-COMP] Busi, I., Belotti, S., Lopez, V., Gonzalez de Dios, O., Sharma, A., Shi, Y., Vilalta, R., Setheraman, K., "Yang model for requesting Path Computation", draft-ietf-teas-yang-path-computation, work in progress.

[PCEP-LS] Dhody, D., Lee, Y., Ceccarelli, D., "PCEP Extensions for Distribution of Link-State and TE Information", draft-dhodylee-pce-pcep-ls, 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.

[sPCE-ID] Dugeon, O. et al., "PCEP Extension for Stateful Inter-Domain Tunnels", draft-ietf-pce-stateful-interdomain, work in progress.

[G.808.1] ITU-T, "Generic protection switching - Linear trail and subnetwork protection", G.808.1, May 2014.

12. Authors' Addresses

Haomian Zheng
Huawei Technologies
H1, Huawei Xiliu Beipo Village, Songshan Lake
Dongguan
Guangdong, 523808 China
Email: zhenghaomian@huawei.com

Xianlong Luo
Huawei Technologies
G1, Huawei Xiliu Beipo Village, Songshan Lake
Dongguan
Guangdong, 523808 China
Email: luoxianlong@huawei.com

Yunbin Xu
CAICT
Email: xuyunbin@caict.ac.cn

Yang Zhao
China Mobile
Email: zhaoyangyjy@chinamobile.com

Sergio Belotti
Nokia
Email: sergio.belotti@nokia.com

Dieter Beller
Nokia
Email: Dieter.Beller@nokia.com

Yi Lin
Huawei Technologies
H1, Huawei Xiliu Beipo Village, Songshan Lake
Dongguan
Guangdong, 523808 China
Email: yi.lin@huawei.com