Network Working Group Kohei Shiomoto (NTT)
Internet Draft Dimitri Papadimitriou (Alcatel)
Jean-Louis Le Roux (France Telecom)
Martin Vigoureux (Alcatel)
Deborah Brungard (AT&T)

Expires: July 2006 January 2006

Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN)

draft-ietf-ccamp-gmpls-mln-reqs-00.txt

Status of this Memo

Copyright Notice

Abstract

Most of the initial efforts on Generalized MPLS (GMPLS) have been

related to environments hosting devices with a single switching capability. The complexity raised by the control of such data planes is similar to that seen in classical IP/MPLS networks.

By extending MPLS to support multiple switching technologies, GMPLS provides a comprehensive framework for the control of a multi-

layered network of either a single switching technology or multiple switching technologies. In GMPLS, a switching technology domain defines a region, and a network of multiple switching types is referenced in this document as a multi-region network (MRN). When referring in general to a layered network, which may consist of either a single or multiple regions, this document uses the term, Multi-layer Network (MLN). This draft defines a framework for GMPLS based multi-region/multi-layer networks and lists a set of functional requirements.

Table of Contents

1. Introduction..2
2. Conventions used in this document...............................4
3. Positioning...4
3.1. Data plane layers and control plane regions..................5
3.2. Services...5
3.3. Vertical and Horizontal interaction and integration..........6
4. Key concepts of GMPLS-based MLNs and MRNs.......................6
4.1. Interface Switching Capability...............................7
4.2. Multiple Interface Switching Capabilities....................7
4.2.1. Networks with multi-switching capable hybrid nodes.........8
4.3. Integrated Traffic Engineering (TE) and Resource Control.....9
4.3.1. Triggered signaling..9
4.3.2. FA-LSP..10
4.3.3. Virtual network topology (VNT)............................10
5. Service networks provided over MRN/MLN.........................11
6. Requirements...11
6.1. Scalability...11
6.2. LSP resource utilization....................................12
6.2.1. FA-LSP release and setup..................................12
6.2.2. Virtual TE-Link...12
6.3. LSP Attribute inheritance...................................14
6.4. Verification of the LSP.....................................14
6.5. Disruption minimization.....................................14
6.6. Stability...14

6.7. Computing paths with and without nested signaling...........15
6.8. Handling single-switching and multi-switching type capable
nodes...16
6.9. Advertisement of the available adaptation resource..........16
7. Security Considerations...17
8. References..17
8.1. Normative Reference...17
8.2. Informative References..18
9. Author's Addresses..18
10. Intellectual Property Considerations...........................19
11. Full Copyright Statement.......................................20

1. Introduction

Generalized MPLS (GMPLS) extends MPLS to handle multiple switching technologies: packet switching, layer-two switching, TDM switching, wavelength switching, and fiber switching (see [RFC3945]). The Interface Switching Capability (ISC) concept is introduced for these switching technologies and is designated as follows: PSC (packet switch capable), L2SC (Layer-2 switch capable), TDM (Time Division Multiplex capable), LSC (lambda switch capable), and FSC (fiber switch capable).

Service providers may operate networks where multiple different switching technologies exist. The representation, in a GMPLS control plane, of a switching technology domain is referred to as a region [HIER].

A switching type describes the ability of a node to forward data of a particular data plane technology, and uniquely identifies a network region. A layer describes a data plane switching granularity level (e.g. VC4, VC-12). A data plane layer is associated with a region in the control plane (e.g. VC4 associated to TDM, IP associated to PSC). However, more than one data plane layer can be associated to the same region (e.g. both VC4 and VC12 are associated to TDM). Thus, a control plane region, identified by its switching type value (e.g. TDM), can itself be sub-divided into smaller granularity based on the bandwidth that defines the "data plane switching layers" e.g. from VC-11 to VC4-256c. The Interface Switching Capability Descriptor (ISCD) [GMPLS-RTG], identifying the interface switching type, the encoding type and the switching bandwidth granularity, enable the characterization of the associated layers.

A network comprising transport nodes with multiple data plane layers of either the same ISC or different ISCs, controlled by a single GMPLS control plane instance, is called a Multi-Layer Network (MLN). To differentiate a network supporting LSPs of different switching technologies (ISCs) from a single region network, a network supporting more than one switching technology is called a Multi-Region Network (MRN).

MRNs can be categorized according to the distribution of the switching type values amongst the LSRs:
- Network elements are single switching capable LSRs and different types of LSRs form the network.
- Network elements are multi-switching capable LSRs i.e. nodes hosting at least more than one switching capability. Multi-switching capable LSRs are further classified as "simplex" and "hybrid" nodes (see Section 4.2).
- Any combination of the above two elements. A network composed of both single and multi-switching capable LSRs.

Since GMPLS provides a comprehensive framework for the control of different switching capabilities, a single GMPLS instance may be used to control the MRNs/MLNs enabling rapid service provisioning and efficient traffic engineering across all switching capabilities. In such networks, TE Links are consolidated into a single Traffic Engineering Database (TED). Since this TED contains the information relative to all the different regions/layers existing in the network, a path across multiple regions/layers can be computed using this TED. Thus optimization of network resources can be achieved across multiple regions/layers.

Consider, for example, a MRN consisting of IP/MPLS routers and TDM cross-connects. Assume that a packet LSP is routed between source and destination IP/MPLS routers, and that the LSP can be routed across the PSC-region (i.e. utilizing only resources of the IP/MPLS level topology). If the performance objective for the LSP is not satisfied, new TE links may be created between the IP/MPLS routers across the TDM-region (for example, VC-12 links) and the LSP can be routed over those links. Further, even if the LSP can be successfully established across the PSC-region, TDM hierarchical

LSPs across the TDM region between the IP/MPLS routers may be established and used if doing so enables meeting an operator's objectives on network resources available (e.g. link bandwidth, and adaptation port between regions) across the multiple regions. The same considerations hold when VC4 LSPs are provisioned to provide extra flexibility for the VC12 and/or VC11 layers in a MLN.

This document describes the requirements to support multi-region/multi-layer networks. There is no intention to specify solution specific elements in this document. The applicability of existing GMPLS protocols and any protocol extensions to the MRN/MLN will be addressed in separate documents [MRN-EVAL].

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Positioning

A multi-region network (MRN) is always a multi-layer network (MLN) since the network devices on region boundaries bring together different ISCs. A MLN, however, is not necessarily a MRN since multiple layers could be fully contained within a single region. For example, VC12, VC4, and VC4-4c are different layers of the TDM region.

3.1. Data plane layers and control plane regions

A data plane layer is a collection of network resources capable of terminating and/or switching data traffic of a particular format. These resources can be used for establishing LSPs or connectionless traffic delivery. For example, VC-11 and VC4-64c represent two different layers.

From the control plane viewpoint, an LSP region is defined as a set of one or several data plane layers that share the same type of switching technology, that is, the same switching type. The currently defined regions are: PSC, L2SC, TDM, LSC, and FSC regions.

Hence, an LSP region is a technology domain (identified by the ISC type) for which data plane resources (i.e. data links) are represented into the control plane as an aggregate of TE information associated with a set of links (i.e. TE links). For example VC-11 and VC4-64c capable TE links are part of the same TDM region. Multiple layers can thus exist in a single region network.

Note also that the region is a control plane only concept. That is, layers of the same region share the same switching technology and, therefore, need the same set of technology specific signaling objects.

3.2. Services

A service provider's network may be divided into different service layers. The customer's network is considered from the provider's perspective as the highest service layer. It interfaces to the highest service layer of the service provider's network. Connectivity across the highest service layer of the service provider's network may be provided with support from successively lower service layers. Service layers are realized via a hierarchy of network layers located generally in several regions and commonly arranged according to the switching capabilities of network devices.

For instance some customers purchase Layer 1 (i.e. transport) services from the service provider, some Layer 2 (e.g. ATM), while others purchase Layer 3 (IP/MPLS) services. The service provider realizes the services by a stack of network layers located within one or more network regions. The network layers are commonly arranged according to the switching capabilities of the devices in the networks. Thus, a customer network may be provided on top of the GMPLS-based multi-region/multi-layer network. For example, a Layer One service (realized via the network layers of TDM, and/or LSC, and/or FSC regions) may support a Layer Two network (realized via ATM VP/VC) which may itself support a Layer Three network (IP/MPLS region). The supported data plane relationship is a data-plane client-server relationship where the lower layer provides a

service for the higher layer using the data links realized in the lower layer.

Services provided by a GMPLS-based multi-region/multi-layer network are referred to as "Multi-region/Multi-layer network services". For example legacy IP and IP/MPLS networks can be supported on top of multi-region/multi-layer networks. It has to be emphasized that delivery of such diverse services is a strong motivator for the deployment of multi-region/multi-layer networks.

3.3. Vertical and Horizontal interaction and integration

Vertical interaction is defined as the collaborative mechanisms within a network element that is capable of supporting more than one switching capability and of realizing the client/server relationships between them. Protocol exchanges between two network controllers managing different regions are also a vertical interaction. Integration of these interactions as part of the control plane is referred to as vertical integration. The latter refers thus to the collaborative mechanisms within a single control plane instance driving multiple switching capabilities. Such a concept is useful in order to construct a framework that facilitates efficient network resource usage and rapid service provisioning in carrier's networks that are based on multiple switching technologies.

In a strict sense, horizontal interaction is defined as the protocol exchange between network controllers that manage transport nodes within a given region (i.e. nodes with the same switching capability). For instance, the control plane interaction between two LSC network elements is an example of horizontal interaction. GMPLS protocol operations handle horizontal interactions within the same routing area. The case where the interaction takes place across a domain boundary, such as between two routing areas within the same network layer, is currently being evaluated as part of the inter-domain work [Inter-domain], and is referred to as horizontal integration. Thus horizontal integration refers to the collaborative mechanisms between network partitions and/or administrative divisions such as routing areas or autonomous systems. This distinction gets blurred when administrative domains match layer boundaries. Horizontal interaction is extended to cover such case. For example, the collaborative mechanisms in place between two lambda switching capable areas relate to horizontal integration. On the other hand, the collaborative mechanisms in place in a network that supports IP/MPLS over TDM switching could be described as vertical and horizontal integration in the case where each network belongs to a separate area.

4. Key concepts of GMPLS-based MLNs and MRNs

A network comprising transport nodes with multiple data plane layers of either the same ISC or different ISCs, controlled by a single GMPLS control plane instance, is called a Multi-Layer Network (MLN). A sub-set of MLNs consists of networks supporting LSPs of different switching technologies (ISCs). A network supporting more than one switching technology is called a Multi-Region Network (MRN).

4.1. Interface Switching Capability

The Interface Switching Capability (ISC) is introduced in GMPLS to support various kinds of switching technology in a unified way [GMPLS-ROUTING]. An ISC is identified via a switching type.

A switching type (also referred to as the switching capability types) describes the ability of a node to forward data of a particular data plane technology, and uniquely identifies a network region. The following ISC types (and, hence, regions) are defined: PSC, L2SC, TDM, LSC, and FSC. Each end of a data link (more precisely, each interface connecting a data link to a node) in a GMPLS network is associated with an ISC.

The ISC value is advertised as a part of the Interface Switching Capability Descriptor (ISCD) attribute (sub-TLV) of a TE link end associated with a particular link interface. Apart from the ISC, the ISCD contains information, such as the encoding type, the bandwidth granularity, and the unreserved bandwidth on each of eight priorities at which LSPs can be established. The ISCD does not "identify" network layers, it uniquely characterizes information associated to one or more network layers.

TE link end advertisements may contain multiple ISCDs. This can be interpreted as advertising a multi-layer (or multi-switching) TE link end.

4.2. Multiple Interface Switching Capabilities

In a MLN, network elements may be single-switching or multi-switching type capable nodes. Single-switching type capable nodes advertise the same ISC value as part of their ISCD sub-TLV(s) to describe the termination capabilities of their TE Link(s). This case is described in [GMPLS-ROUTING].

Multi-switching capable LSRs are classified as "simplex" and "hybrid" nodes. Simplex and Hybrid nodes are categorized according to the way they advertise these multiple ISCs:

- A simplex node can terminate links with different switching capabilities each of them connected to the node by a single link interface. So, it advertises several TE Links each with a single

ISC value as part of its ISCD sub-TLVs. For example, an LSR with PSC and TDM links each of which is connected to the LSR via single interface.

- A hybrid node can terminate links with different switching capabilities terminating on the same interface. So, it advertises at least one TE Link containing more than one ISCDs with different ISC values. For example, a node comprising of PSC and TDM links, which are interconnected via internal links. The external interfaces connected to the node have both PSC and TDM capability.

Additionally TE link advertisements issued by a simplex or a hybrid node may need to provide information about the node's internal adaptation capabilities between the switching technologies supported. That is, the node's capability to perform layer border node functions.

4.2.1. Networks with multi-switching capable hybrid nodes

The network contains at least one hybrid node, zero or more simplex nodes, and a set of single switching capable nodes.

Figure 5a shows an example hybrid node. The hybrid node has two switching elements (matrices), which support, for instance, TDM and PSC switching respectively. The node terminates two PSC and TDM links (Link1 and Link2 respectively). It also has internal link connecting the two swtching elements.

The two switching elements are internally interconnected in such a way that it is possible to terminate some of the resources of, say, Link2 and provide through them adaptation for PSC traffic received/sent over the PSC interface (#b). This situation is modeled in GMPLS by connecting the local end of Link2 to the TDM switching element via an additional interface realizing the termination/adaptation function. Two ways are possible to set up PSC LSPs. Available resource advertisement e.g. Unreserved and Min/Max LSP Bandwidth should cover both two ways.

```
                       Network element
                  .............................
                  :           --------        :
                  :          |  PSC    |      :
      Link1 -------------<->--|#a       |      :
                  :  +--<->---|#b       |      :
                  :  |         --------       :
        TDM       :  |        ----------      :
        +PSC      :  +--<->--|#c  TDM    |     :
      Link2 ------------<->--|#d         |     :
                  :           ----------      :
```



```
                  :............................
```

Figure 5a. Hybrid node.

4.3. Integrated Traffic Engineering (TE) and Resource Control

In GMPLS-based multi-region/multi-layer networks, TE Links are consolidated into a single Traffic Engineering Database (TED). Since this TED contains the information relative to all the layers of all regions in the network, a path across multiple layers (possibly crossing multiple regions) can be computed using the information in this TED. Thus optimization of network resources across the multiple layers of the same region and multiple regions can be achieved.

These concepts allow for the operation of one network layer over the topology (that is, TE links) provided by other network layer(s) (for example, the use of a lower layer LSC LSP carrying PSC LSPs). In turn, a greater degree of control and inter-working can be achieved, including (but not limited too):
- dynamic establishment of Forwarding Adjacency LSPs (see Section 4.3.3)
- provisioning of end-to-end LSPs with dynamic triggering of FA LSPs

Note that in a multi-layer/multi-region network that includes multi-switching type capable nodes, an explicit route used to establish an end-to-end LSP can specify nodes that belong to different layers or regions. In this case, a mechanism to control the dynamic creation of FA LSPs may be required.

There is a full spectrum of options to control how FA LSPs are dynamically established. It can be subject to the control of a policy, which may be set by a management component, and which may require that the management plane is consulted at the time that the FA LSP is established. Alternatively, the FA LSP can be established at the request of the control plane without any management control.

4.3.1. Triggered signaling

When an LSP crosses the boundary from an upper to a lower layer, it may be nested into a lower layer FA LSP that crosses the lower layer. From signaling perspective, there are two alternatives to establish lower layer FA LSP: static and dynamic. Decision will be made either by the operator or automatically using features like TE auto-mesh, for instance. If such a lower layer LSP does not already exist, the LSP may be established dynamically. Such a mechanism is referred to as "triggered signaling".

4.3.2. FA-LSP

Once an LSP is created across a layer, it can be used as a data link in an upper layer.

Furthermore, it can be advertised as a TE-link, allowing other nodes to consider the LSP as a TE link for their path computation [HIER]. An LSP created either statically or dynamically by one instance of the control plane and advertised as a TE link into the same instance of the control plane is called a FA-LSP. The TE-link associated to an FA-LSP is called an FA. An FA has the special characteristic of not requiring a routing adjacency (peering) between its ends yet still guaranteeing control plane connectivity between the FA-LSP ends based on a signaling adjacency. A FA is a useful and powerful tool for improving the scalability of GMPLS Traffic Engineering (TE) capable networks.

The aggregation of LSPs enables the creation of a vertical (nested) LSP Hierarchy. A set of FA-LSPs across or within a lower layer can be used during path selection by a higher layer LSP. Likewise, the higher layer LSPs may be carried over dynamic data links realized via LSPs (just as they are carried over any "regular" static data links). This process requires the nesting of LSPs through a hierarchical process [HIER]. The TED contains a set of LSP advertisements from different layers that are identified by the ISCD contained within the TE link advertisement associated with the LSP [GMPLS-ROUTING].

4.3.3. Virtual network topology (VNT)

A set of one or more of lower-layer LSPs provides information for efficient path handling in upper-layer(s) of the MLN, or, in other words, provides a virtual network topology to the upper-layers. For instance, a set of LSPs, each of which is supported by an LSC LSP, provides a virtual network topology to the layers of a PSC region, assuming that the PSC region is connected to the LSC region. Note that a single lower-layer LSP is a special case of VNT. The virtual network topology is configured by setting up or tearing down the LSC LSPs. By using GMPLS signaling and routing protocols, the virtual network topology can be adapted to traffic demands.

Reconfiguration of the virtual network topology may be triggered by traffic demand change, topology configuration change, signaling request from the upper layer, and network failure. For instance, by reconfiguring the virtual network topology according to the traffic demand between source and destination node pairs, network performance factors, such as maximum link utilization and residual capacity of the network, can be optimized [MAMLTE]. Reconfiguration is performed by computing the new VNT from the traffic demand matrix and optionally from the current VNT. Exact details are

outside the scope of this document. However, this method may be tailored according to the Service Provider's policy regarding network performance and quality of service (delay, loss/disruption, utilization, residual capacity, reliability).

5. Service networks provided over MRN/MLN

A customer network may be provided on top of a server MRN/MLN network (such as a transport network) which is operated by a service provider. For example legacy IP or IP/MPLS networks can be provided on top of GMPLS packet or optical networks [IW-MIG-FW]. The relationship between the networks is a client/server relationship and, such services are referred to as "MRN/MLN services".

The customer network may be provided either as part of the MRN/MLN or in a separate network instance distinct from the MRN/MLN. There could also be an administrative boundary between the customer network and the MRN/MLN operated by the service provider. All requirements described in this document SHOULD be applicable if there is an administrative boundary between the customer network and the MRN/MLN operated by service provider.

Impact on the customer network design, operation, and administration SHOULD be minimized. For instance, the design for address assignment and IGP area division should be kept independent from the underlying MRN/MLN.

The MRN/MLN SHOULD provide mechanisms to allow an administrative boundary between the customer network and the MRN/MLN.

6. Requirements

6.1. Scalability

The MRN/MLN relies on a unified traffic engineering and routing model. The TED in each LSR is populated with TE-links from all layers of all regions. This may lead to a huge amount of information that has to be flooded and stored within the network. Furthermore, path computation times, which may be of great importance during restoration, will depend on the size of the TED.

Thus MRN/MLN routing mechanisms MUST be designed to scale well with an increase of any of the following:

- Number of nodes
- Number of TE-links (including FA-LSPs)
- Number of LSPs
- Number of regions and layers
- Number of ISCDs per TE-link.

6.2. LSP resource utilization

It MUST be possible to utilize network resources efficiently. Particularly, resource usage in all layers SHOULD be optimized as a whole (i.e. across all layers), in a coordinated manner, (ie taking all layers into account). The number of lower-layer LSPs carrying upper-layer LSPs SHOULD be minimized as much as possible (Note that multiple LSPs may be used for load balance) . Unneccesary lower-layer LSPs SHOULD be avoided.

6.2.1. FA-LSP release and setup

Statistical multiplexing can only be employed in PSC and L2SC regions. A PSC or L2SC LSP may or may not consume the maximum reservable bandwidth of the FA LSP that carries it. On the other hand, a TDM, or LSC LSP always consumes a fixed amount of bandwidth as long as it exists (and is fully instantiated) because statistical multiplexing is not available.

If there is low traffic demand, some FA LSPs, which do not carry any LSP may be released so that resources are released. Note that if a small fraction of the available bandwidth is still under use, the nested LSPs can also be re-routed optionally using the make-before-break technique. Alternatively, the FA LSPs may be retained for future usage. Release or retention of underutilized FA LSPs is a policy decision.

As part of the re-optimization process, the solution MUST allow rerouting of FA LSPs while keeping interface identifiers of corresponding TE links unchanged.

Additional FA LSPs MAY also be created based on policy, which might consider residual resources and the change of traffic demand across the region. By creating the new FA LSPs, the network performance such as maximum residual capacity may increase.

As the number of FA LSPs grows, the residual resource may decrease. In this case, re-optimization of FA LSPs MAY be invoked according the policy.

Any solution MUST include measures to protect against network destabilization caused by the rapid set up and tear down of LSPs as traffic demand varies near a threshold.

6.2.2. Virtual TE-Link

It may be considered disadvantageous to fully instantiate (i.e. pre-provision) the set of lower layer LSPs since this may reserve bandwidth that could be used for other LSPs in the absence of the upper-layer traffic.

However, in order to provision upper-layer LSPs across the lower-layer, the LSPs MAY still be advertised into the upper-layer as though they had been fully established. Such TE links that represent the possibility of an underlying LSP are termed "virtual TE-link". Note that this is not a mandatory (MUST) requirement since even if there are no LSPs advertised to the higher layer, it is possible to route an upper layer LSP into a lower layer based on the lower layer's TE-links and making assumptions that proper hierarchical LSPs in the lower layer will be dynamically created as needed.

If an upper-layer LSP that makes use of a virtual TE-Link is set up, the underlying LSP MUST be immediately signaled in the lower layer if it has not been established.

If virtual TE-Links are used in place of pre-established LSPs, the TE links across the upper-layer can remain stable using pre-computed paths while wastage of bandwidth within the lower-layer and unnecessary reservation of adaptation ports at the border nodes can be avoided.

The concept of VNT can be extended to allow the virtual TE-links to form part of the VNT. The combination of the fully provisioned TE-links and the virtual TE-links defines the VNT across the lower layer.

The solution SHOULD provide operations to facilitate the build-up of such virtual TE-links, taking into account the (forecast) traffic demand and available resource in the lower-layer.

Virtual TE-links MAY be modified dynamically (by adding or removing virtual TE links) according to the change of the (forecast) traffic demand and the available resource in the lower-layer.

Any solution MUST include measures to protect against network destabilization caused by the rapid changes in the virtual network topology as traffic demand varies near a threshold.

The VNT can be changed by setting up and/or tearing down virtual TE links as well as by modifying real links (i.e. the fully provisioned LSPs).

The maximum number of virtual TE links that can be configured SHOULD be well-engineered.

How to design the VNT and how to manage it are out of scope of this document and will be treated in a companion document on solution.

6.3. LSP Attribute inheritance

TE-Link parameters SHOULD be inherited from the parameters of the LSP that provides the TE link, and so from the TE links in the lower layer that are traversed by the LSP.
These include:
- Interface Switching Capability
- TE metric
- Maximum LSP bandwidth per priority level
- Unreserved bandwidth for all priority levels
- Maximum Reservable bandwidth
- Protection attribute
- Minimum LSP bandwidth (depending on the Switching Capability)

Inheritance rules MUST be applied based on specific policies. Particular attention should be given to the inheritance of TE metric (which may be other than a strict sum of the metrics of the component TE links at the lower layer) and protection attributes.

6.4. Verification of the LSP

When the LSP is created, it SHOULD be verified that it has been established before it can be used by an upper layer LSP. Note, this is not within the GMPLS capability scope for non-PSC interfaces.

6.5. Disruption minimization

When reconfiguring the VNT according to a change in traffic demand, the upper-layer LSP might be disrupted. Such disruption MUST be minimized.

When residual resource decreases to a certain level, some LSPs MAY be released according to local or network policies. There is a trade-off between minimizing the amount of resource reserved in the lower layer LSPs and disrupting higher layer traffic (i.e. moving the traffic to other TE-LSPs so that some LSPs can be released). Such traffic disruption MAY be allowed but MUST be under the control of policy that can be configured by the operator. Any repositioning of traffic MUST be as non-disruptive as possible (for example, using make-before-break).

6.6. Stability

The path computation is dependent on the network topology and associated link state. The path computation stability of an upper layer may be impaired if the VNT changes frequently and/or if the status and TE parameters (TE metric for instance) of links in the virtual network topology changes frequently.

In this context, robustness of the VNT is defined as the capability to smooth changes that may occur and avoid their propagation into higher layers. Changes of the VNT may be caused by the creation and/or deletion of several LSPs.

Creation and deletion of LSPs MAY be triggered by adjacent layers or through operational actions to meet traffic demand change, topology change, signaling request from the upper layer, and network failure. Routing robustness SHOULD be traded with adaptability with respect to the change of incoming traffic requests.

A full mesh of LSPs MAY be created between every pair of border nodes of the PSC region. The merit of a full mesh of PSC TE-LSPs is that it provides stability to the PSC-level routing. That is, the forwarding table of an PSC-LSR is not impacted by re-routing changes within the lower-layer (e.g., TDM layer). Further, there is always full PSC reachability and immediate access to bandwidth to support PSC LSPs. But it also has significant drawbacks, since it requires the maintenance of n^2 RSVP-TE sessions, which may be quite CPU and memory consuming (scalability impact). Also this may lead to significant bandwidth wasting if LSP with a certain amount of reserved bandwidth is used.
Note that the use of virtual TE-links solves the bandwidth wasting issue, and may reduce the control plane overload.

6.7. Computing paths with and without nested signaling

Path computation MAY take into account LSP region and layer boundaries when computing a path for an LSP. For example, path computation MAY restrict the path taken by an LSP to only the links whose interface switching capability is PSC.

Interface switching capability is used as a constraint in computing the path. A TDM-LSP is routed over the topology composed of TE links of the same TDM layer. In calculating the path for the LSP, the TE database MAY be filtered to include only links where both end include requested LSP switching type. In this way hierarchical routing is done by using a TE database filtered with respect to switching capability (that is, with respect to particular layer).

If triggered signaling is allowed, the path computation mechanism MAY produce a route containing multiple layers/ regions. The path is computed over the multiple layers/regions even if the path is not "connected" in the same layer as the endpoints of the path exist. Note that here we assume that triggered signaling will be invoked to make the path "connected", when the upper-layer signaling request arrives at the boundary node.

The upper-layer signaling request may contain a loose ERO, and the boundary node is responsible for creation of the lower-layer FA-LSP. When the boundary node receives the signaling setup request and determines that it has to expand the loose ERO content by creating the lower-layer FA-LSP, it will create the lower layer FA-LSP accordingly. Once the lower-layer LSP is established, the ERO contents for the upper-layer signaling setup request are expanded to include the lower-layer FA-LSP and signaling setup for the upper-layer LSP are carried in-band of the lower-layer LSP.

The upper-layer signaling request may contain a strict ERO specifying the lower layer FA-LSP route. In this case, the boundary node is responsible for decision as to which it should use the path contained in the strict ERO or it should re-compute the path within in the lower-layer.

Even in case the lower-layer FA-LSPs are already established, a signaling request may also be encoded as loose ERO. In this situation, it is up to the boundary node to decide whether it should a new lower-layer FA-LSP or it should use the existing lower-layer FA-LSPs.

We should note that the lower-layer FA-LSP can be advertised just as an FA-LSP in the upper-layer or an IGP adjacency can be brought up on the lower-layer FA-LSP.

6.8. Handling single-switching and multi-switching type capable nodes

The MRN/MLN can consist of single-switching type capable and multi-switching type capable nodes. The path computation mechanism in the MLN SHOULD be able to compute paths consisting of any combination of such nodes.

Both single switching capable and multi-switching (simplex or hybrid) capable nodes could play the role of layer boundary. MRN/MLN Path computation SHOULD handle TE topologies built of any combination of single switching, simplex and hybrid nodes

6.9. Advertisement of the available adaptation resource

A hybrid node SHOULD maintain resources and advertise the resource information on its internal links, the links required for vertical (layer) integration. Likewise, path computation elements SHOULD be prepared to use the availability of termination/adaptation resources as a constraint in MRN/MLN path computations to reduce the higher layer LSP setup blocking probability because of the lack

of necessary termination/ adaptation resources in the lower layer(s).

The advertisement of the adaptation capability to terminate LSPs of lower-region and forward traffic in the upper-region is REQUIRED, as it provides critical information when performing multi-region path computation.

The mechanism SHOULD cover the case where the upper-layer links which are directly connected to upper-layer switching element and the ones which are connected through internal links between upper-layer element and lower-layer element coexist (See section 4.2.1).

7. Security Considerations

The current version of this document does not introduce any new security considerations as it only lists a set of requirements. In the future versions, new security requirements may be added.

8. References

8.1. Normative Reference

[RFC3979] Bradner, S., "Intellectual Property Rights in IETF Technology", BCP 79, RFC 3979, March 2005.

[GMPLS-ROUTING] K.Kompella and Y.Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching," draft-ietf-ccamp-gmpls-routing-09.txt, October 2003 (work in progress).

[Inter-domain] A.Farrel, J-P. Vasseur, and A.Ayyangar, "A framework for inter-domain MPLS traffic engineering," draft-ietf-ccamp-inter-domain-framework, work in progress.

[HIER] K.Kompella and Y.Rekhter, "LSP hierarchy with generalized MPLS TE," draft-ietf-mpls-lsp-hierarchy-08.txt, work in progress, Sept. 2002.

[STITCH] Ayyangar, A. and Vasseur, JP., "Label Switched Path Stitching with Generalized MPLS Traffic Engineering", draft-ietf-ccamp-lsp-stitching, work in progress.

[LMP] J. Lang, "Link management protocol (LMP)," draft- ietf-ccamp-lmp-10.txt (work in progress), October 2003.

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

8.2. Informative References

[MAMLTE] K. Shiomoto et al., "Multi-area multi-layer traffic engineering using hierarchical LSPs in GMPLS networks", draft-shiomoto-multiarea-te, work in progress.

[MRN-EVAL] Le Roux, J.L., Brungard, D., Oki, E., Papadimitriou, D., Shiomoto, K., Vigoureux, M.,"Evaluation of existing GMPLS Protocols against Multi Layer and Multi Region Networks (MLN/MRN)", draft-leroux-ccamp-gmpls-mrn-eval, work in progress.

[IW-MIG-FW] Shiomoto, K., Papadimitriou, D., Le Roux, J.L., Brungard, D., Oki, E., Inoue, I., " Framework for IP/MPLS-GMPLS interworking in support of IP/MPLS to GMPLS migration ", draft-shiomoto-ccamp-mpls-gmpls-interwork-fmwk-00.txt, work in progress.

9. Author's Addresses

Kohei Shiomoto
NTT Network Service Systems Laboratories
3-9-11 Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: shiomoto.kohei@lab.ntt.co.jp

Dimitri Papadimitriou
Alcatel
Francis Wellensplein 1,
B-2018 Antwerpen, Belgium
Phone : +32 3 240 8491
Email: dimitri.papadimitriou@alcatel.be

Jean-Louis Le Roux
France Telecom R&D,
Av Pierre Marzin,
22300 Lannion, France
Email: jeanlouis.leroux@francetelecom.com

Martin Vigoureux
Alcatel
Route de Nozay, 91461 Marcoussis cedex, France
Phone: +33 (0)1 69 63 18 52

Email: martin.vigoureux@alcatel.fr

Deborah Brungard
AT&T
Rm. D1-3C22 - 200
S. Laurel Ave., Middletown, NJ 07748, USA
Phone: +1 732 420 1573
Email: dbrungard@att.com

Contributors

Eiji Oki (NTT Network Service Systems Laboratories)
3-9-11 Midori-cho, Musashino-shi, Tokyo 180-8585, Japan
Phone: +81 422 59 3441 Email: oki.eiji@lab.ntt.co.jp

Ichiro Inoue (NTT Network Service Systems Laboratories)
3-9-11 Midori-cho, Musashino-shi, Tokyo 180-8585, Japan
Phone: +81 422 59 3441 Email: ichiro.inoue@lab.ntt.co.jp

Emmanuel Dotaro (Alcatel)
Route de Nozay, 91461 Marcoussis cedex, France
Phone : +33 1 6963 4723 Email: emmanuel.dotaro@alcatel.fr

10. Intellectual Property Considerations

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

The IETF has been notified by Tellabs Operations, Inc. of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information, see http://www.ietf.org/ietf/IPR/tellabs-ipr-draft-shiomoto-ccamp-gmpls-mrn-reqs.txt

11. Full Copyright Statement