

CCAMP Working Group
Internet Draft
Intended status: Standard Track
Expires: April 21, 2013

Zafar Ali
George Swallow
Clarence Filsfils
Matt Hartley
Cisco Systems
Kenji Kumaki
KDDI Corporation
Ruediger Kunze
Deutsche Telekom AG

October 22, 2012

Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)
extension for recording TE Metric of a Label Switched Path
draft-ali-ccamp-te-metric-recording-03.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 21, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in

Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Abstract

There are many scenarios in which Traffic Engineering (TE) metrics such as cost, latency and latency variation associated with a Forwarding Adjacency (FA) or Routing Adjacency (RA) Label Switched Path (LSP) are not available to the ingress and egress nodes. This draft provides extensions for the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) for the support of the discovery of cost, latency and latency variation of an LSP.

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].

Table of Contents

Copyright Notice.....	1
1. Introduction.....	3
2. RSVP-TE Requirement.....	3
2.1. Cost, Latency and Latency Variation Collection Indication....	4
2.2. Cost, Latency and Latency Variation Collection.....	4
2.3. Cost, Latency and Latency Variation Update.....	4
3. RSVP-TE signaling extensions.....	4
3.1. Cost Collection Flag.....	4
3.2. Latency Collection Flag.....	5
3.3. Latency Variation Collection Flag.....	5
3.4. Cost subobject.....	5
3.5. Latency subobject.....	6
3.6. Latency Variation subobject.....	7
3.7. Signaling Procedures.....	7
4. Security Considerations.....	9

Ali, Swallow, Filisfils Expires September 2012 [Page 2]

5. IANA Considerations.....	9
5.1. RSVP Attribute Bit Flags.....	9
5.2. New RSVP error sub-code.....	10
6. Acknowledgments.....	10
7. References.....	11
7.1. Normative References.....	11
7.2. Informative References.....	11

1. Introduction

There are many scenarios in packet and optical networks where the route information of an LSP may not be provided to the ingress node for confidentiality reasons and/ or the ingress node may not run the same routing instance as the intermediate nodes traversed by the path. In such scenarios, the ingress node cannot determine the cost, latency and latency variation properties of the LSP's route. Similarly, in Generalized Multi-Protocol Label Switching (GMPLS) networks signaling bidirectional LSP, the egress node cannot determine the cost, latency and latency variation properties of the LSP route. A multi-domain or multi-layer network is an example of such networks. Similarly, a GMPLS User-Network Interface (UNI) [RFC4208] is also an example of such networks.

In certain networks, such as financial information networks, network performance information (e.g. latency, latency variation) is becoming as critical to data path selection as other metrics [DRAFT-OSPF-TE-METRIC], [DRAFT-ISIS-TE-METRIC]. If cost, latency or latency variation associated with an FA or an RA LSP is not available to the ingress or egress node, it cannot be advertised as an attribute of the FA or RA. One possible way to address this issue is to configure cost, latency and latency variation values manually. However, in the event of an LSP being rerouted (e.g. due to re-optimization), such configuration information may become invalid. Consequently, in case where that an LSP is advertised as a TE-Link, the ingress and/ or egress nodes cannot provide the correct latency, latency variation and cost attribute associated with the TE-Link automatically.

In summary, there is a requirement for the ingress and egress nodes to learn the cost, latency and latency variation attributes of an FA or RA LSP. This draft provides extensions to the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) for the support of the automatic discovery of these attributes.

2. RSVP-TE Requirement

This section outlines RSVP-TE requirements for the support of the automatic discovery of cost, latency and latency variation attributes of an LSP. These requirements are very similar to the requirement of discovering the Shared Risk Link Groups (SRLGs)

associated with the route taken by an LSP [DRAFT-SRLG-RECORDING].

2.1. Cost, Latency and Latency Variation Collection Indication

The ingress and egress nodes of the LSP must be capable of indicating whether the cost, latency and latency variation attributes of the LSP should be collected during the signaling procedure of setting up the LSP. No cost, latency or latency variation information is collected without an explicit request being made by the ingress node.

2.2. Cost, Latency and Latency Variation Collection

If requested, cost, latency and latency variation is collected during the setup of an LSP. The endpoints of the LSP may use the collected information and use it for routing, flooding and TE link configuration purposes.

2.3. Cost, Latency and Latency Variation Update

When the cost, latency and latency variation property of a TE link along the route of a LSP for which that property was collected changes, e.g., if the administrator changes cost of a TE link, the node where the change occurred needs to be capable of updating the cost, latency and latency variation information of the path and signaling this to the end-points. Similarly, if a path segment of the LSP is rerouted, the endpoints of the rerouted segment need to be capable of updating the cost, latency and latency variation information of the path. Any node which adds cost, latency or latency variation information to an LSP during initial setup must signal changes to these values to both endpoints.

3. RSVP-TE signaling extensions

3.1. Cost Collection Flag

In order to indicate that cost collection is desired, a new flag in the Attribute Flags TLV which can be carried in an LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Objects is required:

Cost Collection flag (to be assigned by IANA, recommended bit position 11)

The Cost Collection flag is meaningful in a Path message. If the Cost Collection flag is set to 1, the transit nodes SHOULD report the cost information in the Path Record Route Object (RRO) and the Resv RRO.

The procedure for processing the Attribute Flags TLV follows [RFC5420].

3.2. Latency Collection Flag

In order to indicate that latency collection is desired, a new flag in the Attribute Flags TLV which can be carried in an LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Object is required:

Latency Collection flag (to be assigned by IANA, recommended bit position 12)

The Latency Collection flag is meaningful on a Path message. If the Latency Collection flag is set to 1, the transit nodes SHOULD report the latency information in the Path RRO and the Resv RRO.

The procedure for the processing the Attribute Flags TLV follows [RFC5420].

3.3. Latency Variation Collection Flag

In order to indicate that latency variation collection is desired, a new flag in the Attribute Flags TLV which can be carried in an LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Object is required:

Latency Variation Collection flag (to be assigned by IANA, recommended bit position 13)

The Latency Variation Collection flag is meaningful on a Path message. If the Latency Variation Collection flag is set to 1, the transit nodes SHOULD report the latency variation information in the Path RRO and the Resv RRO.

The procedure for the processing the Attribute Flags TLV follows [RFC5420].

3.4. Cost subobject

A new cost subobject is defined for the RRO to record the cost information of the LSP. Its format is similar to the RRO subobjects defined in [RFC3209].

0									1									2									3								
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1				
Type									Length									Reserved (must be zero)																	
																											COST Value								

Type: The type of the subobject, to be assigned by IANA (recommended value 35).

Length: The Length value is set to 8.

Reserved: This field is reserved for future use. It MUST be set to 0 when sent and MUST be ignored when received.

Cost Value: Cost of the link along the route of the LSP. Based on the policy at the recording node, the cost value can be set to the Interior Gateway Protocol (IGP) metric or TE metric of the link in question. This approach has been taken to avoid defining a flag for each cost type in the Attribute-Flags TLV. It is assumed that, based on policy, all nodes reports the same cost-type and that the ingress and egress nodes know the cost type reported in the RRO.

The rules for processing the LSP_ATTRIBUTES and LSP_REQUIRED_ATTRIBUTES Objects and RRO are not changed.

3.5. Latency subobject

A new Latency subobject is defined for RRO to record the latency information of the LSP. Its format is similar the RRO subobjects defined in [RFC3209].

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Type								Length								Reserved (must be zero)															
A Reserved																Delay															

Type: The type of the subobject, to be assigned by IANA (recommended value 36).

Length: The Length value is set to 8.

A-bit: This field represents the Anomalous (A) bit, as defined in [DRAFT-OSPF-TE-METRIC].

Reserved: These fields are reserved for future use. They MUST be set to 0 when sent and MUST be ignored when received.

Delay Value: This 24-bit field carries the average link delay over a configurable interval in micro-seconds, encoded as an integer value. When set to 0, it has not been measured. When set to the maximum value 16,777,215 (16.777215 sec), then the delay is at least that value and may be larger.

Internet-Draft draft-ali-ccamp-te-metric-recording-03.txt

The rules for processing the LSP_ATTRIBUTES and
LSP_REQUIRED_ATTRIBUTES Objects and RRO are not changed.

3.6. Latency Variation subobject

A new Latency Variation subobject is defined for RRO to record the Latency Variation information of the LSP. Its format is similar to the RRO subobjects defined in [RFC3209].

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Type								Length								Reserved (must be zero)															
A Reserved								Delay Variation																							

Type: The type of the subobject, to be assigned by IANA (recommended value 37).

Length: The Length value is set to 8.

A-bit: This field represents the Anomalous (A) bit, as defined in [DRAFT-OSPF-TE-METRIC].

Reserved: These fields are reserved for future use. It MUST be set to 0 when sent and MUST be ignored when received.

Delay Variation Value: This 24-bit field carries the average link delay variation over a configurable interval in micro-seconds, encoded as an integer value. When set to 0, it has not been measured. When set to the maximum value 16,777,215 (16.777215 sec), then the delay is at least that value and may be larger.

The rules for processing the LSP_ATTRIBUTES and
LSP_REQUIRED_ATTRIBUTES Objects and RRO are not changed.

3.7. Signaling Procedures

Typically, the ingress node learns the route of an LSP by adding a RRO in the Path message. If an ingress node also desires cost, latency and/or latency variation recording, it sets the appropriate flag(s) in the Attribute Flags TLV of the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Object. None, all or any of the Cost Collection, Latency Collection or Latency Variation Collection flags may be set in the Attribute Flags TLV of the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Object.

When a node receives a Path message which carries an
LSP_REQUIRED_ATTRIBUTES Object and the Cost, Latency and/or
Ali, Swallow, Filsfils Expires September 2012 [Page 7]

Latency Variation Collection Flag(s) is (are) set, if local policy disallows providing the requested information to the endpoints, the node MUST return a Path Error message with error code "Policy Control Failure (2)" and one of the following error subcodes:

- . "Cost Recoding Rejected" (value to be assigned by IANA, suggest value 105) if Cost Collection Flag is set.
- . "Latency Recording Rejected" (value to be assigned by IANA, suggest value 106) if Latency Collection Flag is set.
- . "Latency Variation Recording Rejected" (value to be assigned by IANA, suggest value 107) if Latency Variation Collection Flag is set.

When a node receives a Path message which carries an LSP_ATTRIBUTES Object and the Cost, Latency and/or Latency Variation Collection Flag(s) is (are) set, if local policy disallows providing the requested information to the endpoints, the node MAY return a Path Error as described for the LSP_REQUIRED_ATTRIBUTES Object.

If local policy permits the provision of the requested information, the processing node SHOULD add the requested subobject(s) with the cost, latency or/ and latency variation metric value(s) associated with the local hop to the Path RRO. It then forwards the Path message to the next node in the downstream direction.

Following the steps described above, the intermediate nodes of the LSP provide the requested metric value(s) associated with the local hop in the Path RRO. When the Path message is received by the egress node, the egress node can calculate the end-to-end cost, latency and/or latency variation properties of the LSP.

Before the Resv message is sent to the upstream node, the egress node adds the requested subobject(s) with the cost, latency or/ and latency variation metric value(s) associated with the local hop to the Resv RRO in a similar manner to that specified above for the addition of Path RRO sub-objects by midpoint nodes.

Similarly, the intermediate nodes of the LSP provide the requested metric value(s) associated with the local hop in the Resv RRO. When the Resv message is received by the ingress node, it can calculate the end-to-end cost, latency or/ and latency variation properties of the LSP.

Typically, cost and latency are additive metrics, but latency variation is not an additive metric. The means by which the ingress and egress nodes compute the end-to-end cost, latency

and latency variation metric from information recorded in the RRO is beyond the scope of this document.

Based on the local policy, the ingress and egress nodes can advertise the calculated end-to-end cost, latency and/or latency variation properties of the FA/ RA LSP in TE link advertisement to the routing instance based on the procedure described in [DRAFT-OSPF-TE-METRIC], [DRAFT-ISIS-TE-METRIC].

Based on the local policy, a transit node (e.g. the edge node of a domain) may edit a Path or Resv RRO to remove route information (e.g. node or interface identifier information) before forwarding it. A node that does this SHOULD summarize the cost, latency and latency variation data removed as a single value for each for the loose hop that is summarized by the transit node. How a transit node calculates the cost, latency or/ and latency variation metric for the segment summarized by the transit node is beyond the scope of this document.

4. Security Considerations

This document does not introduce any additional security issues above those identified in [RFC5920], [RFC5420], [RFC2205], [RFC3209], and [RFC3473].

5. IANA Considerations

5.1. RSVP Attribute Bit Flags

The IANA has created a registry and manages the space of attributes bit flags of Attribute Flags TLV as described in section 11.3 of [RFC5420]. It is requested that the IANA makes assignments from the Attribute Bit Flags defined in this document.

This document introduces the following three new Attribute Bit Flag:

- Bit number: TBD (recommended bit position 11)
- Defining RFC: this I-D
- Name of bit: Cost Collection Flag

- Bit number: TBD (recommended bit position 12)
- Defining RFC: this I-D
- Name of bit: Latency Collection Flag

Internet-Draft draft-ali-ccamp-te-metric-recording-03.txt

- Bit number: TBD (recommended bit position 13)
- Defining RFC: this I-D
- Name of bit: Latency Variation Flag

5.2. ROUTE_RECORD subobject

This document introduces the following three new RRO subobject:

Type	Name	Reference
-----	-----	-----
TBD (35)	Cost subobject	This I-D
TBD (36)	Latency subobject	This I-D
TBD (37)	Latency Variation subobject	This I-D

5.2. New RSVP error sub-code

For Error Code = 2 "Policy Control Failure" (see [RFC2205]) the following sub-code is defined.

Sub-code	Value
-----	-----
Cost Recoding Rejected	To be assigned by IANA. Suggested Value: 105.
Latency Recoding Rejected	To be assigned by IANA. Suggested Value: 106.
Latency Variation Recoding Rejected	To be assigned by IANA. Suggested Value: 107.

6. Acknowledgments

Authors would like to thank Ori Gerstel, Gabriele Maria Galimberti, Luyuan Fang and Walid Wakim for their review comments.

Internet-Draft draft-ali-ccamp-te-metric-recording-03.txt

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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.
- [RFC5420] Farrel, A., Ed., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.
- [DRAFT-OSPF-TE-METRIC] S. Giacalone, D. Ward, J. Drake, A. Atlas, S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", draft-ietf-ospf-te-metric-extensions, work in progress.
- [DRAFT-ISIS-TE-METRIC] S. Previdi, S. Giacalone, D. Ward, J. Drake, A. Atlas, C. Filsfils, "IS-IS Traffic Engineering (TE) Metric Extensions", draft-previdi-isis-te-metric-extensions, work in progress.
- [DRAFT-SRLG-RECORDING] F. Zhang, D. Li, O. Gonzalez de Dios, C. Margaria,, "RSVP-TE Extensions for Collecting SRLG Information", draft-ietf-ccamp-rsvp-te-srlg-collect.txt, work in progress.

7.2. Informative References

- [RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.
- [RFC2209] Braden, R. and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules", RFC 2209, September 1997.
- [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

Authors' Addresses

Internet-Draft draft-ali-ccamp-te-metric-recording-03.txt

Zafar Ali
Cisco Systems, Inc.
Email: zali@cisco.com

George Swallow
Cisco Systems, Inc.
swallow@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
cfilsfil@cisco.com

Matt Hartley
Cisco Systems
Email: mhartley@cisco.com

Kenji Kumaki
KDDI Corporation
Email: ke-kumaki@kddi.com

Rudiger Kunze
Deutsche Telekom AG
Ruediger.Kunze@telekom.de

CCAMP Working Group
Internet Draft
Intended status: Standard Track
Expires: January 15, 2013

Zafar Ali
George Swallow
Clarence Filsfils
Matt Hartley
Ori Gerstel
Gabriele Maria Galimberti
Cisco Systems
Kenji Kumaki
KDDI Corporation
Ruediger Kunze
Deutsche Telekom AG
Julien Meuric
France Telecom Orange
July 16, 2012

Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) LSP
Route Diversity using Exclude Routes

draft-ali-ccamp-xro-lsp-subobject-02.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on September 11, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.
Code Components extracted from this document must include Simplified BSD License text as described in

Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Ali, Swallow, Filsfils Expires January 2013

[Page 1]

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Abstract

[RFC4874] specifies methods by which route exclusions may be communicated during RSVP-TE signaling in networks where precise explicit paths are not computed by the LSP ingress node. This document specifies signaling for additional route exclusions based on LSPs currently existing or expected to exist within the network.

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].

Table of Contents

1. Introduction	3
1.1. Requirements Notation	5
2. RSVP-TE signaling extensions	5
2.1. Terminology	5
2.2. LSP Subobject	5
2.3. Processing rules for the LSP subobject	10
2.4. LSP Subobject in Explicit Exclusion Route Subobject	11
2.4.1. Processing Rules for the EXRS with LSP subobject	12
4. IANA Considerations	12
4.1. New XRO subobject type	12
4.2. New EXRS subobject type	12
4.3. New RSVP error sub-code	12
5. Acknowledgement	13
6. References	13
6.1. Normative References	13
6.2. Informative References	13
1. Introduction	3
1.1. Requirements Notation	5
2. RSVP-TE signaling extensions	5
2.1. Terminology	5
2.2. LSP Subobjects	5
2.2.1. IPv4 Point-to-Point LSP subobject	5
2.2.2. IPv6 Point-to-Point LSP subobject	9
2.3. Processing rules for the LSP subobject	10
2.4. LSP Subobject in Explicit Exclusion Route Subobject (EXRS)	13
2.4.1. Processing Rules for the EXRS with LSP subobject	14

3. Security Considerations	14
4. IANA Considerations	14
4.1. New XRO subobject type	14

4.2. New EXRS subobject type	14
4.3. New RSVP error sub-code	14
5. Acknowledgement	15
6. References	15
6.1. Normative References	15
6.2. Informative References	15

1. Introduction

Label-Switched Path (LSP) diversity is required to ensure LSPs may be established without sharing resources, thus greatly reducing the probability of simultaneous connection failures.

LSP diversity is a well-known requirement from Service Providers. When route computation for LSPs that need to be diverse is performed at ingress node, this requirement can be met by a local decision at that node. However, there are scenarios when route computations are performed by remote nodes, in which case there is a need for relevant diversity requirements to be communicated to those nodes. These include (but are not limited to):

- . LSPs with loose hops in the Explicit Route Object (ERO), e.g. inter-domain LSPs.
- . Generalized Multi-Protocol Label Switching (GMPLS) User-Network Interface (UNI) where route computation may be performed by the (sever layer) core node [RFC4208];

The eXclude Route Object (XRO) and Explicit Exclusion Route Subobject (EXRS) specification [RFC4874] introduces a means of specifying nodes and resources to be excluded from routes, using the XRO and/ or EXRS.

[RFC4874] facilitates the calculation of diverse routes for LSPs based on known properties of those LSPs including addresses of links and nodes traversed, and Shared Risk Link Groups (SRLGs) of traversed links. This requires that these properties of the LSP(s) from which diversity is required be known to the ingress node which initiates signaling. However, there are circumstances under which this may not be possible or desirable, including (but not limited to):

- . Exclusion of the route of a LSP which does not originate, terminate or traverse the ingress node signaling the diverse LSP, in which case the addresses and SRLGs of the LSP from which diversity is required are unknown to the ingress node.

- . Exclusion of the route of a LSP which, while known at the ingress node of the diverse LSP, has incomplete or unavailable route information, e.g. due to confidentiality of the LSP route attributes. In other words, the scenario in which the reference LSP is hosted by the ingress/ requesting node but the properties required to construct an XRO object are not known to ingress/ requesting node. Inter-domain and GMPLS overlay networks may present such restrictions.
- . If the route of the reference LSP from which diversity is required (e.g. LSP1) is known to the ingress node, that node can use this information to construct an XRO and send it in the path message during the signaling of a diverse LSP (LSP2). However, if the route of LSP1 changes (e.g. due to re-optimization or failure in the network), the ingress node would need to change path of LSP2 to ensure that it remains diverse from LSP1. It is preferable to have this decision made by the node that calculated the path for LSP2. For example, in the case of GMPLS-UNI, it is better to have such responsibility at the server layer as opposed to at the client layer so that the diversity requirements are transparent to the client layer. Furthermore, in all networking scenarios, if the node performing the route computation/ expansion is aware of the diversity requirements of LSP1 and LSP2, it may consider joint re-optimization of the diverse LSPs.

This document addresses such scenarios and defines procedures that may be used to exclude the route taken by a particular LSP, or the route taken by all LSPs belonging to a single tunnel. Note that this diversity requirement is different from the diversity requirements of path protection where both the reference and diverse LSPs belong to the same tunnel. The diversity requirements considered in this document do not require that the LSPs in question belonging to the same tunnel or share an ingress node.

The means by which the node calculating or expanding the route of the signaled LSP discovers the route of the LSPs from which the signaled LSP requires diversity is beyond the scope of this document. However, in most cases the LSPs with route diversity requirements may transit the node expanding the route.

This document addresses only the exclusion of point-to-point tunnels; point-to-multipoint tunnels will be addressed in a future version. Similarly, at present only IPv4 addresses are considered; support for IPv6 addresses will be added in a future version.

L	Type	Length	Attribute Flags	Exclusion Flags
+-----				

L

The L-flag is used as for the other XRO subobjects defined in [RFC4874].

0 indicates that the attribute specified MUST be excluded.

1 indicates that the attribute specified SHOULD be avoided.

Type

IPv4 Point-to-Point LSP subobject

(to be assigned by IANA suggested value: 36).

Length

The length contains the total length of the subobject in bytes, including the type and length fields. The length is always 24.

Attribute Flags

The Attribute Flags are used to communicate desirable attributes of the LSP being signaled (In the following, the term LSP2 is used to reference the LSP being signaled; please refer to Section 2.1 for definition of LSP2). The following flags are defined. None, all or multiple attribute flags MAY be set within the same subobject.

0x01 = LSP ID to be ignored

This flag is used to indicate tunnel level exclusion. Specifically, this flag is used to indicate that the lsp-id field of the subobject is to be ignored and the exclusion applies to any LSP matching the rest of the supplied FEC. In other words, if this flag is set, the processing node MUST calculate a route based on exclusions from the routes of all known LSPs matching the tunnel-id, source, destination and extended tunnel-id specified in the subobject.

When this flag is not set, the lsp-id is not ignored and the exclusion applies only to the specified LSP (i.e., LSP level exclusion). In other words, when this flag is not set, route exclusions MUST respect the specified LSP (i.e. the lsp-id, the tunnel-id, source, destination and extended tunnel-id specified needs to be respected during exclusion).

0x02 = Destination node exception

This flag is used to indicate that the destination node may be shared even when sharing of the said node violates the exclusion flags. When this flag is not set, the exclusion flags SHOULD also be respected for the destination node.

0x04 = Processing node exception

This flag is used to indicate that the processing node may be shared even when sharing of the said node violates the exclusion flags. When this flag is not set, the exclusion flags SHOULD also be respected for the processing node.

0x08 = Penultimate node exception

This flag is used to indicate that the penultimate node may be shared even when sharing of the said node violates the exclusion flags. When this flag is not set, the exclusion flags SHOULD also be respected for the penultimate node.

Exclusion Flags

The Exclusion-Flags are used to communicate desirable types of exclusion. The following flags are defined.

0x01 = SRLG exclusion

This flag is used to indicate that the route of the LSP being signaled is requested to be SRLG diverse from the route of the LSP or tunnel specified by the LSP subobject.

0x02 = Node exclusion

This flag is used to indicate that the route of the LSP being signaled is requested to be node diverse from the route of the LSP or tunnel specified by the LSP subobject. The node exclusion is subobject to the setting of the "Processing node exception", the "Penultimate node exception" and the "Destination node exception" Attribute Flags.

0x04 = Link exclusion

This flag is used to indicate that the route of the LSP being signaled is requested to be link diverse from the route of the LSP or tunnel specified by the LSP subobject.

The remaining fields are as defined in [RFC3209].

1 indicates that the attribute specified SHOULD be avoided.

Type

IPv6 Point-to-Point LSP subobject
(to be assigned by IANA suggested value: 36).

Length

The length contains the total length of the subobject in bytes, including the type and length fields. The length is always 48.

The Attribute Flags and Exclusion Flags are as defined for the IPv4 Point-to-Point LSP subobject.

The remaining fields are as defined in [RFC3209].

2.3. Processing rules for the LSP subobject

XRO processing as described in [RFC4874] is unchanged.

If the node is the destination for the LSP being signaled, it SHOULD NOT process a LSP XRO subobject.

If the L-flag is not set, the processing node follows the following procedure:

- The processing node MUST ensure that any route calculated for the signaled LSP (LSP2) respects the requested exclusion flags with respect to the route traversed by the LSP(s) referenced by the LSP subobject (LSP1/TUNNEL1), including local resources.
- If the processing node fails to find a route that meets the requested constraint, the processing node SHOULD return a PathErr with the error code "Routing Problem (24)" and error value "Route blocked by Exclude Route (67)".
- If the route of the LSP or tunnel (LSP1/TUNNEL1) referenced in the LSP subobject is unknown to the processing node, the processing node SHOULD ignore the LSP subobject in the XRO and SHOULD proceed with the signaling request (for LSP2). However,

in this case, after sending Resv for LSP2, the processing node SHOULD return a PathErr with the error code "Notify Error (25)" and error value "Route of XRO LSP unknown (value: to be assigned by IANA, suggest value: 13)" for LSP2.

- If latter, the route of the LSP or tunnel (LSP1/TUNNEL1) referenced in the LSP subobject becomes known (e.g. when LSP1 is signaled) or the TUNNEL1 is re-optimized to a different route, such that the requested exclusion/ diversity constraints are no longer satisfied and a path that can satisfy the requested constraints exists, the node calculating or expanding the path SHOULD send a PathErr message for LSP2 with the error code "Notify Error (25)" and error value "Preferable path exists (6)". An ingress node receiving this error code/value combination MAY try to reoptimize the LSP2 to the new preferred path.
- Route computation for the LSP or tunnel (LSP1/ TUNNEL1) referenced in the LSP subobject for new setup or for re-optimization LSP SHOULD be performed to avoid situation where the requested exclusion/ diversity constraints are no longer satisfied and a path that can satisfy the requested constraints does not exist. However, if such situation arises the node that computed or expanded the route for LSP2 SHOULD send a PathErr message for LSP2 with the error code "Routing Problem (24)" and error value "Route blocked by Exclude Route (67)".

If the L-flag is set, the processing node follows the following procedure:

- The processing node SHOULD respect the requested exclusion flags with respect to the route traversed by the referenced LSP(s) (LSP1/TUNNEL1) as far as possible.
- If the processing node fails to find a route that meets the requested constraint, it SHOULD proceed with a suitable route that best meets the constraint, but after completion of signaling setup, it SHOULD return a PathErr code "Notify Error (25)" and error value "Failed to respect Exclude Route (value: to be assigned by IANA, suggest value: 14)" to the ingress node.
- If the route of the LSP or tunnel (LSP1/TUNNEL1) referenced in the LSP subobject is unknown to the processing node, the processing node SHOULD ignore the LSP subobject in XRO and SHOULD proceed with the signaling request (for LSP2). However, in this case, after sending Resv for LSP2, the processing node

SHOULD return a PathErr with the error code "Notify Error" and error value "Route to XRO LSP unknown" for LSP2.

- If latter, the route of the LSP or tunnel (LSP1/TUNNEL1) referenced in the LSP subobject becomes known (e.g. when LSP1 is signaled) or the TUNNEL1 is re-optimized to a different route, such that the requested exclusion/ diversity constraints are no longer satisfied and a path that can satisfy the requested constraints exists, the node calculating or expanding the path SHOULD send a PathErr message for LSP2 with the error code "Notify Error (25)" and error value "Preferable path exists". An ingress node receiving this error code/value combination MAY try to reoptimize the LSP2 to the new preferred path.
- Route computation for the LSP or tunnel (LSP1/ TUNNEL1) referenced in the LSP subobject for new setup or for re-optimization LSP SHOULD be performed to avoid situation where the requested exclusion/ diversity constraints are no longer satisfied and a path that can satisfy the requested constraints does not exist. However, if such situation arises the node that computed or expanded the route for LSP2 SHOULD send a PathErr message for LSP2 with the error code "Notify Error" and error value "Failed to respect Exclude Route".

The following rules apply equally to L = 0 and L = 1 case:

- XRO object MAY contain multiple LSP subobjects. In this case, the processing node A node receiving a Path message carrying an XRO MAY reject the message if the XRO is too large or complicated for the local implementation or the rules of local policy, as per the roles of XRO defined in [RFC4874]. In this case, the node MUST send a PathErr message with the error code "Routing Error" and error value "XRO Too Complex". An ingress node receiving this error code/value combination MAY reduce the complexity of the XRO or route around the node that rejected the XRO.
- An ingress node receiving PathErr with the error code "Notify Error" and error values "Route to XRO LSP unknown" or "Failed to respect Exclude Route" MAY take no action other than simply logging these notifications.

Note that LSP1 may be signaled with an XRO LSP subobject referencing CircuitID2 (LSP2 FEC) and LSP2 may be signaled with an XRO LSP subobject referencing CircuitID1 (LSP1 FEC). The above-mentioned processing rules cover this case. In fact, if

"LSP ID to be ignored" attribute flag is set when LSP1 is signaled with an XRO LSP subobject referencing CircuitID2, it is RECOMMENDED that LSP2 is signaled with an XRO LSP subobject referencing CircuitID1.

2.4. LSP Subobject in Explicit Exclusion Route Subobject (EXRS)

[RFC4874] defines an ERO subobject called Explicit Exclusion Route Subobject (EXRS). An EXRS is used to identify abstract nodes or resources that must not or should not be used on the path between two inclusive abstract nodes or resources in the explicit route. An EXRS contains one or more subobjects of its own, called EXRS subobjects [RFC4874].

An EXRS MAY include an IPv4 Point-to-Point (P2P) LSP subobject. In this case, EXRS would look as follows:

0										1										2										3																													
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1																												
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
L										Type										Length										Reserved																													
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
L										Type										Length										Attribute Flags										Exclusion Flags																			
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
										IPv4 tunnel end point address																																																	
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
										Must Be Zero										Tunnel ID																																							
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
										Extended Tunnel ID																																																	
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
										IPv4 tunnel sender address																																																	
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													
										Must Be Zero										LSP ID																																							
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																													

The meaning of respective fields in EXRS header is as defined in [RFC4874]. Similarly, the meaning of respective fields in IPv4 P2P LSP subobject is as defined earlier in this document. This is with the exceptions that:

- Processing node exception applies to the node processing the ERO.

- If L bit in the ERO header is not set (ERO.L = 0), the IPv4 P2P LSP subobject is processed against the LSPs for which the processing node is ingress, egress or a transit node.
- Penultimate node exception applies to the penultimate node of the loose hop. This flag is only processed if EXRS.L bit is set, i.e., in the loose ERO hop case.
- Destination node exception applies to the abstract node to which the route is expanded. This flag is only processed if EXRS.L bit is set, i.e., in the loose ERO hop case.

2.4.1. Processing Rules for the EXRS with LSP subobject

Processing rules for the EXRS object are same as processing rules as described in [RFC4874]. When the EXRS contains one or more LSP subobject(s), processing rule specified in Section 2.3 applies to the node processing the ERO with EXRS subobject.

3. Security Considerations

This document does not introduce any additional security issues above those identified in [RFC5920], [RFC2205], [RFC3209], and [RFC3473] and [RFC4874].

4. IANA Considerations

4.1. New XRO subobject type

This document introduces a new subobject for the EXCLUDE_ROUTE object [RFC4874], C-Type 1.

Subobject Type	Subobject Description
-----	-----
To be assigned by IANA (suggest value: 36)	IPv4 P2P LSP subobject

4.2. New EXRS subobject type

IPv4 P2P LSP subobject is also defined as a new EXRS subobject.

4.3. New RSVP error sub-code

For Error Code = 25 "Notify Error" (see [RFC3209]) the following sub-code is defined.

Sub-code -----	Value -----
Route of XRO LSP unknown	To be assigned by IANA. Suggested Value: 13.
Failed to respect Exclude Route	To be assigned by IANA. Suggested Value: 14.

5. Acknowledgement

Authors would like to thanks Luyuan Fang and Walid Wakim for their review comments.

6. References

6.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4874] Lee, CY., Farrel, A., and S. De Cnodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4874, April 2007.

6.2. Informative References

- [RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.
- [RFC2209] Braden, R. and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules", RFC 2209, September 1997.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

Authors' Addresses

Zafar Ali
Cisco Systems.
Email: zali@cisco.com

George Swallow
Cisco Systems
swallow@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
cfilsfil@cisco.com

Matt Hartley
Cisco Systems
Email: mhartley@cisco.com

Ori Gerstel
Cisco Systems
ogerstel@cisco.com

Gabriele Maria Galimberti
Cisco Systems
ggalimbe@cisco.com

Kenji Kumaki
KDDI Corporation
Email: ke-kumaki@kddi.com

Rudiger Kunze
Deutsche Telekom AG
Ruediger.Kunze@telekom.de

Julien Meuric
France Telecom Orange
Email: julien.meuric@orange.com

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: April 25, 2013

D. Ceccarelli, Ed.
F. Fondelli
Ericsson
S. Belotti
D. Papadimitriou, Ed.
Alcatel-Lucent
October 22, 2012

Multi layer implications in GMPLS controlled networks
draft-bcg-ccamp-gmpls-ml-implications-03

Abstract

This document defines requirements and uses cases for the extension of the OTN MLN work to MRNs. It also provides an evaluation of already existing solutions againsts new requirements.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Terminology	3
2. MLN and MRN networks: relationship and rationale	3
3. Applicability Scenarios	5
4. Requirements	7
5. Evaluation	8
6. Missing information	13
7. IANA Considerations	14
8. Contributors	14
9. Acknowledgements	14
10. References	14
10.1. Normative References	14
10.2. Informative References	15
Authors' Addresses	15

1. Introduction

Generalized MPLS (GMPLS) supports the control of multiple switching technologies: packet switching, Layer-2 switching, TDM (Time-Division Multiplexing) switching, wavelength switching, and fiber switching ([RFC3945]).

The Interface Switching Capability concept has been defined for the advertisement of the Switching Capabilities of the different interfaces of a node [RFC4202], while in the context of Multi Region Networks (MRN) the Interface Adjustment Capability concept has been introduced [RFC5339] for the advertisement of adjustment capacity within an hybrid node.

With the introduction of G709v3 networks, a new Switching Capability (OTN-TDM) has been defined [OSPF-OTN] and the ISCD updated in order to cope with the OTN specific multi stage multiplexing capabilities. The new Switching Capability Specific Information (SCSI) field provides information about the bandwidth availability at each layer of the OTN hierarchy and about the operations that can be performed on the different layers, in terms of termination and switching capabilities.

These issues have been addressed in the OTN documents within the OTN multi layer scope but need to be extended to MRNs, where the termination of a hierarchical LSP leads to the need of properly managing different switching capabilities and different adaptation functions.

Scope of this document is describing new requirements derived from the extension of OTN MLN hierarchies to MRNs and evaluating impacts on existing solutions.

1.1. Terminology

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 [RFC2119].

2. MLN and MRN networks: relationship and rationale

As per [RFC5212], the definition of MLNs and MRNs is as follows:

- MLN: "a Traffic Engineering (TE) domain comprising multiple data plane switching layers either of the same ISC (e.g., TDM) or different ISC (e.g., TDM and PSC) and controlled by a single GMPLS control plane instance"

- MRN: "is defined as a TE domain supporting at least two different switching types (e.g., PSC and TDM), either hosted on the same device or on different ones, and under the control of a single GMPLS control plane instance"

A network which is an MLN but not an MRN (i.e. multiple layers but a single switching capability), like for example an OTN domain, can be advertised via the utilization of the Interface Switching Capability Descriptor (ISCD). The ISCD is defined in [RFC4202] and its technology specific extensions (SCSI) are defined in different memos depending on the technology, e.g. the OTN ones in [OSPF-OTN] and the SDH ones in [RFC4203].

On the other side MRNs (i.e. multiple layers with multiple switching capabilities), like for example an OTN data plane (with one or more layers) over a WDM data plane (with one or more layers) controlled by a single GMPLS instance, need the utilization of an ISCD for each technology and an Interface Adjustment Capability Descriptor (IACD) [RFC6001] for the advertisement of the internal links providing adjustment between the switching capabilities. A node able to terminate data links (over the same interface) with different switching capabilities is called hybrid node. [RFC5212]. For more details please see Section 5.

Hybrid nodes have been introduced not only to address the case of nodes able to switch/terminate LSPs from different switching capability but also to perform for instance:

- Traffic-grooming: base GMPLS doesn't enable insertion of traffic at an intermediate point along an established LSP, i.e., the control plane limits the flexibility of nesting LSP only at the head-end of the underlying LSP. MRN extensions enable to multiplex and demultiplex e.g. PSC LSP into LSC LSP even if the LSC LSP does not originate or end at the nodes where the PSC LSP are multiplexed or demultiplexed.
- Transparent regeneration: enables certain nodes equipped with PSC + LSC capability to regenerate the photonic signal without interrupting the LSC LSP. This functionality enables to setup end-to-end LSC even if certain intermediate nodes are being used to regenerate the signal at the PSC level.

This means that MRN extends the node functionality beyond "terminate or switch".

The central notion in MRN is "adjustment capability". Adjustment capability assumes the availability of adjustment capacity. An adjustment capability is the mean by mean which LSP can be adapted

from one SC = X to SC = Y or inserted (e.g. multiplexed or demultiplexed) from SC = X to SC = Y . Advertisement of adjustment capacity by a given node assumes the functionality of adjustment is locally supported.

3. Applicability Scenarios

When moving from OTN MLNs to general MRNs, the multiplexing tree concept introduced in [OSPF-OTN] needs to be extended so to take into account both different switching capabilities within the same muxing tree and adaptations between client hierarchies and server hierarchies.

In the following figure an example of muxing tree supporting TDM, PSC, OTN-TDM and LSC hierarchies mixed together is shown.

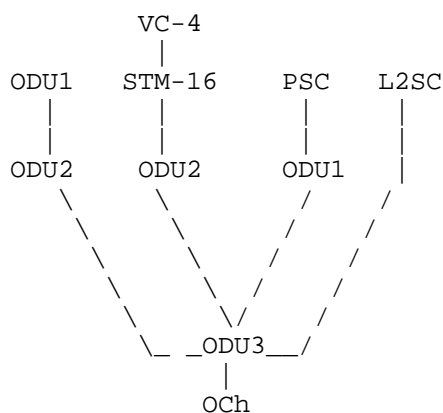


Figure 1: Muxing tree

As it is possible to understand from the figure above, an MRN equipment can host a variety of client-server relationships. Four different scenarios can be identified:

- A signal type X is a client to a Signal type Y (1:1) - e.g. Ethernet over WDM
- A signal type X is a client to a Intra switching technology Hierarchy Y (1:N) - e.g. Ethernet over OTN
- An Intra switching technology Hierarchy X is a client to a Signal Type Y (M:1) - e.g. ODU over WDM

- An Intra switching technology Hierarchy X is a client to an Intra switching technology Hierarchy Y (M:N) - e.g. SDH over OTN

Being the first three scenarios a particular case of the fourth one, in the following only the general case of M:N relationship will be addressed.

This kind of client-server hierarchy can be achieved, depending on the implementation, via single board or a cascade of them. In the latter case boards are connected via internal links, which can be either intra or inter switching capability (e.g. ODU2->ODU3 or PSC->LSC). Those links should not be modeled as external TE links, but there is the need to advertise their characteristics and availability in terms of bandwidth and optical parameters.

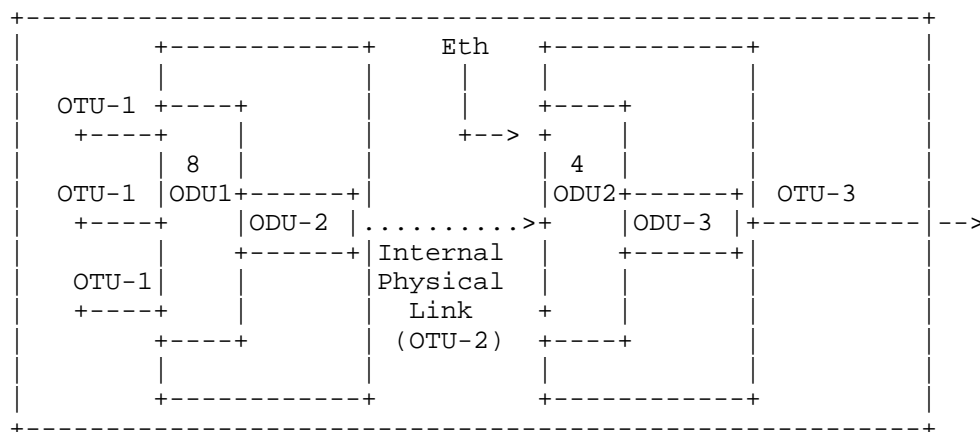


Figure 2: Cascaded muxponder

Moreover, as described in [RFC5212], in a hybrid node there is the need to take into account also the node's internal adjustment capabilities between the switching technologies supported. An example of hybrid node with different switching matrices is shown in the following figure, where both an SDH and OTN matrix are available and the two switching elements are internally interconnected so that it is possible to terminate some resources (e.g. OTN interface Y1) or provide adjustment for the SDH traffic (e.g. OTN interface Y2 toward the SDH matrix). In addition to the internal links between matrices it is possible to have internal links between matrices and cascaded cards for the creation of the muxing hierarchy. In the example below both the SDH and OTN matrices are client to an

be mapped only on interface Y5 of the muxponder board or the 10GbEth on interface Y6. In figure Figure 1 it is also possible to see that the OTN has a hierarchy with 3 branches (i.e. ODU1->ODU2->ODU3, ODU2->ODU3 and ODU1->ODU3) and an SDH signal can be mapped only over the ODU2->ODU3 branch while an Ethernet one can be mapped only on the ODU1->ODU3). So it is not enough to say that SDH can be mapped over ODU or Eth over ODU as further info is needed. Moreover it is also not enough to say that Eth is mapped over ODU1 because in the same example 2 different branches have the ODU1 as the top most layer (i.e. ODU1->ODU2->ODU3 and ODU1->ODU3) and not both of them can support Eth mapping.

2. Adaptation information from X to Y to be used both in case of Y being switched and X mapped over it or in case of both X and Y being switched. Please note that more than one type of adaptation might be available.

3. Amount of available bandwidth in the mapping between X and Y (as per actual IACD definition)

4. It must be possible to advertise intra-switching capability associated to internal links. A typical case is a hierarchy gained through the cascade of multiple cards (e.g. transponders, muxponders) and the link from one board to the other one has a given bandwidth.

5. It must be possible to advertise inter-switching capability associated to internal links. A typical case is a M:N client-layer hierarchy gained through the cascade of multiple cards (e.g. SDH client to a muxponder card) and the link from one board to the other one has a given bandwidth.

5. Evaluation

[RFC6001] defined the Interface Adjustment Capability Descriptor (IACD) for the advertisement of internal adjustment capability of hybrid nodes [RFC5212].

A common adjustment pool is a pool of reservable and sharable resources that are i) allocated on demand/dynamically and ii) either assigned to a single SC (single adjustment pool model) or multiple SC (multiple adjustment pool model) or possibly their combination.

In the former case (single pool model), the "lower SC" value of the IACD sub-TLV (associated to the adjustment pool) is set to the SC value of ISCD sub-TLV of the interface that interfaces with the adjustment pool. The "upper" SC value of the IACD (associated to the

adjustment pool) determines the SC capability of the resource pool itself. In this case the (upper) encoding is set to 0xFF. In other terms, the capacity of the adjustment pool is not directly accessible - over the wire - by other nodes belonging to the same TE domain (assuming homogeneous LSP encoding type along the LSP path). This model (see Example 1) is typically used when the node matrix switching capability is not terminating/initiating any LSP (the node only exposes the capability associated to its I/O) but nodes part of the same TE domain can still take into account the adjustment capacity usage on that node.

In the latter case (multiple pool model), the "lower SC" value of the IACD sub-TLV (associated to the adjustment pool) is set to the SC value of ISCD sub-TLV of the interface(s) that interfaces with the adjustment pool. The "upper" SC value of the IACD sub-TLV (associated to the adjustment pool) determines the SC capability of the adjustment pool itself. However, the (upper) SC value of the IACD sub-TLV shall correspond to at least one of the SC values associated to one of the ISCD sub-TLVs, i.e., the adjustment pool SC value shall be covered by at least one of the SC values associated to the ISCD sub-TLVs. In other terms, the capacity of the adjustment pool is directly accessible compared to the single pool model. This model (see Example 2) is typically used when nodes expose their full (multi-level) grooming and initiation/ termination capacity.

Example of single pool model: in the IACD sub-TLV the "upper" SC type = TDM/HO-SDH, and the "lower" SC type being respectively "L2SC" and "OTH/TDM". In this example, the capacity associated to the IACD represents the "interconnection capacity" between the interface X (L2SC or OTH) to Y = (HO-SDH/TDM). The encoding type associated to the upper SC is set to 0xFF.

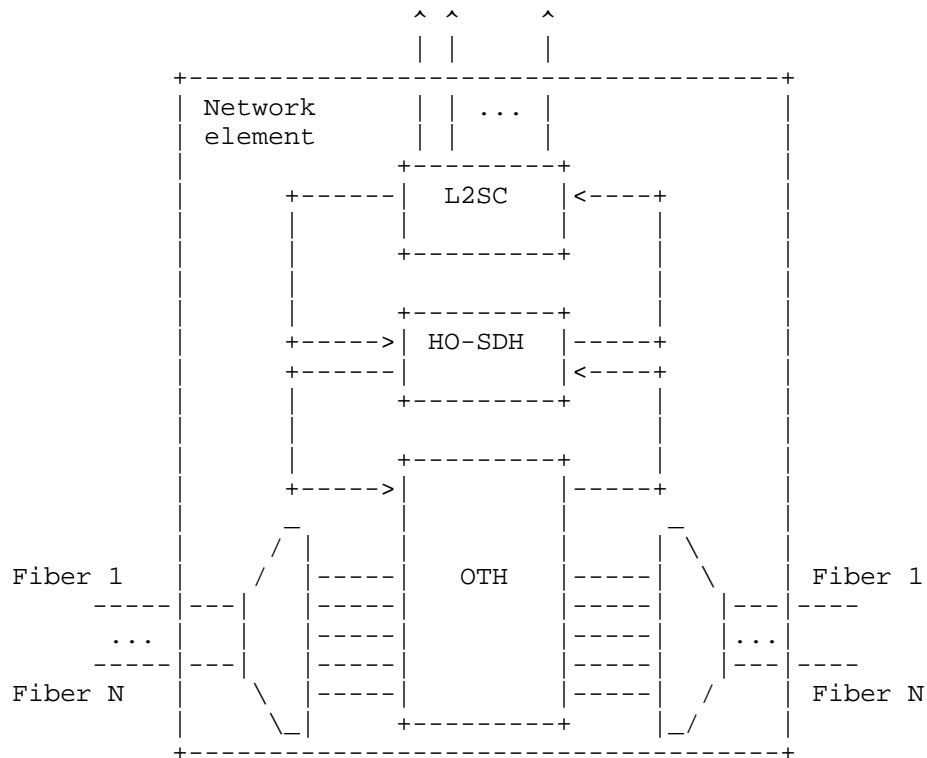


Figure 4: Example of single pool model

The advertisement for the node interfaces will be:

+ L2SC interfaces

- ISCD sub_TLV 1 for L2SC interface
- IACD sub_TLV 1 for L2SC to HO-SDH (1) in figure above

+ OTH interfaces

- ISCD sub_TLV 1 for OTH interface
- IACD sub_TLV 1 for OTH to HO-SDH (2) in figure above

Example of multiple pool model: In this case we will show two examples, the first of which does not foresee any interconnection between the L2SC and the HO-SDH matrices, while the second one does.

In the former case there is at least one ISCD sub-TLV of SC = X corresponding to the lower SC value (HO-SDH/TDM) of the IACD sub-TLV associated to the first adjustment pool (HO-SDH/TDM), and one ISCD sub-TLV of type SC = Y corresponding to the lower SC value (L2SC) of the IACD sub-TLV associated to the second adjustment pool Y (L2SC). In this example, the capacity associated to the IACD represents the "interconnection capacity" between the pool of SC = X (HO-SDH/TDM) to Y (L2SC). Each TE Link 1...N is able to get access to this adjustment capacity.

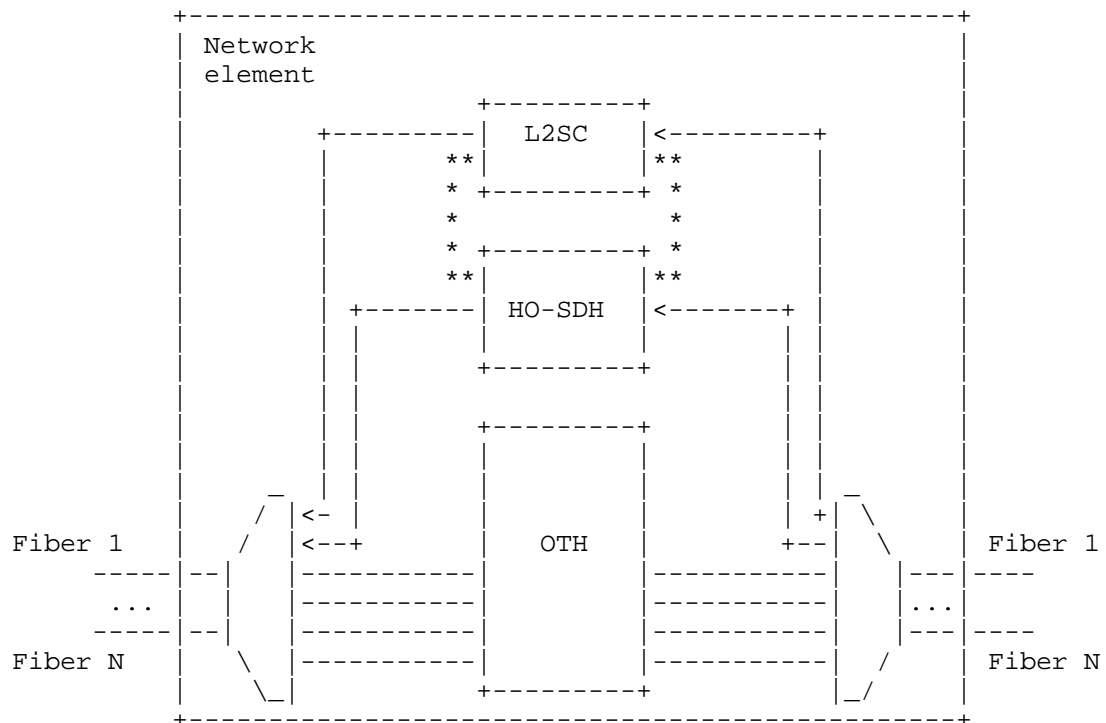


Figure 5: Example of multiple pool model - No interconnection between OTH and HO-SDH

In this case the advertisement, which is the same for each of the N TE Link is:

- ISCD sub_TLV for LSC
- ISCD sub_TLV for HO-SDH

- ISCD sub_TLV for OTH
- IACD sub_TLV for LSC to HO-SDH (starred link)

On the other side, if we consider the same scenario including the inteconnection between the OTH and HO-SDH matrices, as shown in figure below, the advertisement changes as follows.

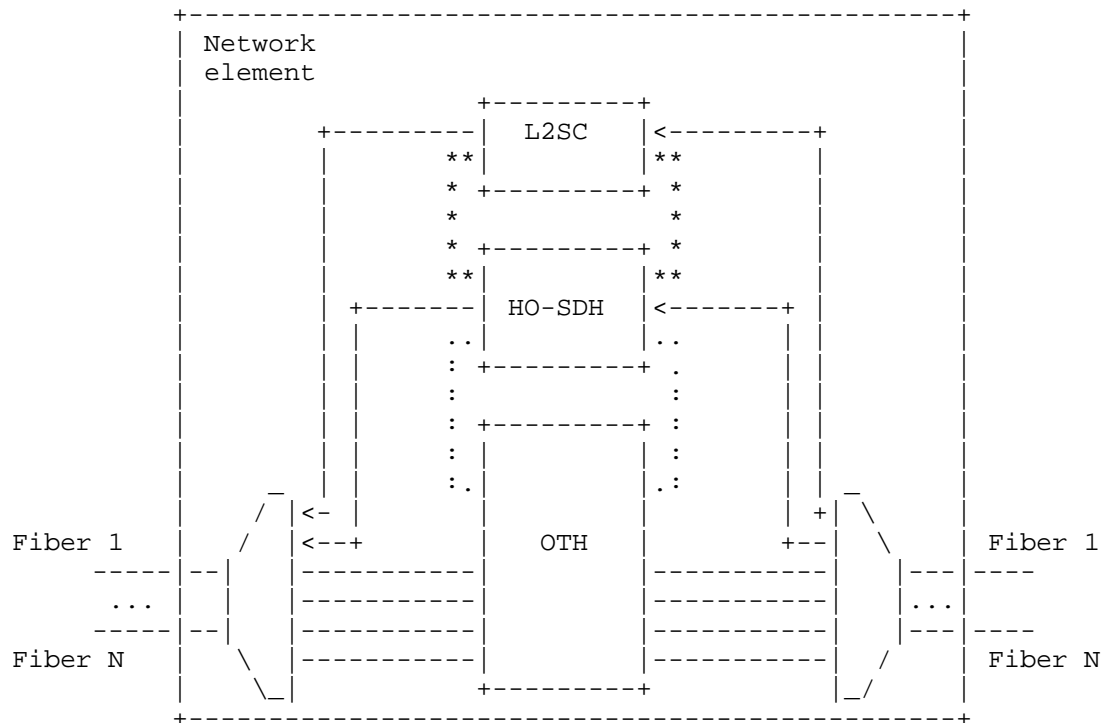


Figure 6: Example of multiple pool model - With interconnection between OTH and HO-SDH

This time the advertisement is modified as follows:

- ISCD sub_TLV 1 for LSC
- ISCD sub_TLV 2 for HO-SDH
- ISCD sub_TLV 3 for OTH

- IACD sub_TLV 1 for LSC to HO-SDH (starred link)
- IACD sub_TLV 2 for HO-SDH to OTH (dotted link)

The IACD is the only object defined in routing for the management of hybrid nodes. It provides the information for the forwarding/switching capability and is used in addition to the ISCD.

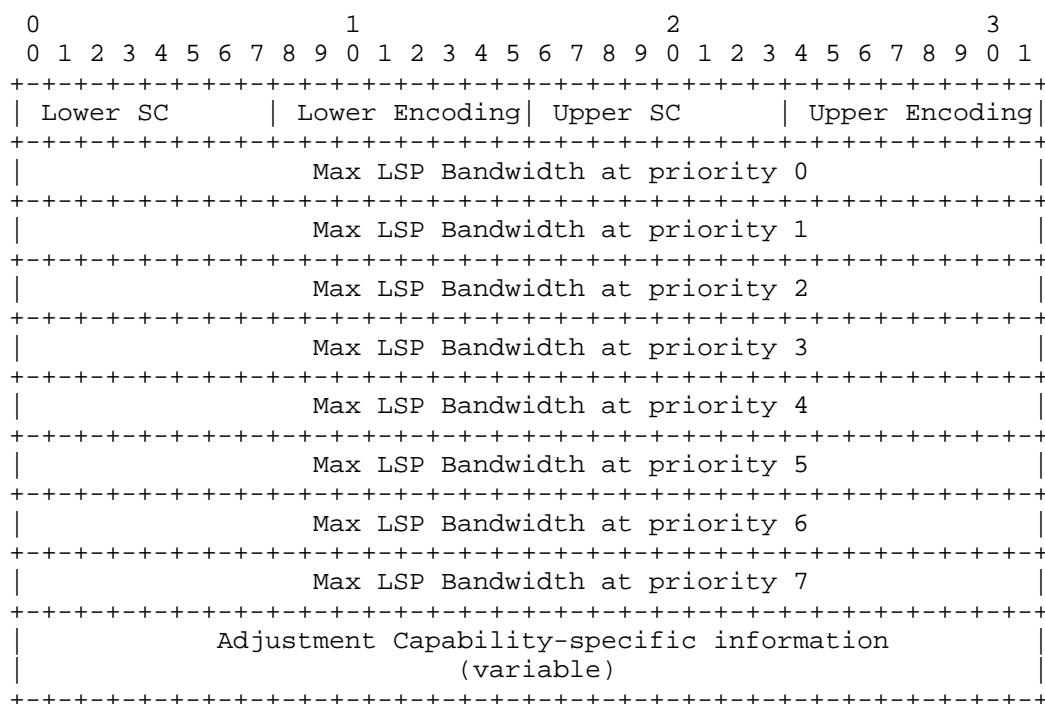


Figure 7: IACD format

6. Missing information

The pieces of information needed for addressing the requirements listed in Section 4 are:

- Mapping information from a client to a server layer. E.g. an ethernet client could be mapped over and OTN hierarchy using a GFP-F or GFP-T adaptation.

- Connectivity constraints: need to describe optical transponder muxing scheme with positioning and restricted connectivity in order to provide end to end connectivity. In the example shown in picture Figure 1, the capability of muxing an SDH hierarchy is shown, but the SDH cannot be injected in any branch of the OTN hierarchy. There is the need to specify that the SDH hierarchy can be only muxed into the ODU->ODU3 branch of the OTN hierarchy and not in all of them.

- Multistage interswitching capability: The IACD already allows advertising the multiplexing of single and multi-stage muxing scenarios like the one in the reference muxing tree, where an SDH hierarchy is muxed over an OTN hierarchy, which is again muxed over an OCh (two levels of muxing).

7. IANA Considerations

TBD

8. Contributors

TBD

9. Acknowledgements

TBD

10. References

10.1. Normative References

[OSPF-OTN]

D.Ceccarelli, D.Caviglia, F.Zhang, D.Li, S.Belotti, P.Grandi, R.Rao, K.Pithewan, J.Drake, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolving G.709 OTN Networks, work in progress draft-ietf-ccamp-gmpls-ospf-g709v3-03", August 2012.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

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

- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC5212] Shiimoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, July 2008.
- [RFC6001] Papadimitriou, D., Vigoureux, M., Shiimoto, K., Brungard, D., and JL. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, October 2010.

10.2. Informative References

- [G.709] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), February 2001.
- [G.709-v3] ITU-T, "Draft revised G.709, version 3", consented by ITU-T on Oct 2009.

Authors' Addresses

Daniele Ceccarelli (editor)
Ericsson
Via Melen 77
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Francesco Fondelli
Ericsson
Via Moruzzi 1
Pisa
Italy

Email: francesco.fondelli@ericsson.com

Sergio Belotti
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

Dimitri Papadimitriou (editor)
Alcatel-Lucent
Copernicuslaan 50
Antwerpen B-2018
Belgium

Email: dimitri.papadimitriou@alcatel-lucent.be

CCAMP Working Group
Internet Draft
Intended status: Standards Track

V.Beeram (Ed)
I.Bryskin
W.Doonan
ADVA Optical Networking
J.Drake (Ed)
G.Grammel
Juniper Networks
Manuel Paul
Ruediger Kunze
Deutsche Telekom
Friedrich Armbruster
Cyril Magaria
NSN
Oscar Gonzalez de Dios
Telefonica

Expires: January 09, 2013

July 09, 2012

Generalized Multiprotocol Label Switching (GMPLS) External Network
Network Interface (E-NNI): Virtual Link Enhancements for the
Overlay Model
draft-beeram-ccamp-gmpls-enni-00.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on January 09, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Abstract

This memo is a companion document to [RFC4208]. It describes how the client domain networking in the overlay model can be enhanced via presenting to the client the network domain as an overlay topology made of Virtual TE Links.

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].

Table of Contents

1. Introduction.....	3
2. Multi-Layered Approach.....	3
3. Traffic Engineering.....	7
3.1. Augmenting the Client layer Topology.....	11
3.1.1. Virtual TE Links.....	13
3.2. Macro SRLGs.....	15
3.3. MELGs.....	17
3.4. Switching Constraints.....	20
4. Connection Setup.....	21
5. Path computation aspects.....	23
6. Access and Virtual TE link addressing.....	24
7. Use cases.....	24
7.1. Service Optimization and Restoration in Multi-layer networks	24
7.2. IP/MPLS Offloading with ENNI automation.....	25
7.3. Use of PCE and VNTM in Multi-layer Network Operation.....	26

8. Security Considerations.....	27
9. IANA Considerations.....	27
10. References.....	27
10.1. Normative References.....	27
10.2. Informative References.....	27
11. Acknowledgments.....	28

1. Introduction

[RFC4208] discusses how GMPLS can be applied to the overlay model, which it defines to be a client network that uses a server network to dynamically instantiate LSPs between the client network's nodes. In the client network such an LSP is a link between two adjacent client nodes, while in the server network the LSP may transit multiple links and nodes; the client network is unaware of the server network topology.

While the client network is unaware of the server network topology, [RFC4208] does suggest that there may be an exchange of routing information, specifically reachability, between the server network and the client network. Building on this premise, this memo describes how introducing a representation of server layer network resources into a client layer network topology enhances client layer networking in the overlay model

This document is designed to be a companion document to [RFC4208], but because routing is generally not considered to be part of the definition of a UNI, this document uses the term 'External Network Network Interface (E-NNI)' to describe this interface between a client and server network, because 'E-NNI' is generally used to indicate a control plane (routing and signaling) exchange of information between two different control plane instances.

2. Multi-Layered Approach

Two adjacent domains in the overlay model represent, generally speaking, regions of dissimilar transport technology. When an end-to-end service crosses a boundary between the domains, it is necessary to execute distinct forms of service activation within each domain/region.

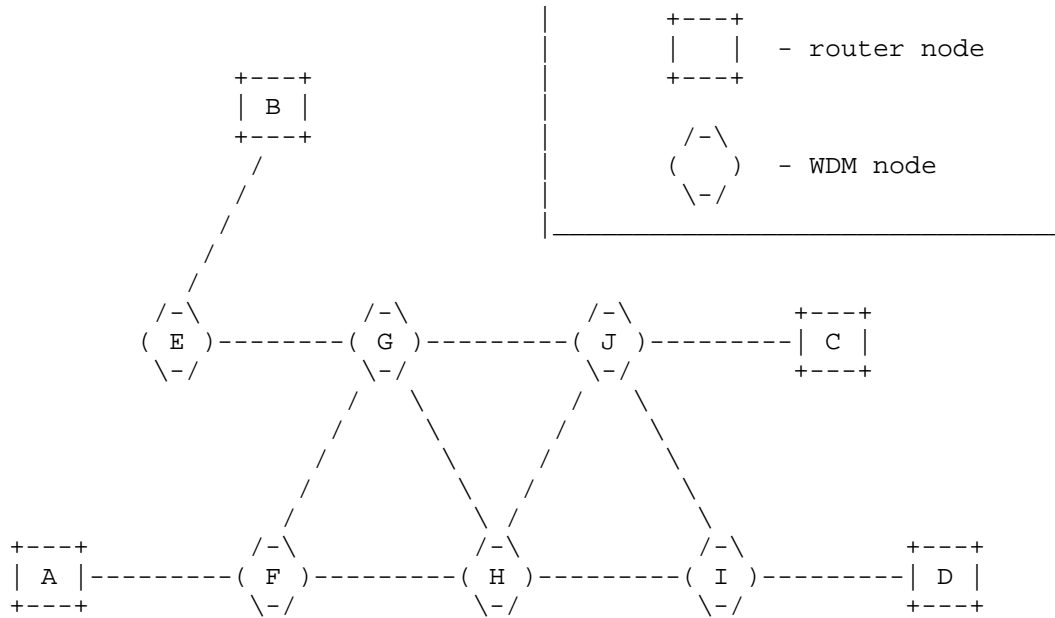


Figure 1: Sample hybrid topology

For example, in the hybrid network illustrated in Fig 1, provisioning a transport service between two GMPLS-enabled IP routers (clients) on either side of the optical WDM transport topology (network domain) requires operations in two distinct layer networks; the client layer network interconnecting the routers themselves, and the server layer network interconnecting the optical transport elements in between the routers.

Activation of the end-to-end service begins with a path determination process, followed by the initiation of a signaling process from the ingress client network element along the determined path, per the example illustrated in Fig 2a-c.

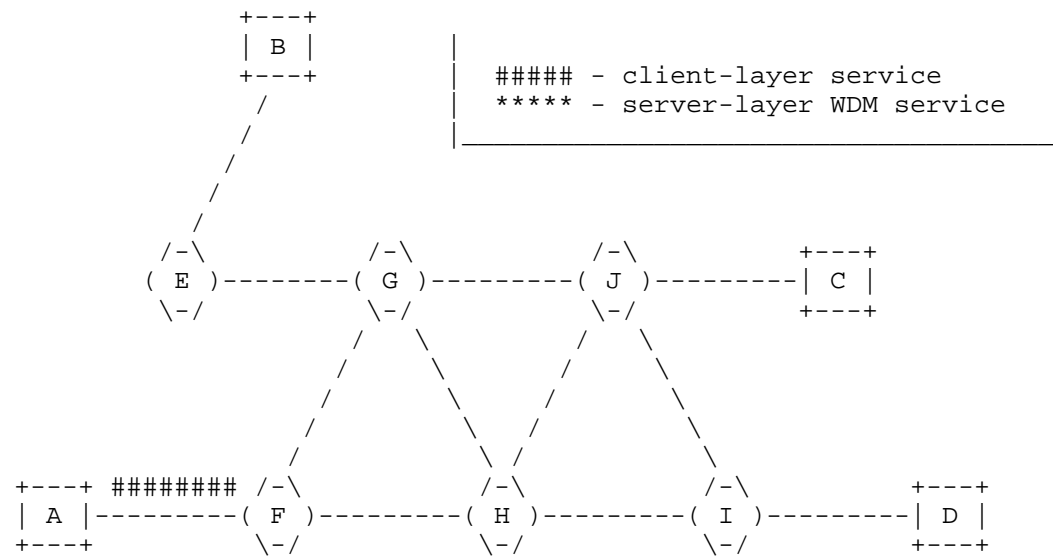


Figure 2a: Hierarchical service activation - Client-layer service setup is initiated.

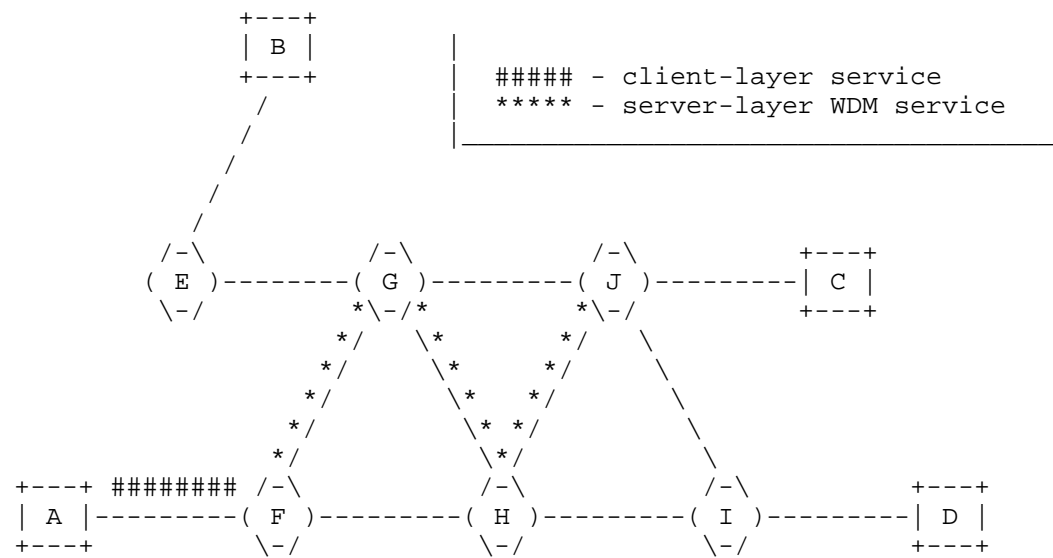


Figure 2b: Hierarchical service activation -
Server-layer WDM service that caters to the
client-layer service is established within the
core.

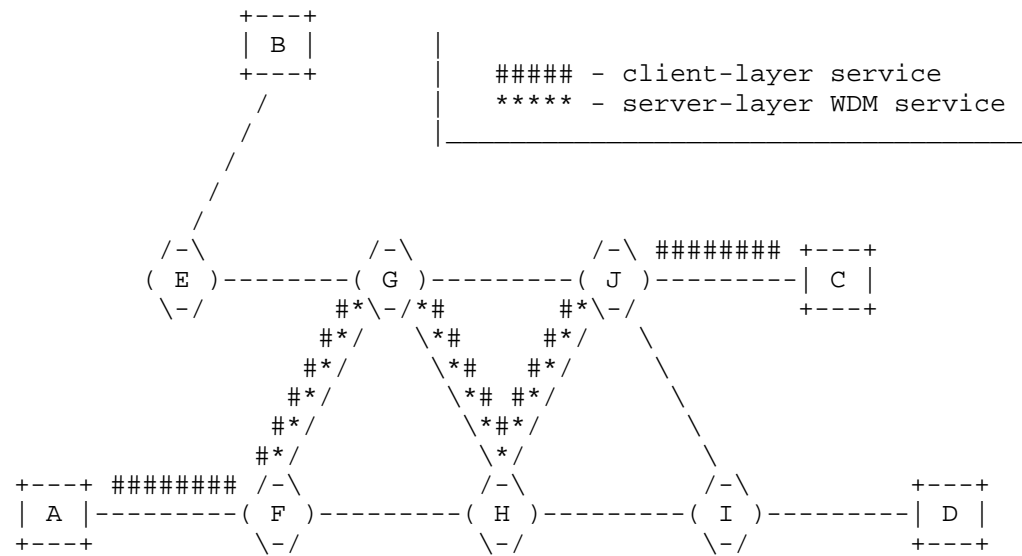


Figure 2c: Hierarchical service activation - Client-layer service setup is resumed and the end-to-end connection is established.

3. Traffic Engineering

The previous section outlines the basic method for activating end-to-end services across a multi-domain/multi-layer network. As a necessary part of that process an initial path selection process is to be performed, whereby an appropriate path between the desired endpoints is to be determined through some means. Further, per expectations set through current practices with regard to service provisioning in homogeneous networks, operators expect that the underlying control plane system provides automated mechanisms for computing the desired path(s) between network endpoints. In particular, operators do not expect under normal circumstances to be required to explicitly specify the end-to-end path; rather, they expect to be able to specify just the endpoints of the path and rely

on an automated computational process to identify and qualify all the elements and links on the path between them. Hence when operating a hybrid multi-layer network such as that described in Fig 1, it is necessary to extend existing traffic engineering and path computation mechanisms to operate in a similar manner.

Path computation and qualification operations occur at the path computation element (PCE - RFC4655) selected by ingress network element of an end-to-end service. In order to be able to compute and qualify paths, the PCE should be provided with information regarding the traffic engineering capabilities of the layer network to which it is associated with, in particular, the topology of the layer network and what layer-specific transport capabilities exist at the various nodes and links in that topology.

It is important to note that topology information is layer-specific; e.g. path computation and qualification operations occur within a given layer, and hence information about topology and resource availability are required for the specific layer to which the connection belongs. The topology and resource availability information required by a path computation element in the client layer is quite distinct from that required by a path computation element in the server layer network. Hence, the actual server layer traffic engineering links are of no importance for the client layer network. In fact, it can be desirable to block their advertisements into the client TE domain by the border nodes.

For example, in the sample hybrid network (Fig 1) there are multiple transport elements supporting client the connection (in this memo terms "connection" and "LSP" are used interchangeably) between the GMPLS-enabled clients A and C, the server layer topology between them includes several nodes and links. However, in this example the optical network elements are not capable of switching traffic with the client layer granularity (i.e. IP/MPLS packets), as the optical network elements are lambda switches, not IP/MPLS switches. Hence, while the intervening server layer network elements may physically exist along the path, they are not a part of the topology required by the client layer nodes for the purposes of traffic engineering in the client layer network.

An example of what the client layer Traffic Engineering topology would look like for the sample hybrid network is shown in the top half of Fig 3.

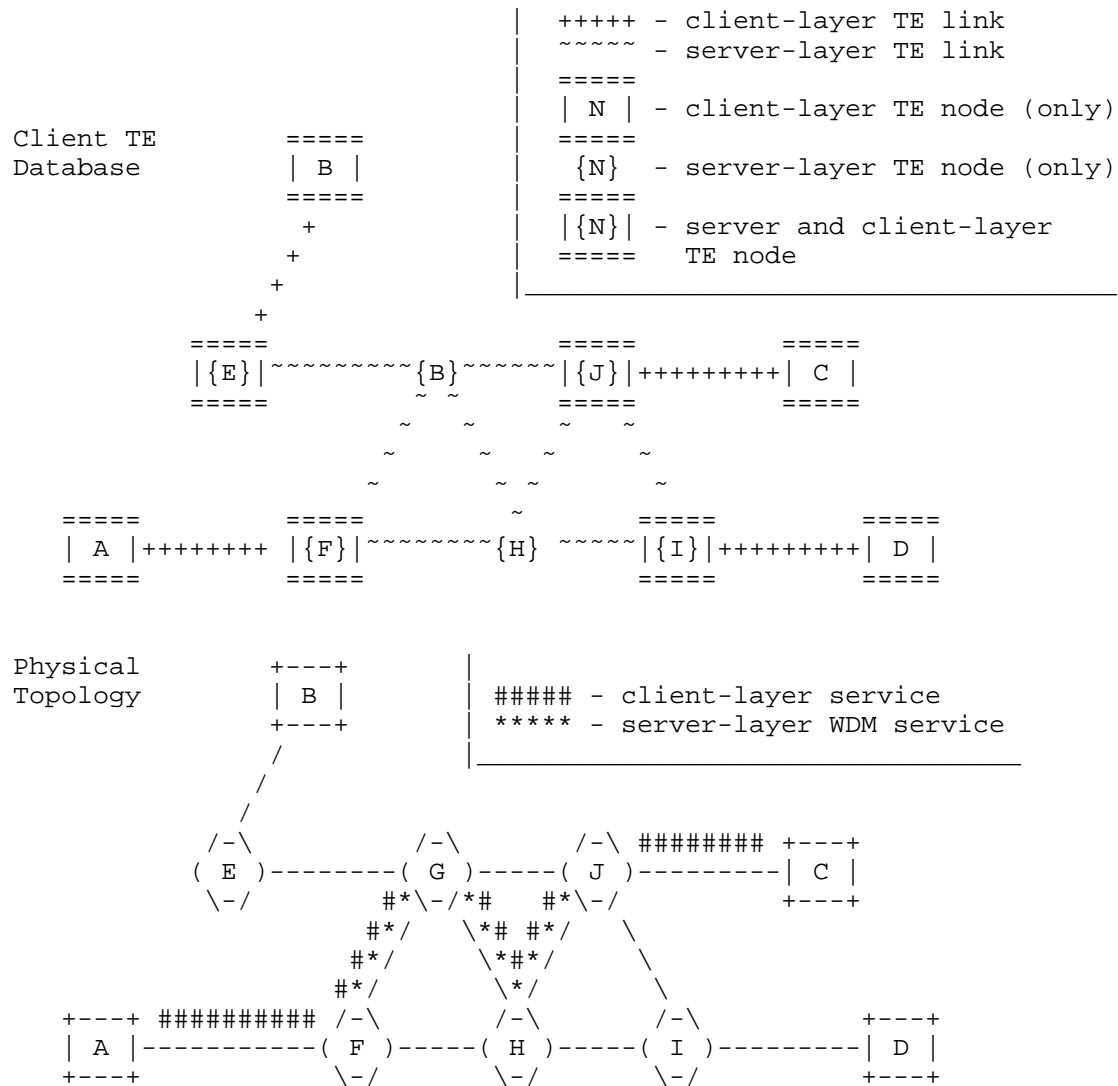


Figure 3: Traffic engineering - ERO with "loose hop"
[Path = {A,F,J,C} (with J loose)].

In this example, the TE topology associated with the client layer network is indicated by the links marked with '+' and nodes marked without brackets, whereas the TE topology associated with the server layer network is indicated by the links marked with '~' and nodes marked in '{}'. The nodes at the edge of the server layer network are visible in both the topologies. The client topology is capable of switching traffic within the client layer, whereas the server topology is capable of switching traffic within the server layer.

In this example, if the "B" router attempts to determine a path to the "D" router it will be unable to do so, as the client topology to which the B and D routers is connected does not include a full path made of just client layer links between them. The only way to setup an end-to-end path in this case is to use an ERO with a "loose hop" across the server layer domain as illustrated in Fig 3. This would cause the server layer to create the necessary link in the client layer topology on the fly. However, this approach has a few drawbacks - [a] the necessity for the operator to specify the ERO with the "loose" hop; [b] potential sub-optimal usage of server layer network resources; [c] unpredictability with regard to the fate-sharing of the new link (that is created on the fly) with other links of the client layer topology.

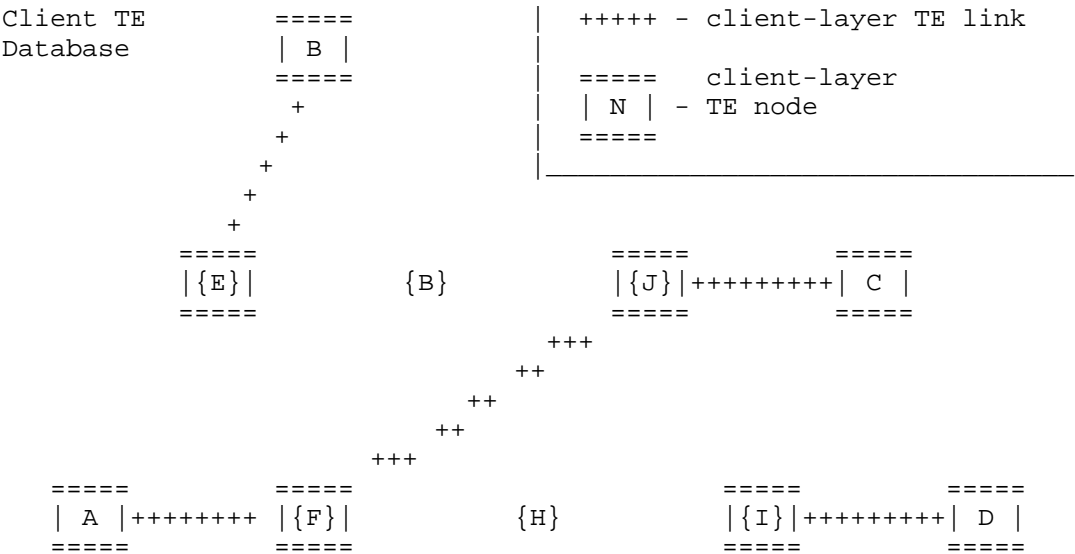
In order to be able to compute an end-to-end path between the two client layer endpoints, the client topology must be sufficiently augmented to indicate where there are paths through the server topology, which can provide connectivity between nodes in the client topology. In other words, in order for a client to compute path(s) across the server layer network to other clients, the feasible paths across the server layer network should be made available (in terms of TE links and nodes that exist in the client layer network) to all the clients. This is discussed in detail in the next section.

As it is mentioned already, in the overlay model the client and network domains, generally speaking, exist in separate layer-networks. One important use case, however, is when the client and network topologies belong to the same layer network. For example, IP routers, connected via GMPLS ENNI to a WDM network, could be capable of terminating optical trails being lambda switched by the network. The method described in the following sections allows also partitioning a single layer network into domains. Those domains do not need to leak the full routing information to their neighboring

domains but rather provide sufficient information for a path computation engine to route connections across a multi-domain network.

3.1. Augmenting the Client layer Topology

In the example hybrid network, shown below in Fig 4, consider a scenario, where each GMPLS-enabled IP router is connected to the optical WDM transport network via a transponder. Further, consider the situation, where the transponder on node F can be connected to the transponder on node J via the optical path F-G-H-J. Suppose, a lambda LSP is provisioned in the server layer along this path and advertised (as a TE link) into the client layer network. With the availability of this TE link, the path computation function at node



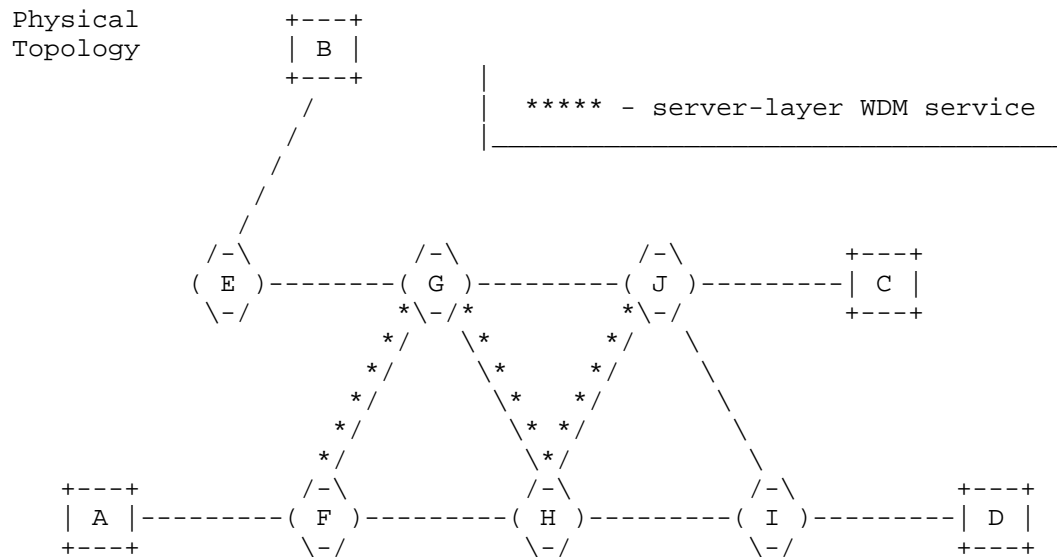


Figure 4: Traffic engineering - end-to-end path computation.[The client layer "TE link" between F and J is produced by creating the underlying server-layer connection; Node A has visibility to end-to-end (A to C) client-layer links and can do CSPF]

A is able to compute an end-to-end path from A to C. In this example, in order for the TE link to be made available in the client layer network topology, the network resources supporting the underlying server layer LSP are fully committed beforehand.

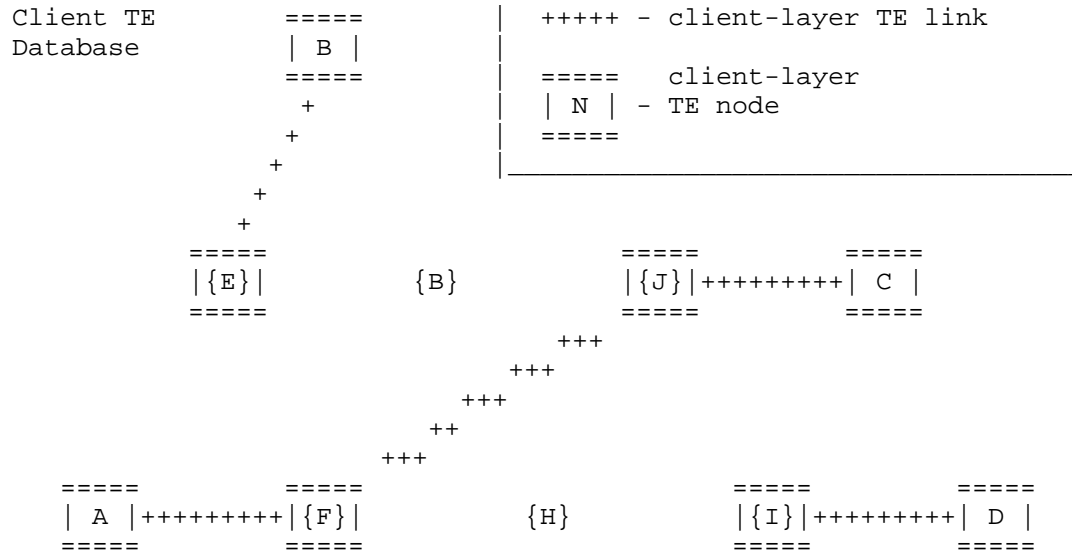
As another scenario, consider a network configuration, where the transponders on nodes E, F, J and I are connected to each other via directionless ROADM technology. In this case it is physically possible to connect any transponder to any other transponder in the server layer network. As there are transport capabilities available in the server layer network between every pair of elements with an adaptation function to the client layer network, the operator in this case would not wish to commit any network resources in the

server layer network until a client LSP is signaled. The next section proposes a method to address this common operational requirement.

3.1.1.1. Virtual TE Links

A "Virtual TE Link" as defined in section 7.3.3 of [RFC4847] is a TE link that is advertised into the client layer network. The advertisement includes information about available but not necessarily reserved/committed resources in the server layer network necessary to support that TE link. In other words, Virtual TE Links represent specific transport capabilities available in the server layer network, which can support the establishment of LSPs in the client layer network.

The two fundamental properties of a Virtual TE Link are: [a] it is advertised just like a real TE link and thus contributes to the buildup of the client layer network topology; and [b] it does not require allocation of resources at the server layer until used, thus allowing the mutually exclusive sharing of server layer network resources with other Virtual TE Links.



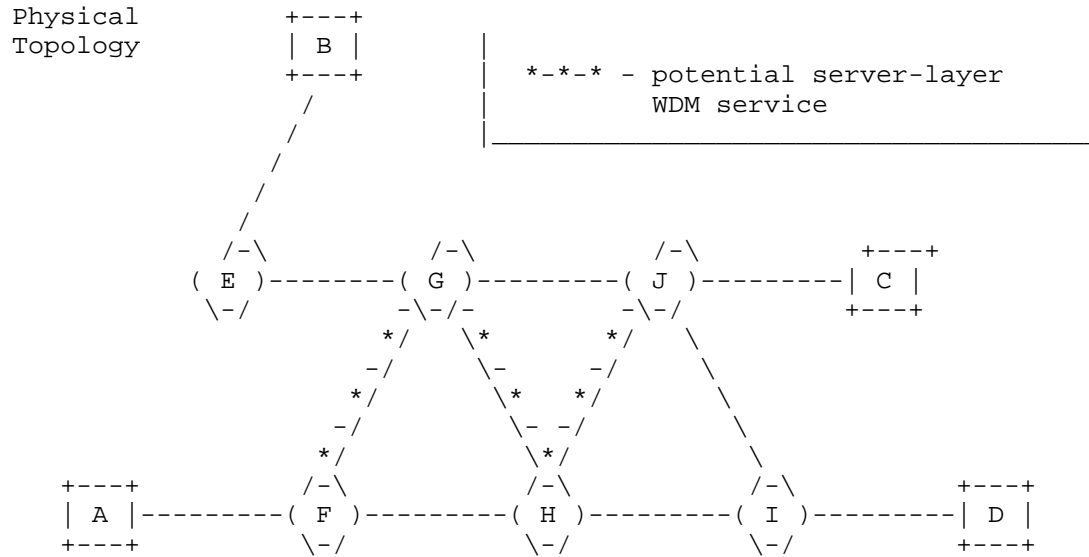


Figure 5: Traffic engineering - end-to-end path computation with "Virtual TE Links". [The "Virtual TE link" between F and J is created in the client layer without actually instantiating the underlying server-layer connection; Node A has visibility to end-to-end client-layer links and can do CSPF]

In the example shown in Fig 5, the availability of a lambda channel along the path F-G-H-J results in the advertisement by nodes F and J of a Virtual TE Link between F and J into the client layer network topology (+++ line). With the advertisement of this Virtual TE Link, the path computation function at node A is able to compute an end-to-end path from A to C.

Whenever a Virtual TE Link gets selected and signaled in the ERO of a client layer LSP, it ceases temporarily to be "virtual" and transforms into a regular TE link. When this transformation takes place, the clients will notice the change in the advertised available bandwidth of this TE link. Also, all other Virtual TE Links that share in a mutual exclusive way some of server layer resources with the TE link in question SHOULD start advertising "zero" available bandwidth. Likewise, the TE network image reverts back to the original form as soon as the last client layer LSP, going through the TE link in question, is released, i.e. Virtual TE Link becomes "virtual" again.

The overlay topology, advertised into the client domain as a set of Virtual TE Links, along with access TE links (the TE links interconnecting client network elements with the network domain) makes up the topology that in the overlay model allows for the client domain path computation function to compute end-to-end paths interconnecting client network elements across the network domain.

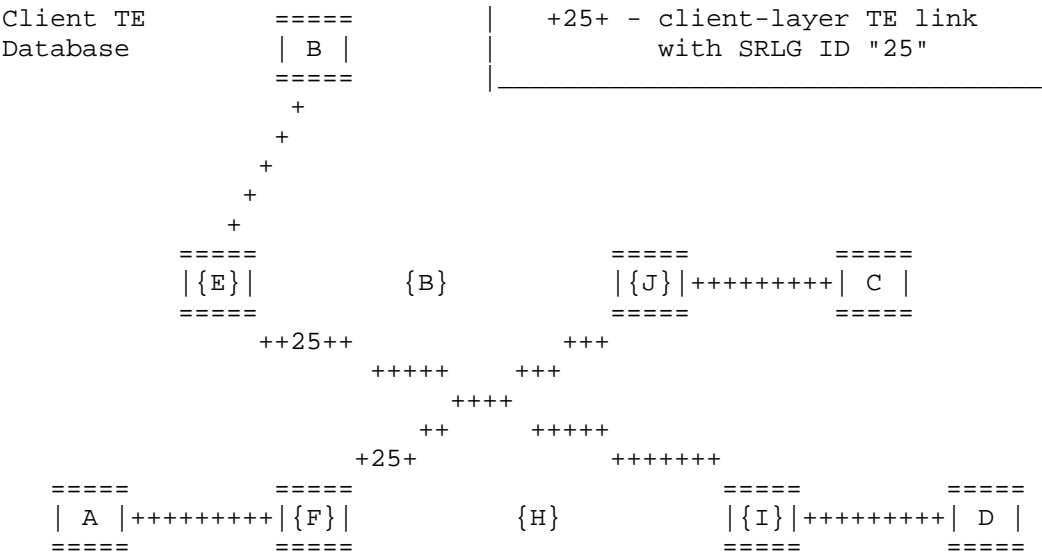
3.2. Macro SRLGs

The Virtual TE Links, which are advertised into the client layer network topology, cannot be assumed to be independent. It is quite possible for a given Virtual TE Link to share fate with one or more other Virtual TE Link(s). This is because the underlying server layer LSPs (established or potential) can traverse the same server layer network link and/or node, and failure of any such shared link/node would make all such LSPs inoperable (along with the Virtual TE Links supported by the LSPs). If diverse end-to-end paths for client layer LSPs are to be computed, the fate sharing information of the Virtual TE Links needs to be taken into account. The standard way of addressing this problem is via the concept of Shared Risk Link Group (SRLG). Specifically, a network resource shared by two or more TE links is identified via a network scope unique number (SRLG ID) and advertised within each such TE link advertisement.

A "traditional" SRLG (per [RFC4202]) represents a shared physical network resource, upon which normal function of a link depends. Such SRLGs can also be referred to as physical SRLGs. Zero, one or more physical SRLGs could be identified and advertised for every TE link in a given layer network. There is a scalability issue with physical SRLGs in multi-layer environments. For example, if a server layer LSP serves a client layer link, every server layer link and node traversed by the LSP must be considered as a separate SRLG. The number of server layer SRLGs to be advertised to client layer per

TE link is directly proportional to the number of hops traversed by the underlying server layer LSP.

This document introduces a notion of Macro SRLGs, which addresses this scaling problem. Macro SRLGs have the same protocol format as their physical counterparts and can be assigned automatically for each TE link that is advertised into the client layer network supported by an underlying server layer LSP (instantiated or otherwise). A Macro SRLG represents a shared path segment that is traversed by two or more of the underlying server layer LSPs. Each shared path segment can be viewed as a set of shared server layer resources. The actual procedure for deriving the Macro SRLGs is beyond the scope of this document.



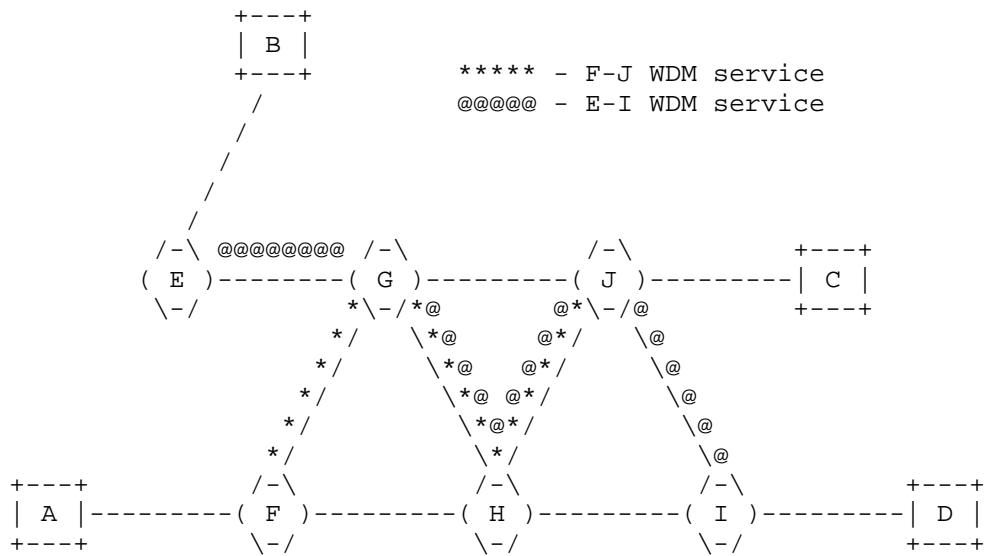


Figure 6: Macro SRLGs - ["TE links" E-I and F-J share fate since the underlying server-layer connections traverse the same path segments [G-H][H-I]. Macro SRLG-ID "25" is assigned to both "TE links"]

3.3. MELGs

If two or more Virtual TE Links share fate, it means that the links could be concurrently activated and used by client LSPs with a caveat that the links could be taken out of service by a single network failure, and, thus, cannot be used in the same protection scheme. There could be a stronger (than fate sharing) relationship between two or more Virtual TE Links. Because a set of Virtual TE Links can depend on the same uncommitted network resources, the situation can arise, when only one Virtual TE Link from the set could be activated at any given time. In other words, two or more Virtual TE Links can be mutually exclusive.

One example of the mutually exclusive relationship of Virtual TE Links is when the paths for the server layer network LSPs supporting the Virtual TE Links not only intersect, but also require usage of the same resource (e.g. lambda channel) on the intersection (see Figure 7). Another example is when the said paths depend on a common physical resource (e.g. transponder, regenerator, wavelength converter, etc.) that could be used only by one LSP at a time.

For a client path computation function (especially a centralized one capable of concurrent computation of multiple paths) it is important to know about such mutually exclusive relationship between Virtual TE Links. This document introduces a concept of Mutually Exclusive Link Group (MELG) and suggests a new sub-TLV - MELGs sub-TLV - to be added to the top level TE Link TLV. The purpose of the MELGs sub-TLV is:

- To indicate via a separate network unique number (MELG ID) an element or a situation that makes the advertised Virtual TE Link to belong to one or more Mutually Exclusive Link Groups. Path computing element will be able to decide on whether two or more Virtual TE Links are mutually exclusive or not by finding an overlap of advertised MELGs (similar to deciding on whether two or more TE links share fate or not by finding common SRLGs)
- To indicate whether the advertised Virtual TE Link is committed or not at the moment of the advertising. Such information is important for a path computation element: committing new Virtual TE links (vs. re-using already committed ones) has a consequence of allocating more server layer resources and disabling other Virtual TE Links that have common MELGs with newly committed Virtual TE Links.

The format of the MELGs sub-TLV is defined as follows:

Name: MELGs

Type: TBD

Length: Variable

0																1																2																3															
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9																								
Sub-TLV Type																Sub-TLV Length																																															
Flags (16 bits)																U	Number of MELGs (16 bits)																																														
MELGID1 (64 bits)																																MELGID2 (64 bits)																															

```

| ..... |
| MELGIDn (64 bits) |
+-----+

```

Number of MELGs: number of MELGS advertised for the Virtual TE Link;

Flags: Virtual TE Link specific flags;

MELGID1,MELGID2,...,MELGIDn: 64-bit network domain unique numbers associated with each of the advertised MELGs

Currently defined Virtual TE Link specific flags are:

- . U bit (bit 1) : Uncommitted ,if set, the Virtual TE Link is uncommitted at the time of the advertising (i.e. the server layer network LSP is not set up); if cleared, the Virtual TE Link is committed (i.e. the server layer LSP is fully provisioned and functioning). All other bits of the "Flags" field are reserved for future use and MUST be cleared.

Note: A Virtual TE Link advertisement MAY include MELGs sub-TLV with zero MELGs for the purpose of communicating to the TE domain whether the Virtual TE Link is currently committed or not.

```

Client TE      ===== | +25/192000+ - client-layer TE link
Database      | B |      | with SRLG ID "25" and
               ===== | Lambda Channel 192000
               +
               +
               +
               +
               +
=====        {B}        =====
| {E} |          | {J} | ++++++ | C |
=====        =====
++25/192000++      +++++
++++              +++
++++              +
++++              +
++                ++++++
+25/192000+        ++++++
=====            =====
| A | ++++++ | {F} |          {H}          | {I} | ++++++ | D |
=====            =====

```

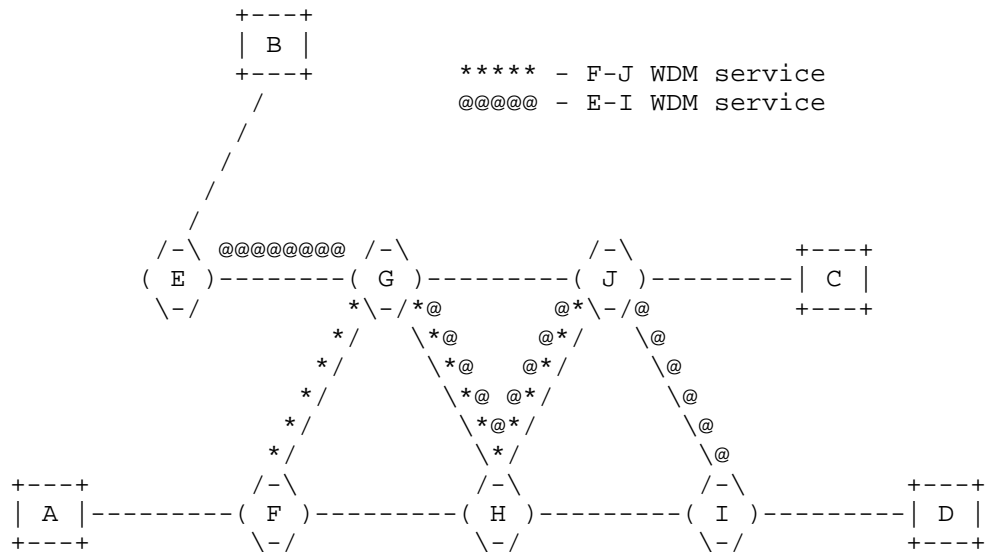


Figure 7: MELGs - ["TE links" E-I and F-J are mutually exclusive (server paths require usage of the same resource: lambda channel 192000). Same MELG ID is assigned to both TE links]

3.4. Switching Constraints

Generally speaking, it SHOULD NOT be assumed that a Virtual TE Link advertised by a given network domain border node can be cross-connected within a client LSP with every access TE link advertised by the said node. This circumstance necessitates the specification of connectivity constraints by network domain border nodes. If such information is not available for client domain path computers, there is a significant risk of provisioning failures of client LSPs, if/when they are signaled with the computed paths (see, Fig 7). This document recommends the use of the advertisements specified in [GEN_CNSTR] and [OSPF_GEN_CNSTR] to address the network element switching limitations problem.

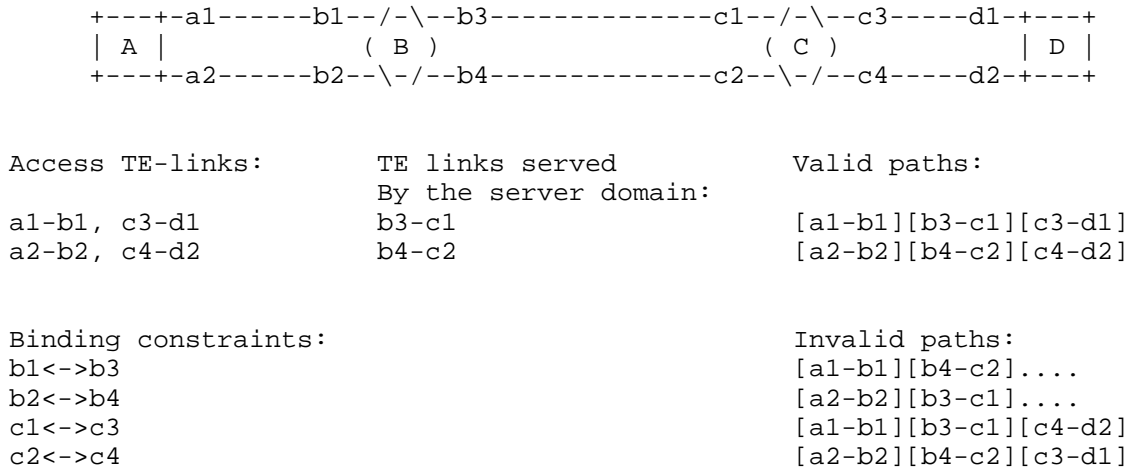


Figure 7: Switching Constraints

4. Connection Setup

Experience with control plane operations in multi-layer networks indicates some benefits in coordinating certain signaling operations of client layer network LSPs and underlying server layer network LSPs in the following manner. Consider the scenario, where the network is a WDM layer topology comprising of ROADMs. The set-up time for a service at the WDM layer can be fairly long, as it can involve time-consuming power-equalization procedures, amongst other layer specific operations. This means that at very least, the setup timers for the client LSPs would need to be somehow coordinated with that of the server LSPs. To avoid this operationally awkward issue, a phased LSP setup process as depicted in Fig 8 is proposed.

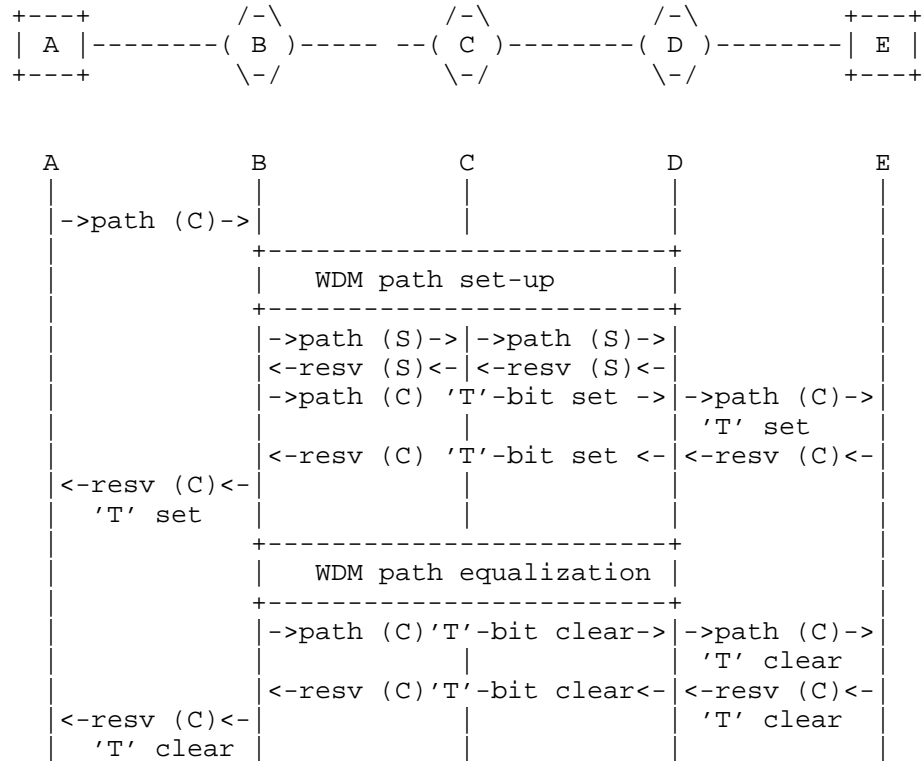


Figure 8: connection set-up

As long as the server LSP is not completely established (i.e. successfully power equalized), the server layer network border nodes, through which the client LSP passes, would signal PATH/RESV messages with the T (Testing) bit set in the ADMIN_STATUS. The T bit would be cleared in these messages only after all server LSPs supporting links taken by the client LSP in question are deemed fully operable.

5. Path computation aspects

It is assumed that a client domain path computation function makes use of advertised access TE links as well as Virtual TE Links, while computing end-to-end paths for client LSPs. The said path computation function could be local (i.e. located on client LSP ingress nodes, as stipulated by [RFC4655] Composite PCE node) or remote (i.e. on network PCEs). Path computations could be triggered by client nodes or NMS. Generally speaking, the responsibility of the client domain path computation function is to (concurrently) compute one or several paths for each source-destination pair (potential client LSP termination points) specified in a single path computation request. The path computation SHOULD be subject to one or more path optimization criterions (such as minimal cost, minimal latency, etc.) and a set of path computation constraints (such as link unreserved bandwidth, link colors, layer-specific constraints, explicit inclusions and exclusions, etc.)

As the overlay topology hides actual server domain/layer links and nodes, it is RECOMMENDED to support SRLG diverse computation of two or more paths.

Furthermore, the path computation SHOULD consider the connectivity/switching limitation constraint (when available) in addition to all other path computation constraints.

The use of the PCE architecture and PCEP protocol is governed by [RFC5440], [RFC5521] and [RFC5541].

As described in section 3.3., two or more Virtual TE Links may not only share risk, but also may exclusively depend on the same server layer resources. Therefore, paths, computed on network topologies containing Virtual TE Links, have an increased probability of LSP setup failures (two LSPs, for example, could be routed over two Virtual TE Links that exclusively depend on the same server layer resource). In such cases concurrent path computation, taking in consideration MELG information, will address this problem. PCEP supports concurrent path computation per [RFC5440]. Specifying MELG diversity constraint in path computation requests is out of scope of this document.

In addition MELG may carry information on the establishment of server-layer resources. A Path computation request MAY constraint the path computation to TE-Links that are fully provisioned only. This information MAY also be used in PCE path computation policies.

6. Access and Virtual TE link addressing

[RFC4208] implies that access TE links could be named from the same address space as network domain TE links or from a separate address space. This memo requires the following:

- It MUST be possible to assign addresses for access TE links from the same address space as the one used for naming network internal TE links (i.e. TE links interconnecting network domain devices);
- It MAY be possible to assign addresses for access TE links from a separate address space, independent from the space used for addressing network internal TE links;
- Virtual TE Links MUST share the address space with any access TE links they are allowed to be cross-connected within a client LSP.

7. Use cases

7.1. Service Optimization and Restoration in Multi-layer networks

Multi-layer networks are a reality today, and they are operated by different groups of people, following different operational procedures. This requires an independent optimization of the client and server layer networks. Such independence may cause a situation, where the re-routing of a client layer LSP fails, because some of resources on the selected alternate path share fate with some of resources on the LSP's failed path. This usually happens due to lack of knowledge of the server layer network by a client layer path computation function at the time when the alternative path is selected.

The high volume and importance of IP traffic in provider networks today requires the client and server layer networks to share sufficient information in order to enable an optimized transport for IP/MPLS services and address existing inefficiencies. From the carrier perspective it is very important that the SRLG information is provided by the server layer TE application and is used by the client layer path computation.

In a typical multi-layer network, where IP/MPLS is the client layer network and WDM/OTN is the server layer network, the client layer network is responsible for the protection of the IP/MPLS traffic from networks failures. This is normally achieved via using

protection schemes, such as FRR and/or LFA. Regardless of the used mechanism, the SRLG information, provided by the server layer network, helps to optimize the client layer network with respect to reduced link utilization and reliable and efficient protection of the user traffic.

Today the SRLGs information is used mainly when calculating diverse alternative paths for the IP/MPLS LSPs. Therefore, the following procedures are performed periodically:

- Building traffic matrix for the server layer network (based on IP links)
- Solving traffic engineering problems in the server layer network
- (Re-)Calculating SRLGs to be propagated into the client layer network
- Simulating failure scenarios
- Making sure that the affected IP/MPLS LSPs function properly after they are replaced onto SRLG diverse alternative paths

GMPLS ENNI reduces the OPEX costs of performing these procedures via the automation as follows:

- server layer network automatically discovers and advertises the SRLG information into client layer network via a common routing protocol;
- client layer network path computer uses the SRLG information when selecting diverse paths.

7.2. IP/MPLS Offloading with ENNI automation

A typical application in multi-layer (IP/MPLS over optical) networks is termed 'IP Offloading', in which the network responds to the increase in traffic of a particular service or across a segment in the IP network by dynamically creating additional IP/MPLS links served by GMPLS LSPs provisioned in the server layer network, and placing the extra IP/MPLS traffic onto said links. Likewise, when the IP/MPLS traffic decreases to a normal pattern, the said GMPLS LSPs are torn down, and the extra IP/MPLS links are removed from the client layer network TE domain. The increase in traffic is typically caused by an elevated number of high traffic flows/services traversing an IP network segment.

The decision process driving IP offloading is complex, and is governed by a set of rules. These rules reduce the cost of running the multi-layer network, while ensuring that it remains stable.

Automation of IP Offloading poses a number of challenges. It includes dynamic provisioning, release and maintenance of GMPLS LSPs in the server layer (e.g. WDM) network as well as automatic advertising/withdrawing them as (numbered or/and unnumbered) TE links into/from the client layer network. In order to pre-plan and manage properly the said dynamic IP/MPLS TE links, it is important to know in advance (and also in real time) the capabilities and resource availability of server layer network. The network domain/layer virtualization procedures described in this document helps to solve this complex operational issue.

7.3. Use of PCE and VNTM in Multi-layer Network Operation

Two key elements have been proposed to help in the management and coordination of multi-layer networks: the Path Computation Element (PCE) and the Virtual Network Topology Manager (VNTM). PCE is responsible for the calculation of paths between endpoints, particularly in complex scenarios involving, for example, WDM layer physical impairments. VNTM is in charge of maintaining the topology of the client layer network by instantiating virtual links, in the server layer network. I.e., it can be used to provide TE links to the client layer network dynamically.

Several cooperation modes between PCE, VNTM and the NMS have been proposed in [RFC5623]. For instance, the operator can request a new MPLS tunnel via the NMS, which communicates with a PCE with information of the multi-layer network. The PCE, in case there are enough resources in the IP/MPLS layer, normally returns a path for the tunnel made of real TE links. On the other hand, if there is a lack of resources in the IP/MPLS layer, the response may contain a path with one or more Virtual TE Links. In this case, the NMS can cooperate with the VNTM to suggest the set-up of a GMPLS LSP(s) in the server layer network. The VNTM, based on the local policies, can accept the suggestion and cause the set-up of the GMPLS LSPs in the server layer network.

In order for the computation to be effective, the PCE needs knowledge of the overlay topology (SRLGs, MELGs, TE metrics of the Virtual TE links), which can be provided via GMPLS ENNI.

8. Security Considerations

TBD

9. IANA Considerations

TBD.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4202] K. Kompella, Y.Rekhter
"Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)",
RFC 4202, October 2005.
- [RFC4208] G. Swallow, J.Drake, H. Ishimatsu, and Y. Rekhter,
"GMPLS UNI: RSVP-TE Support for the Overlay Model",
RFC 4208, October 2005.
- [GEN_CNSTR] G.Bernstein, Y.Lee, D.Li, W.Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks"
[draft-general-constraint-encode-08.txt]
- [OSPF_GEN_CNSTR] F.Zhang, J.Han, Y.Lee, D.Li, G.Bernstein, Y.Hu
"OSPF-TE Extensions for General Network Element Constraints"
[draft-general-constraints-ospf-te-04.txt]

10.2. Informative References

- [RFC4847] T. Takeda, "Framework and Requirements for Layer 1 VPNs", RFC 4847, April 2007.
- [RFC4655] A. Farrel, J.-P. Vasseur, J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

11. Acknowledgments

Chris Bowers [cbowers@juniper.net]

Authors' Addresses

Vishnu Pavan Beeram
ADVA Optical Networking

Email: vbeeram@advaoptical.com

Igor Bryskin
ADVA Optical Networking

Email: ibryskin@advaoptical.com

Wes Doonan
ADVA Optical Networking

Email: wdoonan@advaoptical.com

John Drake
Juniper Networks

Email: jdrake@juniper.net

Gert Grammel
Juniper Networks

Email: ggrammel@juniper.net

Manuel Paul
Deutsche Telekom

Email: Manuel.Paul@telekom.de

Ruediger Kunze
Deutsche Telekom

Email: Ruediger.Kunze@telekom.de

Oscar Gonzalez de Dios
Telefonica

Email: ogondio@tid.es

Cyril Margaria
Nokia Siemens Networks

Email: cyril.margaria@nsn.com

Friedrich Armbruster
Nokia Siemens Networks

Email: friedrich.armbruster@nsn.com

CCAMP Working Group
Internet Draft
Intended status: Standards Track

Igor Bryskin (Ed)
Wes Doonan
ADVA Optical Networking
Vishnu Pavan Beeram (Ed)
John Drake (Ed)
Gert Grammel
Juniper Networks
Manuel Paul
Ruediger Kunze
Deutsche Telekom
Friedrich Armbruster
Cyril Margaria
NSN
Oscar Gonzalez de Dios
Telefonica

Expires: April 22, 2013

October 22, 2012

Generalized Multiprotocol Label Switching (GMPLS) External Network
Network Interface (E-NNI): Virtual Link Enhancements for the
Overlay Model
draft-beeram-ccamp-gmpls-enni-01.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on April 22, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Abstract

This memo is a companion document to [RFC4208]. It describes how the client domain networking in the overlay model can be enhanced via presenting to the client the network domain as an overlay topology made of Virtual TE Links.

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].

Table of Contents

1. Introduction.....	3
2. Hybrid Topology.....	3
3. Traffic Engineering.....	7
3.1. Augmenting the Client layer Topology.....	11
3.1.1. Virtual TE Links.....	13
3.2. Macro SRLGs.....	15
3.3. MELGs.....	17
3.4. Switching Constraints.....	20
4. Connection Setup.....	21
5. GMPLS ENNI and Multiple Server Network Domains.....	23
6. Path computation aspects.....	25
7. Access and Virtual TE link addressing.....	26
8. Use cases.....	26
8.1. Service Optimization and Restoration in Multi-layer networks	26
8.2. IP/MPLS Offloading with ENNI automation.....	27

8.3. Use of PCE and VNTM in Multi-layer Network Operation.....	28
9. Security Considerations.....	29
10. IANA Considerations.....	29
11. References.....	29
11.1. Normative References.....	29
11.2. Informative References.....	29
12. Acknowledgments.....	30

1. Introduction

[RFC4208] discusses how GMPLS can be applied to the overlay model, which it defines to be a client network that uses a server network to dynamically instantiate LSPs between the client network's nodes. In the client network such an LSP is a link between two adjacent client nodes, while in the server network the LSP may transit multiple links and nodes; the client network is unaware of the server network topology.

While the client network is unaware of the server network topology, [RFC4208] does suggest that there may be an exchange of routing information between the server network and the client network. Building on this premise, this memo describes how introducing a representation of server network domain resources into a client network domain topology enhances client networking in the overlay model

This document is designed to be a companion document to [RFC4208], but because routing is generally not considered to be part of the definition of a UNI, this document uses the term 'External Network Network Interface (E-NNI)'. 'E-NNI' is generally used to indicate a control plane (routing and signaling) reference point for exchange of information between two control plane instances. In this document, the term 'ENNI' (as described in [OVERLAY_FWK]) is specifically used to describe the interface between two network domains that allows the exchange of routing and signaling information.

2. Hybrid Topology

Two adjacent domains in the overlay model represent, generally speaking, regions of dissimilar transport technology. When an end-to-end service crosses a boundary between the domains, it is necessary to execute distinct forms of service activation within each domain.

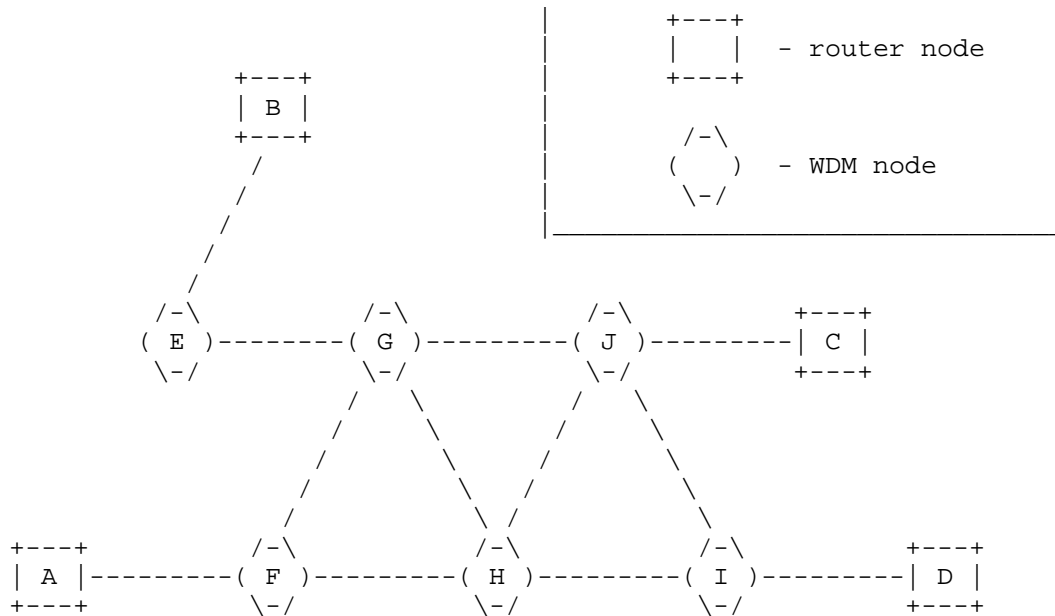


Figure 1: Sample hybrid topology

For example, in the hybrid network illustrated in Fig 1, provisioning a transport service between two GMPLS-enabled IP routers (clients) on either side of the optical WDM transport topology (server network domain) requires operations in two distinct layer networks; the client layer network interconnecting the routers themselves, and the server layer network interconnecting the optical transport elements in between the routers.

The activation of the end-to-end service begins with a path determination process, followed by the initiation of a signaling process from the ingress client network element along the determined path, per the example illustrated in Fig 2a-c.

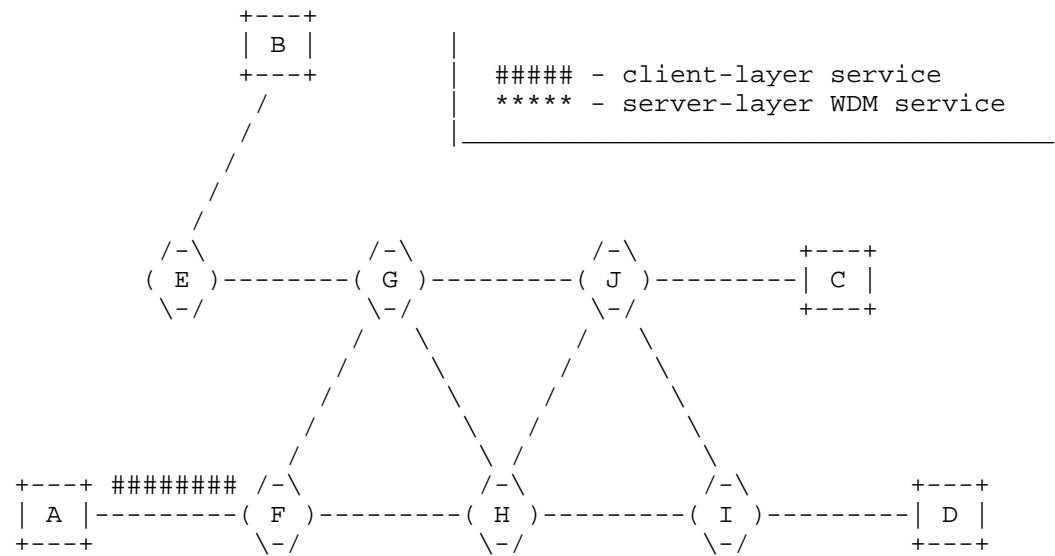


Figure 2a: Hierarchical service activation -
Client-layer service setup is initiated.

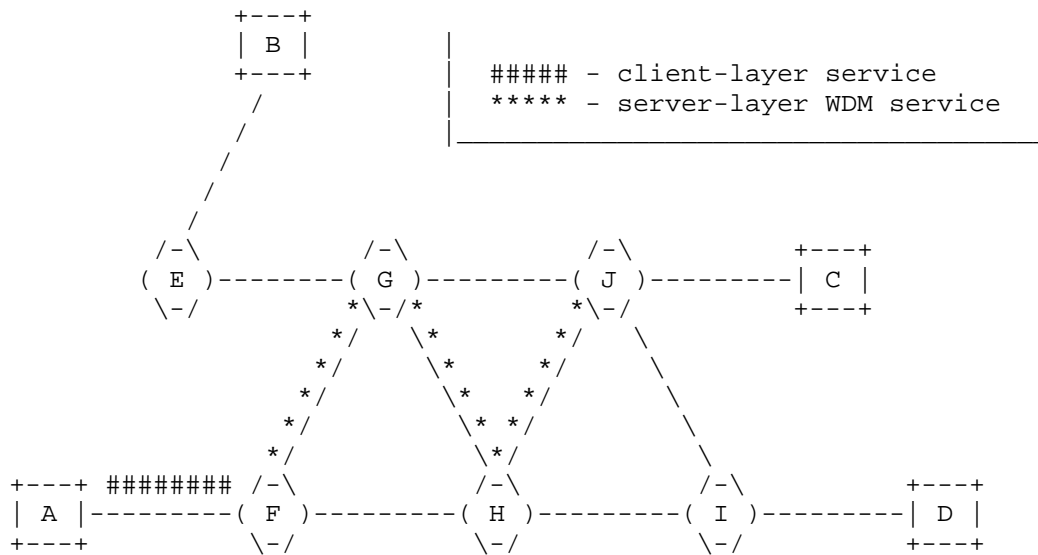


Figure 2b: Hierarchical service activation -
Server-layer WDM service that caters to the
client-layer service is established within the
core.

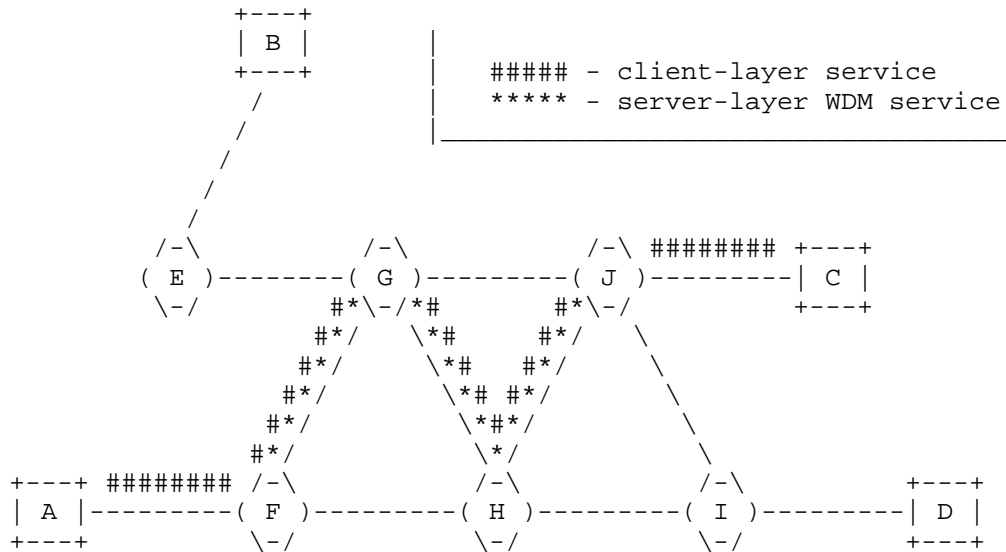


Figure 2c: Hierarchical service activation - Client-layer service setup is resumed and the end-to-end connection is established.

3. Traffic Engineering

The previous section outlines the basic method for activating end-to-end services across a multi-domain/multi-layer network. As a necessary part of that process an initial path selection process is to be performed, whereby an appropriate path between the desired endpoints is to be determined through some means. Further, per expectations set through current practices with regard to service provisioning in homogeneous networks, operators expect that the underlying control plane system provides automated mechanisms for computing the desired path(s) between network endpoints.

In particular, operators do not expect under normal circumstances to be required to explicitly specify the end-to-end path; rather, they expect to be able to specify just the endpoints of the path and rely on an automated computational process to identify and qualify all the elements and links on the path between them. Hence when operating a hybrid multi-layer network such as that described in Fig 1, it is necessary to extend existing traffic engineering and path computation mechanisms to operate in a similar manner.

Path computation and qualification operations occur at the path computation element (PCE - RFC4655) selected by ingress network element of an end-to-end service. In order to be able to compute and qualify paths, the PCE should be provided with information regarding the traffic engineering capabilities of the layer network to which it is associated with, in particular, the topology of the layer network and what layer-specific transport capabilities exist at the various nodes and links in that topology.

It is important to note that topology information is layer-specific; e.g. path computation and qualification operations occur within a given layer, and hence information about topology and resource availability are required for the specific layer to which the connection belongs. The topology and resource availability information required by a path computation element in the client layer is quite distinct from that required by a path computation element in the server layer network. Hence, the actual server layer traffic engineering links are of no importance for the client layer network. In fact, it can be desirable to block their advertisements into the client TE domain by the border nodes.

For example, in the sample hybrid network (Fig 1) there are multiple transport elements supporting client the connection (in this memo terms "connection" and "LSP" are used interchangeably) between the GMPLS-enabled clients A and C, the server layer topology between them includes several nodes and links. However, in this example the optical network elements are not capable of switching traffic with the client layer granularity (i.e. IP/MPLS packets), as the optical network elements are lambda switches, not IP/MPLS switches. Hence, while the intervening server layer network elements may physically exist along the path, they are not a part of the topology required by the client layer nodes for the purposes of traffic engineering in the client layer network.

An example of what the client layer Traffic Engineering topology would look like for the sample hybrid network is shown in the top half of Fig 3.

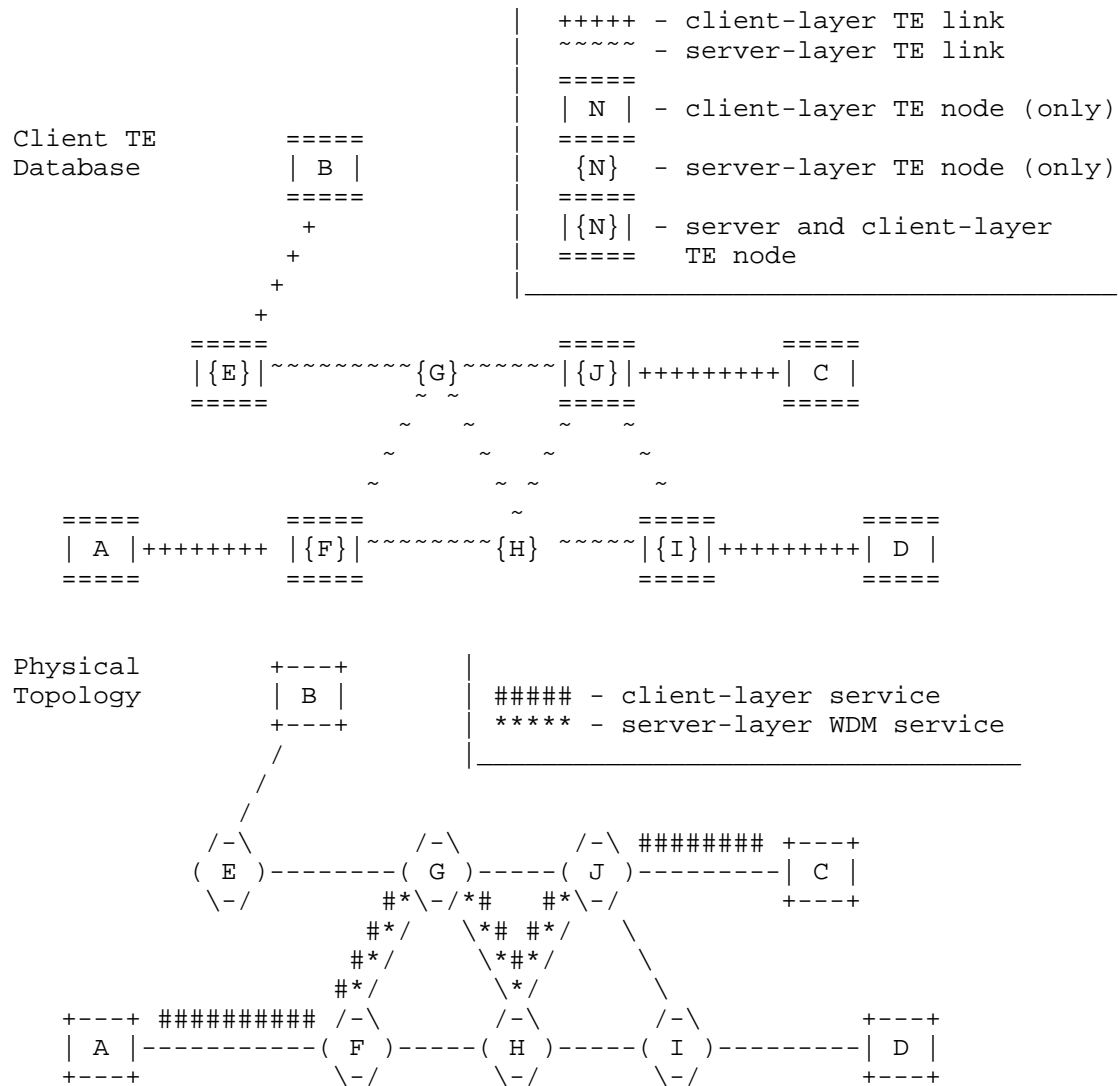


Figure 3: Traffic engineering - ERO with "loose hop"
[Path = {A,F,J,C} (with J loose)].

In this example, the TE topology associated with the client layer network is indicated by the links marked with '+' and nodes marked without brackets, whereas the TE topology associated with the server layer network is indicated by the links marked with '~' and nodes marked in '{}'. The nodes at the edge of the server layer network are visible in both the topologies. The client topology is capable of switching traffic within the client layer, whereas the server topology is capable of switching traffic within the server layer.

In this example, if the "B" router attempts to determine a path to the "D" router it will be unable to do so, as the client topology to which the B and D routers is connected does not include a full path made of just client layer links between them. The only way to setup an end-to-end path in this case is to use an ERO with a "loose hop" across the server layer domain as illustrated in Fig 3. This would cause the server layer to create the necessary link in the client layer topology on the fly. However, this approach has a few drawbacks - [a] the necessity for the operator to specify the ERO with the "loose" hop; [b] potential sub-optimal usage of server layer network resources; [c] unpredictability with regard to the fate-sharing of the new link (that is created on the fly) with other links of the client layer topology.

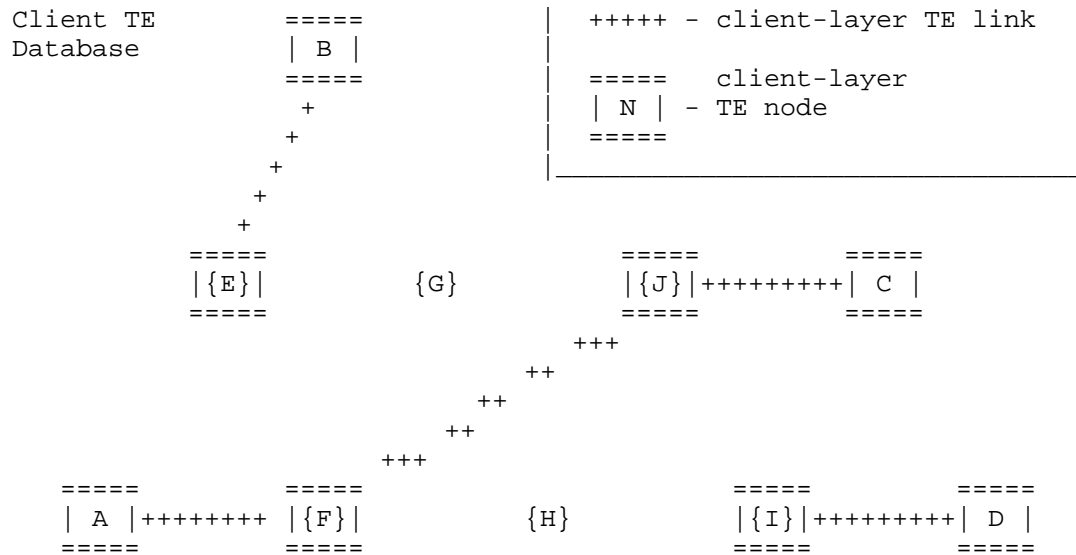
In order to be able to compute an end-to-end path between the two client layer endpoints, the client topology must be sufficiently augmented to indicate where there are paths through the server topology, which can provide connectivity between nodes in the client topology. In other words, in order for a client to compute path(s) across the server layer network to other clients, the feasible paths across the server layer network should be made available (in terms of TE links and nodes that exist in the client layer network) to all the clients. This is discussed in detail in the next section.

As it is mentioned already, in the overlay model the client and network domains, generally speaking, exist in separate layer-networks. One important use case, however, is when the client and network topologies belong to the same layer network. For example, IP routers, connected via GMPLS ENNI to a WDM network, could be capable of terminating optical trails being lambda switched by the network. The method described in the following sections allows also partitioning a single layer network into domains. Those domains do not need to leak the full routing information to their neighboring domains but rather provide sufficient information for a path

computation engine to route connections across a multi-domain network.

3.1. Augmenting the Client layer Topology

In the example hybrid network, shown below in Fig 4, consider a scenario, where each GMPLS-enabled IP router is connected to the optical WDM transport network via a transponder. Further, consider the situation, where the transponder on node F can be connected to the transponder on node J via the optical path F-G-H-J. Suppose, a lambda LSP is provisioned in the server layer along this path and advertised (as a TE link) into the client layer network. With the availability of this TE link, the path computation function at node



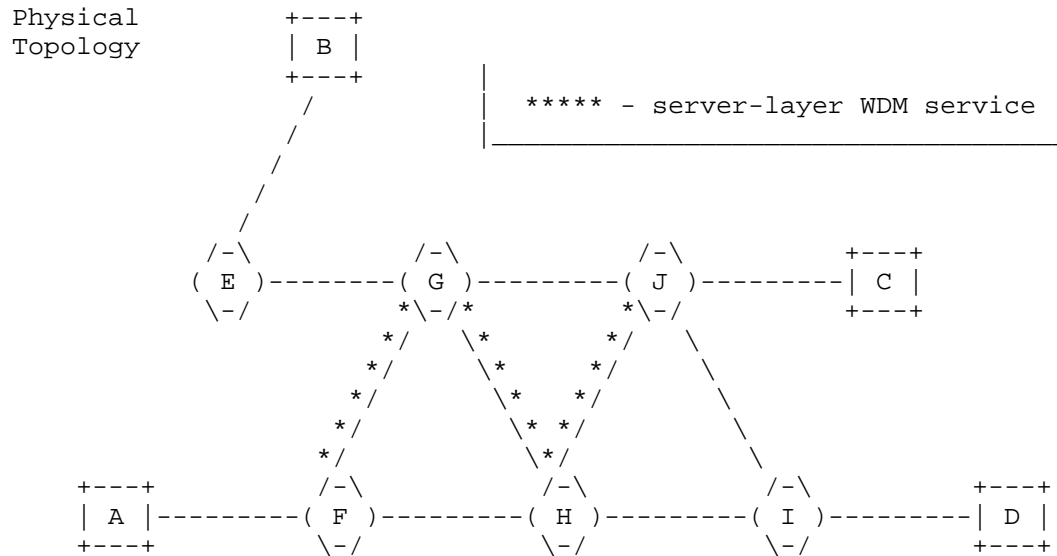


Figure 4: Traffic engineering - end-to-end path computation.[The client layer "TE link" between F and J is produced by creating the underlying server-layer connection; Node A has visibility to end-to-end (A to C) client-layer links and can do CSPF]

A is able to compute an end-to-end path from A to C. In this example, in order for the TE link to be made available in the client layer network topology, the network resources supporting the underlying server layer LSP are fully committed beforehand.

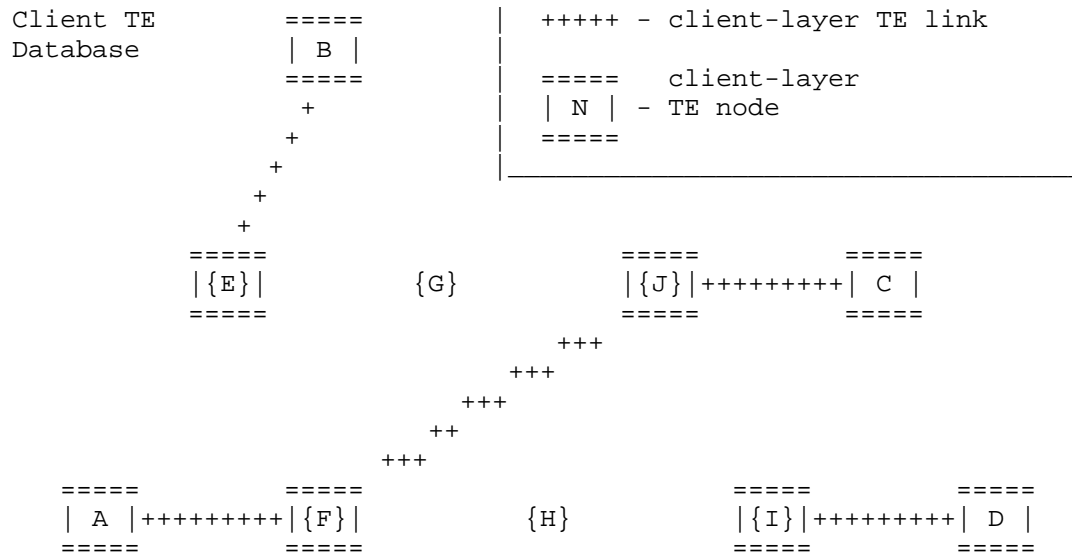
As another scenario, consider a network configuration, where the transponders on nodes E, F, J and I are connected to each other via directionless ROADM technology. In this case it is physically possible to connect any transponder to any other transponder in the server layer network. As there are transport capabilities available in the server layer network between every pair of elements with an adaptation function to the client layer network, the operator in

this case would not wish to commit any network resources in the server layer network until a client LSP is signaled. The next section proposes a method to address this common operational requirement.

3.1.1. Virtual TE Links

A "Virtual TE Link" as defined in section 7.3.3 of [RFC4847] is a TE link that is advertised into the client layer network. The advertisement includes information about available but not necessarily reserved/committed resources in the server layer network necessary to support that TE link. In other words, Virtual TE Links represent specific transport capabilities available in the server layer network, which can support the establishment of LSPs in the client layer network.

The two fundamental properties of a Virtual TE Link are: [a] it is advertised just like a real TE link and thus contributes to the buildup of the client layer network topology; and [b] it does not require allocation of resources at the server layer until used, thus allowing the mutually exclusive sharing of server layer network resources with other Virtual TE Links.



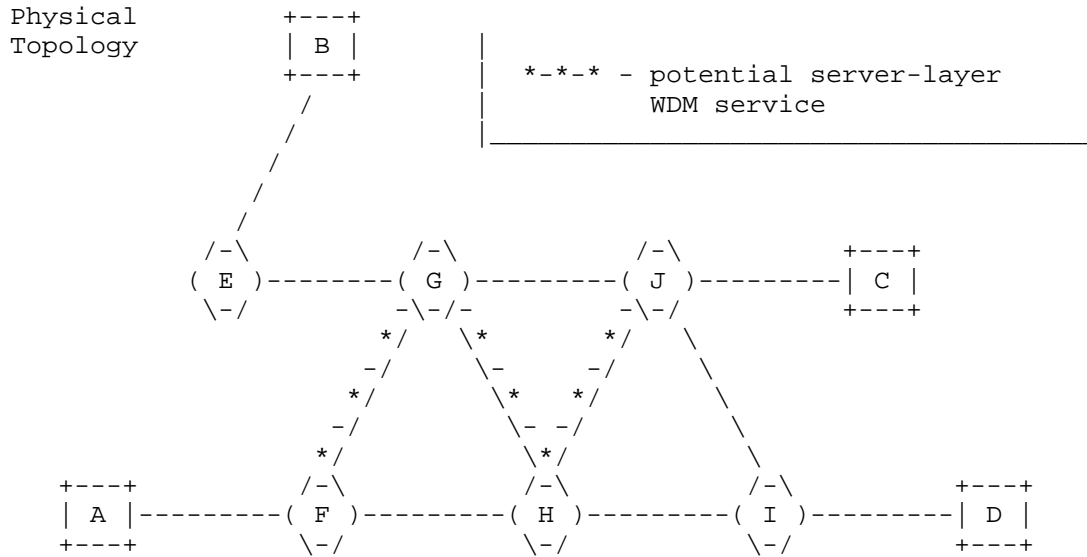


Figure 5: Traffic engineering - end-to-end path computation with "Virtual TE Links". [The "Virtual TE link" between F and J is created in the client layer without actually instantiating the underlying server-layer connection; Node A has visibility to end-to-end client-layer links and can do CSPF]

In the example shown in Fig 5, the availability of a lambda channel along the path F-G-H-J results in the advertisement by nodes F and J of a Virtual TE Link between F and J into the client layer network topology (+++ line). With the advertisement of this Virtual TE Link, the path computation function at node A is able to compute an end-to-end path from A to C.

Whenever a Virtual TE Link gets selected and signaled in the ERO of a client layer LSP, it ceases temporarily to be "virtual" and transforms into a regular TE link. When this transformation takes place, the clients will notice the change in the advertised available bandwidth of this TE link. Also, all other Virtual TE Links that share in a mutual exclusive way some of server layer resources with the TE link in question SHOULD start advertising "zero" available bandwidth. Likewise, the TE network image reverts back to the original form as soon as the last client layer LSP, going through the TE link in question, is released, i.e. Virtual TE Link becomes "virtual" again.

The overlay topology, advertised into the client domain as a set of Virtual TE Links, along with access TE links (the TE links interconnecting client network elements with the network domain) makes up the topology that in the overlay model allows for the client domain path computation function to compute end-to-end paths interconnecting client network elements across the network domain.

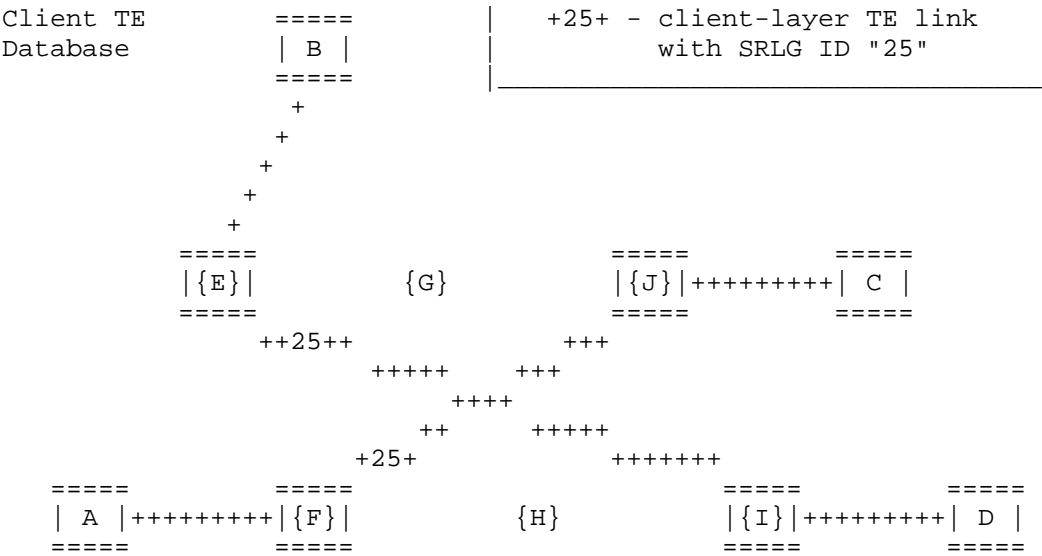
3.2. Macro SRLGs

The Virtual TE Links, which are advertised into the client layer network topology, cannot be assumed to be independent. It is quite possible for a given Virtual TE Link to share fate with one or more other Virtual TE Link(s). This is because the underlying server layer LSPs (established or potential) can traverse the same server layer network link and/or node, and failure of any such shared link/node would make all such LSPs inoperable (along with the Virtual TE Links supported by the LSPs). If diverse end-to-end paths for client layer LSPs are to be computed, the fate sharing information of the Virtual TE Links needs to be taken into account. The standard way of addressing this problem is via the concept of Shared Risk Link Group (SRLG). Specifically, a network resource shared by two or more TE links is identified via a network scope unique number (SRLG ID) and advertised within each such TE link advertisement.

A "traditional" SRLG (per [RFC4202]) represents a shared physical network resource, upon which normal function of a link depends. Such SRLGs can also be referred to as physical SRLGs. Zero, one or more physical SRLGs could be identified and advertised for every TE link in a given layer network. There is a scalability issue with physical SRLGs in multi-layer environments. For example, if a server layer LSP serves a client layer link, every server layer link and node traversed by the LSP must be considered as a separate SRLG. The number of server layer SRLGs to be advertised to client layer per

TE link is directly proportional to the number of hops traversed by the underlying server layer LSP.

This document introduces a notion of Macro SRLGs, which addresses this scaling problem. Macro SRLGs have the same protocol format as their physical counterparts and can be assigned automatically for each TE link that is advertised into the client layer network supported by an underlying server layer LSP (instantiated or otherwise). A Macro SRLG represents a shared path segment that is traversed by two or more of the underlying server layer LSPs. Each shared path segment can be viewed as a set of shared server layer resources. The actual procedure for deriving the Macro SRLGs is beyond the scope of this document.



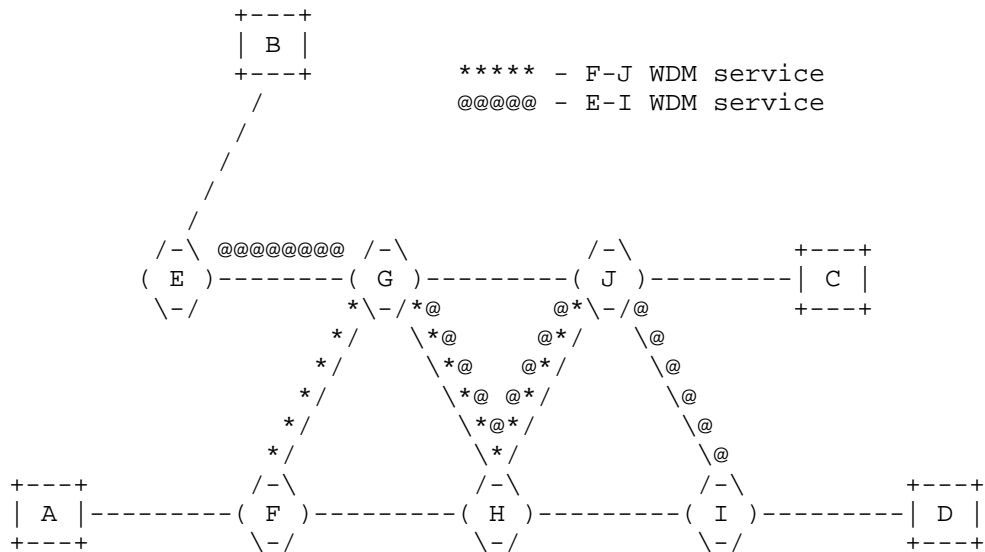


Figure 6: Macro SRLGs - ["TE links" E-I and F-J share fate since the underlying server-layer connections traverse the same path segments [G-H][H-I]. Macro SRLG-ID "25" is assigned to both "TE links"]

3.3. MELGS

If two or more Virtual TE Links share fate, it means that the links could be concurrently activated and used by client LSPs with a caveat that the links could be taken out of service by a single network failure, and, thus, cannot be used in the same protection scheme. There could be a stronger (than fate sharing) relationship between two or more Virtual TE Links. Because a set of Virtual TE Links can depend on the same uncommitted network resources, the situation can arise, when only one Virtual TE Link from the set could be activated at any given time. In other words, two or more Virtual TE Links can be mutually exclusive.

One example of the mutually exclusive relationship of Virtual TE Links is when the paths for the server layer network LSPs supporting the Virtual TE Links not only intersect, but also require usage of the same resource (e.g. lambda channel) on the intersection (see Figure 7). Another example is when the said paths depend on a common

physical resource (e.g. transponder, regenerator, wavelength converter, etc.) that could be used only by one LSP at a time.

For a client path computation function (especially a centralized one capable of concurrent computation of multiple paths) it is important to know about such mutually exclusive relationship between Virtual TE Links. This document introduces a concept of Mutually Exclusive Link Group (MELG) and suggests a new sub-TLV - MELGs sub-TLV - to be added to the top level TE Link TLV. The purpose of the MELGs sub-TLV is:

- To indicate via a separate network unique number (MELG ID) an element or a situation that makes the advertised Virtual TE Link to belong to one or more Mutually Exclusive Link Groups. Path computing element will be able to decide on whether two or more Virtual TE Links are mutually exclusive or not by finding an overlap of advertised MELGs (similar to deciding on whether two or more TE links share fate or not by finding common SRLGs)
- To indicate whether the advertised Virtual TE Link is committed or not at the moment of the advertising. Such information is important for a path computation element: committing new Virtual TE links (vs. re-using already committed ones) has a consequence of allocating more server layer resources and disabling other Virtual TE Links that have common MELGs with newly committed Virtual TE Links.

The format of the MELGs sub-TLV is defined as follows:

Name: MELGs

Type: TBD

Length: Variable

0																1																2																3															
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9																								
Sub-TLV Type																Sub-TLV Length																																															
Flags (16 bits)																U	Number of MELGs (16 bits)																																														
MELGID1 (64 bits)																																																															
MELGID2 (64 bits)																																																															
.....																																																															
MELGIDn (64 bits)																																																															

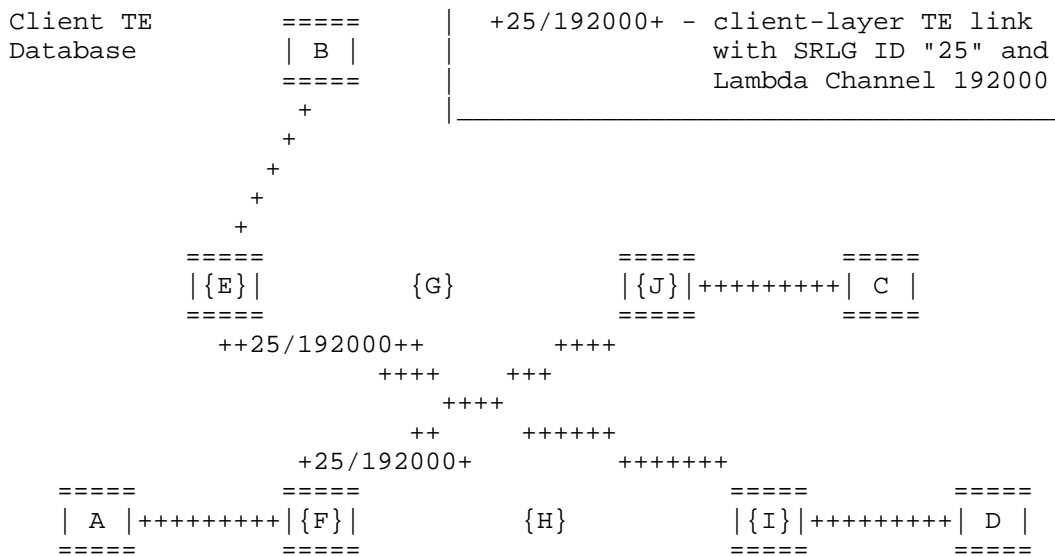
Number of MELGs: number of MELGS advertised for the

Virtual TE Link;
 Flags: Virtual TE Link specific flags;
 MELGID1,MELGID2,...,MELGIDn: 64-bit network domain unique numbers
 associated with each of the advertised
 MELGs

Currently defined Virtual TE Link specific flags are:

U bit (bit 1) : Uncommitted ,if set, the Virtual TE Link is uncommitted at the time of the advertising (i.e. the server layer network LSP is not set up); if cleared, the Virtual TE Link is committed (i.e. the server layer LSP is fully provisioned and functioning). All other bits of the "Flags" field are reserved for future use and MUST be cleared.

Note: A Virtual TE Link advertisement MAY include MELGs sub-TLV with zero MELGs for the purpose of communicating to the TE domain whether the Virtual TE Link is currently committed or not.



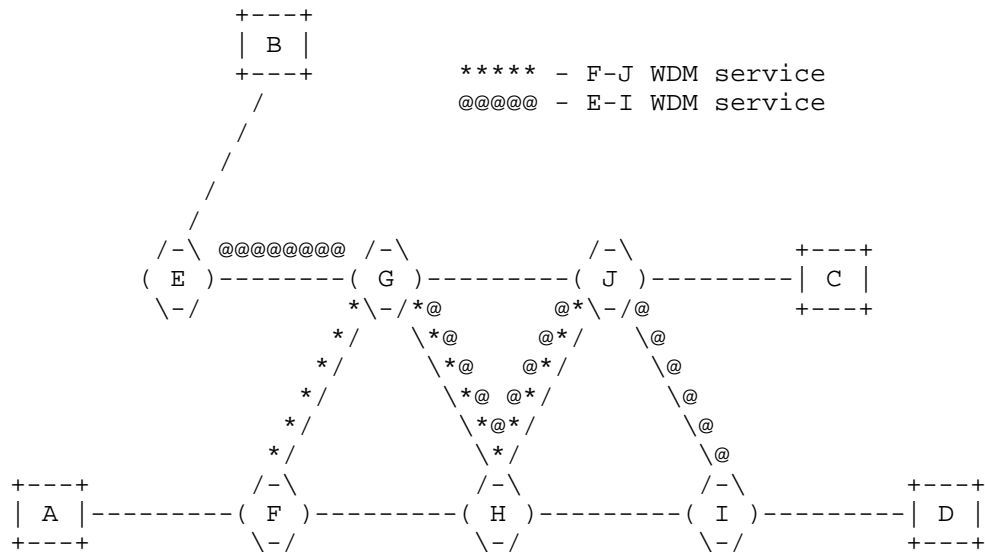


Figure 7: MELGs - ["TE links" E-I and F-J are mutually exclusive (server paths require usage of the same resource: lambda channel 192000). Same MELG ID is assigned to both TE links]

3.4. Switching Constraints

Generally speaking, it SHOULD NOT be assumed that a Virtual TE Link advertised by a given network domain border node can be cross-connected within a client LSP with every access TE link advertised by the said node. This circumstance necessitates the specification of connectivity constraints by network domain border nodes. If such information is not available for client domain path computers, there is a significant risk of provisioning failures of client LSPs, if/when they are signaled with the computed paths (see, Fig 7). This document recommends the use of the advertisements specified in [GEN_CNSTR] and [OSPF_GEN_CNSTR] to address the network element switching limitations problem.

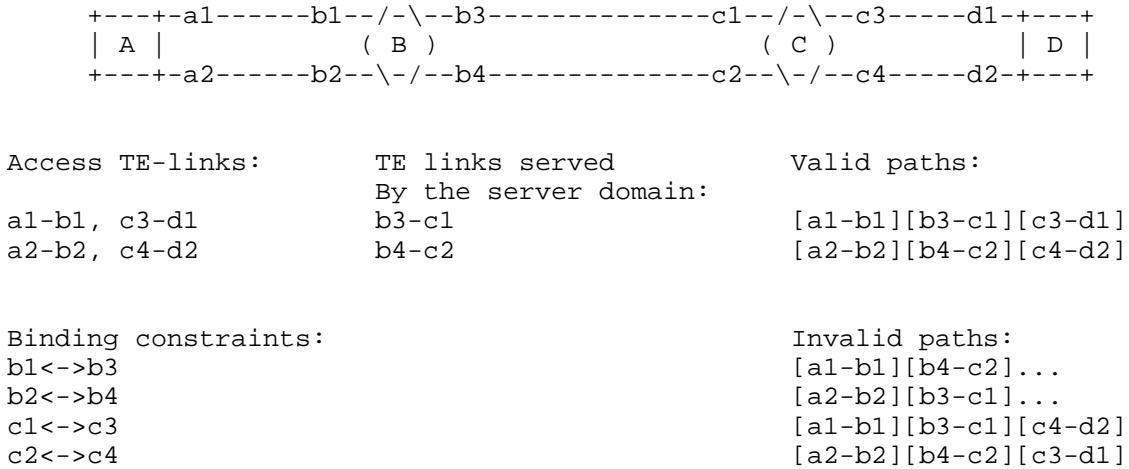


Figure 7: Switching Constraints

4. Connection Setup

Experience with control plane operations in multi-layer networks indicates some benefits in coordinating certain signaling operations of client layer network LSPs and underlying server layer network LSPs in the following manner. Consider the scenario, where the network is a WDM layer topology comprising of ROADMs. The set-up time for a service at the WDM layer can be fairly long, as it can involve time-consuming power-equalization procedures, amongst other layer specific operations. This means that at very least, the setup timers for the client LSPs would need to be somehow coordinated with that of the server LSPs. To avoid this operationally awkward issue, a phased LSP setup process as depicted in Fig 8 is proposed.

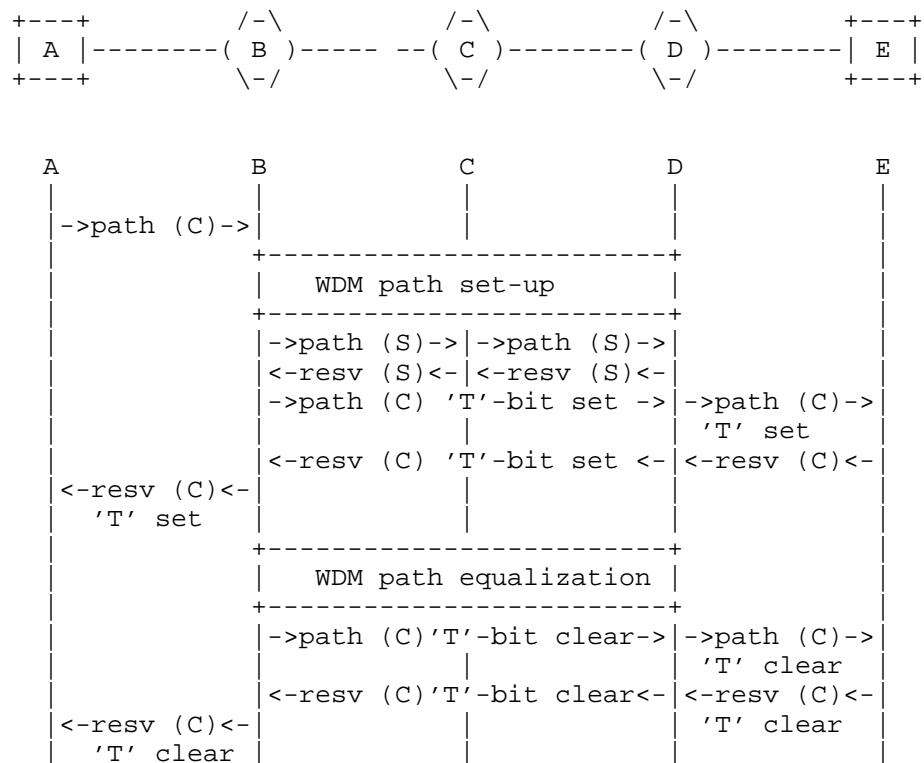


Figure 8: connection set-up

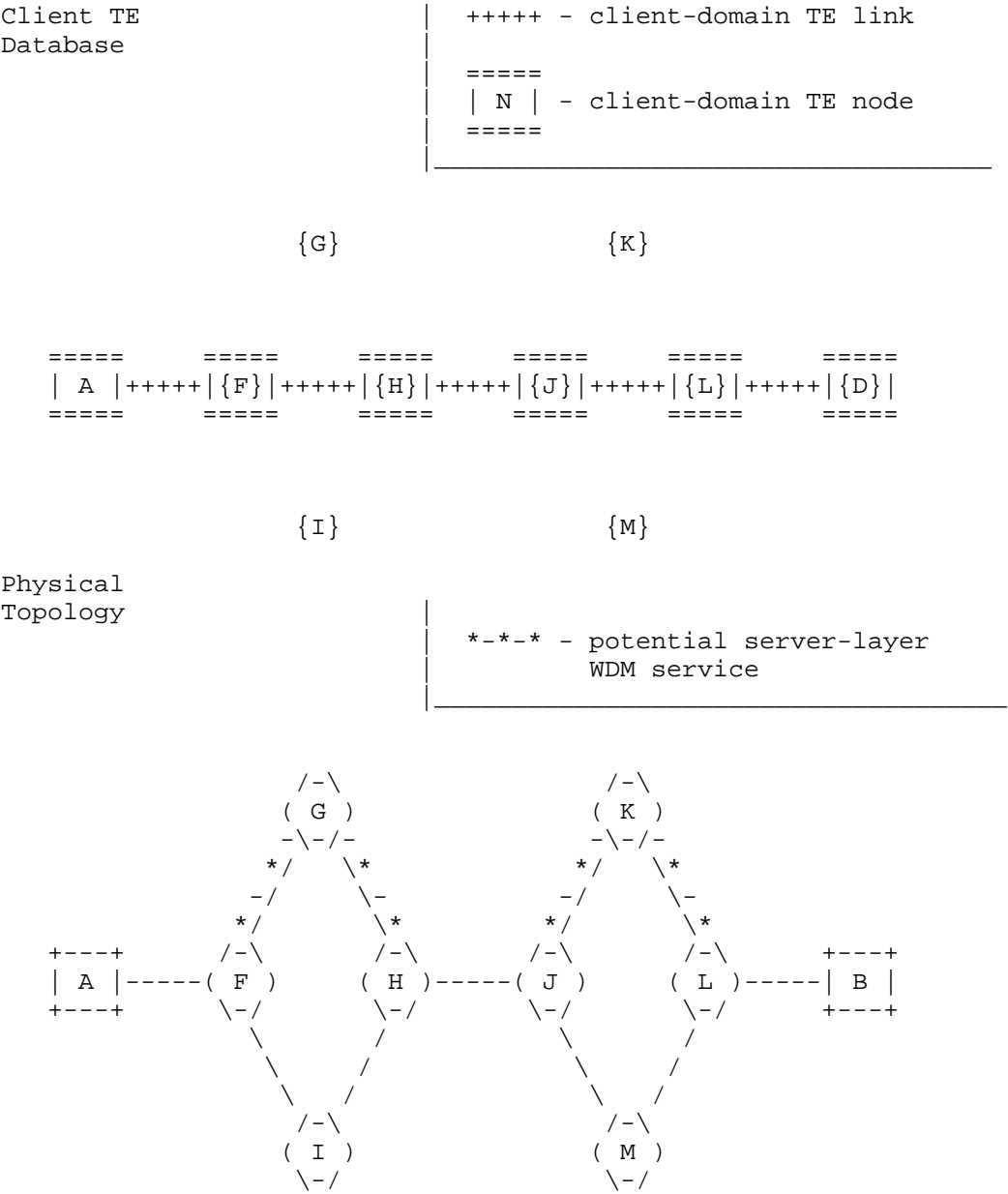
As long as the server LSP is not completely established (i.e. successfully power equalized), the server layer network border nodes, through which the client LSP passes, would signal PATH/RESV messages with the T (Testing) bit set in the ADMIN_STATUS. The T bit would be cleared in these messages only after all server LSPs supporting links taken by the client LSP in question are deemed fully operable.

5. GMPLS ENNI and Multiple Server Network Domains

In the previous sections a single server network domain GMPLS ENNI configuration was considered. The said configuration is modeled as a set of client nodes, belonging to one or more client domains, connected to a single server network domain. The connectivity is realized via access links in the data plane and GMPLS ENNI interfaces in the control plane. The server domain is independent from the client domain(s) (administratively and from the Traffic Engineering and control/management plane point of view). The network domain exposes its resources to the clients in a form of Virtual TE Links, thus, enabling the clients to influence the way their LSPs are routed across the network domain.

There are important use cases that require client LSPs to traverse more than one server network domains. In such use cases the server domains, generally speaking, are independent from each other and from each of the client domains. In such configurations the clients would still want to control the routing of their LSPs in each of the server domains, the LSPs are going through, for the same reasons it is necessary to do so in the single server domain configuration (e.g. diversity, fate sharing, MELG considerations, etc.). Fortunately, the Virtual TE Link approach allows for exposing of the resources of multiple network domains in the same way as in the single server domain case, and, thus, provides the same tools for dynamic provisioning of client LSPs across either single or multiple server network domains.

Multiple server network domains are modeled as separate independent networks interconnected with each other on their respective border nodes via inter-domain links in the data plane and GMPLS ENNI interfaces in the control plane. A network border node sees no difference between an access link and an inter-domain link terminated on the node (nor can it tell whether it is connected via a given link to a client node or a border node of a neighboring server network domain). Just like in the single-domain case, each server domain exposes its resources to other server and client network domains via Virtual TE Links configured in accordance with local domain policies. It is responsibility of server domain border nodes to advertise into the neighboring domains all access, inter-domain and Virtual TE Links it locally terminates, as well as imposed on them switching limitations. The said advertisements are flooded into the client domains and populate the client path computer's TEDs. Successful path computations produce end-to-end paths in the form of access, Virtual and inter-domain TE link chains.



6. Path computation aspects

It is assumed that a client domain path computation function makes use of advertised access TE links as well as Virtual TE Links, while computing end-to-end paths for client LSPs. The said path computation function could be local (i.e. located on client LSP ingress nodes, as stipulated by [RFC4655] Composite PCE node) or remote (i.e. on network PCEs). Path computations could be triggered by client nodes or NMS. Generally speaking, the responsibility of the client domain path computation function is to (concurrently) compute one or several paths for each source-destination pair (potential client LSP termination points) specified in a single path computation request. The path computation SHOULD be subject to one or more path optimization criterions (such as minimal cost, minimal latency, etc.) and a set of path computation constraints (such as link unreserved bandwidth, link colors, layer-specific constraints, explicit inclusions and exclusions, etc.)

As the overlay topology hides actual server domain/layer links and nodes, it is RECOMMENDED to support SRLG diverse computation of two or more paths.

Furthermore, the path computation SHOULD consider the connectivity/switching limitation constraint (when available) in addition to all other path computation constraints.

The use of the PCE architecture and PCEP protocol is governed by [RFC5440], [RFC5521] and [RFC5541].

As described in section 3.3., two or more Virtual TE Links may not only share risk, but also may exclusively depend on the same server layer resources. Therefore, paths, computed on network topologies containing Virtual TE Links, have an increased probability of LSP setup failures (two LSPs, for example, could be routed over two Virtual TE Links that exclusively depend on the same server layer resource). In such cases concurrent path computation, taking in consideration MELG information, will address this problem. PCEP supports concurrent path computation per [RFC5440]. Specifying MELG diversity constraint in path computation requests is out of scope of this document.

In addition MELG may carry information on the establishment of server-layer resources. A Path computation request MAY constraint the path computation to TE-Links that are fully provisioned only. This information MAY also be used in PCE path computation policies.

7. Access and Virtual TE link addressing

[RFC4208] implies that access TE links could be named from the same address space as network domain TE links or from a separate address space. This memo requires the following:

- It MAY be possible to assign addresses for access TE links from the same address space as the one used for naming network internal TE links (i.e. TE links interconnecting network domain devices);
- It MUST be possible to assign addresses for access TE links from a separate address space, independent from the space used for addressing network internal TE links;
- Virtual TE Links MUST share the address space with any access TE links they are allowed to be cross-connected within a client LSP.

8. Use cases

8.1. Service Optimization and Restoration in Multi-layer networks

Multi-layer networks are a reality today, and they are operated by different groups of people, following different operational procedures. This requires an independent optimization of the client and server layer networks. Such independence may cause a situation, where the re-routing of a client layer LSP fails, because some of resources on the selected alternate path share fate with some of resources on the LSP's failed path. This usually happens due to lack of knowledge of the server layer network by a client layer path computation function at the time when the alternative path is selected.

The high volume and importance of IP traffic in provider networks today requires the client and server layer networks to share sufficient information in order to enable an optimized transport for IP/MPLS services and address existing inefficiencies. From the carrier perspective it is very important that the SRLG information is provided by the server layer TE application and is used by the client layer path computation.

In a typical multi-layer network, where IP/MPLS is the client layer network and WDM/OTN is the server layer network, the client layer network is responsible for the protection of the IP/MPLS traffic from networks failures. This is normally achieved via using

protection schemes, such as FRR and/or LFA. Regardless of the used mechanism, the SRLG information, provided by the server layer network, helps to optimize the client layer network with respect to reduced link utilization and reliable and efficient protection of the user traffic.

Today the SRLGs information is used mainly when calculating diverse alternative paths for the IP/MPLS LSPs. Therefore, the following procedures are performed periodically:

- Building traffic matrix for the server layer network (based on IP links)
- Solving traffic engineering problems in the server layer network
- (Re-)Calculating SRLGs to be propagated into the client layer network
- Simulating failure scenarios
- Making sure that the affected IP/MPLS LSPs function properly after they are replaced onto SRLG diverse alternative paths

GMPLS ENNI reduces the OPEX costs of performing these procedures via the automation as follows:

- server layer network automatically discovers and advertises the SRLG information into client layer network via a common routing protocol;
- client layer network path computer uses the SRLG information when selecting diverse paths.

8.2. IP/MPLS Offloading with ENNI automation

A typical application in multi-layer (IP/MPLS over optical) networks is termed 'IP Offloading', in which the network responds to the increase in traffic of a particular service or across a segment in the IP network by dynamically creating additional IP/MPLS links served by GMPLS LSPs provisioned in the server layer network, and placing the extra IP/MPLS traffic onto said links. Likewise, when the IP/MPLS traffic decreases to a normal pattern, the said GMPLS LSPs are torn down, and the extra IP/MPLS links are removed from the client layer network TE domain. The increase in traffic is typically caused by an elevated number of high traffic flows/services traversing an IP network segment.

The decision process driving IP offloading is complex, and is governed by a set of rules. These rules reduce the cost of running the multi-layer network, while ensuring that it remains stable.

Automation of IP Offloading poses a number of challenges. It includes dynamic provisioning, release and maintenance of GMPLS LSPs in the server layer (e.g. WDM) network as well as automatic advertising/withdrawing them as (numbered or/and unnumbered) TE links into/from the client layer network. In order to pre-plan and manage properly the said dynamic IP/MPLS TE links, it is important to know in advance (and also in real time) the capabilities and resource availability of server layer network. The network domain/layer virtualization procedures described in this document helps to solve this complex operational issue.

8.3. Use of PCE and VNTM in Multi-layer Network Operation

Two key elements have been proposed to help in the management and coordination of multi-layer networks: the Path Computation Element (PCE) and the Virtual Network Topology Manager (VNTM). PCE is responsible for the calculation of paths between endpoints, particularly in complex scenarios involving, for example, WDM layer physical impairments. VNTM is in charge of maintaining the topology of the client layer network by instantiating virtual links, in the server layer network. I.e., it can be used to provide TE links to the client layer network dynamically.

Several cooperation modes between PCE, VNTM and the NMS have been proposed in [RFC5623]. For instance, the operator can request a new MPLS tunnel via the NMS, which communicates with a PCE with information of the multi-layer network. The PCE, in case there are enough resources in the IP/MPLS layer, normally returns a path for the tunnel made of real TE links. On the other hand, if there is a lack of resources in the IP/MPLS layer, the response may contain a path with one or more Virtual TE Links. In this case, the NMS can cooperate with the VNTM to suggest the set-up of a GMPLS LSP(s) in the server layer network. The VNTM, based on the local policies, can accept the suggestion and cause the set-up of the GMPLS LSPs in the server layer network.

In order for the computation to be effective, the PCE needs knowledge of the overlay topology (SRLGs, MELGs, TE metrics of the Virtual TE links), which can be provided via GMPLS ENNI.

9. Security Considerations

TBD

10. IANA Considerations

TBD.

11. References

11.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4202] K. Kompella, Y.Rekhter
"Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)",
RFC 4202, October 2005.
- [RFC4208] G. Swallow, J.Drake, H. Ishimatsu, and Y. Rekhter,
"GMPLS UNI: RSVP-TE Support for the Overlay Model",
RFC 4208, October 2005.
- [GEN_CNSTR] G.Bernstein, Y.Lee, D.Li, W.Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks"
[draft-general-constraint-encode-10.txt]
- [OSPF_GEN_CNSTR] F.Zhang, J.Han, Y.Lee, D.Li, G.Bernstein, Y.Hu
"OSPF-TE Extensions for General Network Element Constraints"
[draft-general-constraints-ospf-te-04.txt]

11.2. Informative References

- [RFC4847] T. Takeda, "Framework and Requirements for Layer 1 VPNs", RFC 4847, April 2007.
- [RFC4655] A. Farrel, J.-P. Vasseur, J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [OVERLAY_FWK] D.Ceccarelli et al "Multi-domain network integration framework in the context of overlay

model"
[draft-many-ccamp-gmpls-overlay-model-00.txt]

12. Acknowledgments

Chris Bowers [cbowers@juniper.net]
Daniele Ceccarelli [daniele.ceccarelli@ericsson.com]

Authors' Addresses

Igor Bryskin
ADVA Optical Networking

Email: ibryskin@advaoptical.com

Wes Doonan
ADVA Optical Networking

Email: wdoonan@advaoptical.com

Vishnu Pavan Beeram
Juniper Networks

Email: vbeeram@juniper.net

John Drake
Juniper Networks

Email: jdrake@juniper.net

Gert Grammel
Juniper Networks

Email: ggrammel@juniper.net

Manuel Paul
Deutsche Telekom

Email: Manuel.Paul@telekom.de

Ruediger Kunze

Deutsche Telekom

Email: Ruediger.Kunze@telekom.de

Oscar Gonzalez de Dios
Telefonica

Email: ogondio@tid.es

Cyril Margaria
Nokia Siemens Networks

Email: cyril.margaria@nsn.com

Friedrich Armbruster
Nokia Siemens Networks

Email: friedrich.armbruster@nsn.com

Network Working Group
Internet Draft
Category: Standards Track

Fatai Zhang
Xian Zhang
Huawei
D. Ceccarelli
D. Caviglia
Ericsson
Guoying Zhang
CATR
D. Beller
S. Belotti
Alcatel-Lucent
October 12, 2012

Expires: April 11, 2013

Link Management Protocol (LMP) extensions for G.709
Optical Transport Networks

draft-cecczhang-ccamp-gmpls-g709v3-lmp-00.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on April 11, 2013.

Abstract

Recent progress of the Optical Transport Network (OTN) has introduced new signal types (i.e., ODU0, ODU4, ODU2e and ODUFlex) and new Tributary Slot granularity (1.25Gbps).

Since equipments deployed prior to recently defined ITU-T recommendations only support 2.5 Gbps Tributary Slot granularity and ODU1, ODU2 and ODU3 containers, the compatibility problem should be considered. In addition, a Higher Order ODU (HO ODU) link may not support all the types of Lower Order ODU (LO ODU) signals defined by the new OTN standard because of the limitation of the devices at the two ends of a link. In these cases, the control plane is required to run the capability discovering functions for the evolutive OTN.

This document describes the extensions to the Link Management Protocol (LMP) needed to discover the capability of HO ODU link, including the granularity of Tributary Slot to be used and the LO ODU signal types that the link can support. Moreover, extensions of LMP test messages detailing the OTN technology specific information in order to cover also G.709v3 signal types and containers are also provided.

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 [RFC2119].

Table of Contents

1. Introduction	3
2. Terminology	4
3. Overview of the Evolutive G.709	4
4. Link Capability Discovery	5
4.1. Link Capabaplity Discovery Requirements	5
4.1.1. Discovering the Granularity of the TS	5
4.1.2. Discovering the Supported LO ODU Signal Types.....	6
4.2. Extensions: LMP Link Summary Message	7
4.2.1. Message Extension	7
4.2.1.1. LinkSummary Message	7
4.2.1.2. LinkSummaryAck Message	8
4.2.1.3. LinkSummaryNack Message	8
4.2.2. Object Definitions	8
4.2.3. Procedures	10
5. Verifying Link Connectivity	12
5.1. Encoding Type	13
5.2. Verify Transport Mechanism	13
5.3. Transmission Rate	14
6. Trace Monitoring	15
6.1. TRACE Object for evolutive OTN	15
6.2. Discovery Response Message for Layer Adjacency Discovery	17

7. LMP Behavior Negotiation Update	18
8. Security Considerations	18
9. IANA Considerations	19
10. Acknowledgments	19
11. References	19
11.1. Normative References	19
11.2. Informative References	20
12. Authors' Addresses	20
13. Contributors	22

1. Introduction

The Link Management Protocol (LMP) defined in [RFC4204] is being developed as part of the Generalized MPLS (GMPLS) protocol suite to manage Traffic Engineering (TE) links.

Recently, great progress has been made for the Optical Transport Networking (OTN) technologies in ITU-T. New ODU containers (i.e., ODU0, ODU4, ODU2e and ODUFlex) and a new Tributary Slot (TS) granularity (1.25Gbps) have been introduced by the [G709-V3], enhancing the flexibility of OTNs.

With the evolution and deployment of G.709 technology, the backward compatibility problem requires to be considered. In data plane, the equipment supporting 1.25Gbps TS can combine the specific Tributary Slots together (e.g., combination of TS#i and TS#i+4 on a HO ODU2 link) so that it can interwork with other equipments which support 2.5Gbps TS. From the control plane point of view, it is necessary to discover which type of TS is supported at both ends of a link, so that it can choose and reserve the TS resources correctly in this link for the connection.

Additionally, the requirement of discovering the signal types of Lower Order ODU (LO ODU) that can be supported by a Higher Order ODU (HO ODU) should be taken into account. Equipment at one end of a HO ODU link may not support to transport some types of LO ODU signals (e.g., may not support the ODUFlex). In this case, this HO ODU link should not be selected for those types of LO ODU connections.

From the perspective of control plane, it is necessary to discover the capability of a HO ODUK or OTUK link including the granularity of TS to be used and the LO ODU signal types that the link can support. Note that this capability information can be, in principle, discovered by routing. Since in certain case, routing is not present (e.g., UNI case) we need to extend link management protocol capabilities to cover this aspect. Obviously, in case of routing presence, the discovering procedure by LMP could also be optional.

A further enhancement needed with respect to LMP covers the link verification and link property correlation functionalities and the G.709 test procedures they are based on. Such procedures require the definition of a G.709 specific TRACE object. After data links have been verified, it is possible to group them into the TE links.

This document extends the LMP and describes the solution of discovering HO ODU link capability and operating link verification and link property correlation.

2. Terminology

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 [RFC2119].

3. Overview of the Evolutive G.709

The traditional OTN standard [ITU-T-G709] describes the optical transport hierarchy (OTH) and introduces three ODU signal types (i.e., ODU1, ODU2 and ODU3). The ODU_j can be mapped into one or more Tributary Slots (with a granularity of 2.5Gbps) of OPU_k where $j < k$. The ODU_j can also be mapped into OTU_j ($j=1, 2$ or 3) directly.

Recent revisions of ITU-T Recommendation G.709 have introduced new features for the evolutive Optical Transport Networks (OTN). New ODU signals, including ODU0, ODU4, ODU2e and ODUFlex, are described in [G709-V3]. This document also defines the new multiplexing hierarchy for the evolutive OTN. In this multiplexing hierarchy, LO ODU_j can be mapped into an OTU_j, or multiplexed into a HO ODU_k (where $j < k$) by occupying several tributary slots.

In case of LO ODU_j mapping into OTU_j, the following mappings are defined:

- ODU1 into OTU1 mapping
- ODU2 into OTU2 mapping
- ODU3 into OTU3 mapping
- ODU4 into OTU4 mapping

In case of LO ODU_j multiplexing into HO ODU_k, a new Tributary Slot granularity (i.e., 1.25Gbps) is introduced in [G709-V3]. For the

evolutive OTN, the multiplexing of ODU_j ($j = 0, 1, 2, 2e, 3, flex$) into an ODU_k ($k > j$) signal can be depicted as follows:

- ODU0 into ODU1 multiplexing (with 1,25Gbps TS granularity)
- ODU0, ODU1, ODUflex into ODU2 multiplexing (with 1.25Gbps TS granularity)
- ODU1 into ODU2 multiplexing (with 2.5Gbps TS granularity)
- ODU0, ODU1, ODU2, ODU2e and ODUflex into ODU3 multiplexing (with 1.25Gbps TS granularity)
- ODU1, ODU2 into ODU3 multiplexing (with 2.5Gbps TS granularity)
- ODU0, ODU1, ODU2, ODU2e, ODU3 and ODUflex into ODU4 multiplexing (with 1.25Gbps TS granularity)

Note that If TS auto-negotiation is supported, a node supporting 1.25Gbps TS can interwork with the other nodes that supporting 2.5Gbps TS by combining specific TSs together in data plane, as descirbied in [OTN-frwk].

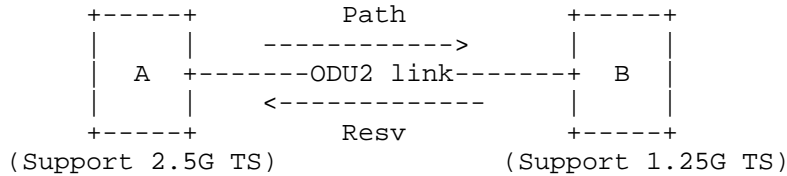
4. Link Capability Discovery

4.1. Link Capabaplity Discovery Requirements

4.1.1. Discovering the Granularity of the TS

As described in section 3.1, if the two ends of a link use different granularities of TS, The LO ODU must be mapped into specific combined Tributary Slots in the end of link with TS of 1.25Gbps.

From the perspective of control plane, when creating a LO ODU connection, the node MUST select and reserve specific TS for the connection if the two ends of a link use different granularities of TS. For example, for an ODU2 link, we suppose that node A only supports the 2.5Gbps TS while node B supports the 1.25Gbps TS. When node B receives a Path message from node A requesting an ODU1 connection, node B MUST reserve the TS#*i* and TS#*i*+4 (where $i \leq 4$) (with the granularity of 1.25Gbps) and tell node A via the label carried in the Resv message that the TS#*i* (with the granularity of 2.5Gbps) among the 4 slots has been reserved for the ODU1 connection. Otherwise, the reservation procedure will fail.

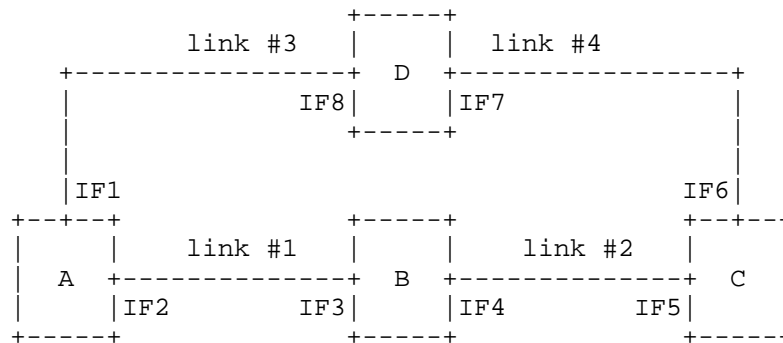


Therefore, for an ODU2 or ODU3 link, in order to reserve TS resources correctly for a LO ODU connection, the control plane of the two ends MUST know which granularity the other end can support before creating the LO ODU connection.

4.1.2. Discovering the Supported LO ODU Signal Types

Many new ODU signal types are introduced by [G709-V3], such as ODU0, ODU4, ODU2e and ODUFlex. It is possible that equipment does not always support all the LO ODU signal types introduced by [G709-V3]. If one end of a HO ODU link can not support a certain LO ODU signal type and there is no HO ODU FA LSP able to support this LO ODU signal, the HO ODU link/FA LSP can not be selected to carry such type of LO ODU connection.

For example, in the following figure, if the interfaces IF1, IF2, IF8, IF7, IF5 and IF6 can support ODUFlex signals, while the interfaces IF3 and IF4 can not support ODUFlex signals. In this case, if one ODUFlex connection from A to C is requested, and there is no HO ODU FA LSP from node A to C through node B, link #1 and #2 should be excluded, link #3 and link #4 are the candidates (the possible path could be A-D-C through link #3 and link #4).



Therefore, it is necessary for the two ends of a HO ODU link to discover which types of LO ODU can be supported by the HO ODU link.

After discovering, the capability information can be flooded by IGP, so that the correct path for an ODU connection can be calculated.

4.2. Extensions: LMP Link Summary Message

[RFC4204] defines the Link Management Protocol (LMP) which consists of four main procedures: control channel management, link property correlation, link connectivity verification, and fault management. As part of LMP, the link property correlation is used to verify the consistency of the TE and data link information on both sides of a link. This document extends the link property correlation procedure to discover the capability of both sides of a HO ODU link.

The designated HO ODU overhead bytes (e.g., the GCC1 and GCC2 overhead bytes) can be used as the control channel to carry the LMP message after the HO ODU link is created. The out-of-band Data Communication Network (DCN) can also be used.

4.2.1. Message Extension

Three messages are used for link property correlation: LinkSummary, LinkSummaryAck and LinkSummaryNack Message. This document does not change the basic procedure of LMP but just add a new subobject (HO ODU Link Capability) in the DATA_LINK object to carry the capability of one end of a HO ODU link.

The formats of LinkSummary, LinkSummaryAck and LinkSummaryNack messages are defined in [RFC4204].

4.2.1.1. LinkSummary Message

The local end of a TE link can send a LinkSummary message to the remote end to start the negotiation about the capability that the TE link can support.

One new Subobject named HO ODU Link Capability Subobject in the DATA_LINK object is introduced by this document. This new subobject is used to tell the remote end of the HO ODU link which TS granularity and which LO ODU signal types that the local end can support. When the DATA_LINK object carries the new HO ODU Link Capability Subobject, the N flag SHOULD be set to 1 which means that the subobject is negotiable.

OD(T)Uk field	Signal type of HO ODUk or OTUk
-----	-----
0	Reserved (for future use)
1	HO ODU1 or OTU1
2	HO ODU2 or OTU2
3	HO ODU3 or OTU3
4	HO ODU4 or OTU4
5-15	Reserved (for future use)

T (2 bits):

The T bits are used to indicate the granularity of the TS of the HO ODU link.

T field	TS type
-----	-----
00	Meaningless
01	1.25Gbps TS granularity
10	2.5Gbps TS granularity
11	Reserved (for future use)

In case that an OTUk link only support ODUj (j=k) into OTUk mapping and does not support any ODUj into ODUk (j<k) multiplexing, then the T field is not meaningful and MUST be filled with 0 and be ignored on receipt.

LO ODU flags (A|B|C|D|E|F|G) (16 bits):

These flags are used to indicate which LO ODU signal types that one end or the both end can support. The flags will be set to 1 if the corresponding LO ODU signal types are supported to be mapped or multiplexed into the OTUk or HO ODUk link.

This rule imposes that:

- At least one flag is set to 1.
- When the ODUj (j=k) flag corresponding to the signal type HO ODUk/OTUk is set to 1, then the signal type OD(T)Uk has to be intended as LO ODUk and direct mapping over OTUk is supported.
- * Furthermore, if only the ODUj(j=k) flag is set to 1, it means that the HO ODUk/OTUk link only supports ODUj(j=k) into OTUk mapping. In other words, the link does not support any ODUj

into ODU_k (j<k) multiplexing (i.e., payload type != 20/21), but may support carrying various non-ODU client signals listed in Table 15-8 of [G709-V3].

- When an ODU_j (j<k) flag not corresponding to the signal type HO ODU_k/OTU_k is set to 1 then the signal type OD(T)U_k has to be intended as HO ODU_k and multiplexing of LO ODU_j over HO ODU_k is supported.

Flag A: indicates whether LO ODU₀ is supported.

Flag B: indicates whether LO ODU₁ is supported.

Flag C: indicates whether LO ODU₂ is supported.

Flag D: indicates whether LO ODU₃ is supported.

Flag E: indicates whether LO ODU₄ is supported.

Flag F: indicates whether LO ODU_{2e} is supported.

Flag G: indicates whether LO ODUflex is supported.

For example, if one end of an OTU₂ link supports LO ODU₀, LO ODU₁, LO ODUflex into HO ODU₂ multiplexing and supports LO ODU₂ into OTU₂ mapping, the flags A, B, C, and G will be set to 1.

As a further example, if one end of an OTU₂ link supports only LO ODU₂ into OTU₂ mapping but no multiplexing, only flag C will be set to 1.

The remaining flags are reserved for future use and MUST be set to 0.

4.2.3. Procedures

The Link Summary messages used for capability discovery for HO ODU_k or OTU_k link are sent between adjacent nodes after the HO ODU link is created or driven by some events (e.g., an operator command). The procedure is described below:

- o The local end of the HO ODU link sends a LinkSummary message including one or more DATA_LINK objects, each of which contains the Local_Interface_Id, the Remote_Interface_Id, and the HO ODU link capability subobject. This subobject carries the capability that the local end can support, i.e., the granularity of TS and the set of LO ODU signal types that the local end can support. The LinkSummary message is sent to the remote end.

- o On receipt of the LinkSummary message, the remote end of the HO ODU link firstly determines whether the local/remote Interface_Id mappings match those that are stored locally as described in [RFC4204], and then obtains the HO ODU link capability subobject and determines the capability of the HO ODU link that both ends can support. The detail procedures are as follow:
 - Only if both ends support the 1.25Gbps TS, the remote end would choose the 1.25Gbps as the negotiated granularity for the HO ODU link. In other cases, the 2.5Gbps TS MUST be used (e.g., if the local end can support 1.25Gbps, and the remote end can support 2.5Gbps, and then the local end should imitate 2.5Gbps).
 - The remote end compares the two sets of LO ODU signal types that the local end and the remote end can support, and calculates the intersection of them, i.e., extracts all the LO ODU signal types that both two ends can support. This intersection is the set of LO ODU signal types that the HO ODU link can support.
- o If both the two ends support the same capability, i.e., they support the same granularity of TS and the same LO ODU signal types, the remote end replies a LinkSummaryAck message to the local end. So the both ends know what capability the HO ODU link can support.
- o If the two ends support different capabilities, i.e., they support different granularities of TS or different LO ODU signal types, the remote end replies a LinkSummaryNack message to the local end. The LinkSummaryNack message carries an ERROR_CODE object and one or more DATA_LINK objects. The ERROR_CODE "Renegotiate LINK_SUMMARY parameters" (see [RFC4204]) indicates that the two ends of the HO ODU link support different capabilities, and the DATA_LINK object carries the HO ODU link capability subobject which contains the negotiated granularity of TS and the set of LO ODU signal types that both ends can support. The local end can learn the HO ODU link capability after receiving the LinkSummaryNack message.
- o If the remote end does not support the HO ODU link capability negotiation procedure, the LinkSummaryNack message MUST be responded with an ERROR_CODE "Not support of HO ODU Link Capability subobject" (TBA) indicating the reason of rejection.

5. Verifying Link Connectivity

[RFC4204] defines a link verification procedure based on the in-band transmission of Test messages over the data links. It is used to verify the physical connectivity of such links, to discover data plane resources and to exchange the Interface_Ids. It is also possible to use a single procedure to verify multiple data links and correlate the information collected by means of the Verify_Id assigned to the procedure.

The link verification procedure works as follows:

- BeginVerify message: the local node sends a BeginVerify message over a control channel. It includes a BEGIN_VERIFY object which contains all the parameters characterizing the data link like, for example, the number of data links that must be verified, the transmission interval of the Test messages or the wavelength over which the Test messages will be sent.
- BeginVerifyAck: if the remote node, upon receiving a BeginVerify message, is ready to begin the procedure, it replies with a BeginVerifyAck message. Such message specifies the desired transport mechanism for the Test messages and the Verify_Id of the procedure assigned by the remote node.
- Data link Testing: the local node, upon receiving the BeginVerifyAck message, can begin testing the data links repeatedly sending Test messages over them. The remote node will reply either with a TestStatsSuccess or a TestStatusFailure for each data link. As a consequence the local node will send a TestStatusAck.
- End of testing: The local node can terminate the Test procedure at anytime just sending an EndVerifyMessage towards the remote node.

Evolutionary OTNs need the support from LMP for the testing of all the possible data links defined by ITU-T. This document provides, at present, support to the data links defined by G.709 and G.709 amendment 3 recommendations and to G.Sup43 temporary document.

The BEGIN_VERIFY class is defined in Section 13.8 of [RFC4204]. The following fields are extended: Encoding Type, Verify Transport Mechanism and Transmission Rate.

5.1. Encoding Type

The Encoding Type identifies the type of encoding supported by the interface. LMP encoding type is consistent with the LSP encoding types defined for RSVP-TE [RFC3471]. In particular, the value to be used for G.709 hierarchy ODU and OTU signals is "Digital Wrapper".

5.2. Verify Transport Mechanism

This field defines the transport mechanism for the Test messages and its scope depends on each encoding type. It is a 16 bit mask set by the local node where each bit identifies the various mechanisms it can support for LMP test messages transmission. This document defines the field values with respect to the G.709 digital encoding (they are expressed in network byte order).

- 0x01 OTUk TTI: 64 byte Test Message

Capability of transmitting Test messages using OTUk Trail Trace Identifier (TTI) overhead with frame length of 64 bytes. See ITU G.709 Section 15.2 and Section 15.7 for the structure and definition. The Test message is sent according to [RFC4204].

- 0x02 ODUk TTI: 64 byte Test Message

Capability of transmitting Test messages using ODUk Trail Trace Identifier (TTI) overhead with frame length of 64 bytes. See ITU G.709 Section 15.2 and Section 15.8 for the structure and definition. The Test message is sent according to [RFC4204].

- 0x04 GCC0: Test Message over the GCC0

Capability of transmitting Test messages using the OTUk Overhead General Communications Channel (GCC0). See ITU G.709 Section 15.7 for the structure and definition. The Test message is sent according to [RFC4204] using bit-oriented HDLC framing format [RFC1662].

- 0x08 GCC1/2: Test Message over the GCC1/2

Capability of transmitting Test messages using the ODUk Overhead General Communications Channels (GCC1/2). See ITU G.709 Section 15.8 for the structure and definition. The Test message is sent according to [RFC4204] using bit-oriented HDLC framing format [RFC1662].

- 0x10 OTUk TTI - Section Trace Correlation

Capability of transmitting OTUk Trail Trace Identifier (TTI) as defined in ITU-T G.709. The Test message is not transmitted using the OTUk TTI overhead bytes (i.e. data link), but is sent over the control channel and correlated for consistency to the received pattern. The correlation between the Interface_Id and the in-band pattern is achieved using the TRACE Object as defined in Section 4 of [RFC4207]. No modification to TestStatusSuccess or TestStatusFailure messages is required.

- 0x20 ODUk TTI - Path Trace Correlation

Capability of transmitting ODUk Trail Trace Identifier (TTI) as defined in ITU-T G.709. The Test message is not transmitted using the OTUk TTI overhead bytes (i.e. data link), but is sent over the control channel and correlated for consistency to the received pattern. The correlation between the Interface_Id the Testmessage is sent from and the pattern sent in-band is achieved using the TRACE Object as defined in Section 4 of [RFC4207]. No modification to TestStatusSuccess or TestStatusFailure messages is required.

5.3. Transmission Rate

The transmission rate of the data links where the link verification procedure can be performed is defined into the TransmissionRate field of the BEGIN_VERIFY class ([RFC4204] Section 13.8). Values are expressed in IEEE floating point format using a 32-bit number field and expressed in bytes per second. The following table defines the values to be used in OTNs:

+-----+-----+-----+			
Signal Type	Bit-rate (kbps)	Value (Bytes/Sec)	
+-----+-----+-----+			
ODU0	1 244 160	0x4D1450C0	
+-----+-----+-----+			
ODU1	2 498 775	0x4D94F048	
OTU1	2 666 057	0x4D9EE8CD	
+-----+-----+-----+			
ODU2	10 037 274	0x4E959129	

	OTU2		10 709 226		0x4E9F9475	
+-----+-----+-----+						
	ODU2e		10 399 525		0x4E9AF70A	
+-----+-----+-----+						
	ODU3		40 319 219		0X4F963367	
	OTU3		43 018 416		0X4FA0418F	
+-----+-----+-----+						
	ODU4		104 794 445		0x504331E3	
	OTU4		111 809 973		0x50504326	
+-----+-----+-----+						

Transmission Rate values (Bytes/Sec)

6. Trace Monitoring

[RFC4207] describes the set of trace monitoring procedures that allow a node to do trace monitoring by using the G.709 hierarchy capabilities.

This document defines a new C-Type of the TRACE Object class used for Trace Monitoring features as defined in [RFC4207].

6.1. TRACE Object for evolutive OTN

The TRACE Object Class assigned by IANA is 21. A new C-Type is TBA and value 2 is suggested. The TRACE Object format is the same as defined in [RFC4207] and is shown in the following:

[illegible]

```

+-----+-----+
|          Trace Type          |          Trace Length          |
+-----+-----+
|                               |
//                               //
|                               |
+-----+-----+

```

TRACE Object Class

Trace Type: 16 bits

The Trace Type field is used to identify the type of the trace message. The following values are defined and all other values are reserved and should be sent as zero and ignored on receipt.

- 1 = OTUk TTI
- 2 = ODUk TTI
- 3 = Level 1 ODUkT TTI
- 4 = Level 2 ODUkT TTI
- 5 = Level 3 ODUkT TTI
- 6 = Level 4 ODUkT TTI
- 7 = Level 5 ODUkT TTI
- 8 = Level 6 ODUkT TTI (default for layer adjacency discovery)

It shall be noted that an Amendment to ITU-T G.7714.1 has been approved in September 2010 that defines an extension for OTN layer adjacency discovery based on the ODUk TCM function (ODUkT) providing 6 TCM levels. By default the TCM level 6 SHALL be used.

Trace Length: 16 bits

Expresses the length of the trace message in bytes (as specified by the Trace Type).

Trace Message:

This field includes the value of the expected message to be received in-band. The valid length and value combinations are determined by the ITU-T G.709 recommendation. The message MUST be padded with zeros to a 32-bit boundary, if necessary. Trace Length does not include padding zeroes.

This object is non negotiable.

6.2. Discovery Response Message for Layer Adjacency Discovery

ITU-T Recommendation G.7714.1 [ITUT-G.7714.1] describes an automatic layer adjacency discovery procedure that can be applied to the ITU-T G.709 OTN technology. The discovery message can be sent to the adjacent node via the Trail Trace Identifier (TTI) and Appendix III of G.7714.1 describes how the discovery response message can be sent back to the originator of the discovery message (discovery agent in G.7714.1 terminology) using the LMP protocol.

As defined in [ITUT-G.7714.1], the TraceMonitor message [RFC4207] is used to convey the discovery response message. The following mapping table shows how the discovery response message attributes are mapped to TraceMonitor message objects or other fields of the LMP message (see G.7714.1, section 11 for the description of the attributes):

G.7714.1 discovery response		TraceMonitor/LMP message field
message attribute		
-----+-----		
<Received DA DCN ID>		<TRACE>: received discovery message
<Received TCP-ID>		<TRACE>: received discovery message
<Sent DA DCN ID>		IP source address in the IP header
<Sent Tx TCP-ID>		identical to <Sent Rx TCP-ID>

```

|
<Sent Rx TCP-ID> | <LOCAL_INTERFACE_ID>
|

```

The received TTI, more specifically the discovery message in the SAPI field contains the <Received DA DCN ID> and the <Received TCP-ID>. These attributes are included in the discovery response message by copying the received TTI into the <TRACE> field of the TraceMonitor message.

The IP address of the node sending the discovery response message corresponds to the <Sent_DA_DCN_ID> and is the IP source address in the IP header of the LMP TraceMonitor message.

Typically, the Trail Connection Point (TCP-)IDs in transmit and receive direction are identical for OTN equipment, i.e., the <Sent Rx TCP-ID> is identical to the <Sent Tx TCP-ID>. The <Sent Rx TCP-ID> identifies the TCP on which the Discovery Message was received and corresponds to the <LOCAL_INTERFACE_ID> object in the TraceMonitor message.

7. LMP Behavior Negotiation Update

This document also introduces an update to the BehaviorConfig C-Type defined in [LMP-NEG]. A new flag in the BehaviorConfig is needed for the indication of the support for OTN Test Messages:

```

0                               1                               2                               3

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|S|D|C|O|                               Must Be Zero (MBZ)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
- O: 1 bit

```

This bit indicates support for the TEST behavior of OTN technology-specific defined in this document.

8. Security Considerations

TBD.

9. IANA Considerations

TBD.

10. Acknowledgments

TBD.

11. References

11.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4204] J. Lang, Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [OTN-frwk] Zhang, F. et al, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", draft-ietf-ccamp-gmpls-g709-framework-08.txt, June 19, 2012.
- [ITUT-G709] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation, March 2003.
- [G709-V3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709 Recommendation, December 2009.
- [ITUT-G.7714.1] ITU-T, "Protocol for automatic discovery in SDH and OTN networks, G.7714.1 Recommendation", April 2003.
- [RFC1662] Simpson, W., "PPP in HDLC-like Framing", STD 51, RFC 1662, July 1994.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC4207] Lang, J. and D. Papadimitriou, "Synchronous Optical Network (SONET)/Synchronous Digital Hierarchy (SDH) Encoding for Link Management Protocol (LMP) Test Messages", RFC 4207, October 2005.

11.2. Informative References

- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, Jan 2006.
- [LMP-NEG] D.Li, D.Ceccarelli, L.Berger, "Link Management Protocol Behavior Negotiation and Configuration Modifications," July 2012, draft-ietf-ccamp-lmp-behavior-negotiation-06.

12. Authors' Addresses

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Xian Zhang
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base,
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28972913
Email: zhang.xian@huawei.com

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Francesco Fondelli
Ericsson
Via Moruzzi 1
Pisa
Italy

Email: francesco.fondelli@ericsson.com

Guoying Zhang
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.China

Phone: +86-10-68094272
Email: zhangguoying@mail.ritt.com.cn

Pietro Grandi
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6864930

Email: pietro_vittorio.grandi@alcatel-lucent.it

Sergio Belotti
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6863033

Email: sergio.belotti@alcatel-lucent.it

Dan Li

Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28973237

Email: danli@huawei.com

Dieter Beller
Alcatel-Lucent
Lorenzstrasse 10
Stuttgart 70435
Germany

Email: dieter.beller@alcatel-lucent.com

13. Contributors

Yi Lin
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base,
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28972914
Email: yi.lin@huawei.com

Intellectual Property

The IETF Trust 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 any IETF 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.

Copies of Intellectual Property 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 any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of RFC 5378. No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under RFC 5378, shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Full Copyright Statement

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

D. Hiremagalur, Ed.
G. Grammel, Ed.
J. Drake, Ed.
Juniper
G. Galimberti, Ed.
Z. Ali, Ed.
Cisco
R. Kunze, Ed.
Deutsche Telekom
October 22, 2012

Extension to the Link Management Protocol (LMP/DWDM -rfc4209) for Dense
Wavelength Division Multiplexing (DWDM) Optical Line Systems to manage
black-link optical interface parameters of DWDM application
draft-dharinigert-ccamp-g-698-2-lmp-01

Abstract

This memo defines extensions to LMP(rfc4209) for managing Optical parameters associated with Wavelength Division Multiplexing (WDM) systems or characterized by the Optical Transport Network (OTN) in accordance with the Black-Link approach defined in ITU-T Recommendation G.698.2[ITU.G698.2].[ITU.G698.2], G.694.1[ITU.G694.1].[ITU.G694.1] and its extensions./>

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

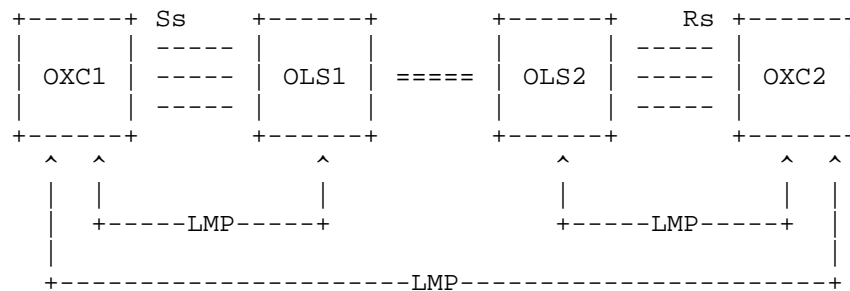
Table of Contents

1. Introduction	3
2. Extensions to LMP-WDM Protocol	3
3. Black Link General Parameters - BL_General	4
4. Black Link ApplicationCode - BL_ApplicationCode	5
5. Black Link - BL_Ss	6
6. Black Link - BL_SsRs	7
7. Black Link - BL_Rs	8
8. Black Link - OLS_Status	10
9. Security Considerations	12
10. IANA Considerations	12
11. References	13
11.1. Normative References	13
11.2. Informative References	13
Authors' Addresses	14

1. Introduction

This extension is based on "draft-galikunze-ccamp-g-698-2-snmp-mib-01" and "draft-kunze-g-698-2-management-control-framework-02", for the relevant optical parameters mainly (but not restricted to) described in recommendations like ITU-T G.698.2 [ITU.G698.2]. The LMP Model from RFC4902 is extended to provide link property correlation between a client and an OLS device. By using LMP, the capabilities of either end of this link are exchanged where the term 'link' refers to the attachment link between OXC and OLS (see Figure 1). By performing link property correlation, both ends of the link can agree on a common parameter window that can be supported and supervised by each device. The actual selection of a specific parameter value within the parameter window is outside the scope of LMP. In GMPLS the parameter selection (e.g. wavelength) is performed by RSVP-TE and Wavelength routing by IGP.

Figure 1 Extended LMP Model (from [RFC4209])



OXC : is an entity that contains transponders
 OLS : generic optical system, it can be -
 Optical mux , Optical demux, Optical Add Drop Mux
 , et
 OLS to OLS : represents the black-Link itself
 Rs/Ss : inbetween the OXC and the OLS

Figure 1: Extended LMP Model

2. Extensions to LMP-WDM Protocol

This document defines extensions to [RFC4209] to allow the Black Link (BL) parameters of G.698.2, as described in the draft draft-kunze-g-698-2-management-control-framework-02, to be exchanged

between a router or optical switch and the optical line system to which it is attached. In particular, this document defines additional Data Link sub-objects to be carried in the LinkSummary message defined in [RFC4204]. The OXC and OLS systems may be managed by different Network management systems and hence may not know the capability and status of their peer. The intent of this draft is to enable the OXC and OLS systems to exchange this information. These messages and their usage are defined in subsequent sections of this document.

The following new messages are defined for the WDM extension for ITU-T G.698.2. [ITU.G698.2]/ITU-T G.698.1. [ITU.G698.1]/ITU-T G.959.1. [ITU.G959.1]

- BL_General (sub-object Type = TBA)
- BL_ApplicationCode (sub-object Type = TBA)
- BL_Ss (sub-object Type = TBA)
- BL_SsRs (sub-object Type = TBA)
- BL_Rs (sub-object Type = TBA)
- BL_OLS_Status (sub-object Type = TBA)

3. Black Link General Parameters - BL_General

These are the general parameters as described in [G698.2]. Please refer to the "draft-galikunze-ccamp-g-698-2-snmp-mib-00". for more details about these parameters.

The general parameters are

1. Bit-Rate/line coding of optical tributary signals
2. Wavelength - (Tera Hertz) 4 bytes
3. Min Wavelength Range - (Tera Hertz) 4 bytes
4. Max Wavelength Range - (Tera Hertz) 4 bytes
5. Channel Spacing -(Giga Hertz) 4 bytes
6. BER mantisa - 4 bytes
7. BER exponent - 4 bytes
8. FEC Coding - 1 byte
9. Administrative state - 1 byte
10. Operation state - 1 byte

Figure 2: The format of the this sub-object (Type = TBA, Length = TBA) is as follows:

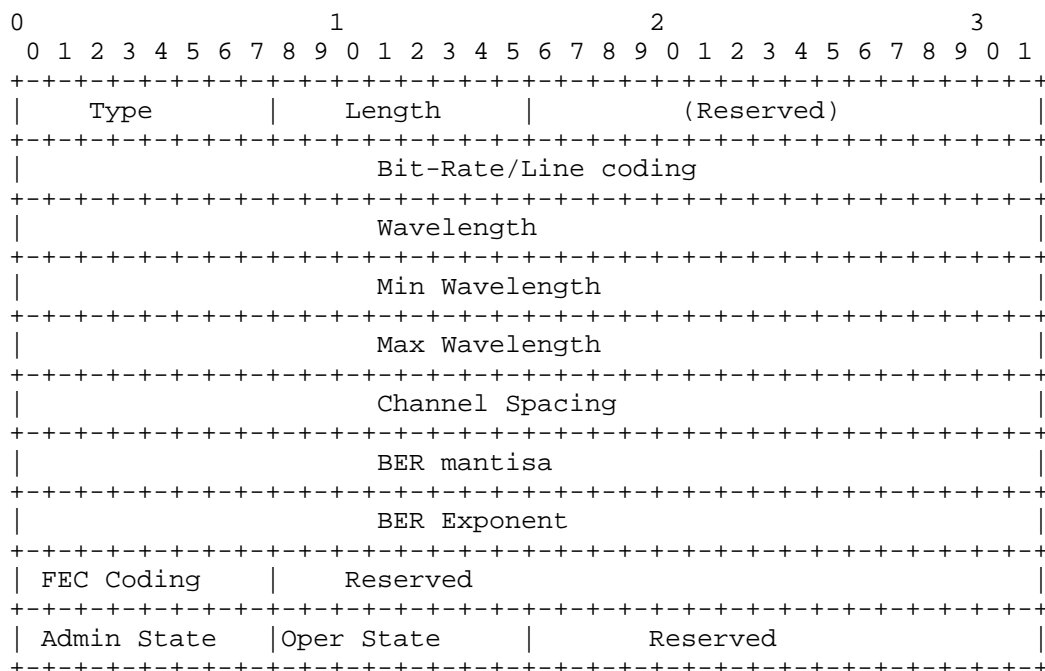


Figure 2: BL_General

4. Black Link ApplicationCode - BL_ApplicationCode

These are the general parameters as described in [G698.2]. Please refer to the "draft-galikunze-ccamp-g-698-2-snmp-mib-00". for more details about these parameters.

The general parameters are

1. Single-channel application codes -- 32 bytes
(from [G698.1]/[G698.2]/[G959.1])
2. Vendor Transceiver Class -- 32 bytes

When Single-channel application code (which is defined in G.698.2) is used in the message, then the vendor transaction class need not be used (i.e. all 0) and the optional parameters except nominal central frequency need not be used. When vendor transaction class is used in the message, then the Single-channel application code need not be used (i.e. all 0) and the optional parameters needs to be used.

Figure 3: The format of the this sub-object (Type = TBA, Length = TBA) is as follows:

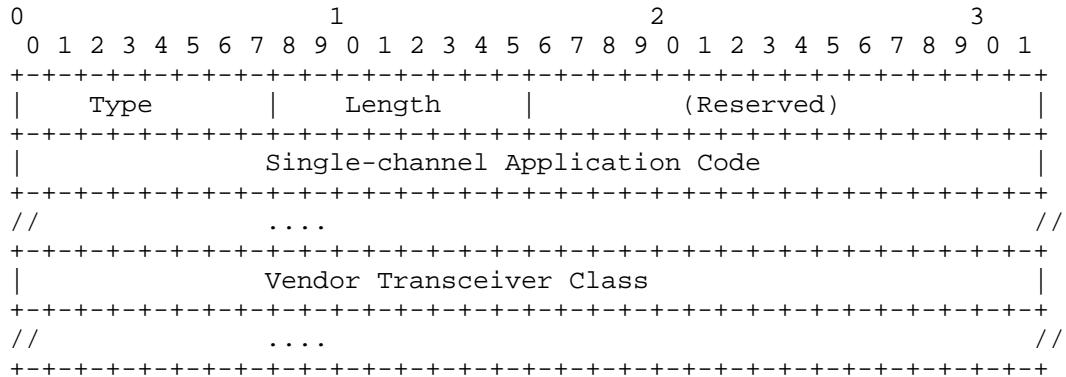


Figure 3: BL_ApplicationCode

5. Black Link - BL_Ss

These are the G.698.2 parameters at the Source(Ss reference points). Please refer to "draft-galikunze-ccamp-g-698-2-snmp-mib-00" for more details about these parameters.

1. Minimum Mean Channel Output Power -(0.1 dbm) 4 bytes
2. Maximum Mean Channel Output Power -(0.1 dbm) 4 bytes
3. Minimum Central Frequency - (0.01 THz) 4 bytes
4. Maximum Central Frequency - (0.01 THz) 4 bytes
5. Maximum Spectral Excursion - (0.1 GHz) 4 bytes
6. Maximum Tx Dispersion OSNR Penalty - (0.1 dbm) 4 bytes
7. Current Output Power - (0.1dbm) 4 bytes
8. Status of TX - Status of the Transmit link at OXC - 4 bytes

Figure 4: The format of the Black link sub-object (Type = TBA, Length = TBA) is as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Type										Length										(Reserved)																		
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Minimum Mean Channel Output Power																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Maximum Mean Channel Output Power																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Minimum Central Frequency																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Maximum Central Frequency																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Maximum Spectral Excursion																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Maximum Tx Dispersion OSNR Penalty																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Current Output Power																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								
	Status of TX																																						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-								

Figure 4: Black Link - BL_Ss

6. Black Link - BL_SsRs

These are the G.698.2 parameters for the path (Ss-Rs). Please refer to the "draft-galikunze-ccamp-g-698-2-snmplib-00" for more details about these parameters.

1. Minimum Chromatic Dispersion - (ps/nm) 4 bytes
2. Maximum Chromatic Dispersion -(ps/nm) 4 bytes
3. Minimum Src Optical ReturnLoss -(0.1 db) 4 bytes
4. Maximum Discrete Reflectance Src To Sink - (0.1 db) 4 bytes
5. Maximum Differential Group Delay - (ps) 4 bytes
6. Maximum Polarisation Dependent Loss - (0.1 db) 4 bytes
7. Maximum Inter Channel Crosstalk - (0.1 db) 4 bytes
8. Interferometric Crosstalk - (0.1 db) 4 bytes
9. Optical Path OSNR Penalty - (0.1 db) 4 bytes
10. Fiber type - 1 byte

Figure 5: The format of the Black link sub-object (Type = TBA, Length = TBA) is as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type										Length										(Reserved)																			
										Minimum Chromatic Dispersion																													
										Maximum Chromatic Dispersion																													
										Minimum Src Optical ReturnLoss																													
										Maximum Discrete Reflectance Src To Sink																													
										Maximum Differential Group Delay																													
										Maximum Polarisation Dependent Loss																													
										Maximum Inter Channel Crosstalk																													
										Interferometric Crosstalk																													
										Optical Path OSNR Penalty																													
Fiber Type										Reserved																													

Figure 5: Black Link - BL_SsRs

7. Black Link - BL_Rs

These are the G.698.2 parameters at the Sink (Rs reference points). Please refer to the "draft-galikunze-ccamp-g-698-2-snmp-mib-00" for more details about these parameters.

1. Minimum Mean Input Power - (0.1dbm) 4bytes
2. Maximum Mean Input Power - (0.1dbm) 4bytes
3. Minimum OSNR - (0.1dB) 4bytes
4. OSNR Tolerance - (0.1dB) 4bytes
5. Current Input Power at the OXC - (0.1dbm) 4bytes
6. Threshold of the input power at OLS
 - The power level above which the OLS will not function (0.1dbm) 4bytes
7. Current Optical OSNR (0.1dB)
8. Q factor
9. Post FEC BER Mantissa
10. Post FEC BER Exponent
11. Status of RX - Status of the Receive link at OXC - 2bytes

Figure 6: The format of the Black link sub-object (Type = TBA, Length = TBA) is as follows:

The format of the Black Link/OLS Sink sub-object (Type = TBA, Length = TBA) is as follows:

0	1	2	3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1			
+	+	+	+
	Type		Length
+	+	+	+
	Minimum Mean Input Power		
+	+	+	+
	Maximum Mean Input Power		
+	+	+	+
	Minimum OSNR		
+	+	+	+
	OSNR Tolerance		
+	+	+	+
	Current Input Power		
+	+	+	+
	Threshold for the input power at OXC		
+	+	+	+
	Current Optical OSNR		
+	+	+	+
	Current Q Factor		
+	+	+	+
	Post FEC BER Mantissa		
+	+	+	+
	Post FEC BER Exponent		
+	+	+	+
	Status of the RX		
+	+	+	+

Figure 6: Black Link - BL_Rs

8. Black Link - OLS_Status

This message is sent by the OLS to the OXC. It includes the wavelength information and the status of the OLS.

1. Wavelength
The wavelength which has been accepted by the OLS (Tera Hertz) 4 bytes
2. Length of the Wavelength Availability Map 1 byte
3. Wavelength Availability bits - variable bits depending on the number of wavelengths available (For eg 96 bits for C-band 50GHz)
(Allocation is in multiples of 1byte - 96 bits - 10 bytes)
0 - wavelength is available, 1 - used - variable length
4. Current Input Power (0.1dbm) 4 bytes
- This is the current input power at OLS
5. Delta between output power at the Src(OXC) and Input Power at OLS (0.1dbm) 4 bytes
- This is the delta between the input power and the transmitted output power at the OXC (from message 2.2 BL_Src)
6. Threshold of the input power at OLS 4 bytes
- This is the power level above which the OLS will not function.
7. Current Output Power (0.1dbm) 4 bytes
- This is the transmitted output power at the OLS.
8. Status of Rx link at OLS 2 bytes
- Status of the Receive link at the OLS
9. Status of Tx link at OLS 2 bytes
- Status of the Transmit link at the OLS

Figure 7: The format of the Black link sub-object (Type = TBA, Length = TBA) is as follows:

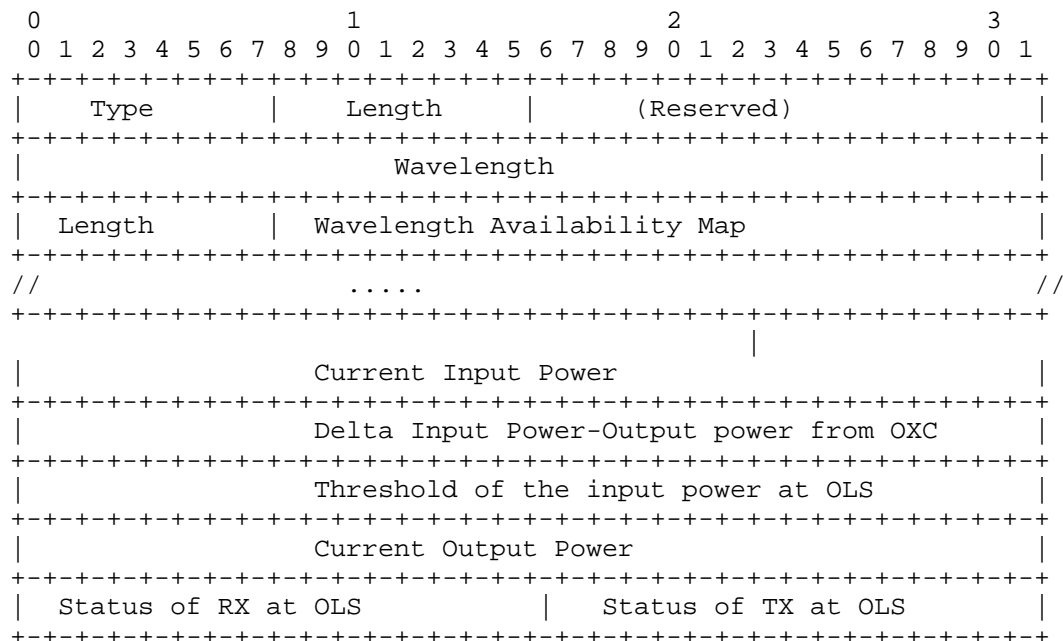


Figure 7: Black Link - OLS_Status

9. Security Considerations

LMP message security uses IPsec, as described in [RFC4204]. This document only defines new LMP objects that are carried in existing LMP messages, similar to the LMP objects in [RFC:4209]. This document does not introduce new security considerations.

10. IANA Considerations

LMP [RFC4204] defines the following name spaces and the ways in which IANA can make assignments to these namespaces: - LMP Message Type - LMP Object Class - LMP Object Class type (C-Type) unique within the Object Class - LMP Sub-object Class type (Type) unique within the Object Class This memo introduces the following new assignments: LMP Sub-Object Class names: under DATA_LINK Class name (as defined in [RFC4204]) - BL_General (sub-object Type = TBA) - BL_Applicationcode (sub-object Type = TBA) - BL_Ss (sub-object Type = TBA) - BL_SsRs (sub-object Type = TBA) - BL_Rs (sub-object Type = TBA) - OLS_Status

(sub-object Type = TBA)

11. References

11.1. Normative References

- [RFC4204] Lang, J., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [RFC4209] Fredette, A. and J. Lang, "Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems", RFC 4209, October 2005.
- [ITU.G698.2] International Telecommunications Union, "Amplified multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.
- [ITU.G694.1] International Telecommunications Union, "'Spectral grids for WDM applications: DWDM frequency grid'", ITU-T Recommendation G.698.2, June 2002.
- [ITU.G709] International Telecommunications Union, "Interface for the Optical Transport Network (OTN)", ITU-T Recommendation G.709, March 2003.
- [ITU.G872] International Telecommunications Union, "Architecture of optical transport networks", ITU-T Recommendation G.872, November 2001.

11.2. Informative References

- [I-D.kunze-g-698-2-management-control-framework] Kunze, R., "A framework for Management and Control of optical interfaces supporting G.698.2", draft-kunze-g-698-2-management-control-framework-00 (work in progress), July 2011.
- [I-D.galimbe-kunze-g-698-2-snmp-mib] Kunze, R. and D. Hiremagalur, "A SNMP MIB to manage black-link optical interface parameters of DWDM applications", draft-galimbe-kunze-g-698-2-snmp-mib-02 (work in progress), March 2012.

Authors' Addresses

Dharini Hiremagalur (editor)
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 California
USA

Phone: +1408
Email: dharinih@juniper.net

Gert Grammel (editor)
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 California
USA

Phone: +1408
Email: ggrammel@juniper.net

John E. Drake (editor)
Juniper
1194 N Mathilda Avenue
HW-US, Pennsylvania
USA

Phone: +1408
Email: jdrake@juniper.net

Gabriele Galimberti (editor)
Cisco
Via Philips,12
20052 - Monza
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

Zafar Ali (editor)
Cisco
3000 Innovation Drive
KANATA
ONTARIO K2K 3E8

Phone:
Email: zali@cisco.com

Ruediger Kunze (editor)
Deutsche Telekom
Dddd, xx
Berlin
Germany

Phone: +49xxxxxxxxxx
Email: RKunze@telekom.de

CCAMP Working Group
Internet-Draft
Intended status: Experimental
Expires: April 6, 2013

D. Dhody
U. Pallé
V. Kondreddy
Huawei Technologies India Pvt
Ltd
R. Casellas
CTTC - Centre Tecnologic de
Telecomunicacions de Catalunya
October 3, 2012

Domain Subobjects for Resource ReserVation Protocol - Traffic
Engineering (RSVP-TE)
draft-dhody-ccamp-rsvp-te-domain-subobjects-00

Abstract

The RSVP-TE specification, [RFC3209] and GMPLS extensions to RSVP-TE, [RFC3473] allow abstract nodes and resources to be explicitly included in a path setup. Further Exclude Routes Extension [RFC4874] allow abstract nodes and resources to be explicitly excluded in a path setup.

The use of Autonomous Number (AS) (2-Byte) as an abstract node representing domain is already defined in [RFC3209] and [RFC4874].

This document specifies new subobjects to include or exclude domains during path setup where domain is a collection of network elements within a common sphere of address management or path computational responsibility such as an IGP area or an Autonomous Systems (4-Byte).

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 6, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Terminology	3
3. Subobjects for Domains	4
3.1. Domains	4
3.2. Explicit Route Object (ERO)'s Subobjects	4
3.2.1. Autonomous system	5
3.2.2. IGP Area	6
3.2.3. Mode of Operation	7
3.3. Exclude Route Object (XRO)'s Subobjects	7
3.3.1. Autonomous system	8
3.3.2. IGP Area	8
3.3.3. Mode of Operation	9
3.4. Explicit Exclusion Route Subobject	9
4. Interaction with Path Computation Element (PCE)	10
5. IANA Considerations	10
5.1. New Subobjects	10
6. Security Considerations	10
7. Acknowledgments	11
8. References	11
8.1. Normative References	11
8.2. Informative References	11

1. Introduction

The RSVP-TE specification [RFC3209] and GMPLS extensions [RFC3473] allow abstract nodes and resources to be explicitly included in a path setup, using the Explicit Route Object (ERO).

Further RSVP-TE specification [RFC4874] allows abstract nodes or resources to be excluded from the whole path using the Exclude Route object (XRO). To exclude certain abstract nodes or resources between a specific pair of abstract nodes present in an ERO, a subobject Explicit Exclusion Route Subobject (EXRS) is used.

[RFC3209] already describes the notion of abstract nodes, where an abstract node is a group of nodes whose internal topology is opaque to the ingress node of the LSP. It further defines a subobject for Autonomous Systems (AS) (2-Byte).

This document extends the notion of abstract nodes by adding new subobjects for IGP Areas and 4-byte AS numbers. These subobjects MAY be included in Explicit Route Object (ERO), Exclude Route object (XRO) or Explicit Exclusion Route Subobject (EXRS).

1.1. Requirements Language

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 [RFC2119].

2. Terminology

The following terminology is used in this document.

AS: Autonomous System.

Domain: Any collection of network elements within a common sphere of address management or path computational responsibility. Examples of domains include Interior Gateway Protocol (IGP) areas and Autonomous Systems (ASs).

ERO: Explicit Route Object

EXRS: Explicit Exclusion Route Subobject

IGP: Interior Gateway Protocol. Either of the two routing protocols, Open Shortest Path First (OSPF) or Intermediate System to Intermediate System (IS-IS).

IS-IS: Intermediate System to Intermediate System.

OSPF: Open Shortest Path First.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

PCEP: Path Computation Element Protocol.

RSVP: Resource Reservation Protocol

TE LSP: Traffic Engineering Label Switched Path.

XRO: Exclude Route Object

3. Subobjects for Domains

3.1. Domains

[RFC4726] and [RFC4655] define domain as a separate administrative or geographic environment within the network. A domain may be further defined as a zone of routing or computational ability. Under these definitions a domain might be categorized as an Autonomous System (AS) or an Interior Gateway Protocol (IGP) area.

3.2. Explicit Route Object (ERO)'s Subobjects

As stated in [RFC3209], an explicit route is a particular path in the network topology. In addition to the ability to identify specific nodes along the path, an explicit route can identify a group of nodes (abstract nodes) that must be traversed along the path.

Some subobjects are defined in [RFC3209], [RFC3473], [RFC3477], [RFC4874] and [RFC5553] but new subobjects related to domains are needed.

The following subobject types are used in ERO.

Type	Subobject
1	IPv4 prefix
2	IPv6 prefix
3	Label
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
33	Explicit Exclusion (EXRS)
34	SRLG
64	IPv4 Path Key
65	IPv6 Path Key

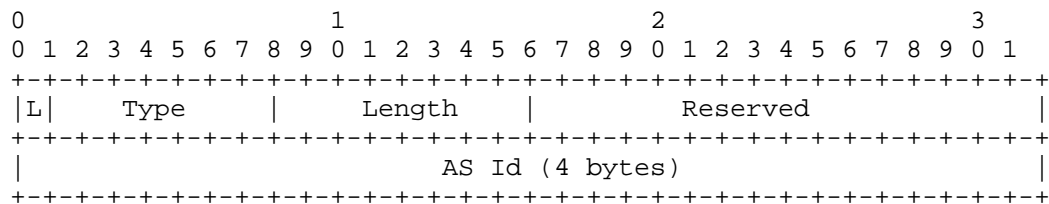
This document extends the above list to support 4-Byte AS numbers and IGP Areas.

Type	Subobject
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

3.2.1. Autonomous system

[RFC3209] already defines 2-Byte AS number.

To support 4-Byte AS number as per [RFC4893] following subobject is defined:



L: The L bit is an attribute of the subobject as define in [RFC3209].

Type: (TBA by IANA) indicating a 4-Byte AS Number.

Length: 8 (Total length of the subobject in bytes).

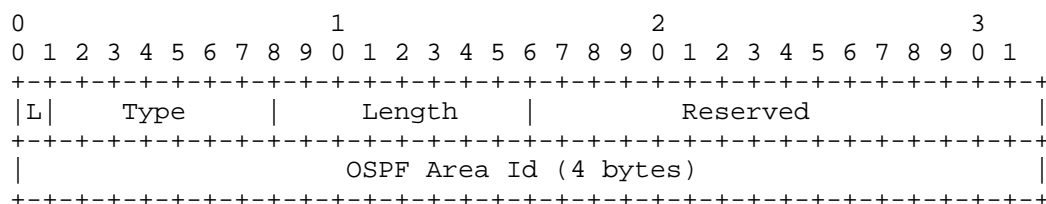
Reserved: Zero at transmission, Ignored at receipt.

AS-ID: The 4-Byte AS Number. Note that if 2-Byte AS numbers are in use, the low order bits (16 through 31) should be used and the high order bits (0 through 15) should be set to zero.

3.2.2. IGP Area

Since the length and format of Area-id is different for OSPF and ISIS, the following two subobjects are defined:

For OSPF, the area-id is a 32 bit number. The subobject is encoded as follows:



L: The L bit is an attribute of the subobject as define in [RFC3209].

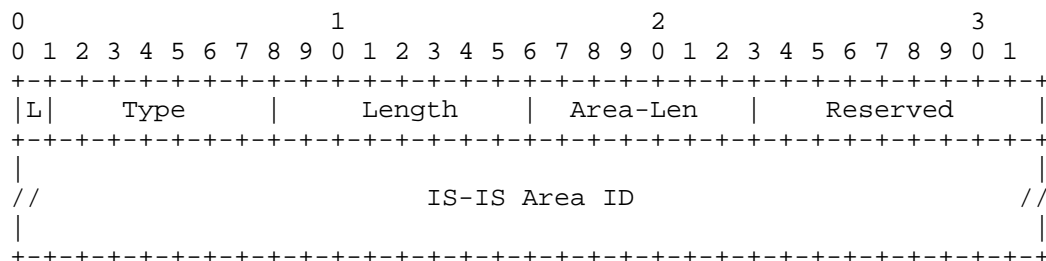
Type: (TBA by IANA) indicating 4-Byte OSPF Area ID.

Length: 8 (Total length of the subobject in bytes).

Reserved: Zero at transmission, Ignored at receipt.

OSPF Area Id: The 4-Byte OSPF Area ID.

For IS-IS, the area-id is of variable length and thus the length of the subobject is variable. The Area-id is as described in IS-IS by ISO standard [ISO 10589]. The subobject is encoded as follows:



L: The L bit is an attribute of the subobject as define in [RFC3209].

Type: (TBA by IANA) indicating IS-IS Area ID.

Length: Variable (Total length of the subobject in bytes including padding). The Length MUST be at least 4, and MUST be a multiple of

4.

Area-Len: Variable (Length of the actual (non-padded) IS-IS Area Identifier in bytes; Valid values are from 2 to 11 inclusive).

Reserved: Zero at transmission, Ignored at receipt.

IS-IS Area Id: The variable-length IS-IS area identifier. Padded with trailing zeroes to a four-byte boundary.

3.2.3. Mode of Operation

The new subobjects to support 4-Byte AS and IGP (OSPF / ISIS) Area MAY also be used in the ERO to specify an abstract node (a group of nodes whose internal topology is opaque to the ingress node of the LSP).

All the rules of processing (for example Next Hop Selection, L bit processing, unrecognized subobjects etc) are as per the [RFC3209].

3.3. Exclude Route Object (XRO)'s Subobjects

As stated in [RFC4874], the exclude route identifies a list of abstract nodes that should not be traversed along the path of the LSP being established.

Some subobjects are defined in [RFC3209], [RFC3477], [RFC4874] and [RFC6001] but new subobjects related to domains are needed.

The following subobject types are used in XRO.

Type	Subobject
1	IPv4 prefix
2	IPv6 prefix
3	Label
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
34	SRLG

This document extends the above list to support 4-Byte AS numbers and IGP Areas.

Type	Subobject
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id

TBD ISIS Area id

3.3.1. Autonomous system

[RFC3209] and [RFC4874] already defines a 2-Byte AS number.

To support 4-Byte AS number as per [RFC4893], the following subobject is defined:

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|L|      Type      |      Length      |      Reserved      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     AS Id (4 bytes)          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The meaning of the L bit, similar to [RFC4874], is as follows:

0 indicates that the abstract node (AS) specified MUST be excluded.

1 indicates that the abstract node (AS) specified SHOULD be avoided.

The meaning of all the other elements (Type, Length, Reserved and 4-Byte AS Id) is same as explained above in Section 3.2.1.

3.3.2. IGP Area

Since the length and format of Area-id is different for OSPF and ISIS, following two subobjects are defined:

For OSPF, the area-id is a 32 bit number. The subobject is encoded as follows:

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|L|      Type      |      Length      |      Reserved      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     OSPF Area Id (4 bytes)    |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The meaning of the L bit, similar to [RFC4874], is as follows:

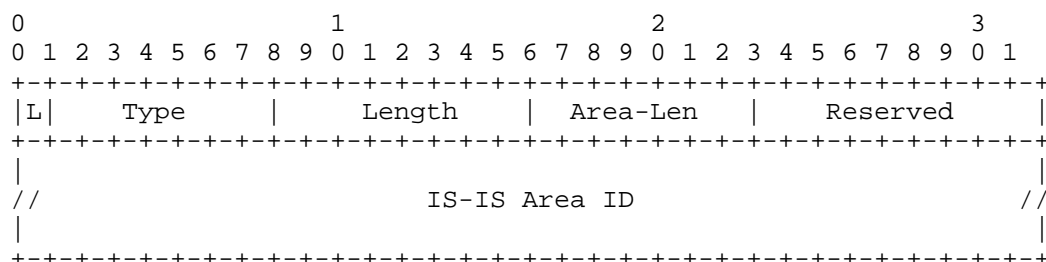
0 indicates that the abstract node (OSPF Area) specified MUST be

excluded.

1 indicates that the abstract node (OSPF Area) specified SHOULD be avoided.

The meaning of all the other elements (Type, Length, Reserved and OSPF Area Id) is same as explained above in Section 3.2.2.

For IS-IS, the area-id is of variable length and thus the length of the subobject is variable. The Area-id is as described in IS-IS by ISO standard [ISO 10589]. The subobject is encoded as follows:



The meaning of the L bit, similar to [RFC4874], is as follows:

0 indicates that the abstract node (IS-IS Area) specified MUST be excluded.

1 indicates that the abstract node (IS-IS Area) specified SHOULD be avoided.

The meaning of all the other elements (Type, Length, Reserved and IS-IS Area Id) is same as explained above in Section 3.2.2.

3.3.3. Mode of Operation

The new subobjects to support 4-Byte AS and IGP (OSPF / ISIS) Area MAY also be used in the XRO to specify exclusion of an abstract node (a group of nodes whose internal topology is opaque to the ingress node of the LSP).

All the rules of processing are as per the [RFC4874].

3.4. Explicit Exclusion Route Subobject

As stated in [RFC4874], the Explicit Exclusion Route defines abstract nodes or resources that must not or should not be used on the path between two inclusive abstract nodes or resources in the explicit

route. EXRS is an ERO subobject that contains one or more subobjects of its own, called EXRS subobjects.

The EXRS subobject may carry any of the subobjects defined for XRO, thus the new subobjects to support 4-Byte AS and IGP (OSPF / ISIS) Area MAY also be used in the EXRS. The meanings of the fields of the new XRO subobjects are unchanged when the subobjects are included in an EXRS, except that scope of the exclusion is limited to the single hop between the previous and subsequent elements in the ERO.

All the rules of processing are as per the [RFC4874].

4. Interaction with Path Computation Element (PCE)

The domain subobjects to be used in Path Computation Element Protocol (PCEP) are referred to in [PCE-DOMAIN]. Note that the new domain subobjects follow the principle that subobjects used in PCEP [RFC5440] are identical to the subobjects used in RSVP-TE.

5. IANA Considerations

5.1. New Subobjects

IANA registry: RSVP PARAMETERS

Subsection: Class Names, Class Numbers, and Class Types

IANA is requested to add further subobjects to the existing entry for:

20	EXPLICIT_ROUTE
232	EXCLUDE_ROUTE

Subobject	Type	Reference
TBA	4-Byte AS number	[This I.D.]
TBA	OSPF Area ID	[This I.D.]
TBA	IS-IS Area ID	[This I.D.]

6. Security Considerations

Security considerations for MPLS-TE and GMPLS signaling are covered in [RFC3209] and [RFC3473]. This document does not introduce any new messages or any substantive new processing, and so those security considerations continue to apply.

The route exclusion security consideration are covered in [RFC4874]

and continue to apply.

7. Acknowledgments

We would like to thank Reeja Paul and Sandeep Boina for their useful comments and suggestions.

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative 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., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, November 2006.
- [RFC4874] Lee, CY., Farrel, A., and S. De Chodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4874, April 2007.
- [RFC4893] Vohra, Q. and E. Chen, "BGP Support for Four-octet AS Number Space", RFC 4893, May 2007.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

- [RFC5553] Farrel, A., Bradford, R., and JP. Vasseur, "Resource Reservation Protocol (RSVP) Extensions for Path Key Support", RFC 5553, May 2009.
- [RFC6001] Papadimitriou, D., Vigoureux, M., Shiimoto, K., Brungard, D., and JL. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, October 2010.
- [PCE-DOMAIN] Dhody, D., Palle, U., and R. Casellas, "Standard Representation Of Domain Sequence. (draft-ietf-pce-pcep-domain-sequence-01)", July 2012.
- [ISO 10589] ISO, "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002.

Authors' Addresses

Dhruv Dhody
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: dhruv.dhody@huawei.com

Udayasree Palle
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: udayasree.palle@huawei.com

Venugopal Reddy Kondreddy
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: venugopalreddyk@huawei.com

Ramon Casellas
CTTC - Centre Tecnologic de Telecomunicacions de Catalunya
Av. Carl Friedrich Gauss n7
Castelldefels, Barcelona 08860
SPAIN

EMail: ramon.casellas@cttc.es

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

J. Dong
M. Chen
Huawei Technologies
Z. Li
China Mobile
October 22, 2012

RSVP-TE Extensions for Lock Instruct and Loopback in MPLS Transport
Profile
draft-dong-ccamp-rsvp-te-mpls-tp-li-lb-04

Abstract

This document specifies extensions to RSVP-TE to support lock instruct and loopback mechanism for MPLS-TP LSPs. The mechanisms are intended to be applicable to other technologies which use GMPLS/RSVP-TE as control plane.

Requirements Language

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].

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. Extensions to RSVP-TE	3
3. Operations	4
3.1. Lock Instruct	4
3.2. Loopback	5
4. IANA Considerations	6
5. Security Considerations	6
6. Acknowledgements	6
7. References	6
7.1. Normative References	6
7.2. Informative References	7
Authors' Addresses	7

Reflect (R): 1 bit - see [RFC3471]
Handover (H): 1 bit - see [RFC5852]
Lockout (L): 1 bit - see [RFC4872]
Inhibit Alarm Indication (I): 1 bit - see [RFC4783]
Call Control (C): 1 bit - see [RFC4974]
Testing (T): 1 bit - see [RFC3471]
Administratively down (A): 1 bit - see [RFC3471], reused for Lock
Deletion in progress (D): 1 bit - see [RFC3471]

A new bit is defined in Attribute Flags TLV [RFC5420] to indicate the loopback mode. The bit number is TBA.

Bit Number	Name and Usage
TBA	Loopback mode desired. This flag indicates a particular node on the LSP is required to enter loopback mode. This MAY also be used for specifying the loopback state of the node.

3. Operations

3.1. Lock Instruct

When a MEP wants to put an LSP into lock mode, it MUST send a Path message with the Administratively down (A) bit and the Reflect (R) bit in ADMIN_STATUS Object set. The intermediate nodes SHOULD forward the message with the A bit unchanged to the downstream .

On receipt of this Path message, the receiving MEP node SHOULD try to take the LSP out of service. If the receiving MEP locks the LSP successfully, it SHOULD send a Resv message with the A bit in ADMIN_STATUS object set. Otherwise, it SHOULD send a PathErr message with the Error Code "OAM Problem" and the new Error Value "Lock Failure", and the following Resv message SHOULD be sent with the A bit cleared. During this procedure, the intermediate nodes would be aware of whether the LSP is in Lock mode or not.

When an LSP is put in lock mode, the subsequent Path and Resv messages SHOULD keep the A bit in ADMIN_STATUS Object set.

When a MEP wants to take the LSP out of the lock mode, it MUST send a Path message with the A bit in ADMIN_STATUS Object cleared. The intermediate nodes SHOULD forward this message with the A bit unchanged to the downstream.

On receipt of this Path message, the receiving MEP node SHOULD try to bring the LSP back to service. If the receiving MEP unlocks the LSP

successfully, it SHOULD send a Resv message with the A bit in ADMIN_STATUS Object cleared. Otherwise, it SHOULD send a PathErr message with the Error Code "OAM Problem" and the new Error Value "Unlock Failure", and the following Resv message SHOULD be sent with the A bit set.

When an LSP is taken out of lock mode, the subsequent Path and Resv messages SHOULD keep the A bit in ADMIN_STATUS Object cleared.

3.2. Loopback

The loopback request can be sent either to the remote MEP or to a particular MIP node. The mechanism defined in [I-D.margaria-ccamp-lsp-attribute-ero] is used for addressing the loopback request to a particular node on the LSP. The loopback request is acceptable only when the LSP is in lock mode.

When a MEP wants to put a particular LSR on the LSP into loopback mode, it MUST send a Path message with the Loopback bit in the Attribute Flags TLV set. The mechanism defined in [I-D.margaria-ccamp-lsp-attribute-ero] is used to address the loopback request to the particular LSR. The Administratively down (A) bit in ADMIN_STATUS object SHOULD be set to keep the LSP in lock mode.

On receipt of this Path message, the target LSR of the loopback request SHOULD try to put the LSP into loopback mode. If the node puts the LSP into loopback mode successfully, it SHOULD set the the Loopback (B) Bit in the RRO Attribute subobject [RFC5420] and push this subobject onto the RRO object in the corresponding Resv message. The Administratively down (A) bit in ADMIN_STATUS object SHOULD be set in the Resv message. If the node cannot put the LSP into loopback mode, it SHOULD send a PathErr message with the Error Code "OAM Problem" and the new Error Value "Loopback Failure".

When a MEP wants to take the LSP out of loopback mode, it MUST send a Path message with the Loopback bit in the Attribute Flags TLV cleared. The mechanism defined in [I-D.margaria-ccamp-lsp-attribute-ero] is used to indicate that the particular LSR SHOULD exit loopback mode for this LSP. The Administratively down (A) bit in ADMIN_STATUS object SHOULD be set.

On receipt of this Path message, the target node SHOULD try to take the LSP out of loopback mode. If the node takes the LSP out of loopback mode successfully, it SHOULD clear the the Loopback (B) Bit in the RRO Attribute subobject and push this subobject onto the RRO object in the corresponding Resv message. The Administratively down (A) Bit in ADMIN_STATUS Object SHOULD be set. Otherwise, the node

SHOULD send a PathErr message with the Error Code "OAM Problem" and the new Error Value "Exit Loopback Failure".

4. IANA Considerations

One bit number Loopback needs to be assigned in the Attribute Flags registry.

Four new Error Values need to be allocated for Error Code "OAM Problem": "Lock Failure", "Unlock Failure", "Loopback Failure", "Exit Loopback Failure".

5. Security Considerations

This document does not introduce any new security issues above those identified in [RFC3209] and [RFC3473].

6. Acknowledgements

The authors would like to thank Greg Mirsky, Lou Berger and Francesco Fondelli for their comments and suggestions.

7. References

7.1. Normative References

- [I-D.ietf-ccamp-oam-configuration-fwk]
Takacs, A., Fedyk, D., and H. Jia, "GMPLS RSVP-TE extensions for OAM Configuration",
draft-ietf-ccamp-oam-configuration-fwk-08 (work in progress), July 2012.
- [I-D.margaria-ccamp-lsp-attribute-ero]
Margaria, C., Schroetter, D., Martinelli, G., Balls, S.,
and B. Wright, "LSP Attribute in ERO",
draft-margaria-ccamp-lsp-attribute-ero-01 (work in progress), July 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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.

- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.
- [RFC5860] Vigoureux, M., Ward, D., and M. Betts, "Requirements for Operations, Administration, and Maintenance (OAM) in MPLS Transport Networks", RFC 5860, May 2010.
- [RFC6371] Busi, I. and D. Allan, "Operations, Administration, and Maintenance Framework for MPLS-Based Transport Networks", RFC 6371, September 2011.

7.2. Informative References

- [I-D.ietf-ccamp-rsvp-te-mpls-tp-oam-ext]
Bellagamba, E., Andersson, L., Skoldstrom, P., Ward, D., and A. Takacs, "Configuration of Pro-Active Operations, Administration, and Maintenance (OAM) Functions for MPLS-based Transport Networks using RSVP-TE", draft-ietf-ccamp-rsvp-te-mpls-tp-oam-ext-10 (work in progress), October 2012.
- [RFC6435] Boutros, S., Sivabalan, S., Aggarwal, R., Vigoureux, M., and X. Dai, "MPLS Transport Profile Lock Instruct and Loopback Functions", RFC 6435, November 2011.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Mach Chen
Huawei Technologies
Huawei Building, No.156 Beiqing Rd.
Beijing 100095
China

Email: mach.chen@huawei.com

Zhenqiang Li
China Mobile
Unit2, Dacheng Plaza, No. 28 Xuanwumenxi Ave.
Beijing 100053
China

Email: lizhenqiang@chinamobile.com

Network Working Group
Internet Draft
Intended status: Standards Track

D. Fedyk
D. Beller
Lieven Levrau
Alcatel-Lucent

Expires: January 2013

July 9, 2012

Layer 1 VPN Enhanced Mode - Overlay Extension Service Model
draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on January 9, 2009.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
carefully, as they describe your rights and restrictions with respect
to this document. Code Components extracted from this document must
include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the Simplified BSD License.

Abstract

This document builds on the L1VPN framework [RFC4847] to extend the L1VPN from the basic mode to the enhanced mode by including additional constraints, focusing upon the overlay extension service model. Route Diversity for customer LSPs are common requirement applicable to L1VPNs. This document describes L1VPN compatible mechanisms to achieve diversity for sets of customer LSPs. The extended overlay service model can support other extensions for L1VPN signaling, for example, those related to latency requirements.

Table of Contents

1. Introduction.....	3
2. Conventions used in this document.....	3
3. Contributors.....	3
4. LSP Diversity in the Overlay Extension Service Model.....	4
4.1. LSP diversity for dual-homed customer edge (CE) devices...	5
4.1.1. Exchanging SRLG information between the PEs via the CE device.....	7
4.1.1.1. Operational Procedures.....	8
4.1.1.2. Error handling procedures.....	8
4.1.2. Using Path Affinity Set extension.....	9
4.1.2.1. Operational Procedures.....	12
4.1.2.2. Error handling procedures.....	13
4.1.2.3. Distribution of the Path Affinity Set information	13
5. Latency signaling.....	14
6. Security Considerations.....	14
7. IANA Considerations.....	15
8. References.....	15
8.1. Normative References.....	15
8.2. Informative References.....	15
9. Acknowledgments.....	16

1. Introduction

This document builds on the L1VPN framework [RFC4847] to extend the L1VPN from the basic mode to the enhanced mode by including additional constraints, focusing upon the overlay extension service model.

The overlay model assumes a UNI interface between the edge nodes of the respective transport domains. Route diversity for LSPs from single homed CE and dual-home CEs is a common requirement in optical transport networks. This document describes two signaling variations that may be used for supporting LSP diversity within the overlay extension service model considering dual-homing. Dual-homing is typically used to avoid a single point of failure (UNI link, PE) or if two disjoint connections are forming a protection group. While both methods are similar in that they utilize common mechanisms in the PE network to achieve diversity, they are distinguished according to whether the CE is permitted to retrieve provider SRLG diversity information for an LSP from a PE1 and pass it on to a PE2 (SRLG information is shared with the CE), or whether a new attribute is used that allows the PE2 that receives this attribute to derive the SRLG information for an LSP based on this attribute value.

The extended overlay service model can support other extensions for L1VPN signaling, for example, those related to latency. When requesting diverse LSPs latency may also be an additional requirement.

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. Contributors

The Authors would like to thank Eve Varma and Sergio Belotti for their review and contributions to this document.

4. LSP Diversity in the Overlay Extension Service Model

The L1VPN Framework [RFC4847] (Enhanced Mode) describes the overlay extension service model, which builds upon the UNI Overlay [RFC4208] serving as the interface between the CE edge node and the PE edge node. In this service model, a CE receives a list of CE-PE TE link addresses to which it can request a L1VPN connection (i.e., membership information) and may include additional information concerning these TE links. This document further builds on the overlay extension service model by adding shared constraint information for path diversity in the optical transport network.

This document describes two signaling variations that may be used for supporting LSP diversity within the overlay extension service model considering dual-homing. While both methods are similar in that they utilize common mechanisms in the PE network to achieve diversity, they are distinguished according to whether the CE is permitted to retrieve provider SRLG diversity information for an LSP from a PE1 and pass it on to a PE2 (SRLG information is shared with the CE or whether a new attribute is used that allows the PE2 that receives this attribute to derive the SRLG information for an LSP based on this attribute value. The selection between these methods is governed by both PE-network specific policies and approaches taken (i.e., in terms of how the provider chooses to perform routing internal to their network).

The first method (see 3.1.1) assumes that provider Shared Resource Link Group (SRLG) Identifier information is both available and shareable (policy decision) with the CE. Since SRLG IDs can then be used (passed transparently between PEs via the dual-homed CE) as signaled information on a UNI message, a mechanism supporting LSP diversity for the overlay extension service model can be provided via straightforward signaling extensions.

The second method (see 3.1.2) assumes that provider SRLG IDs are either not available or not shareable (based on provider network operator policy) with the CE. For this case, a mechanism is provided where information signaled to the PE on UNI messages does not require shared knowledge of provider SRLG IDs to support LSP diversity for the overlay extension model.

Both approaches follow the L1VPN framework.

While both methods could be implemented in the same PE network, it is likely that an L1VPN CE network would use only one mechanism at a time.

Single-homed CE devices are connected to a single PE device via a single UNI link (could be a bundle of parallel links which are typically using the same fiber cable). This single UNI link may constitute a single point of failure. Such a single point of failure can be avoided when the CE device is connected to two PE devices via two UNI interfaces as depicted for CE1 in Figure 1 below.

For the dual-homing case, it is possible to establish two connections from the source CE device to the same destination CE device where one connection is using one UNI link to, for example, PE1 and the other connection is using the UNI link to PE2. In order to avoid single points of failure within the provider network, it is necessary to also ensure path (LSP) diversity within the provider network in order to achieve end-to-end diversity for the two LSPs between the two CE devices. This document describes how it is possible to enable such path diversity to be achieved within the provider network (which is subject to additional routing constraints). [RFC4202] defines SRLG information that can be used to allow GMPLS to provide path diversity in a GMPLS controlled transport network. As the two connections are entering the provider network at different PE devices, the PE device that receives the connection request for the second connection needs to be capable of determining the additional path computation constraints such that the path of the second LSP is disjoint with respect to the already established first connection entering the network at a different PE device. The methods described in this document allow a PE device to determine the SRLG information for a connection in the provider network that is entering the network on a different PE device.

PE SRLG information can be used directly by a CE if the CE understands the context, and the CE view is limited to its L1VPN context. In this case, there is a dependency on the provider information and there is a need to be able to query the SRLG in the provider network.

It may, on the other hand, be preferable to avoid this dependency and to decouple the SRLG identifier space used in the provider network from the SRLG space used in the client network. This is possible with both methods detailed below. Even for the method where provider SRLG information is passing through the CE device (note the CE device does not need to process and decode this information) the two SRLG identifier spaces can remain fully decoupled and the operator of the client network is free to assign SRLG identifiers from the client

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
 SRLG identifier space to the CE to CE connection that is passing
 through the provider network.

Referring to Figure 1, the UNI signaling mechanism must support at
 least one of the two mechanisms described in this document for CE
 dual homing to achieve LSP diversity in the provider network.

The described mechanisms can also be applied to a scenario where two
 CE devices are connected to two different PE devices. In this case,
 the additional information that is exchanged across the UNI
 interfaces also needs to be exchanged between the two CE devices in
 order to achieve the desired diversity in the provider network.

This information may be configured or exchanged by some automated
 mechanism not described in this document.

In the dual-homing example, CE1 can locally correlate the LSP
 requests. For the slightly more complicated example involving CE2 and
 CE3, both requiring a path that shall be diverse to a connection
 initiated by the other CE device, CE2 and CE3 need to have a common
 view of the SRLG information to be signaled. In this document, we
 detail the required diversity information and the signaling of this
 diversity information; however, the means for distributing this
 information within the PE domain or the CE domain is out of scope.

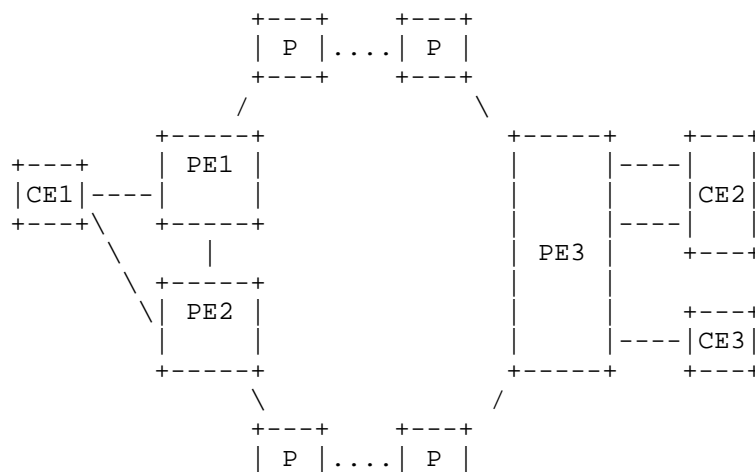


Figure 1: Generalized Layer 1 VPN Reference Model

Figure 1 Overlay Reference Diagram

In an overlay model, the information exchanged between the CE and the PE is kept to a minimum.

How diversity is achieved, in terms of configuration, distribution and usage in each part of the transport networks should be kept independent and separate from how diversity is signaled at the UNI between the two transport networks.

Signaling parameters discussed in this document are:

- o SRLG information (see [RFC4202])
- o Path Affinity Set

4.1.1.1. Exchanging SRLG information between the PEs via the CE device

SRLG information is defined in [RFC4202] and if the SRLG information of an LSP is known, it can be used to calculate a path for another LSP that is SRLG diverse with respect to an existing LSP. SRLG information is an unordered list of SRLGs. SRLG information is normally not shared between the transport network and the client network; i.e., not shared with the CEs of a L1VPN in the L1VPN context. However, this becomes more challenging when a CE is dual-homed. For example, CE1 in Figure 1 may have requested an LSP1 from CE1 to CE2 via PE1 and PE3. CE1 could subsequently request an LSP2 to CE2 via PE2 and PE3 with the requirement that it should be maximally SRLG disjoint with respect to LSP1. Since PE2 does not have any information about LSP1, PE2 would need to know the SRLG information associated with LSP1. If CE1 could request the SRLG information of LSP1 from PE1, it could then transparently pass this information to PE2 as part of the LSP2 setup request, and PE2 would now be capable of calculating a path for LSP2 that is SRLG disjoint with respect to LSP1.

The exchange of SRLG information is achieved on a per L1VPN LSP basis using the existing RSVP-TE signaling procedures. It can be exchanged in the PATH (exclusion information) or RESV message in the original request or it can be requested by the CE at any time the path is active.

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
It shall be noted that SRLG information is an unordered list of SRLG identifiers and the encoding of SRLG information for RSVP signaling is already defined in [SRLG_info]. Even if SRLG information is known for several LSPs it is not possible for the CEs to derive the provider network topology from this information.

4.1.1.1. Operational Procedures

Retrieving SRLG information from a PE for an existing LSP:

When a dual-homed UNI-C intends to establish an LSP to the same destination UNI-C via another PE node, it can request the SRLG information for an already established LSP by setting the SRLG information flag in the LSP attributes sub-object of the RSVP PATH message (IANA to assign the new SRLG flag). As long as the SRLG information flag is set in the PATH message, the PE node inserts the SRLG sub-object as defined in [SRLG_info] into the RSVP RESV message that contains the current SRLG information for the LSP. If the provider network's policy has been configured so as not to share SRLG information with the client network, the SRLG sub-object is not inserted in the PATH message even if the SRLG information flag is set. The PE passes on the SRLG information for the LSP. Note the SRLG information is expected to be up-to-date.

Establishment of a new LSP with SRLG diversity constraints:

When a dual-homed CE device sends an LSP setup requests to a PE device for a new LSP that is required to be SRLG diverse with respect to an existing LSP that is entering the network via another PE device, the UNI-C sets the SRLG diversity flag (note: IANA to assign the new SRLG diversity flag) in the LSP attributes sub-object of the PATH message that initiates the setup of this new LSP. When the PE device receives this request it calculates a path to the given destination and uses the received SRLG information as path computation constraints.

4.1.1.2. Error handling procedures

To be added in the next version of the document.

The Path Affinity Set (PAS) is used to signal diversity in a pure CE context by abstracting SRLG information. There are two types of diversity information in the PAS. The first type of information is a single PAS identifier. Optionally, more detailed PATH information of an exclude path or set of paths can be specified. The motive behind the PAS information is to have as little exchange of diversity information as possible between the L1VPN CE and PE elements.

Rather than a detailed CE or PE SRLG list, the Path Affinity Set contains an abstract SRLG identifier that associates the given path as diverse. Logically the identifier is in an L1VPN context and therefore only unique with respect to a particular L1VPN.

How the CE determines the PAS identifier is a local matter for the CE administrator. A CE may signal PAS as a diversity object in the PATH message. This identifier is a suggested identifier and may be overridden by a PE under some conditions.

For example, PAS can be used with no prior exchange of PAS information between the CE and the PE. Upon reception of the PAS information the PE can infer the CE's requirements. The actual PAS identifier used will be returned in the RESV message. Optionally an empty PAS identifier allows the PE to pick the PAS identifier. Similar to the section 4.1.1 on SRLG information, a PE can return PAS identifier as the response to a Query allowing flexibility.

A PE interprets the specific PAS identifier, for example, "123" as meaning to exclude that identifier and by association any PE related SRLG information, for any LSPs associated with the resources assigned to the L1VPN. For example, if a Path exists for the LSP with the identifier "123", the PE would use local knowledge of the PE SRLGs associated with the "123" LSPs and exclude those SRLGs in the path request. In other words, two LSPs that need to be diverse both signal "123" and the PEs interpret this as meaning not to use shared resources. Alternatively, a PE could use the PAS identifier to select from already established LSPs. Once the path is established it becomes associated with the "123" identifier or optionally another PAS identifier for that L1VPN.

The PAS Source and Destination Address tuple represents one or more source addresses and destination addresses associated with the CE Path Affinity Set identifier. These associated address tuples represent paths that use resources that should be excluded for the establishment of the current LSP. The address tuple information gives both finer grain details on the path diversity request and

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
 serves as an alternative identifier in the case when the PAS
 identifier is not known by the PE. The address tuples used in
 signaling is within a CE context and its interpretation is local to a
 PE that receives a Path request from a CE. The PE can use the address
 information to relate to PE Addresses and PE SRLG information. When
 a PE satisfies a connection setup for a (SRLG) diverse signaled path,
 the PE may optionally record the PE SRLG information for that
 connection in terms of PE based parameters and associate that with
 the CE addresses in the Path message.

The L1VPN Port Information table (PIT) [RFC5251] can be leveraged to
 translate between CE based addresses and PE based addresses. The Path
 Affinity Set and associated PE addresses with PE SRLG information can
 be distributed via the IGP in the provider transport network (or by
 other means such as configuration); they can be utilized by other PEs
 when other CE Paths are setup that would require path/connection
 diversity. This information is distributed on a L1VPN basis and
 contains a PAS identifier, PE addresses and SRLG information.

The CE Path Affinity Set may be used to signal paths without CE
 Source and Destination addresses; however, the PE will always
 associate the CE SRLG Group with a list of PE SRLG plus the PE
 addresses associated with this LSP.

If diversity is not signaled, the assumption is that no diversity is
 required and the Provider network is free to route the LSP to
 optimize traffic. No Path affinity set information needs to be
 recorded for these LSPs. If a diversity object is included in the
 connection request, the PE in the Provider Network should be able to
 look-up the existing Provider SRLG information from the provider
 network and choose an LSP that is maximally diverse from other LSPs.
 The mechanisms to achieve this are outside the scope of this
 document.

A new L1VPN Diverse LSP LABEL object is specified:

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               |                               |
|      Length                  |      Type (TBA) | 0 | C-type (TBA) |
+-----+-----+-----+-----+-----+-----+-----+-----+

      1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| ADDR Length | Number of PAS | D |                               reserved |

```

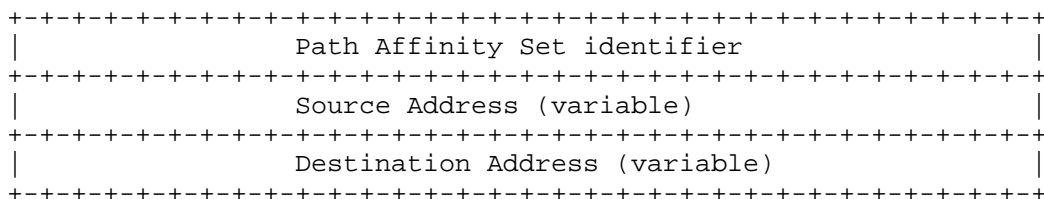


Figure 2 Diverse LSP information

1. The Address Length field (8 bits) is the number of bytes for both the source address and destination address. The address may be in any format from 1 to 32 bytes but the key point is the customers can maintain their existing addresses. A value of zero indicates there are no addresses included.
2. The Number of Path Affinity (8 bits)sets is included in the object. This is typically 1. Addition of other sets is for further study.
3. The Path affinity Set identifier (4 bytes) is a single number that represents a summarized SRLG for this path. Paths with that same Path Affinity set should be set up with diverse paths and associated with the path affinity set. A value of all zeros allows the PE to pick a PAS identifier to return. A PAS identifier of an established path may be different than the requested path identifier.
4. The diversity Bit (D) (one Bit) indicates if the diversity must be satisfied when set as a one. If a PE finds an established path with a Path Affinity set matching the signaled Path Affinity Set or the signaled Address tuple it should attempt find a diverse path.

5. The Diverse Path Source address/destination address tuple is that of an established LSP in the PE network that belongs to the same Path Affinity Set identifier. If the path for these addresses is not setup or cannot be determined by the PE edge processing the UNI then the path is only with the Path Affinity set constraint. If the path(s) for these address tuples are known by the PE the PE uses the SRLG information associated with these addresses. If in any case a diverse path cannot be setup then the Diverse bit controls whether a path is established anyway. The PE must use a mechanism to translate CE Addresses into provider addresses when correlating with provider SRLG information. How SRLG information and network address tuples are distributed is for future study.

4.1.2.1. Operational Procedures

When a UNI-C constructs a PATH message it may optionally specify and insert a Path Affinity Set in the PATH message. This Path Affinity Set may optionally include the address of an LSP that that could belong to the same Path Affinity Set. The Path Affinity Set identifier is a value (0 through 2³²-255) that is independent of the mechanism the CE or the PE use for diversity. The Path Affinity Set is a single identifier that can be used to request diversity and associate diversity.

When processing a CE PATH message in a L1VPN Overlay, the PE first looks up the PE based addresses in the Provider Index Table (PIT). If the Path Affinity Set is included in the PATH message, the PE must look up the SRLG information (or equivalent) in the PE network that has been allocated by LSPs associated with a Path Affinity Set and exclude those resources from the path computation for this LSP if it is a new path. The PE may alternatively choose from an existing path with a disjoint set of resources. If a path that is disjoint cannot be found, the value of the PAS diversity bit determines whether a path should be setup anyway. If the PAS diversity bit is clear, one can still attempt to setup the LSP. A PE should still attempt to minimize shared resources but that is an implementation issue, and is outside the scope of this document.

Optionally the CE may use a value of all zeros in the PAS identifier allowing the PE to select an appropriate PAS identifier. Also the PE may to override the PAS identifier allowing the PE to re-assign the identifier if required. A CE should not assume that the PAS identifier used for setup is the actual PAS identifier.

4.1.2.2. Error handling procedures

The PAS object must be understood by the PE device. Otherwise, the CE should not use the PAS object. Path Message processing of the PAS object SHOULD follow CTYPE 0. An Error code of IANA (TBD) indicates that the PAS object is not understood.

When a PAS identifier is not recognized by a PE it must assume this LSP defines that PAS identifier however the PE may override PAS identifier under certain conditions.

If the identifier is recognized but the Source Address-Destination address pair(s) are not recognized, this LSP must be set up using the PAS identifier only.

If the identifier is recognized and the Source Address-Destination address pair(s) are also recognized, then the PE SHOULD use the PE SRLG information associated with the LSPs identified by the address pairs to select a disjoint path.

The Following are the additional error codes:

- 1) Route Blocked by Exclude Route Value IANA (TBA).

4.1.2.3. Distribution of the Path Affinity Set information

Information about SRLG is already available in the IGP TE database. A PE network can be designed to have additional opaque records for Provider paths that distribute PE paths and SRLG on a L1VPN basis. When a PE path is setup, the following information allows a PE to lookup the PE diversity information:

- L1 VPN Identifier 8 bytes
- Path Affinity Set Identifier
- Source PE Address
- Destination PE Address
- List of PE SRLG (variable)

The source PE address and destination PE address are the same addresses in the L1VPN PIT and correspond to the respective CE address identifiers.

Note that all of the information is local to the PE context and is not shared with the CE. The L1VPN Identifier is associated with a

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
CE. The only value that is signaled from the CE is the Path Affinity Set and optionally the addresses of an existing LSP. The PE stores source and destination PE addresses of the LSP in their native format along with the SRLG information. This information is internal to the PE network and is always known.

PE paths may be setup on demand or they may be pre-established. When paths are pre-established, the Path Affinity Set is set to unassigned 0x0000 and is ignored. When a CE uses a pre-established path the PE may set the Path SRLG Path Affinity Set value if the CE signals one otherwise the Path Affinity Set remains unassigned 0x0000.

5. Latency signaling

A latency requirement can be added to signaling in the form of a constraint [DRAFT OBJECTIVE FUNCTION]. The constraint can take the form of:

- Minimize latency
- Maximum acceptable

While some systems may be able to compute routes based on delay metrics it is usual that minimizing hops subject to bandwidth reservation are satisfied as the object function and delay is not considered. When considering diversity latency falls after diversity constraints have been satisfied.

Recording the latency of existing paths [DRAFT_TE_METRIC RECORD] to ensure they meet a maximum acceptable latency can be utilized to ensure latency constraint is met.

When a low latency path is required, the minimize latency subject to other constraints criteria should be signaled. A CE device can use the record latency to ensure that the maximum acceptable latency has been met.

More detail to be added in a future revision.

6. Security Considerations

Security for L1VPNs is covered in [RFC4847], [RFC5251] and [RFC5253]. In this document, the model follows the L1VPN control plane model where CE addresses are completely distinct from the PE addresses.

The use of a private network assumes that entities outside the network cannot spoof or modify control plane communications between CE and PE. Furthermore, all entities in the private network are assumed to be trusted. Thus, no security mechanisms are required by the protocol exchanges described in this document.

However, an operator that is concerned about the security of their private control plane network may use the authentication and integrity functions available in RSVP-TE [RFC3473] or utilize IPsec ([RFC4301], [RFC4302], [RFC4835], [RFC5996], and [RFC6071]) for the point-to-point signaling between PE and CE. See [RFC5920] for a full discussion of the security options available for the GMPLS control plane.

7. IANA Considerations

TBD

8. References

Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4202] Kompella, K., Rekhter, Y., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC5251] Fedyk, D., Rekhter, Y., Editors "Layer 1 VPN Basic Mode", RFC 5251, July 2008.
- [SRLG_info] Zhang, F., Gonzalez de Dios, O., Margaria, C., "RSVP-TE Extensions for Collecting SRLG Information", draft-zhang-ccamp-srlg-fa-configuration-05.txt, March 2012.

8.2. Informative References

- [RFC6071] S. Frankel, S. Krishnan, " IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap", RFC 6071, February 2011.
- [RFC3473] Berger, L. (editor), "Generalized MPLS Signaling - RSVP-TE Extensions", RFC 3473, January 2003.

- Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
- [RFC4208] G. Swallow, J. Drake, H. Ishimatsu, Y. Rekhter,
"Generalized Multiprotocol Label Switching (GMPLS) User-
Network Interface (UNI): Resource ReserVation Protocol-
Traffic Engineering (RSVP-TE) Support for the Overlay
Model", RFC 4208, October 2005.
- [RFC4301] S. Kent, K. Seo, "Security Architecture for the Internet
Protocol," December 2005.
- [RFC4302] S. Kent, "IP Authentication Header," December 2005.
- [RFC5996] C. Kaufman, P. Hoffman, Y. Nir, P. Eronen " Internet Key
Exchange Protocol Version 2 (IKEv2)", September 2010.
- [RFC4835] V. Manral, "Cryptographic Algorithm Implementation
Requirements for Encapsulating Security Payload (ESP) and
Authentication Header (AH)", April 2007.
- [RFC4847] Takeda, T., Editor "Framework and Requirements for Layer 1
Virtual Private Networks", RFC 4847, April 2007.
- [RFC5253] Takeda, T., Editor "Applicability Statement for Layer 1
Virtual Private Network (L1VPN) Basic Mode", RFC 5253, July
2008.
- [RFC5920] L. Fang, Ed., "Security Framework for MPLS and GMPLS
Networks", RFC 5920, July 2010.
- [DRAFT OBJECTIVE FUNCTION] Ali, Z., Swallow, G., Filsfils, C., Fang,
L., Kunze, R., "Resource ReserVation Protocol-Traffic
Engineering (RSVP-TE) extension for signaling Objective
Function and Metric Bound", draft-ali-ccamp-rc-objective-
function-metric-bound-01.txt, March 2012.
- [DRAFT_TE_METRIC RECORD] Ali, Z., Swallow, G., Filsfils, C., Kunze,
R., "Resource ReserVation Protocol-Traffic Engineering
(RSVP-TE) extension for recording TE Metric of a Label
Switched Path", draft-ali-ccamp-te-metric-recording-01.txt,
March 2012.

9. Acknowledgments

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

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
Copyright (c) 2012 IETF Trust and the persons identified as authors
of the code. All rights reserved.

Redistribution and use in source and binary forms, with or without
modification, is permitted pursuant to, and subject to the license
terms contained in, the Simplified BSD License set forth in Section
4.c of the IETF Trust's Legal Provisions Relating to IETF Documents
(<http://trustee.ietf.org/license-info>).

Internet-Draft draft-fedyk-ccamp-llvpn-extnd-overlay-00.txt July 2012
Authors' Addresses

Don Fedyk
Alcatel-Lucent
Groton, MA, 01450
Email: donald.fedyk@alcatel-lucent.com

Dieter Beller
Alcatel-Lucent
Email: Dieter.Beller@alcatel-lucent.com

Lieven Levrau
Alcatel-Lucent
Email: Lieven.Levrau@alcatel-lucent.com

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

G.Galimberti, Ed.
Cisco
R.Kunze, Ed.
Deutsche Telekom
Kam Lam, Ed.
Alcatel-Lucent
D. Hiremagalur, Ed.
Juniper
October 22, 2012

An SNMP MIB extension to RFC3591 to manage optical interface parameters
of DWDM applications
draft-galikunze-ccamp-g-698-2-snmp-mib-01

Abstract

This memo defines a module of the Management Information Base (MIB) used by Simple Network Management Protocol (SNMP) in TCP/IP- based internets. In particular, it defines objects for managing Optical parameters associated with Dense Wavelength Division Multiplexing (DWDM) interfaces or characterized by the Optical Transport Network (OTN). This is an extension of the RFC3591 to support the optical parameters mainly but not only described in recommendations like ITU-T G.698.2. [ITU.G698.2]

The MIB module defined in this memo can be used for Optical Parameters monitoring and/or configuration of the endpoints of Black Links.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	4
2. The Internet-Standard Management Framework	5
3. Conventions	5
4. Overview	5
4.1. Optical Parameters Description	6
4.1.1. General	7
4.1.2. Parameters at Ss	9
4.1.3. Optical path from point Ss to Rs	10
4.1.4. Interface at point Rs	12
4.1.5. Alarms and Threshold definition	13
4.1.6. Performance Monitoring (PM) description	14
4.1.7. Generic Parameter description	16
4.2. Use of ifTable	16
4.2.1. Use of ifTable for OPS Layer	18
4.2.2. Use of ifTable for OCh Layer	19
4.2.3. Use of ifStackTable	19
5. Structure of the MIB Module	20
6. Object Definitions	20
7. Relationship to Other MIB Modules	77
7.1. Relationship to the [TEMPLATE TODO] MIB	77
7.2. MIB modules required for IMPORTS	77
8. Definitions	77
9. Security Considerations	77
10. IANA Considerations	78
11. Contributors	79
12. References	80
12.1. Normative References	80
12.2. Informative References	81
Appendix A. Change Log	82
Appendix B. Open Issues	82
Authors' Addresses	82

1. Introduction

This memo defines a portion of the Management Information Base (MIB) used by Simple Network Management Protocol (SNMP) in TCP/IP- based internets. In particular, it defines objects for managing Optical parameters associated with Wavelength Division Multiplexing (WDM) systems or characterized by the Optical Transport Network (OTN) in accordance with but not limited to the optical interface defined in G.698.2 [ITU.G698.2]

Black Link approach allows supporting an optical transmitter/receiver pair of one vendor to inject a DWDM channel and run it over an optical network composed of amplifiers, filters, add-drop multiplexers from a different vendor. From architectural point of view, the "Black Link" is a set of pre-configured/qualified network connections between the G.698.2 reference points S and R. The black links will be managed at the edges (i.e. the transmitters and receivers attached to the S and R reference points respectively) for the relevant parameters specified in G.698.2 [ITU.G698.2], G.798 [ITU.G798], G.874 [ITU.G874], and the performance parameters specified G.7710/Y.1701 [ITU-T G.7710] and and G.874.1 [ITU.G874.1].

The G.698.2 [ITU.G698.2] provides optical parameter values for physical layer interfaces of Dense Wavelength Division Multiplexing (DWDM) systems primarily intended for metro applications which include optical amplifiers. Applications are defined in G.698.2 [ITU.G698.2] using optical interface parameters at the single-channel connection points between optical transmitters and the optical multiplexer, as well as between optical receivers and the optical demultiplexer in the DWDM system. This Recommendation uses a methodology which does not specify the details of the optical link, e.g. the maximum fibre length, explicitly. The Recommendation currently includes unidirectional DWDM applications at 2.5 and 10 Gbit/s (with 100 GHz and 50 GHz channel frequency spacing). Work is still underway for 40 and 100 Gbit/s interfaces. There is possibility for extensions to a lower channel frequency spacing.

This draft refers and supports also the
draft-kunze-g-698-2-management-control-framework

The building of an SNMP MIB describing the optical parameters defined in G.698.2 [ITU.G698.2] G.798 [ITU.G798], G.874 [ITU.G874], parameters specified G.7710/Y.1701 [ITU-T G.7710] allows the different vendors and operator to retrieve, provision and exchange information related to Optical black links in a standardized way. This facilitates interworking in case of using optical interfaces from different vendors at the end of the link.

The MIB, reporting the Optical parameters and their values, characterizes the features and the performances of the optical components and allow a reliable black link design in case of multivendor optical networks.

Although RFC 3591 [RFC3591] describes and defines the SNMP MIB of a number of key optical parameters, alarms and Performance Monitoring, a more complete description of optical parameters and processes can be found in the ITU-T Recommendations. Appendix A of this document provides an overview about the extensive ITU-T documentation in this area. The same considerations can be applied to the RFC 4054 [RFC4054]

2. The Internet-Standard Management Framework

For a detailed overview of the documents that describe the current Internet-Standard Management Framework, please refer to section 7 of RFC 3410 [RFC3410].

Managed objects are accessed via a virtual information store, termed the Management Information Base or MIB. MIB objects are generally accessed through the Simple Network Management Protocol (SNMP). Objects in the MIB are defined using the mechanisms defined in the Structure of Management Information (SMI). This memo specifies a MIB module that is compliant to the SMIv2, which is described in STD 58, RFC 2578 [RFC2578], STD 58, RFC 2579 [RFC2579] and STD 58, RFC 2580 [RFC2580].

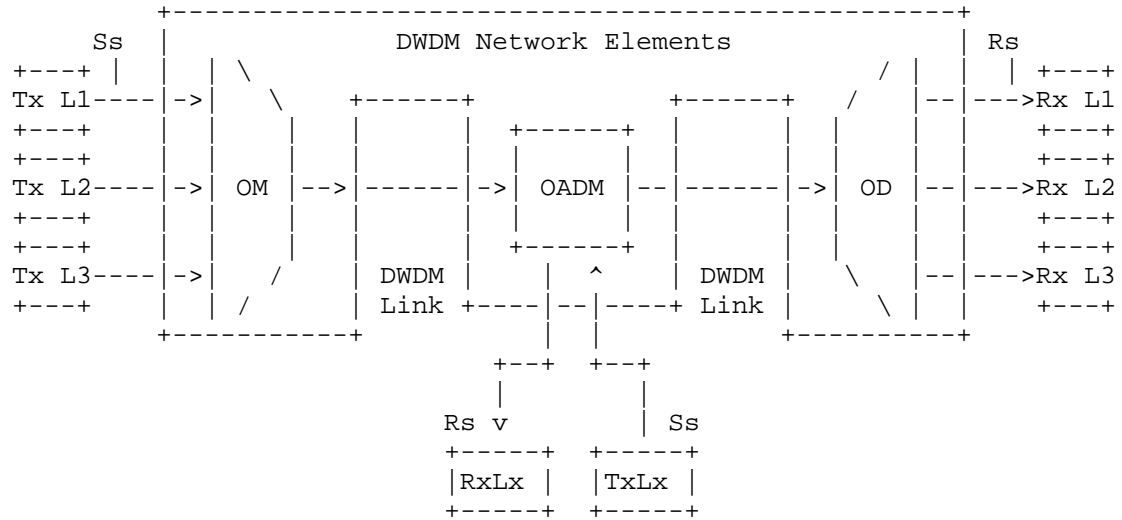
3. Conventions

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] In the description of OIDs the convention: Set (S) Get (G) and Trap (T) conventions will describe the action allowed by the parameter.

4. Overview

In this document, the term OTN (Optical Transport Network) system is used to describe devices that are compliant with the requirements specified in the ITU-T Recommendations G.872 [ITU.G872], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1] while refers to G.698.2 [ITU.G698.2] for the Black Link and DWDM parameter description.

Figure 1 shows a set of reference points, for the linear "black link" approach, for single-channel connection (Ss and Rs) between transmitters (Tx) and receivers (Rx). Here the DWDM network elements include an OM and an OD (which are used as a pair with the opposing element), one or more optical amplifiers and may also include one or more OADMs.



Ss = reference point at the DWDM network element tributary output
Rs = reference point at the DWDM network element tributary input
Lx = Lambda x
OM = Optical Mux
OD = Optical Demux
OADM = Optical Add Drop Mux

from Fig. 5.1/G.698.2

Figure 1: Linear Black Link

G.698.2 [ITU.G698.2] defines also Ring Black Link configurations [Fig. 5.2/G.698.2] and Bidirectional Black Link configurations [Fig. 5.3/G.698.2]

4.1. Optical Parameters Description

The black links are managed at the edges, i.e. at the transmitters (Tx) and receivers (Rx) attached to the S and R reference points respectively. The parameters that could be managed at the black link edges are specified in G.698.2 [ITU.G698.2] for the optical

interface, in G.798 [ITU.G798] for the equipment aspect, and in G.7710/Y.1701 [ITU.G7710] and G.874 [ITU.G874] for fault management and performance monitoring.

The definitions of the optical parameters are provided below to increase the readability of the document, where the definition is ended by (G) the parameter can be retrieve with a GET, when (S) it can be provisioned by a SET, (G,S) can be either GET and SET.

To support the management of these parameters, the SNMP MIB in RFC 3591 [RFC3591] is extended with a new MIB module defined in section 6 of this document. This new MIB module includes the definition of new configuration table of the OCh Layer for the parameters at Tx (S) and Rx (R).

4.1.1. General

The following general parameters from G.698.2 [ITU.G698.2] and G.694.1 [ITU.G694.1] provide general information at the optical interface reference points.

Minimum channel spacing:

This is the minimum nominal difference in frequency (in GHz) between two adjacent channels (G).

Bit rate/line coding of optical tributary signals:

Optical tributary signal class NRZ 2.5G (from nominally 622 Mbit/s to nominally 2.67 Gbit/s) or NRZ 10G nominally 2.4 Gbit/s to nominally 10.71 Gbit/s. (nominally 2.4 Gbit/s to nominally 10.71 Gbit/s). 40Gbit/s and 100Gbit/s are under study (G, S).

FEC Coding:

This parameter indicate what Forward Error Correction (FEC) code is used at Ss and Rs (G, S) (not mentioned in G.698). EDITOR NOTE: Need to check whether this parameter is to be put in "vendor specific" parameter or can be a standard parameter as defined in G.698.2. Is this the various adaptations (FEC encoding types) specified in G.798 clauses 12.3.1.1 (with FEC), 12.3.1.2 (without FEC), and 12.3.1.5 (vendor-specific FEC) .

Maximum bit error ratio (BER):

This parameter indicate the maximum Bit error rate can be supported by the application at the Receiver. In case of FEC applications it is intended after the FEC correction (G) .

Fiber type:

Fiber type as per fibre types are chosen from those defined in ITU-T Recs G.652, G.653, G.654 and G.655 (G,S) .

Wavelength Range (see G.694.1): [ITU.G694.1]

This parameter indicate minimum and maximum wavelength spectrum (G) in a definite wavelength Band (L, C and S).

Wavelength Value (see G.694.1 Table 1):

This parameter indicates the wavelength value that Ss and Rs will be set to work (in THz) se in particular Section 6/G.694.1 (G, S).

Vendor Transceiver Class:

Other than specifying all the Transceiver parameter, it might be convenient for the vendors to summarize a set of parameters in a single proprietary parameter: the Class of transceiver. The Transceiver classification will be based on the Vendor Name and the main TX and RX parameters (i.e. Trunk Mode, Framing, Bit rate, Trunk Type, Channel Band, Channel Grid, Modulation Format, Channel Modulation Format, FEC Coding, Electrical Signal Framing at Tx, Minimum maximum Chromatic Disperion (CD) at Rx, Maximum Polarization Mode Dispersion (PMD) at Rx, Maximum differential group delay at Rx, Loopbacks, TDC, Pre-FEC BER, Q-factor, Q-margin,etc.). If this parameter is used, the MIB parameters specifying the Transceiver characteristics may not be significant and the vendor will be responsible to specify the Class contents and values. The Vendor can publish the parameters of its Classes or declare to be compatible with published Classes.(G) Optional for compliance. (not mentioned in G.698)

Single-channel application codes (see G.698.2):

This parameter indicates the transceiver application code at Ss and Rs as defined in [ITU.G698.2] Chapter 5.4 - this parameter can be called Optical Interface Identifier OII as per [draft-martinelli-wson-interface-class] (G).

PARAMETERS	Get/Set	Reference
Minimum channel spacing	G	G.698.2 S.7.1.1
Bit rate/line coding of opt. trib. signals	G,S	G.698.2 S.7.1.2
FEC Coding	G,S	G.975
Maximum bit error ratio (BER)	G	G.698.2 S.7.1.3
Fiber type	G,S	G.698.2 S.7.1.4
Wavelength Range	G	G.694.1 S.6
Wavelength Value	G,S	G.694.1 S.6
Vendor Transceiver Class	G	N.A.
Single-channel application codes	G	G.698.2 S.5.3

Table 1: General parameters

4.1.2. Parameters at Ss

The following parameters for the interface at point S are defined in G.698.2 [ITU.G698.2].

Maximum and minimum mean channel output power:

The mean launched power at Ss is the average power (in dBm) of a pseudo-random data sequence coupled into the DWDM link. It is defined as the rrange (Max and Min) of the parameter (G, S)

Minimum and maximum central frequency:

The central frequency is the nominal single-channel frequency (in THz) on which the digital coded information of the particular optical channel is modulated by use of the NRZ line code. The central frequencies of all channels within an application lie on the frequency grid for the minimum channel spacing of the application given in ITU-T Rec. G.694.1. This parameter give the Maximum and minimum frequency interval the channel must be modulated (G)

Maximum spectral excursion:

This is the maximum acceptable difference between the nominal central frequency (in GHz) of the channel and the minus 15 dB points of the transmitter spectrum furthest from the nominal central frequency measured at point Ss. (G)

Maximum transmitter (residual) dispersion OSNR penalty (B.3/G.959.1) [ITU.G959.1]

Defines a reference receiver that this penalty is measured with. Lowest OSNR at Ss with worst case (residual) dispersion minus the Lowest OSNR at Ss with no dispersion. Lowest OSNR at Ss with no

dispersion (G)

Minimum side mode suppression ratio, Minimum channel extinction ratio, Eye mask:

Although are defined in G.698.2 are not supported by this draft (G).

Current Laser Output power:

This parameter report the current Transceiver Output power, it can be either a setting and measured value (G, S) NEED TO DISCUSS ON THIS.

PARAMETERS	Get/Set	Reference
MAX and min mean channel output power	G,S	G.698.2 S.7.2.1
Min and MAX central frequency	G	G.698.2 S.7.2.2
MAX spectral excursion	G	G.698.2 S.7.2.3
MAX transmitter (residual) disper.	G	G.698.2 S.7.2.7
OSNR penalty		
MAX side mode suppression ratio, min	G	G.698.2 S.7.2.6
channel extinction ratio, Eye mask		
Current Laser Output power	G,S	N.A.

Table 2: parameters at Ss

4.1.3. Optical path from point Ss to Rs

The following parameters for the optical path from point S and R are defined in G.698.2 [ITU.G698.2].

Maximum and minimum (residual) chromatic dispersion:

These parameters define the maximum and minimum value of the optical path "end to end chromatic dispersion" (in ps/nm) that the system shall be able to tolerate. (G)

Minimum optical return loss at Ss:

These parameter defines minimum optical return loss (in dB) of the cable plant at the source reference point (Ss), including any connectors (G)

Maximum discrete reflectance between Ss and Rs:

Optical reflectance is defined to be the ratio of the reflected optical power present at a point, to the optical power incident to that point. Control of reflections is discussed extensively in ITU-T Rec. G.957 (G)

Maximum differential group delay:

Differential group delay (DGD) is the time difference between the fractions of a pulse that are transmitted in the two principal states of polarization of an optical signal. For distances greater than several kilometres, and assuming random (strong) polarization mode coupling, DGD in a fibre can be statistically modelled as having a Maxwellian distribution. (G)

Maximum polarization dependent loss:

The polarization dependent loss (PDL) is the difference (in dB) between the maximum and minimum values of the channel insertion loss (or gain) of the black link from point SS to RS due to a variation of the state of polarization (SOP) over all SOPs. (G)

Maximum inter-channel crosstalk:

Inter-channel crosstalk is defined as the ratio of total power in all of the disturbing channels to that in the wanted channel, where the wanted and disturbing channels are at different wavelengths. The parameter specifies the isolation of a link conforming to the "black link" approach such that under the worst-case operating conditions the inter-channel crosstalk at any reference point RS is less than the maximum inter-channel crosstalk value (G)

Maximum interferometric crosstalk:

This parameter places a requirement on the isolation of a link conforming to the "black link" approach such that under the worst case operating conditions the interferometric crosstalk at any reference point RS is less than the maximum interferometric crosstalk value. (G)

Maximum optical path OSNR penalty:

The optical path OSNR penalty is defined as the difference between the Lowest OSNR at Rs and Lowest OSNR at Ss that meets the BER requirement (G)

Maximum ripple:

Although is defined in G.698.2, this parameter is not supported by this draft.

PARAMETERS	Get/Set	Reference
MAX and min (residual) chromatic dispersion	G	G.698.2 S.7.3.2
Min optical return loss at Ss	G	G.698.2 S.7.3.3
MAX discrete reflectance between Ss and Rs	G	G.698.2 S.7.3.4
MAX differential group delay	G	G.698.2 S.7.3.5
MAX polarization dependent loss	G	G.698.2 S.7.3.6
MAX inter-channel crosstalk	G	G.698.2 S.7.3.7
MAX interferometric crosstalk	G	G.698.2 S.7.3.8
MAX optical path OSNR penalty	G	G.698.2 S.7.3.9
MAX ripple	G	G.698.2 S.7.3.1

Table 3: parameters between Ss and Rs

4.1.4. Interface at point Rs

The following parameters for the interface at point R are defined in G.698.2.

4.1.4.1. Mandatory parameters

Maximum and minimum mean input power:

The maximum and minimum values of the average received power (in dBm) at point Rs. (G)

Minimum optical signal-to-noise ratio (OSNR):

The minimum optical signal-to-noise ratio (OSNR) is the minimum value of the ratio of the signal power in the wanted channel to the highest noise power density in the range of the central frequency plus and minus the maximum spectral excursion (G)

Receiver OSNR tolerance:

The receiver OSNR tolerance is defined as the minimum value of OSNR at point Rs that can be tolerated while maintaining the maximum BER of the application. (G)

Maximum reflectance at receiver:

Although is defined in G.698.2, this parameter is not supported by this draft (G).

PARAMETERS	Get/Set	Reference
MAX and min mean input power	G	G.698.2 S.7.4.1
Min optical signal-to-noise ratio (OSNR)	G	G.698.2 S.7.4.2
Receiver OSNR tolerance	G	G.698.2 S.7.4.3
MAX reflectance at receiver	G	G.698.2 S.7.4.4

Table 4: mandatory parameters

4.1.4.2. Optional parameters

Current Chromatic Dispersion (CD):

Residual Chromatic Dispersion measuread at Rx Transceiver port (G).

Current Optical Signal to Noise Ratio (OSNR):

Current Optical Signal to Noise Ratio (OSNR) estimated at Rx Transceiver port (G).

Current Quality factor (Q):

"Q" factor estimated at Rx Transceiver port (G).

PARAMETERS	Get/Set	Reference
Current Chromatic Dispersion (CD)	G	N.A.
Current Opt. Signal to Noise Ratio (OSNR)	G	N.A.
Current Quality factor (Q)	G	N.A.

Table 5: optional parameters

4.1.5. Alarms and Threshold definition

This section describes the Alarms and the Thresholds at Ss and Rs points according to ITU-T Recommendations G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1].

OTN alarms defined in RFC3591:

Threshold Crossing Alert (TCA Alarm)

LOW-TXPOWER

HIGH-TXPOWER

LOW-RXPOWER

HIGH-RXPOWER

Loss of Signal (LOS)

Loss of Frame (LOF)

Server Signal Failure-P (SSF-P)

Loss of Multiframe (LOM)

OTN Thresholds (for TCA) defined in RFC3591

LOW-TXPOWER

HIGH-TXPOWER

LOW-RXPOWER

HIGH-RXPOWER

As the above parameters/alarms are already defined in RFC3591, they are out of scope of this document and the RFC3591 will continue to be the only reference for them

The list below reports the new Alarms and Thresholds not managed in RFC3591

4.1.6. Performance Monitoring (PM) description

This section describes the Performance Monitoring parameters and their thresholds at Ss and Rs points (Near -End and Far-End) according to ITU-T Recommendations G.826 [ITU.G826], G.8201 [ITU.G8201], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1].

Failure Counts (fc) :

Number of Failures occurred in an observation period (G)

Errored Second (es) :

It is a one-second period in which there is one or more errored blocks or during which a defect (e.g. Loss of Signal (LOS)) is detected. The number of errored seconds is summed over 15-minute and 24-hour intervals. (G)

Severely Errored Seconds (ses) :

It is a one-second period in which the errored block ratio exceeds a threshold or during which a defect is detected. See ITU-T Recommendation G.8201 Table 7-1 for details. The number of severely errored seconds is summed over 15-minute and 24-hour intervals.(G)

Unavailable Seconds (uas) :

It is a one-second period in the unavailable time. A period of unavailable time begins at the onset of ten consecutive SES events. These ten seconds are considered to be part of unavailable time. A new period of available time begins at the onset of ten consecutive non-SES events. These ten seconds are considered to be part of available time. (G)

Background Block Errors (bbe) :

An errored block not occurring as part of an SES(G)

Error Seconds Ratio (esr) :

The ratio of ES in available time to total seconds in available time during a fixed measurement interval(G)

Severely Errored Seconds Ratio (sesr) :

The ratio of SES in available time to total seconds in available time during a fixed measurement interval(G)

Background Block Errored Seconds Ratio (bber) :

The ratio of Background Block Errors (BBE) to total blocks in available time during a fixed measurement interval. The count of total blocks excludes all blocks during SESS.(G)

FEC corrected Bit Error (FECcorrErr):

The number of bits corrected by the FEC are counted over one second (G)

FEC un-corrected Bit Error :

The number of bits un-corrected by the FEC are counted over one second (G)

Pre-FEC Bit Error :

The number of Errored bits at receiving side before the FEC function counted over one second (G)

OTN Valid Intervals :

The number of contiguous 15 minute intervals for which valid OTN performance monitoring data is available for the particular interface (G)

FEC Valid Intervals :

The number of contiguous 15 minute intervals for which valid FEC PM data is available for the particular interface.(G)

4.1.7. Generic Parameter description

This section describes the Generic Parameters at Ss and Rs points according to ITU-T Recommendations G.872 [ITU.G872], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1].

Interface Admin Status :

The Administrative Status of an Interface: Up/Down - In Service/Out of Service (can be Automatic in Service) (G/S)

Interface Operational Status :

The Operational Status of an Interface: Up/Down - In Service/Out of Service (G)

4.2. Use of ifTable

This section specifies how the MIB II interfaces group, as defined in RFC 2863 [RFC2863], is used for the link ends of a black link. Only the ifGeneralInformationGroup will be supported for the ifTable and the ifStackTable to maintain the relationship between the OCh and OPS layers. The OCh and OPS layers are managed in the ifTable using IfEntries that correlate to the layers depicted in Figure 1.

For example, a device with TX and/or RX will have an Optical Physical Section (OPS) layer, and an Optical Channel (OCh) layer. There is a one to n relationship between the OPS and OCh layers.

EDITOR NOTE: Reason for changing from OChr to OCh: Work on revised G.872 in the SG15 December 2011 meeting agreed to remove OChr from the architecture and to update G.709 to account for this architectural change. The meeting also agreed to consent the revised text of G.872 and G.709 at the September 2012 SG15 meeting.

Figure 2 In the following figures, opticalChannel and opticalPhysicalSection are abbreviated as och and ops respectively.

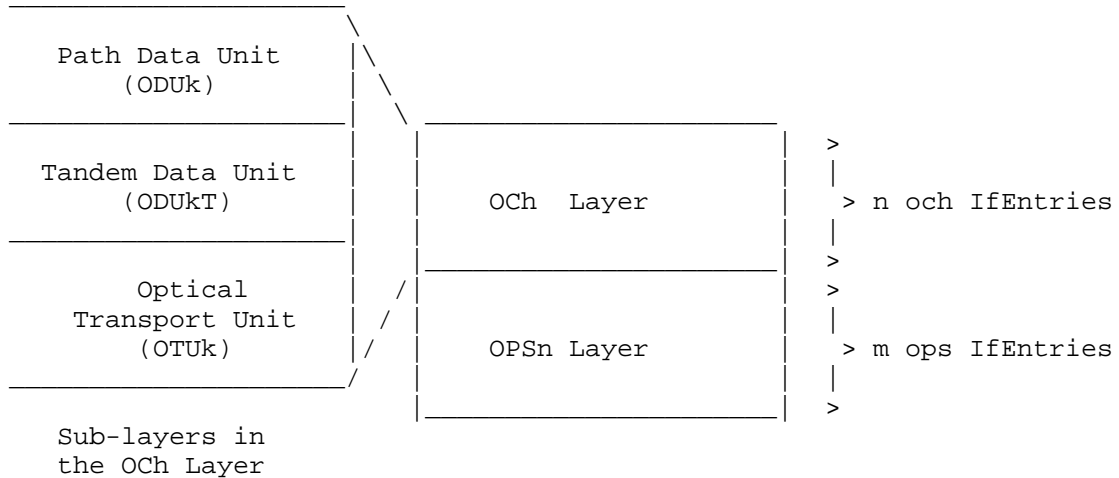


Figure 2: OTN Layers for OPS and OCh

Each opticalChannel IfEntry is mapped to one of the m opticalPhysicalSection IfEntries, where m is greater than or equal to 1. Conversely, each opticalTransPhysicalSection port entry is mapped to one of the n opticalChannel IfEntries, where n is greater than or equal to 1.

The design of the Optical Interface MIB provides the option to model an interface either as a single bidirectional object containing both sink and source functions or as a pair of unidirectional objects, one containing sink functions and the other containing source functions.

If the sink and source for a given protocol layer are to be modelled as separate objects, then there need to be two ifTable entries, one that corresponds to the sink and one that corresponds to the source, where the directionality information is provided in the configuration tables for that layer via the associated Directionality objects. The agent is expected to maintain consistent directionality values between ifStackTable layers (e.g., a sink must not be stacked in a 1:1 manner on top of a source, or vice-versa), and all protocol layers that are represented by a given ifTable entry are expected to have the same directionality.

When separate ifTable entries are used for the source and sink functions of a given physical interface, association between the two uni-directional ifTable entries (one for the source function and the other for the sink functions) should be provided. It is recommended that identical ifName values are used for the two ifTable entries to indicate such association. An implementation shall explicitly state what mechanism is used to indicate the association, if ifName is not used.

4.2.1. Use of ifTable for OPS Layer

Only the ifGeneralInformationGroup needs to be supported.

ifTable Object	Use for OTN OPS Layer
=====	
ifIndex	The interface index.
ifDescr	Optical Transport Network (OTN) Optical Physical Section (OPS)
ifType	opticalPhysicalSection (xxx)
<<<Editor Note: Need new IANA registration value for xxx. >>>	
ifSpeed	Actual bandwidth of the interface in bits per second. If the bandwidth of the interface is greater than the maximum value of 4,294,967,295, then the maximum value is reported and ifHighSpeed must be used to report the interface's speed.
ifPhysAddress	An octet string with zero length. (There is no specific address associated with the interface.)
ifAdminStatus	The desired administrative state of the interface. Supports read-only access.
ifOperStatus	The operational state of the interface. The value lowerLayerDown(7) is not used, since there is no lower layer interface. This object is set to notPresent(6) if a component is missing, otherwise it is set to down(2) if either of the objects optIfOPSnCurrentStatus indicates that any defect is present.

ifLastChange	The value of sysUpTime at the last change in ifOperStatus.
ifName	Enterprise-specific convention (e.g., TL-1 AID) to identify the physical or data entity associated with this interface or an OCTET STRING of zero length. The enterprise-specific convention is intended to provide the means to reference one or more enterprise-specific tables.
ifLinkUpDownTrapEnable	Default value is enabled(1). Supports read-only access.
ifHighSpeed	Actual bandwidth of the interface in Mega-bits per second. A value of n represents a range of 'n-0.5' to 'n+0.499999'.
ifConnectorPresent	Set to true(1).
ifAlias	The (non-volatile) alias name for this interface as assigned by the network manager.

4.2.2. Use of ifTable for OCh Layer

Use of ifTable for OCh Layer See RFC 3591 [RFC3591] section 2.4

4.2.3. Use of ifStackTable

Use of the ifStackTable and ifInvStackTable to associate the opticalPhysicalSection and opticalChannel interface entries is best illustrated by the example shown in Figure 3. The example assumes an ops interface with ifIndex i that carries two multiplexed och interfaces with ifIndex values of j and k, respectively. The example shows that j and k are stacked above (i.e., multiplexed into) i. Furthermore, it shows that there is no layer lower than i and no layer higher than j and/or k.

Figure 3

HigherLayer	LowerLayer
0	j
0	k
j	i
k	i
i	0

Figure 3: Use of ifStackTable for an OTN port

For the inverse stack table, it provides the same information as the interface stack table, with the order of the Higher and Lower layer interfaces reversed.

5. Structure of the MIB Module

EDITOR NOTE: text will be provided based on the MIB module in Section 6

6. Object Definitions

EDITOR NOTE: Once the scope in Section 1 and the parameters in Section 4 are finalized, a MIB module will be defined. It could be an extension to the OPT-IF-MIB module of RFC 3591. >>>

OPT-IF-EXT-MIB DEFINITIONS ::= BEGIN

IMPORTS

MODULE-IDENTITY,
OBJECT-TYPE,
Gauge32,
Integer32,
Unsigned32,
transmission,
NOTIFICATION-TYPE
FROM SNMPv2-SMI
TEXTUAL-CONVENTION,
RowPointer,
RowStatus,
TruthValue,
DateAndTime
FROM SNMPv2-TC
SnmAdminString
FROM SNMP-FRAMEWORK-MIB
MODULE-COMPLIANCE, OBJECT-GROUP
FROM SNMPv2-CONF
ifIndex
FROM IF-MIB
optIfMibModule,
optIfOChConfigEntry,
optIfOChSinkCurrentEntry,
OptIfDirectionality
FROM OPT-IF-MIB;

-- This is the MIB module for the optical parameters associated with
-- the black link end points.

optIfExtMibModule MODULE-IDENTITY
LAST-UPDATED "201204250000Z"
ORGANIZATION "IETF OPSAWG/CCAMP Working Group"
CONTACT-INFO
"WG charter:
http://www.ietf.org/html.charters/

Mailing Lists:
Editor: Gabrielle Galimberti
Email: ggalimbe@cisco.com"

DESCRIPTION

"The MIB module to describe Black Link extension to rfc3591.

Copyright (C) The Internet Society (2012). This version
of this MIB module is part of ; see the RFC
itself for full legal notices."

REVISION "201204250000Z"

DESCRIPTION

"Draft version 1.0"

::={ optIfMibModule 3 }

OptIfChannelSpacing ::= TEXTUAL-CONVENTION

STATUS current

DESCRIPTION

"Channel spacing

- 1 - 100 Ghz
- 2 - 50GHz
- 3 - 25GHz
- 4 - 12.5GHz
- 5 - 6.25Ghz

"

SYNTAX INTEGER {
 spacing100Ghz(1),
 spacing50Ghz(2),
 spacing25Ghz(3),
 spacing12point5Ghz(4),
 spacing6point25Ghz(5)
}

OptIfBitRateLineCoding ::= TEXTUAL-CONVENTION

STATUS current

DESCRIPTION

"Optical tributary signal class

- 1 - NRZ 2.5G (from nominally 622 Mbit/s to nominal 2.67 Gbit/s)
- 2 - NRZ 10G nominally 2.4 Gbit/s to nominally 10.71 Gbit/s.
- 3 - 40Gbits/s
- 4 - 100Gbits/s
- 5 - 400Gbits/s

40Gbits/s and above are under study. "

SYNTAX INTEGER {
 rate2point5G(1),
 rate10G(2),
 rate40G(3),
 rate100G(4),
 rate400G(5)
}

```
OptIfFiberTypeRecommendation ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        " Fiber Types - ITU-T Recs G.652, G.653, G.654 and G.655
          One for recommendation and one for category.
          G.652 A, B, C, D
          G.653 A, B
          G.654 A, B, C
          G.655 C, D, E
          G.656
          G.657 A, B
        "
    SYNTAX INTEGER {
        g652(1),
        g653(2),
        g654(3),
        g655(4),
        g656(5),
        g657(6)
    }
```

```
OptIfFiberTypeCategory ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        " Fiber Types - ITU-T Recs G.652, G.653, G.654 and G.655
          G.652 A, B, C, D
          G.653 A, B
          G.654 A, B, C
          G.655 C, D, E
          G.656
          G.657 A, B
          Categories - A, B, C, D and E
        "
    SYNTAX INTEGER {
        categoryA(1),
        categoryB(2),
        categoryC(3),
        categoryD(4),
        categoryE(5)
    }
```

```
OptIfPerformanceDataType ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        "
          This parameter indicates the parameters for the table are for
          the Near End or Far End performance data.
          1 - Near End
```

```

        2 - Far End
    "
    SYNTAX INTEGER {
        nearEnd(1),
        farEnd(2)
    }
}

OptIfOTNLayer ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        "
        This parameter indicates the parameters for the table are for OTUk,
        ODUk, TCM performance data.
        1 - OTUk
        2 - ODUk
        3 - TCM
        The ODUk layer and TCM sublayer PM is not related to the black
        link PM management, but since this could be a common PM model for
        the ODUk layer and TCM sublayers, they are included here so it may
        be used for simple scenarios where only lower order ODUk or higher
        order ODUk is present. For scenarios where both lower order ODUk
        and higher order ODUk are present, further extension to the MIB
        model is required, in particular for the indexing for these layers.
        "
    SYNTAX INTEGER {
        optIfOTUkLayer(1),
        optIfODUkLayer(2),
        optIfTCMSubLayer(3)
    }

--
-- Alarm for the OCh and OTUk layer
--
OptIfOTNOChAlarms ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        " This is the possible alarms from the OCh and OTUk layer."
    SYNTAX INTEGER {
        -- OTN Loss of signal alarm
        optIfOtnLosAlarm(1),
        -- OTN Loss of frame alarm
        optIfOtnLofAlarm(2),
        -- OTN Loss of multi framealarm
        optIfOtnLomAlarm(3),
        -- OTN SSF alarm
        optIfOtnOtuSsfAlarm(4),
        -- OTN OTU BDI alarm
        optIfOtnOtuBdiAlarm(5),
    }
```

```

    -- OTN OTU Trail Trace mismatch alarm
    optIfOtnOtuTimAlarm(6),
    -- OTN OTU IAE alarm
    optIfOtnOtuIaeAlarm(7),
    -- OTN OTU Degraded alarm,
    optIfOtnOtuDegAlarm(8),
    -- OTN OTU Fec ExcessiveErrors alarm
    optIfOptIfOtnOtuFecExcessiveErrsAlarm(9),
    -- OTN OTU BBE Thresholdalarm
    optIf15MinThreshBBETCA(10),
    -- OTN OTU ES Thresholdalarm
    optIf15MinThreshESTCA(11),
    -- OTN OTU SES Threshold alarm
    optIf15MinThreshSESTCA(12),
    -- OTN OTU UAS Threshold alarm
    optIf15MinThreshUASTCA(13),
    -- OTN OTU Fcs Thresholdalarm alarm
    optIf15MinThreshFcstCA(14),
    -- OTN FEC uncorrectedwords TCA
    optIf15MinThreshFECUnCorrectedWordsTCA(15),
    -- OTN Pre FEC BER TCA
    optIf15MinThreshPreFECBERTCA(16)
}

OptIfOTNODUKTcmAlarms ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        " This is the alarms from the ODUk and TCM layer."
    SYNTAX INTEGER {
        -- OTN ODU/TCM OCI alarm
        optIfOTNodukTcmOciAlarm(1),
        -- OTN ODU/TCM LCK alarm
        optIfOTNodukTcmLckAlarm(2),
        -- OTN ODU/TCM BDI alarm
        optIfOTNodukTcmBdiAlarm(3),
        -- OTN ODU/TCM Trail Trace mismatch alarm
        optIfOTNodukTcmTimAlarm(4),
        -- OTN ODU/TCM Degraded alarm,
        optIfOTNodukTcmDegAlarm(5),
        -- OTN ODU/TCM SSF alarm,
        optIfOTNodukTcmSSfAlarm(6),
        -- OTN OTU BBE Threshold alarm
        optIfOTNodukTcm15MinThreshBBETCA(7),
        -- OTN OTU ES Threshold alarm
        optIfOTNodukTcm15MinThreshESTCA(8),
        -- OTN OTU SES Threshold alarm
        optIfOTNodukTcm15MinThreshSESTCA(9),
        -- OTN OTU UAS Threshold alarm

```

```

        optIfOTNOduktCm15MinThreshUASTCA(10),
        -- OTN OTU Fcs Threshold alarm
        optIfOTNOduktCm15MinThreshFcsTCA(11)
    }

-- Addition to the RFC 3591 objects
optIfOPSmEntry          OBJECT IDENTIFIER ::= { optIfExtMibModule 1 }
optIfOChSrcSinkGroup    OBJECT IDENTIFIER ::= { optIfExtMibModule 2 }
optIfOTNPMGroup         OBJECT IDENTIFIER ::= { optIfExtMibModule 3 }
optIfOTNAlarm           OBJECT IDENTIFIER ::= { optIfExtMibModule 4 }
optIfOTNNotifications   OBJECT IDENTIFIER ::= { optIfExtMibModule 5 }


-- OPS - Optical Physical Section
optIfOPSmConfigTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOPSmConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A table of OPS General config parameters."
    ::= { optIfObjects 10 }

optIfOPSmConfigEntry OBJECT-TYPE
    SYNTAX OptIfOPSmConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "An conceptual row of OPS General config parameters."
    INDEX { ifIndex }
    ::= { optIfOPSmConfigTable 1 }

    OptIfOPSmConfigEntry ::=
        SEQUENCE {
            optIfOPSmDirectionality          OptIfDirectionality,
            optIfOPSmFiberTypeRecommendation OptIfFiberTypeRecommendation,
            optIfOPSmFiberTypeCategory        OptIfFiberTypeCategory
        }

optIfOPSmDirectionality OBJECT-TYPE
    SYNTAX OptIfDirectionality
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "Indicates the directionality of the entity."
```



```
 ::= { optIfOPSmConfigEntry 1 }

optIfOPSmFiberTypeRecommendation OBJECT-TYPE
    SYNTAX      OptIfFiberTypeRecommendation
    MAX-ACCESS   read-write
    STATUS      current
    DESCRIPTION
        "Fiber type as per fibre types are chosen from those defined in
        ITU-T Recs G.652, G.653, G.654, G.655, G.656 and G.657."
    ::= { optIfOPSmConfigEntry 2 }

optIfOPSmFiberTypeCategory OBJECT-TYPE
    SYNTAX      OptIfFiberTypeCategory
    MAX-ACCESS   read-write
    STATUS      current
    DESCRIPTION
        "Fiber type as per fibre types are chosen from those defined in
        ITU-T Recs G.652, G.653, and G.655.
        The categories are A, B, C, D and E."
    ::= { optIfOPSmConfigEntry 3 }

-- OCh config table
-- modified the OCh Table group
-- General parameters for the Black Link Ss-Rs will be added to
-- the OchConfigTable

optIfOChConfigExtTable OBJECT-TYPE
    SYNTAX      SEQUENCE OF OptIfOChConfigExtEntry
    MAX-ACCESS   not-accessible
    STATUS      current
    DESCRIPTION
        "A table of OCh General config extension parameters"
    ::= { optIfOChSrcSinkGroup 1 }

optIfOChConfigExtEntry OBJECT-TYPE
    SYNTAX      OptIfOChConfigExtEntry
    MAX-ACCESS   not-accessible
    STATUS      current
    DESCRIPTION
        " A conceptual row that contains OCh configuration extension
        information of an interface. "
    AUGMENTS { optIfOChConfigEntry }
    ::= { optIfOChConfigExtTable 1 }

OptIfOChConfigExtEntry ::=
    SEQUENCE {
```

optIfOChMimumChannelSpacing	OptIfChannelSpacing,
optIfOChBitRateLineCoding	OptIfBitRateLineCoding,
optIfOChFEC	Unsigned32,
optIfOChSinkMaximumBERMantisa	Unsigned32,
optIfOChSinkMaximumBERExponent	Unsigned32,
optIfOChMinWavelength	Unsigned32,
optIfOChMaxWavelength	Unsigned32,
optIfOChWavelength	Unsigned32,
optIfOChVendorTransceiverClass	OCTET STRING,
optIfOChOpticalInterfaceApplicationCode	OCTET STRING,
optIfOChLaserAdminState	INTEGER,
optIfOChLaserOperationalState	INTEGER,
optIfOChAdminState	INTEGER,
optIfOChOperationalState	INTEGER

}

optIfOChMimumChannelSpacing OBJECT-TYPE

SYNTAX OptIfChannelSpacing

UNITS "Gigahertz"

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"A minimum nominal difference in frequency (GHz) between two adjacent channels."

::= { optIfOChConfigExtEntry 1 }

optIfOChBitRateLineCoding OBJECT-TYPE

SYNTAX OptIfBitRateLineCoding

MAX-ACCESS read-write

STATUS current

DESCRIPTION

" Optical tributary signal class

NRZ 2.5G (from nominally 622 Mbit/s to nominally 2.67 Gbit/s)

NRZ 10G (nominally 2.4 Gbit/s to nominally 10.71 Gbit/s)

"

::= { optIfOChConfigExtEntry 2 }

optIfOChFEC OBJECT-TYPE

SYNTAX Unsigned32

MAX-ACCESS read-write

STATUS current

DESCRIPTION

" This parameter indicates what Forward Error Correction (FEC) code is used at Source and Sink.

GFEC (from G709) and the I.x EFEC's

(G.975 - Table I.1 super FEC).

1 - No FEC

```

    2 - GFEC
    3 - I.2 EFEC
    4 - I.3 EFEC
    5 - I.4 EFEC
    6 - I.5 EFEC
    7 - I.6 EFEC
    8 - I.7 EFEC
    9 - I.8 EFEC
   10 - I.9 EFEC
   99 - Vendor Specific
"
 ::= { optIfOChConfigExtEntry 3 }

optIfOChSinkMaximumBERMantisa OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"   This parameter indicate the maximum Bit(mantisa) error rate
    can be supported by the application at the Receiver. In case
    of FEC applications it is intended after the FEC correction.
"
 ::= { optIfOChConfigExtEntry 4 }

optIfOChSinkMaximumBERExponent OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"   This parameter indicate the maximum Bit(exponent) error rate
    can be supported by the application at the Receiver. In case
    of FEC applications it is intended after the FEC correction.
"
 ::= { optIfOChConfigExtEntry 5 }

optIfOChMinWavelength OBJECT-TYPE
SYNTAX Unsigned32
UNITS "0.01 Ghz"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    This parameter indicate minimum wavelength spectrum in a
    definite wavelength Band (L, C and S).
"
 ::= { optIfOChConfigExtEntry 6 }
```

```
optIfOchMaxWavelength OBJECT-TYPE
    SYNTAX Unsigned32
    UNITS "hertz"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            This parameter indicate maximum wavelength spectrum in a
            definite wavelength Band (L, C and S)
        "
    ::= { optIfOchConfigExtEntry 7 }

optIfOchWavelength OBJECT-TYPE
    SYNTAX Unsigned32
    UNITS "hertz"
    MAX-ACCESS read-write
    STATUS current
    DESCRIPTION
        "
            This parameter indicates the wavelength value in Hertz
            Specified in Table 1 of G.694.1 e.g. 195.8875
        "
    ::= { optIfOchConfigExtEntry 8 }

optIfOchVendorTransceiverClass OBJECT-TYPE
    SYNTAX OCTET STRING
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            As defined in G.698
            Vendors can summarize a set of parameters in a
            single proprietary parameter: the Class of transceiver. The
            Transceiver classification will be based on the Vendor Name and
            the main TX and RX parameters (i.e. Trunk Mode, Framing, Bit
            rate, Trunk Type etc).
            If this parameter is used, the MIB parameters
            specifying the Transceiver characteristics may not be
            significant and the vendor will be responsible to specify the
            Class contents and values. The Vendor can publish the
            parameters of its Classes or declare to be compatible with
            published Classes.(G) Optional for compliance. (not
            mentioned in G.698)
        "
    ::= { optIfOchConfigExtEntry 9 }

optIfOchOpticalInterfaceApplicationCode OBJECT-TYPE
    SYNTAX OCTET STRING
    MAX-ACCESS read-write
    STATUS current
```

```
DESCRIPTION
"   This parameter indicates the transceiver application code at Ss
   and Rs as defined in [ITU.G698.2] Chapter 5.3
"
 ::= { optIfOChConfigExtEntry 10 }

optIfOChLaserAdminState OBJECT-TYPE
SYNTAX  INTEGER {
    off(0),
    on(1),
    autoInService(2)
}
MAX-ACCESS  read-write
STATUS  current
DESCRIPTION
"
    The configured State of the laser: 0 - Off
    1 - On
    2 - Automatic - Inservice
"
 ::= { optIfOChConfigExtEntry 11 }

optIfOChLaserOperationalState OBJECT-TYPE
SYNTAX  INTEGER {
    off(0),
    on(1)
}
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"
    The Operational Status of Laser : 0 - Off
    1 - On
"
 ::= { optIfOChConfigExtEntry 12 }

optIfOChAdminState OBJECT-TYPE
SYNTAX  INTEGER {
    off(0),
    on(1),
    autoInService(2)
}
MAX-ACCESS  read-write
STATUS  current
DESCRIPTION
"
    The Administrative Status of an Interface:
    0 - Out of Service
```

```

        1 - In Service
        2 - Automatic in Service.
    "
 ::= { optIfOChConfigExtEntry 13 }

optIfOChOperationalState OBJECT-TYPE
    SYNTAX      INTEGER {
                        off(0),
                        on(1)
                    }
    MAX-ACCESS   read-only
    STATUS       current
    DESCRIPTION
        "
            The Operational Status of an Interface:
            0 - Off
            1 - On
        "
 ::= { optIfOChConfigExtEntry 14 }

-- Parameters at OCh Src (Ss)
-- OptIfOChSrcConfigEntry

optIfOChSrcConfigTable OBJECT-TYPE
    SYNTAX      SEQUENCE OF OptIfOChSrcConfigEntry
    MAX-ACCESS   not-accessible
    STATUS       current
    DESCRIPTION
        "A configuration table of OCh Src (Ss) parameters."
 ::= { optIfOChSrcSinkGroup 2 }

optIfOChSrcConfigEntry OBJECT-TYPE
    SYNTAX      OptIfOChSrcConfigEntry
    MAX-ACCESS   not-accessible
    STATUS       current
    DESCRIPTION
        " A conceptual row that contains the Src (Ss) configuration
          parameters for a given interface."
    INDEX       { ifIndex }
 ::= { optIfOChSrcConfigTable 1 }

OptIfOChSrcConfigEntry ::=
    SEQUENCE {
        optIfOChMinimumMeanChannelOutputPower      Integer32,
        optIfOChMaximumMeanChannelOutputPower       Integer32,
        optIfOChMinimumCentralFrequency             Unsigned32,

```

```

        optIfOchMaximumCentralFrequency          Unsigned32,
        optIfOchMaximumSpectralExcursion         Unsigned32,
        optIfOchMaximumTxDispersionOSNRPenalty   Integer32
    }

optIfOchMinimumMeanChannelOutputPower OBJECT-TYPE
    SYNTAX      Integer32
    UNITS       "0.1 dbm"
    MAX-ACCESS  read-write
    STATUS      current
    DESCRIPTION
        "
        The minimum mean launched power at Ss is the average power (in dBm)
        of a pseudo-random data sequence coupled into the DWDM link.
        "
    ::= { optIfOchSrcConfigEntry 1}

optIfOchMaximumMeanChannelOutputPower OBJECT-TYPE
    SYNTAX      Integer32
    UNITS       "0.1 dbm"
    MAX-ACCESS  read-write
    STATUS      current
    DESCRIPTION
        "
        The maximum mean launched power at Ss is the average power (in dBm)
        of a pseudo-random data sequence coupled into the DWDM link.
        "
    ::= { optIfOchSrcConfigEntry 2}

optIfOchMinimumCentralFrequency OBJECT-TYPE
    SYNTAX      Unsigned32
    UNITS       "0.0001 THz"
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "
        The minimum central frequency is the nominal single-channel
        frequency (in THz) on which the digital coded information of
        the particular optical channel is modulated by use of the NRZ
        line code. Eg 191.5THz will be represented as 1915000
        "
    ::= { optIfOchSrcConfigEntry 3}

optIfOchMaximumCentralFrequency OBJECT-TYPE
    SYNTAX      Unsigned32
    UNITS       "0.0001 THz"
    MAX-ACCESS  read-only
    STATUS      current
```

DESCRIPTION

"
The maximum central frequency is the nominal single-channel frequency (in THz) on which the digital coded information of the particular optical channel is modulated by use of the NRZ line code. Eg 191.5THz will be represented as 1915000.
"

::= { optIfOChSrcConfigEntry 4}

optIfOChMaximumSpectralExcursion OBJECT-TYPE

SYNTAX Unsigned32

UNITS "0.1 GHz"

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"
This is the maximum acceptable difference between the nominal central frequency (in GHz) of the channel and the minus 15 dB points of the transmitter spectrum furthest from the nominal central frequency measured at point Ss.
"

::= { optIfOChSrcConfigEntry 5}

optIfOChMaximumTxDispersionOSNRPenalty OBJECT-TYPE

SYNTAX Integer32

UNITS "0.1 dB"

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"
Defines a reference receiver that this penalty is measured with. Lowest OSNR at Ss with worst case (residual) dispersion minus the Lowest OSNR at Ss with no dispersion. Lowest OSNR at Ss with no dispersion
"

::= { optIfOChSrcConfigEntry 6}

-- Optical Path from Point Src (Ss) to Sink (Rs)

-- Alternatively this can be optIfOChSsRsTable

optIfOChSrcSinkConfigTable OBJECT-TYPE

SYNTAX SEQUENCE OF OptIfOChSrcSinkConfigEntry

MAX-ACCESS not-accessible

STATUS current

DESCRIPTION

"A table of paramters for the optical path from Src to Sink (Ss to Rs)."


```

 ::= { optIfOChSrcSinkGroup 3 }

optIfOChSrcSinkConfigEntry OBJECT-TYPE
    SYNTAX      OptIfOChSrcSinkConfigEntry
    MAX-ACCESS  not-accessible
    STATUS      current
    DESCRIPTION
        "A conceptual row that contains the optical path Src-Sink (Ss-Rs)
        configuration parameters for a given interface."
    INDEX       { ifIndex }
    ::= { optIfOChSrcSinkConfigTable 1 }

OptIfOChSrcSinkConfigEntry ::=
    SEQUENCE {
        optIfOChSrcSinkMinimumChromaticDispersion      Integer32,
        optIfOChSrcSinkMaximumChromaticDispersion      Integer32,
        optIfOChSrcSinkMinimumSrcOpticalReturnLoss      Integer32,
        optIfOChSrcSinkMaximumDiscreteReflectanceSrcToSink Integer32,
        optIfOChSrcSinkMaximumDifferentialGroupDelay    Integer32,
        optIfOChSrcSinkMaximumPolarizationDependentLoss Integer32,
        optIfOChSrcSinkMaximumInterChannelCrosstalk     Integer32,
        optIfOChSrcSinkMaximumInterFerometricCrosstalk  Integer32,
        optIfOChSrcSinkMaximumOpticalPathOSNRPenalty    Integer32
    }

optIfOChSrcSinkMinimumChromaticDispersion OBJECT-TYPE
    SYNTAX      Integer32
    UNITS        "ps/nm"
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "
        These parameters define the minimum value of the
        optical path 'end to end chromatic dispersion' (in ps/nm) that the
        system shall be able to tolerate."
    ::= { optIfOChSrcSinkConfigEntry 1}

optIfOChSrcSinkMaximumChromaticDispersion OBJECT-TYPE
    SYNTAX      Integer32
    UNITS        "ps/nm"
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "
        These parameters define the maximum value of the
        optical path 'end to end chromatic dispersion' (in ps/nm) that the
        system shall be able to tolerate."
    ::= { optIfOChSrcSinkConfigEntry 2 }

```

optIfOChSrcSinkMinimumSrcOpticalReturnLoss OBJECT-TYPE
SYNTAX Integer32
UNITS ".1 db"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
These parameter defines minimum optical return loss (in dB) of the
cable plant at the source reference point (Src/Ss), including any
connectors."
::= { optIfOChSrcSinkConfigEntry 3 }

optIfOChSrcSinkMaximumDiscreteReflectanceSrcToSink OBJECT-TYPE
SYNTAX Integer32
UNITS ".1 db"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
Optical reflectance is defined to be the ratio of the reflected
optical power present at a point, to the optical power incident
to that point. Control of reflections is discussed extensively in
ITU-T Rec. G.957."
::= { optIfOChSrcSinkConfigEntry 4 }

optIfOChSrcSinkMaximumDifferentialGroupDelay OBJECT-TYPE
SYNTAX Integer32
UNITS "ps"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
Differential group delay (DGD) is the time difference between the
fractions of a pulse that are transmitted in the two principal
states of polarization of an optical signal. For distances
greater than several kilometres, and assuming random (strong)
polarization mode coupling, DGD in a fibre can be statistically
modelled as having a Maxwellian distribution."
::= { optIfOChSrcSinkConfigEntry 5 }

optIfOChSrcSinkMaximumPolarizationDependentLoss OBJECT-TYPE
SYNTAX Integer32
UNITS "0.1 db"
MAX-ACCESS read-only
STATUS current
DESCRIPTION

```
"
    The polarization dependent loss (PDL) is the difference (in dB)
    between the maximum and minimum values of the channel insertion
    loss (or gain) of the black link from point SS to RS due to a
    variation of the state of polarization (SOP) over all SOPs."
::= { optIfOChSrcSinkConfigEntry 6}

optIfOChSrcSinkMaximumInterChannelCrosstalk OBJECT-TYPE
    SYNTAX Integer32
    UNITS "0.1 db"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            Inter-channel crosstalk is defined as the ratio of total power in
            all of the disturbing channels to that in the wanted channel,
            where the wanted and disturbing channels are at different
            wavelengths. The parameter specify the isolation of a link
            conforming to the 'black link' approach such that under the
            worst-case operating conditions the inter-channel crosstalk at
            any reference point RS is less than the maximum inter-channel
            crosstalk value."
        ::= { optIfOChSrcSinkConfigEntry 7}

optIfOChSrcSinkMaximumInterFerometricCrosstalk OBJECT-TYPE
    SYNTAX Integer32
    UNITS "0.1 db"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            This parameter places a requirement on the isolation of a link
            conforming to the 'black link' approach such that under the worst
            case operating conditions the interferometric crosstalk at any
            reference point RS is less than the maximum interferometric
            crosstalk value.."
        ::= { optIfOChSrcSinkConfigEntry 8}

optIfOChSrcSinkMaximumOpticalPathOSNRPenalty OBJECT-TYPE
    SYNTAX Integer32
    UNITS "0.1 db"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The optical path OSNR penalty is defined as the difference between
            the Lowest OSNR at Rs and Lowest OSNR at Ss that meets the BER
```

```

        requirement."
 ::= { optIfOChSrcSinkConfigEntry 9}

-- Parameters at Sink (Rs)
-- optIfOChSinkConfigTable
optIfOChSinkConfigTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOChSinkConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A table of OCh Sink (Rs) configuration parameters."
    ::= { optIfOChSrcSinkGroup 4 }

optIfOChSinkConfigEntry OBJECT-TYPE
    SYNTAX OptIfOChSinkConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A conceptual row that contains the Sink (Rs) configuration
        parameters for a given interface."
    INDEX { ifIndex }
    ::= { optIfOChSinkConfigTable 1 }

OptIfOChSinkConfigEntry ::=
    SEQUENCE {
        optIfOChSinkMinimumMeanIntputPower Integer32,
        optIfOChSinkMaximumMeanIntputPower Integer32,
        optIfOChSinkMinimumOSNR Integer32,
        optIfOChSinkOSNRTolerance Integer32
    }

optIfOChSinkMinimumMeanIntputPower OBJECT-TYPE
    SYNTAX Integer32
    UNITS "0.1 dBm"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        " The minimum values of the average received power (in dBm
        at point the Sink (Rs))."
    ::= { optIfOChSinkConfigEntry 1}

optIfOChSinkMaximumMeanIntputPower OBJECT-TYPE
    SYNTAX Integer32
    UNITS "0.1 dBm"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        " The maximum values of the average received power (in dBm)

```

```
        at point the Sink (Rs)."  
 ::= { optIfOChSinkConfigEntry 2}  
  
optIfOChSinkMinimumOSNR OBJECT-TYPE  
    SYNTAX  Integer32  
    UNITS   "0.1 dB"  
    MAX-ACCESS  read-only  
    STATUS   current  
    DESCRIPTION  
        " The minimum optical signal-to-noise ratio (OSNR) is the minimum  
        value of the ratio of the signal power in the wanted channel to  
        the highest noise power density in the range of the central  
        frequency plus and minus the maximum spectral excursion."  
 ::= { optIfOChSinkConfigEntry 3}  
  
optIfOChSinkOSNRTolerance OBJECT-TYPE  
    SYNTAX  Integer32  
    UNITS   "0.1 dB"  
    MAX-ACCESS  read-only  
    STATUS   current  
    DESCRIPTION  
        " The receiver OSNR tolerance is defined as the minimum value of  
        OSNR at point Sink (Rs) that can be tolerated while maintaining  
        the maximum BER of the application. Sink (Rs)."  
 ::= { optIfOChSinkConfigEntry 4}  
  
-- Performance Monitoring  
  
-- The OptIfOChSinkCurrentExtEntry table is an extension to the  
-- optIfOChSinkCurrentExtEntry  
-- following optional parameters for current status  
-- OptIfOChSinkCurrentExtEntry  
  
optIfOChSinkCurrentExtTable OBJECT-TYPE  
    SYNTAX  SEQUENCE OF OptIfOChSinkCurrentExtEntry  
    MAX-ACCESS  not-accessible  
    STATUS   current  
    DESCRIPTION  
        "A table of OCh sink etxension to the performance monitoring  
        information for the current 15-minute interval."  
 ::= { optIfOTNPMGroup 1 }  
  
optIfOChSinkCurrentExtEntry OBJECT-TYPE
```

```
SYNTAX  OptIfOChSinkCurrentExtEntry
MAX-ACCESS  not-accessible
STATUS  current
DESCRIPTION
    "A conceptual row that contains OCh sink performance
    monitoring information for an interface for the current
    15-minute interval."
AUGMENTS { optIfOChSinkCurrentEntry }
::= { optIfOChSinkCurrentExtTable 1 }
```

```
OptIfOChSinkCurrentExtEntry ::=
    SEQUENCE {
        optIfOChSinkCurrentChromaticDispersion      Integer32,
        optIfOChSinkCurrentOSNR                     Integer32,
        optIfOChSinkCurrentQ                         Integer32
    }
```

```
optIfOChSinkCurrentChromaticDispersion OBJECT-TYPE
SYNTAX  Integer32
UNITS   "ps/nm"
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
    "    Residual Chromatic Dispersion measured at Rx Transceiver port."
::= { optIfOChSinkCurrentExtEntry 1 }
```

```
optIfOChSinkCurrentOSNR OBJECT-TYPE
SYNTAX  Integer32
UNITS   "0.1 db"
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
    "    Current Optical Signal to Noise Ratio (OSNR) estimated at Rx
    Transceiver port ."
::= { optIfOChSinkCurrentExtEntry 2 }
```

```
optIfOChSinkCurrentQ OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
    "    'Q' factor estimated at Rx Transceiver port."
::= { optIfOChSinkCurrentExtEntry 3 }
```

```
-- Performance Monitoring
-- OTN PM Config Table
```

```
--
optIfOTNPMConfigTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A table of performance monitoring configuration for the type
        'optIfOTNPMConfigLayer' layer."
    ::= { optIfOTNPMGroup 2 }

optIfOTNPMConfigEntry OBJECT-TYPE
    SYNTAX OptIfOTNPMConfigEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        " A conceptual entry in the performance monitoring configuration
        for the type
        'optIfOTNPMConfigLayer' layer.
        "
    INDEX { ifIndex, optIfOTNPMConfigType, optIfOTNPMConfigLayer,
    optIfOTNPMConfigTCMLLevel }
    ::= { optIfOTNPMConfigTable 1 }

OptIfOTNPMConfigEntry ::=
    SEQUENCE {
        optIfOTNPMConfigType                OptIfPerformanceDataType,
        optIfOTNPMConfigLayer                OptIfOTNLayer,
        optIfOTNPMConfigTCMLLevel            Unsigned32,
        optIfOTNPM15MinFcsThreshold          Unsigned32,
        optIfOTNPM15MinESsThreshold          Unsigned32,
        optIfOTNPM15MinSESSsThreshold        Unsigned32,
        optIfOTNPM15MinUASsThreshold         Unsigned32,
        optIfOTNPM15MinBBEsThreshold         Unsigned32
    }

optIfOTNPMConfigType OBJECT-TYPE
    SYNTAX OptIfPerformanceDataType
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        This parameter indicates the parameters for the table are for the
        Near End or Far End performance data.
        1 - Near End
        2 - Far End
        "
    ::= { optIfOTNPMConfigEntry 1}
```

```
optIfOTNPMConfigLayer OBJECT-TYPE
    SYNTAX  OptIfOTNLayer
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the parameters for the table are for OTUk,
            ODUk, TCMn performance data.
            1 - OTUk
            2 - ODUk
            3 - TCM
            The ODUk/TCM Layer PM is not related to the black link PM
            management, but since this is a common PM model for the ODU/TCM
            layer, we may include it here.
        "
    ::= { optIfOTNPMConfigEntry 2}

optIfOTNPMConfigTCMLevel OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the TCM level (1-6)
            if the PM is of the type TCM. This will be 0 for OTUK/ODUK.
        "
    ::= { optIfOTNPMConfigEntry 3}

optIfOTNPM15MinFcsThreshold OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            The number of Fcs encountered by the interface within any
            given 15 minutes performance data collection period, which causes
            the SNMP agent to send optIf15MinThreshFcsTCA. One notification
            will be sent per interval per interface. A value of '0' will
            disable the notification.
        "
    ::= { optIfOTNPMConfigEntry 4 }

optIfOTNPM15MinESsThreshold OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
```



```

    The number of ES encountered by the interface within any
    given 15 minutes performance data collection period, which causes
    the SNMP agent to send optIf15MinThreshEstTCA. One notification
    will be sent per interval per interface if the threshold is
    exceeded. A value of '0' will disable the notification.
    "
 ::= { optIfOTNPMConfigEntry 5 }

optIfOTNPM15MinSESSsThreshold OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of SES encountered by the interface within any
            given 15 minutes performance data collection period, which causes
            the SNMP agent to send optIf15MinThreshSESTCA. One notification
            will be sent per interval per interface if the threshold is
            exceeded. A value of '0' will disable the notification.
        "
    ::= { optIfOTNPMConfigEntry 6 }

optIfOTNPM15MinUASsThreshold OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of UAS encountered by the interface within any
            given 15 minutes performance data collection period, which causes
            the SNMP agent to send optIf15MinThreshUASTCA. One notification
            will be sent per interval per interface if the threshold is
            exceeded. A value of '0' will disable the notification.
        "
    ::= { optIfOTNPMConfigEntry 7 }

optIfOTNPM15MinBBEsThreshold OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of UAS encountered by the interface within any
            given 15 minutes performance data collection period, which causes
            the SNMP agent to send optIf15MinThreshBBETCA. One notification
            will be sent per interval per interface if the threshold is
            exceeded. A value of '0' will disable the notification.
        "
```

```

"
 ::= { optIfOTNPMConfigEntry 8 }

--
-- PM Current Entry at either the OTU/ODUK/TCM
--
optIfOTNPMCurrentTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMCurrentEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A table for the Performance monitoring Current Table.
        "
    ::= { optIfOTNPMGroup 3}

optIfOTNPMCurrentEntry OBJECT-TYPE
    SYNTAX      OptIfOTNPMCurrentEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A conceptual entry in the Near end or Far End performance
        monitoring Current table for the type 'optIfOTNPMCurrentLayer'
        layer.
        "
    INDEX { ifIndex, optIfOTNPMCurrentType ,
            optIfOTNPMCurrentLayer, optIfOTNPMCurrentTCMLevel  }
    ::= { optIfOTNPMCurrentTable 1 }

OptIfOTNPMCurrentEntry ::=
    SEQUENCE {
        optIfOTNPMCurrentType          OptIfPerformanceDataType,
        optIfOTNPMCurrentLayer          OptIfOTNLayer,
        optIfOTNPMCurrentTCMLevel       Unsigned32,
        optIfOTNPMCurrentSuspectedFlag TruthValue,
        optIfOTNPMCurrentInterval       Unsigned32,
        optIfOTNPMCurrentValidIntervals Unsigned32,
        optIfOTNPMCurrentFcs            Unsigned32,
        optIfOTNPMCurrentESs            Unsigned32,
        optIfOTNPMCurrentSESSs          Unsigned32,
        optIfOTNPMCurrentUASs           Unsigned32,
        optIfOTNPMCurrentBBES           Unsigned32,
        optIfOTNPMCurrentESR            Unsigned32,
        optIfOTNPMCurrentSESR           Unsigned32,
        optIfOTNPMCurrentBBER           Unsigned32
    }

optIfOTNPMCurrentType          OBJECT-TYPE

```

```
SYNTAX OptIfPerformanceDataType
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    This parameter indicates the parameters for the table are for the
    Near End or Far End performance data.
    1 - Near End
    2 - Far End
"
 ::= { optIfOTNPMCurrentEntry 1}

optIfOTNPMCurrentLayer OBJECT-TYPE
SYNTAX OptIfOTNLayer
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    This parameter indicates the parameters for the table are for OTUk,
    ODUk, TCMn performance data.
    1 - OTUk (OCh which is used for the black link)
    2 - ODUk
    3 - TCM
    The ODUk/TCM Layer PM is not related to the black link PM
    management, but since this is a common PM model for the ODU/TCM
    layer, we may include it here.
"
 ::= { optIfOTNPMCurrentEntry 2}

optIfOTNPMCurrentTCMLevel OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    This parameter indicates the TCM level (1-6)
    if the PM is of the type TCM. This will be 0 for OTUK/ODUK.
"
 ::= { optIfOTNPMCurrentEntry 3}

optIfOTNPMCurrentSuspectedFlag OBJECT-TYPE
SYNTAX TruthValue
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    If true, the data in this entry may be unreliable.
```

```
"
 ::= { optIfOTNPMCurrentEntry 4}

optIfOTNPMCurrentInterval OBJECT-TYPE
    SYNTAX Unsigned32
    UNITS "seconds"
    MAX-ACCESS read-write
    STATUS current
    DESCRIPTION
        " This parameter indicates the measurement interval
          for calculation of the ratios.
        "
    ::= { optIfOTNPMCurrentEntry 5}

optIfOTNPMCurrentValidIntervals OBJECT-TYPE
    SYNTAX Unsigned32
    UNITS "seconds"
    MAX-ACCESS read-write
    STATUS current
    DESCRIPTION
        " The number of contiguous 15 minute intervals for which valid
          PM data is available for the particular interface.
        "
    ::= { optIfOTNPMCurrentEntry 6}

optIfOTNPMCurrentFcS OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
          Number of Failures occurred in an observation period.
        "
    ::= { optIfOTNPMCurrentEntry 7 }

optIfOTNPMCurrentESS OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
          This is the number of seconds in which one or more blocks are in
          error or during which a defect (e.g. Loss of Signal (LOS)) is
          detected.
        "
    ::= { optIfOTNPMCurrentEntry 8 }

optIfOTNPMCurrentSESS OBJECT-TYPE
```

```
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " The number of seconds which have a severe error.
      This is the number of seconds in which the errored block ratio
        exceeds the threshold or during which a defect (e.g. Loss of
          Signal (LOS)) is detected.
    "
 ::= { optIfOTNPMCurrentEntry 9 }

optIfOTNPMCurrentUASS OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " It is the number of unavailable seconds.
      A period of unavailable time begins at the onset of ten
        consecutive SES events. These ten seconds are considered to be
          part of unavailable time. A new period of available time begins
            at the onset of ten consecutive non-SES events. These ten
              seconds are considered to be part of available time.
    "
 ::= { optIfOTNPMCurrentEntry 10 }

optIfOTNPMCurrentBBES OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    "
      An errored block not occurring as part of an SES.
    "
 ::= { optIfOTNPMCurrentEntry 11 }

optIfOTNPMCurrentESR OBJECT-TYPE
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    "
      The ratio of ES in available time to total seconds in available
        time during a fixed measurement interval.
    "
 ::= { optIfOTNPMCurrentEntry 12 }
```

```
optIfOTNPMCurrentSESR OBJECT-TYPE
    SYNTAX      Unsigned32
    UNITS       ".001"
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
        The ratio of SES in available time to total seconds in available
        time during a fixed measurement interval.
        "
    ::= { optIfOTNPMCurrentEntry 13 }

optIfOTNPMCurrentBBER OBJECT-TYPE
    SYNTAX      Unsigned32
    UNITS       ".001"
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
        The ratio of BER in available time to total seconds in available
        time during a fixed measurement interval.
        "
    ::= { optIfOTNPMCurrentEntry 14 }

--
-- OTN PM Interval Table
-- Upto 96 15-minute intervals
--
optIfOTNPMIntervalTable OBJECT-TYPE
    SYNTAX      SEQUENCE OF OptIfOTNPMIntervalEntry
    MAX-ACCESS   not-accessible
    STATUS      current
    DESCRIPTION
        "A Performance monitoring Interval Table.
        "
    ::= { optIfOTNPMGroup 4 }

optIfOTNPMIntervalEntry OBJECT-TYPE
    SYNTAX      OptIfOTNPMIntervalEntry
    MAX-ACCESS   not-accessible
    STATUS      current
    DESCRIPTION
        "A conceptual entry in the Near end or Far End performance
        monitoring Interval table for the type 'optIfOTNPMIntervalLayer'
        layer.
        "
    INDEX { ifIndex, optIfOTNPMIntervalType, optIfOTNPMIntervalLayer,
            optIfOTNPMIntervalTCMLLevel, optIfOTNPMIntervalNumber }
```

```

 ::= { optIfOTNPMIntervalTable 1 }

OptIfOTNPMIntervalEntry ::=
    SEQUENCE {
        optIfOTNPMIntervalType          OptIfPerformanceDataType,
        optIfOTNPMIntervalLayer         OptIfOTNLayer,
        optIfOTNPMIntervalTCMLevel      Unsigned32,
        optIfOTNPMIntervalNumber        Unsigned32,
        optIfOTNPMIntervalSuspectedFlag TruthValue,
        optIfOTNPMIntervalFcs           Unsigned32,
        optIfOTNPMIntervalESS           Unsigned32,
        optIfOTNPMIntervalSESS          Unsigned32,
        optIfOTNPMIntervalUASS          Unsigned32,
        optIfOTNPMIntervalBBES          Unsigned32,
        optIfOTNPMIntervalESR           Unsigned32,
        optIfOTNPMIntervalSESR          Unsigned32,
        optIfOTNPMIntervalBBER          Unsigned32
    }

optIfOTNPMIntervalType          OBJECT-TYPE
    SYNTAX  OptIfPerformanceDataType
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the parameters for the table are for the
            Near End or Far End performance data.
            1 - Near End
            2 - Far End
        "
    ::= { optIfOTNPMIntervalEntry 1}

optIfOTNPMIntervalLayer        OBJECT-TYPE
    SYNTAX  OptIfOTNLayer
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the parameters for the table are for OTUk,
            ODUk, TCMn performance data.
            1 - OTUk
            2 - ODUk
            3 - TCM
            The ODUk/TCM Layer PM is not related to the black link PM
            management, but since this is a common PM model for the ODU/TCM
            layer, we may include it here."
    ::= { optIfOTNPMIntervalEntry 2}

```

```
optIfOTNPMIntervalTCMLevel    OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the TCM level (1-6)
            if the PM is of the type TCM. This will be 0 for OTUK/ODUK.
        "
    ::= { optIfOTNPMIntervalEntry 3}

optIfOTNPMIntervalNumber    OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            A number between 1 and 96, where 1 is the most
            recently completed 15 minute interval and 96 is
            the 15 minutes interval completed 23 hours and 45
            minutes prior to interval 1.
        "
    ::= { optIfOTNPMIntervalEntry 4}

optIfOTNPMIntervalSuspectedFlag    OBJECT-TYPE
    SYNTAX  TruthValue
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            If true, the data in this entry may be unreliable.
        "
    ::= { optIfOTNPMIntervalEntry 5}

optIfOTNPMIntervalFcs    OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            Number of Failures occurred in an observation period.
        "
    ::= { optIfOTNPMIntervalEntry 6}

optIfOTNPMIntervalESs    OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
```


DESCRIPTION

```
"
    It is a one-second period which has one or more errored blocks
    or during which a defect (e.g. Loss of Signal (LOS)) is detected.
"
 ::= { optIfOTNPMIntervalEntry 7}
```

optIfOTNPMIntervalSESS OBJECT-TYPE

```
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " The number of seconds which have a severe error.
      It is a one-second period in which the errored block ratio exceeds
      the threshold or during which a defect (e.g. Loss of Signal(LOS))
      is detected.
    "
 ::= { optIfOTNPMIntervalEntry 8}
```

optIfOTNPMIntervalUASS OBJECT-TYPE

```
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " It is the number of unavailable seconds in this 15 minute
      interval.
      A period of unavailable time begins at the onset of ten
      consecutive SES events. These ten seconds are considered to be
      part of unavailable time. A new period of available time begins
      at the onset of ten consecutive non-SES events. These ten
      seconds are considered to be part of available time.
    "
 ::= { optIfOTNPMIntervalEntry 9}
```

optIfOTNPMIntervalBBES OBJECT-TYPE

```
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    "
      An errored block not occurring as part of an SES.
    "
 ::= { optIfOTNPMIntervalEntry 10}
```

optIfOTNPMIntervalESR OBJECT-TYPE

```
SYNTAX Unsigned32
UNITS ".001"
```

```
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The ratio of ES in available time to total seconds in available
    time during a fixed measurement interval.
"
::= { optIfOTNPMIntervalEntry 11}

optIfOTNPMIntervalSESR OBJECT-TYPE
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The ratio of SES in available time to total seconds in available
    time during a fixed measurement interval.
"
::= { optIfOTNPMIntervalEntry 12}

optIfOTNPMIntervalBBER OBJECT-TYPE
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The ratio of BBE in available time to total seconds in available
    time during a fixed measurement interval.
"
::= { optIfOTNPMIntervalEntry 13}

--
-- PM Current Day Entry
--
optIfOTNPMCurrentDayTable OBJECT-TYPE
SYNTAX SEQUENCE OF OptIfOTNPMCurrentDayEntry
MAX-ACCESS not-accessible
STATUS current
DESCRIPTION
" A Performance monitoring Current Day Table.
"
::= { optIfOTNPMGroup 5 }

optIfOTNPMCurrentDayEntry OBJECT-TYPE
SYNTAX OptIfOTNPMCurrentDayEntry
MAX-ACCESS not-accessible
```

```

STATUS    current
DESCRIPTION
    "A conceptual entry in the Near end or Far End performance
    monitoring Current day table for the type
    'optIfOTNPMCurrentDayLayer' layer.
    "
INDEX { ifIndex, optIfOTNPMCurrentDayType, optIfOTNPMCurrentDayLayer,
        optIfOTNPMCurrentDayTCMLLevel }
 ::= { optIfOTNPMCurrentDayTable 1 }

OptIfOTNPMCurrentDayEntry ::=
    SEQUENCE {
        optIfOTNPMCurrentDayType          OptIfPerformanceDataType,
        optIfOTNPMCurrentDayLayer          OptIfOTNLayer,
        optIfOTNPMCurrentDayTCMLLevel      Unsigned32,
        optIfOTNPMCurrentDaySuspectedFlag  TruthValue,
        optIfOTNPMCurrentDayFcs            Unsigned32,
        optIfOTNPMCurrentDayESS            Unsigned32,
        optIfOTNPMCurrentDaySESS           Unsigned32,
        optIfOTNPMCurrentDayUASs           Unsigned32,
        optIfOTNPMCurrentDayBBEs           Unsigned32,
        optIfOTNPMCurrentDayESR            Unsigned32,
        optIfOTNPMCurrentDaySESR           Unsigned32,
        optIfOTNPMCurrentDayBBER           Unsigned32
    }

optIfOTNPMCurrentDayType          OBJECT-TYPE
SYNTAX  OptIfPerformanceDataType
MAX-ACCESS  read-only
STATUS    current
DESCRIPTION
    "
    This parameter indicates the parameters for the table are for
    the Near End or Far End performance data.
    1 - Near End
    2 - Far End
    "
 ::= { optIfOTNPMCurrentDayEntry 1}

optIfOTNPMCurrentDayLayer          OBJECT-TYPE
SYNTAX  OptIfOTNLayer
MAX-ACCESS  read-only
STATUS    current
DESCRIPTION
    "
    This parameter indicates the parameters for the table are for OTUk,
    ODUk, TCMn performance data.
    1 - OTUk

```

```
2 - ODUk
3 - TCM
The ODUk/TCM Layer PM is not related to the black link PM
management, but since this is a common PM model for the ODU/TCM
layer, we may include it here."
 ::= { optIfOTNPMCurrentDayEntry 2}

optIfOTNPMCurrentDayTCMLevel OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    This parameter indicates the TCM level (1-6)
    if the PM is of the type TCM. This will be 0 for OTUK/ODUK.
"
 ::= { optIfOTNPMCurrentDayEntry 3}

optIfOTNPMCurrentDaySuspectedFlag OBJECT-TYPE
SYNTAX TruthValue
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    If true, the data in this entry may be unreliable.
"
 ::= { optIfOTNPMCurrentDayEntry 4}

optIfOTNPMCurrentDayFcs OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    Number of Failures occurred in an observation period.
"
 ::= { optIfOTNPMCurrentDayEntry 5}

optIfOTNPMCurrentDayESS OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" The number of seconds which have an error.
    It is a one-second period which has one or more errored blocks
```

```
    or during which a defect(e.g., Loss of Signal (LOS)) is detected.
"
 ::= { optIfOTNPMCurrentDayEntry 6}

optIfOTNPMCurrentDaySESS OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        " The number of seconds which have a severe error.
          A severely errored second, is a one-second period in which the
          errored block ratio exceeds the threshold or during which a
          defect (e.g. Loss of Signal (LOS)) is detected.
        "
    ::= { optIfOTNPMCurrentDayEntry 7}

optIfOTNPMCurrentDayUASS OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        " It is the number of unavailable seconds in the cunrrent day.
          A period of unavailable time begins at the onset of ten
          consecutive SES events. These ten seconds are considered to be
          part of unavailable time. A new period of available time begins
          at the onset of ten consecutive non-SES events. These ten seconds
          are considered to be part of available time.
        "
    ::= { optIfOTNPMCurrentDayEntry 8}

optIfOTNPMCurrentDayBBES OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
          An errored block not occurring as part of an SES.
        "
    ::= { optIfOTNPMCurrentDayEntry 9}

optIfOTNPMCurrentDayESR OBJECT-TYPE
    SYNTAX Unsigned32
    UNITS ".001"
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
          The ratio of ES in available time to total seconds in available
```

```
        time during a fixed measurement interval.
    "
    ::= { optIfOTNPMCurrentDayEntry 10}

optIfOTNPMCurrentDaySESR OBJECT-TYPE
    SYNTAX  Unsigned32
    UNITS   ".001"
    MAX-ACCESS  read-only
    STATUS   current
    DESCRIPTION
        "
        The ratio of SES in available time to total seconds in available
        time during a fixed measurement interval.
        "
    ::= { optIfOTNPMCurrentDayEntry 11}

optIfOTNPMCurrentDayBBER OBJECT-TYPE
    SYNTAX  Unsigned32
    UNITS   ".001"
    MAX-ACCESS  read-only
    STATUS   current
    DESCRIPTION
        "
        The ratio of BBE in available time to total seconds in available
        time during a fixed measurement interval.
        "
    ::= { optIfOTNPMCurrentDayEntry 12}

--
-- PM Prev Day Entry
--
optIfOTNMPPrevDayTable OBJECT-TYPE
    SYNTAX  SEQUENCE OF OptIfOTNMPPrevDayEntry
    MAX-ACCESS  not-accessible
    STATUS   current
    DESCRIPTION
        "A Performance monitoring Previous Day Table.
        "
    ::= { optIfOTNPMGroup 6 }

optIfOTNMPPrevDayEntry OBJECT-TYPE
    SYNTAX  OptIfOTNMPPrevDayEntry
    MAX-ACCESS  not-accessible
    STATUS   current
    DESCRIPTION
        "A conceptual entry in the Near end or Far End performance
        monitoring previous day table for the type
        'optIfOTNMPPrevDayLayer' layer.
```

```
"
INDEX { ifIndex, optIfOTNPMPPrevDayType
        optIfOTNPMPPrevDayLayer, optIfOTNPMPPrevDayTCMLLevel }
 ::= { optIfOTNPMPPrevDayTable 1 }

OptIfOTNPMPPrevDayEntry ::=
    SEQUENCE {
        optIfOTNPMPPrevDayType          OptIfPerformanceDataType,
        optIfOTNPMPPrevDayLayer          OptIfOTNLayer,
        optIfOTNPMPPrevDayTCMLLevel      Unsigned32,
        optIfOTNPMPPrevDaySuspectedFlag  TruthValue,
        optIfOTNPMPPrevDayFcS            Unsigned32,
        optIfOTNPMPPrevDayESS            Unsigned32,
        optIfOTNPMPPrevDaySESS           Unsigned32,
        optIfOTNPMPPrevDayUASS           Unsigned32,
        optIfOTNPMPPrevDayBBES           Unsigned32,
        optIfOTNPMPPrevDayESR            Unsigned32,
        optIfOTNPMPPrevDaySESR           Unsigned32,
        optIfOTNPMPPrevDayBBER           Unsigned32
    }

optIfOTNPMPPrevDayType          OBJECT-TYPE
SYNTAX  OptIfPerformanceDataType
MAX-ACCESS  read-only
STATUS    current
DESCRIPTION
    "
        This parameter indicates the parameters for the table are for the
        Near End or Far End performance data.
        1 - Near End
        2 - Far End
    "
 ::= { optIfOTNPMPPrevDayEntry 1}

optIfOTNPMPPrevDayLayer          OBJECT-TYPE
SYNTAX  OptIfOTNLayer
MAX-ACCESS  read-only
STATUS    current
DESCRIPTION
    "
        This parameter indicates the parameters for the table are for
        OTUk, ODUk, TCMn performance data.
        1 - OTUk
        2 - ODUk
        3 - TCM
        The ODUk/TCM Layer PM is not related to the black link PM
        management, but since this is a common PM model for the ODU/TCM
        layer, we may include it here."

```

```
 ::= { optIfOTNPMPPrevDayEntry 2}

optIfOTNPMPPrevDayTCMLLevel OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        This parameter indicates the TCM level (1-6)
        if the PM is of the type TCM.
        "
 ::= { optIfOTNPMPPrevDayEntry 3}

optIfOTNPMPPrevDaySuspectedFlag OBJECT-TYPE
    SYNTAX TruthValue
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        If true, the data in this entry may be unreliable.
        "
 ::= { optIfOTNPMPPrevDayEntry 4}

optIfOTNPMPPrevDayFcs OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        Number of pre FEC failures occurred in an observation period.
        "
 ::= { optIfOTNPMPPrevDayEntry 5}

optIfOTNPMPPrevDayESSs OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        The number of seconds which have an error.
        It is a one-second period which has one or more errored block
        or during which a defect (e.g. Loss of Signal (LOS)) is detected.
        "
 ::= { optIfOTNPMPPrevDayEntry 6}

optIfOTNPMPPrevDaySESSs OBJECT-TYPE
    SYNTAX Unsigned32
```



```
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " The number of seconds which have a severe error.
      A severely errored second, is a one-second period in which the
      errored block ratio exceeds the threshold or during which a defect
      (e.g. Loss of Signal (LOS)) is detected.
    "
 ::= { optIfOTNPMPPrevDayEntry 7}

optIfOTNPMPPrevDayUASS OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    " It is the number of unavailable seconds in the previous day.
      A period of unavailable time begins at the onset of ten
      consecutive SES events. These ten seconds are considered to be
      part of unavailable time. A new period of available time begins
      at the onset of ten consecutive non-SES events. These ten seconds
      are considered to be part of available time.
    "
 ::= { optIfOTNPMPPrevDayEntry 8}

optIfOTNPMPPrevDayBBES OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    "
      An errored block not occurring as part of an SES.
    "
 ::= { optIfOTNPMPPrevDayEntry 9}

optIfOTNPMPPrevDayESR OBJECT-TYPE
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
    "
      The ratio of ES in available time to total seconds in available
      time during a fixed measurement interval.
    "
 ::= { optIfOTNPMPPrevDayEntry 10}

optIfOTNPMPPrevDaySESR OBJECT-TYPE
```

```
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The ratio of SES in available time to total seconds in available
    time during a fixed measurement interval.
"
 ::= { optIfOTNPMPrevDayEntry 11}

optIfOTNPMPrevDayBBER OBJECT-TYPE
SYNTAX Unsigned32
UNITS ".001"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The ratio of BBE in available time to total seconds in available
    time during a fixed measurement interval.
"
 ::= { optIfOTNPMPrevDayEntry 12}

--
-- OTN FEC PM Config Table
--
optIfOTNPMFECConfigTable OBJECT-TYPE
SYNTAX SEQUENCE OF OptIfOTNPMFECConfigEntry
MAX-ACCESS not-accessible
STATUS current
DESCRIPTION
    "A table of performance monitoring FEC configuration. "
 ::= { optIfOTNPMGroup 7 }

optIfOTNPMFECConfigEntry OBJECT-TYPE
SYNTAX OptIfOTNPMFECConfigEntry
MAX-ACCESS not-accessible
STATUS current
DESCRIPTION
    " A conceptual entry in the performance monitoring FEC
    configuration layer."
INDEX { ifIndex, optIfOTNPMFECConfigType }
 ::= { optIfOTNPMFECConfigTable 1 }

OptIfOTNPMFECConfigEntry ::=
    SEQUENCE {
        optIfOTNPMFECConfigType OptIfPerformanceDataType,
        optIfOTNPMFECConfigValidIntervals Unsigned32,
```

```

    optIfOTNPM15MinFECUncorrectedWordsThreshold
                                                Unsigned32,
    optIfOTNPM15MinPreFECBERThresholdMantissa
                                                Unsigned32,
    optIfOTNPM15MinPreFECBERThresholdExponent
                                                Unsigned32
}

optIfOTNPMFECConfigType OBJECT-TYPE
    SYNTAX  OptIfPerformanceDataType
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the parameters for the table are for the
            Near End or Far End performance data.
            1 - Near End
            2 - Far End
        "
    ::= { optIfOTNPMFECConfigEntry 1}

optIfOTNPMFECValidIntervals OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            The number of contiguous 15 minute intervals for which valid FEC
            PM data is available for the particular interface.
        "
    ::= {optIfOTNPMFECConfigEntry 2}

optIfOTNPM15MinFECUncorrectedWordsThreshold OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            The number of Uncorrected words encountered by the interface within
            any given 15 minutes performance data collection period, which
            causes the SNMP agent to send
            optIf15MinThreshFECUncorrectedWordsTCA.
            One notification will be sent per interval per interface if the
            thresholds is exceeded.. A value of '0' will disable
            the notification.
        "
    ::= {optIfOTNPMFECConfigEntry 3}
```

```
optIfOTNPM15MinPreFECBERThresholdMantissa OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        The Pre FEC BER (mantissa) by the interface within any
        given 15 minutes performance data collection period, which causes
        the SNMP agent to send optIf15MinThreshPreFECBERTCA. One
        notification will be sent per interval per interface. A value of
        '0' will disable the notification.
        "
    ::= { optIfOTNPMFECConfigEntry 4 }

optIfOTNPM15MinPreFECBERThresholdExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        The Pre FEC BER (exponent) by the interface within any
        given 15 minutes performance data collection period, which causes
        the SNMP agent to send optIf15MinThreshPreFECBERTCA. One
        notification will be sent per interval per interface. A value of
        '0' will disable the notification.
        "
    ::= { optIfOTNPMFECConfigEntry 5 }

--
-- FEC PM Table
--
optIfOTNPMFECCurrentTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMFECCurrentEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A Performance monitoring FEC Current Table.
        "
    ::= { optIfOTNPMGroup 8 }

optIfOTNPMFECCurrentEntry OBJECT-TYPE
    SYNTAX OptIfOTNPMFECCurrentEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        " A conceptual entry in the Near end or Far End performance
        monitoring FEC current table.
```

```

"
INDEX { ifIndex, optIfOTNPMFECCurrentType}
 ::= { optIfOTNPMFECCurrentTable 1 }

OptIfOTNPMFECCurrentEntry ::=
    SEQUENCE {
        optIfOTNPMFECCurrentType          OptIfPerformanceDataType,
        optIfOTNPMFECCurrentSuspectedFlag TruthValue,
        optIfOTNPMCurrentFECCorrectedErr   Unsigned32,
        optIfOTNPMCurrentFECUncorrectedWords Unsigned32,
        optIfOTNPMCurrentFECBERMantissa    Unsigned32,
        optIfOTNPMCurrentFECBERExponent    Unsigned32
    }

optIfOTNPMFECCurrentType          OBJECT-TYPE
SYNTAX  OptIfPerformanceDataType
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"
    This parameter indicates the parameters for the table are for the
    Near End or Far End performance data.
    1 - Near End
    2 - Far End
"
 ::= { optIfOTNPMFECCurrentEntry 1}

optIfOTNPMFECCurrentSuspectedFlag OBJECT-TYPE
SYNTAX  TruthValue
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"
    If true, the data in this entry may be unreliable.
"
 ::= { optIfOTNPMFECCurrentEntry 2}

optIfOTNPMCurrentFECCorrectedErr OBJECT-TYPE
SYNTAX  Unsigned32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"
    The number of bits corrected by the FEC are counted in the
    interval.
"
 ::= { optIfOTNPMFECCurrentEntry 3}

```

```
optIfOTNPMCurrentFECUncorrectedWords OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of un-corrected words by the FEC are counted over the
            interval.
        "
    ::= { optIfOTNPMFECCurrentEntry 4}

optIfOTNPMCurrentFECBERMantissa OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of Errored bits at receiving side before the FEC
            function counted over one second .. mantisa.
        "
    ::= { optIfOTNPMFECCurrentEntry 5}

optIfOTNPMCurrentFECBERExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The number of Errored bits at receiving side before the FEC
            function counted over one second .. exponent (eg -1).
        "
    ::= { optIfOTNPMFECCurrentEntry 6}

--
-- FEC PM Interval Table
--
optIfOTNPMFECIntervalTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMFECIntervalEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A Performance monitoring FEC Interval Table.
        "
    ::= { optIfOTNPMGroup 9 }

optIfOTNPMFECIntervalEntry OBJECT-TYPE
    SYNTAX OptIfOTNPMFECIntervalEntry
    MAX-ACCESS not-accessible
```

```

STATUS    current
DESCRIPTION
    "A conceptual entry in the Near end or Far End performance
      monitoring FEC interval table.
    "
INDEX { ifIndex, optIfOTNPMFECIntervalType, optIfOTNPMFECIntervalNumber }
::= { optIfOTNPMFECIntervalTable 1 }

OptIfOTNPMFECIntervalEntry ::=
    SEQUENCE {
        optIfOTNPMFECIntervalType          OptIfPerformanceDataType,
        optIfOTNPMFECIntervalNumber        Unsigned32,
        optIfOTNPMFECIntervalSuspectedFlag TruthValue,
        optIfOTNPMFECIntervalFECCorrectedErr Unsigned32,
        optIfOTNPMFECIntervalFECUncorrectedWords Unsigned32,
        optIfOTNPMFECIntervalMinFECBERMantissa Unsigned32,
        optIfOTNPMFECIntervalMinFECBERExponent Unsigned32,
        optIfOTNPMFECIntervalMaxFECBERMantissa Unsigned32,
        optIfOTNPMFECIntervalMaxFECBERExponent Unsigned32,
        optIfOTNPMFECIntervalAvgFECBERMantissa Unsigned32,
        optIfOTNPMFECIntervalAvgFECBERExponent Unsigned32
    }

optIfOTNPMFECIntervalType          OBJECT-TYPE
SYNTAX    OptIfPerformanceDataType
MAX-ACCESS read-only
STATUS    current
DESCRIPTION
    "
        This parameter indicates the parameters for the table are for the
        Near End or Far End performance data.
        1 - Near End
        2 - Far End
    "
    ::= { optIfOTNPMFECIntervalEntry 1}

optIfOTNPMFECIntervalNumber        OBJECT-TYPE
SYNTAX    Unsigned32
MAX-ACCESS read-only
STATUS    current
DESCRIPTION
    "
        A number between 1 and 96, where 1 is the most
        recently completed 15 minute interval and 96 is
        the 15 minutes interval completed 23 hours and 45
        minutes prior to interval 1.
    "
    ::= { optIfOTNPMFECIntervalEntry 2}

```

```
optIfOTNPMFECIntervalSuspectedFlag    OBJECT-TYPE
    SYNTAX      TruthValue
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
            If true, the data in this entry may be unreliable.
        "
    ::= { optIfOTNPMFECIntervalEntry 3}

optIfOTNPMIntervalFECCorrectedErr      OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
            The number of bits corrected by the FEC are counted in the
            interval.
        "
    ::= { optIfOTNPMFECIntervalEntry 4}

optIfOTNPMIntervalFECUncorrectedWords  OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
            The number of words un-corrected words by the FEC are counted over
            the interval.
        "
    ::= { optIfOTNPMFECIntervalEntry 5}

optIfOTNPMIntervalMinFECBERMantissa    OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS   read-only
    STATUS      current
    DESCRIPTION
        "
            The minimum bit error rate at receiving side before the FEC
            function counted over one second .. mantissa. This is the minimum
            Pre FEC BER in the current 24hour period.
        "
    ::= { optIfOTNPMFECIntervalEntry 6}

optIfOTNPMIntervalMinFECBERExponent    OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS   read-only
    STATUS      current
```


DESCRIPTION
"
The minimum bit error rate at receiving side before the FEC
function counted over one second .. exponent. This is the minimum
Pre FEC BER in the current 24hour period.
"
::= { optIfOTNPMFECIntervalEntry 7}

optIfOTNPMIntervalMaxFECBERMantissa OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
The maximum bit error rate at receiving side before the FEC
function counted over one second .. mantissa. This is the maximum
Pre FEC BER in the current 24hour period.
"
::= { optIfOTNPMFECIntervalEntry 8}

optIfOTNPMIntervalMaxFECBERExponent OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
The maximum bit error rate at receiving side before the FEC
function counted over one second .. exponent. This is the maximum
Pre FEC BER in the current 24hour period.
"
::= { optIfOTNPMFECIntervalEntry 9}

optIfOTNPMIntervalAvgFECBERMantissa OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
The average bit error rate at receiving side before the FEC
function counted over one second .. mantissa. This is the average
Pre FEC BER in the current 24hour period.
"
::= { optIfOTNPMFECIntervalEntry 10}

optIfOTNPMIntervalAvgFECBERExponent OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current

```

DESCRIPTION
"
    The average bit error rate at receiving side before the FEC
    function counted over one second .. exponent. This is the average
    Pre FEC BER in the current 24hour period.
"
 ::= { optIfOTNPMFECIntervalEntry 11}

--
-- FEC PM Current Day day Table
--
optIfOTNPMFECCurrentDayTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMFECCurrentDayEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A Performance monitoring FEC current day table.
        "
    ::= { optIfOTNPMGroup 10 }

optIfOTNPMFECCurrentDayEntry OBJECT-TYPE
    SYNTAX OptIfOTNPMFECCurrentDayEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A conceptual entry in the Near end or Far End performance
        monitoring FEC current day table.
        "
    INDEX { ifIndex, optIfOTNPMFECCurrentDayType }
    ::= { optIfOTNPMFECCurrentDayTable 1 }

OptIfOTNPMFECCurrentDayEntry ::=
    SEQUENCE {
        optIfOTNPMFECCurrentDayType          OptIfPerformanceDataType,
        optIfOTNPMFECCurrentDaySuspectedFlag TruthValue,
        optIfOTNPMCurrentDayFECCorrectedErr  Unsigned32,
        optIfOTNPMCurrentDayFECUncorrectedWords Unsigned32,
        optIfOTNPMCurrentDayMinFECBERMantissa Unsigned32,
        optIfOTNPMCurrentDayMinFECBERExponent Unsigned32,
        optIfOTNPMCurrentDayMaxFECBERMantissa Unsigned32,
        optIfOTNPMCurrentDayMaxFECBERExponent Unsigned32,
        optIfOTNPMCurrentDayAvgFECBERMantissa Unsigned32,
        optIfOTNPMCurrentDayAvgFECBERExponent Unsigned32
    }

optIfOTNPMFECCurrentDayType OBJECT-TYPE
    SYNTAX OptIfPerformanceDataType
    MAX-ACCESS read-only

```

```
STATUS current
DESCRIPTION
"
    This parameter indicates the parameters for the table are for the
    Near End or Far End performance data.
    1 - Near End
    2 - Far End
"
::= { optIfOTNPMFECCurrentDayEntry 1}

optIfOTNPMFECCurrentDaySuspectedFlag OBJECT-TYPE
SYNTAX TruthValue
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    If true, the data in this entry may be unreliable.
"
::= { optIfOTNPMFECCurrentDayEntry 2}

optIfOTNPMCurrentDayFECCorrectedErr OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The number of bits corrected by the FEC are counted in the
    interval.
"
::= { optIfOTNPMFECCurrentDayEntry 3}

optIfOTNPMCurrentDayFECUncorrectedWords OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"
    The number of words un-corrected by the FEC are counted over the
    Day.
"
::= { optIfOTNPMFECCurrentDayEntry 4}

optIfOTNPMCurrentDayMinFECBERMantissa OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
```

```
"
    The minimum bit error rate at receiving side before the FEC
    function counted over one second .. mantissa. This is the minimum
    PreFEC BER in the current 24hour period.
"
 ::= { optIfOTNPMFECCurrentDayEntry 5}

optIfOTNPMCurrentDayMinFECBERExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The minimum bit error rate at receiving side before the FEC
            function counted over one second .. exponent. This is the minimum
            PreFEC BER in the current 24hour period.
        "
    ::= { optIfOTNPMFECCurrentDayEntry 6}

optIfOTNPMCurrentDayMaxFECBERMantissa OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The maximum bit error rate at receiving side before the FEC
            function counted over one second .. mantissa. This is the maximum
            PreFEC BER in the current 24hour period.
        "
    ::= { optIfOTNPMFECCurrentDayEntry 7}

optIfOTNPMCurrentDayMaxFECBERExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The maximum bit error rate at receiving side before the FEC
            function counted over one second .. exponent. This is the maximum
            PreFEC BER in the current 24hour period..
        "
    ::= { optIfOTNPMFECCurrentDayEntry 8}

optIfOTNPMCurrentDayAvgFECBERMantissa OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
```

```

"
    The average bit error rate at receiving side before the FEC
    function counted over one second .. mantissa. This is the average
    PreFEC BER in the current 24hour period. .
"
 ::= { optIfOTNPMFECCurrentDayEntry 9}

optIfOTNPMCurrentDayAvgFECBERExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
            The average bit error rate at receiving side before the FEC
            function counted over one second .. exponent. This is the average
            PreFEC BER in the current 24hour period.
        "
    ::= { optIfOTNPMFECCurrentDayEntry 10}
--
-- FEC PM Prev day Table
--
optIfOTNPMFECPrevDayTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNPMFECPrevDayEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A Performance monitoring FEC previous day table.
        "
    ::= { optIfOTNPMGroup 11 }

optIfOTNPMFECPrevDayEntry OBJECT-TYPE
    SYNTAX OptIfOTNPMFECPrevDayEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A conceptual entry in the Near end or Far End performance
        monitoring FEC previous day table
        "
    INDEX { ifIndex, optIfOTNPMFECPrevDayType }
    ::= { optIfOTNPMFECPrevDayTable 1 }

OptIfOTNPMFECPrevDayEntry ::=
    SEQUENCE {
        optIfOTNPMFECPrevDayType OptIfPerformanceDataType,
        optIfOTNPMFECPrevDaySuspectedFlag TruthValue,
        optIfOTNPMPrevDayFECCorrectedErr Unsigned32,
        optIfOTNPMPrevDayFECUncorrectedWords Unsigned32,
        optIfOTNPMPrevDayMinFECBERMantissa Unsigned32,

```

```

        optIfOTNPMPrevDayMinFECBERExponent      Unsigned32,
        optIfOTNPMPrevDayMaxFECBERMantissa      Unsigned32,
        optIfOTNPMPrevDayMaxFECBERExponent      Unsigned32,
        optIfOTNPMPrevDayAvgFECBERMantissa      Unsigned32,
        optIfOTNPMPrevDayAvgFECBERExponent      Unsigned32
    }

optIfOTNPMFECPrevDayType      OBJECT-TYPE
    SYNTAX  OptIfPerformanceDataType
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            This parameter indicates the parameters for the table are for the
            Near End or Far End performance data.
            1 - Near End
            2 - Far End
        "
    ::= { optIfOTNPMFECPrevDayEntry 1}

optIfOTNPMFECPrevDaySuspectedFlag  OBJECT-TYPE
    SYNTAX  TruthValue
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            If true, the data in this entry may be unreliable.
        "
    ::= { optIfOTNPMFECPrevDayEntry 2}

optIfOTNPMPrevDayFECCorrectedErr  OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
            The number of bits corrected by the FEC are counted in the
            previous day.
        "
    ::= { optIfOTNPMFECPrevDayEntry 3}

optIfOTNPMPrevDayFECUncorrectedWords  OBJECT-TYPE
    SYNTAX  Unsigned32
    MAX-ACCESS  read-only
    STATUS  current
    DESCRIPTION
        "
```

The number of un-corrected words by the FEC are counted over the previous Day.

"

::= { optIfOTNPMFECPrevDayEntry 4}

optIfOTNPMPrevDayMinFECBERMantissa OBJECT-TYPE

SYNTAX Unsigned32

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"

The maximum bit error rate at receiving side before the FEC function counted over one second .. mantissa. This is the maximum Pre FEC BER in the previous 24hour period.

"

::= { optIfOTNPMFECPrevDayEntry 5}

optIfOTNPMPrevDayMinFECBERExponent OBJECT-TYPE

SYNTAX Unsigned32

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"

The minimum bit error rate at receiving side before the FEC function counted over one second .. exponent. This is the maximum Pre FEC BER in the previous 24hour period

"

::= { optIfOTNPMFECPrevDayEntry 6}

optIfOTNPMPrevDayMaxFECBERMantissa OBJECT-TYPE

SYNTAX Unsigned32

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"

The maximum bit error rate at receiving side before the FEC function counted over one second .. mantissa. This is the maximum Pre FEC BER in the previous 24hour period (mantissa).

"

::= { optIfOTNPMFECPrevDayEntry 7}

optIfOTNPMPrevDayMaxFECBERExponent OBJECT-TYPE

SYNTAX Unsigned32

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"

The maximum bit error rate at receiving side before the FEC

```
function counted over one second .. exponent (eg -3).
This is the maximum Pre FEC BER in the previous 24hour period.
"
 ::= { optIfOTNPMFECPrevDayEntry 8}

optIfOTNMPPrevDayAvgFECBERMantissa OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        The average bit error rate at receiving side before the FEC
        function counted over one second .. mantissa. This is the average
        Pre FEC BER during the previous 24hour period (mantissa).
        "
    ::= { optIfOTNPMFECPrevDayEntry 9}

optIfOTNMPPrevDayAvgFECBERExponent OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        The average bit error rate at receiving side before the FEC
        function counted over one second .. exponent (eg -3).
        This is the average Pre FEC BER during the previous 24hour period.
        "
    ::= { optIfOTNPMFECPrevDayEntry 10}

--
-- OTN Alarm Table
--
optIfOTNAlarmTable OBJECT-TYPE
    SYNTAX SEQUENCE OF OptIfOTNAlarmEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A table of alarm entries."

    ::= { optIfOTNAlarm 1 }

optIfOTNAlarmEntry OBJECT-TYPE
    SYNTAX OptIfOTNAlarmEntry
    MAX-ACCESS not-accessible
    STATUS current
    DESCRIPTION
        "A conceptual entry in the alarm table."
    INDEX { ifIndex, optIfOTNAlarmIndex }
```



```

 ::= { optIfOTNAlarmTable 1 }

OptIfOTNAlarmEntry ::= SEQUENCE {
    optIfOTNAlarmIndex          Unsigned32,
    optIfOTNAlarmLayer          OptIfOTNLayer,
    optIfOTNAlarmTCMLLevel      Unsigned32,
    optIfOTNAlarmType           Unsigned32,
    optIfOTNAlarmDate           DateAndTime,
    optIfOTNAlarmStatus         TruthValue
}

optIfOTNAlarmIndex OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "An index that uniquely identifies an entry in the
        alarm table."
    ::= { optIfOTNAlarmEntry 1 }

optIfOTNAlarmLayer OBJECT-TYPE
    SYNTAX      OptIfOTNLayer
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "This specifies which layer this alarm is for."
    ::= { optIfOTNAlarmEntry 2 }

optIfOTNAlarmTCMLLevel OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "TCM level 1-6 of the alarm. It will be 0 if alarm layer is
        OCh, OTUk or ODUk."
    ::= { optIfOTNAlarmEntry 3 }

optIfOTNAlarmType OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS  read-only
    STATUS      current
    DESCRIPTION
        "This specifies the type of alarm of the layer
        'optIfOTNAlarmLayer' ."
    ::= { optIfOTNAlarmEntry 4 }

optIfOTNAlarmDate OBJECT-TYPE

```

```

        SYNTAX      DateAndTime
        MAX-ACCESS   read-only
        STATUS       current
        DESCRIPTION
        "This specifies the date and time when this alarm occurred."
        ::= { optIfOTNAlarmEntry 5 }

optIfOTNAlarmStatus OBJECT-TYPE
    SYNTAX      TruthValue
    MAX-ACCESS   read-only
    STATUS       current
    DESCRIPTION
    "This specifies the state of the alarm -- cleared(0) or set(1)."
```

--

-- OTN Notifications

--

```

optIfOTNAlarmSet NOTIFICATION-TYPE
    OBJECTS { optIfOTNAlarmLayer,
               optIfOTNAlarmTCMLLevel,
               optIfOTNAlarmType,
               optIfOTNAlarmDate }
    STATUS      current
    DESCRIPTION
        "Notification of a recently set OTN alarm of layer
         and Type."
    ::= { optIfOTNNotifications 1 }

optIfOTNAlarmClear NOTIFICATION-TYPE
    OBJECTS { optIfOTNAlarmLayer,
               optIfOTNAlarmTCMLLevel,
               optIfOTNAlarmType,
               optIfOTNAlarmDate }
    STATUS      current
    DESCRIPTION
        "Notification of a recently clear OTN alarm of layer
         and Type."
    ::= { optIfOTNNotifications 2 }

END
```

7. Relationship to Other MIB Modules

7.1. Relationship to the [TEMPLATE TODO] MIB

7.2. MIB modules required for IMPORTS

8. Definitions

[TEMPLATE TODO]: put your valid MIB module here.
A list of tools that can help automate the process of checking MIB definitions can be found at <http://www.ops.ietf.org/mib-review-tools.html>

9. Security Considerations

There are a number of management objects defined in this MIB module with a MAX-ACCESS clause of read-write and/or read-create. Such objects may be considered sensitive or vulnerable in some network environments. The support for SET operations in a non-secure environment without proper protection can have a negative effect on network operations. These are the tables and objects and their sensitivity/vulnerability:

o

Some of the readable objects in this MIB module (i.e., objects with a MAX-ACCESS other than not-accessible) may be considered sensitive or vulnerable in some network environments. It is thus important to control even GET and/or NOTIFY access to these objects and possibly to even encrypt the values of these objects when sending them over the network via SNMP.

SNMP versions prior to SNMPv3 did not include adequate security. Even if the network itself is secure (for example by using IPsec), even then, there is no control as to who on the secure network is allowed to access and GET/SET (read/change/create/delete) the objects in this MIB module.

It is RECOMMENDED that implementers consider the security features as provided by the SNMPv3 framework (see [RFC3410], section 8), including full support for the SNMPv3 cryptographic mechanisms (for authentication and privacy).

Further, deployment of SNMP versions prior to SNMPv3 is NOT RECOMMENDED. Instead, it is RECOMMENDED to deploy SNMPv3 and to

enable cryptographic security. It is then a customer/operator responsibility to ensure that the SNMP entity giving access to an instance of this MIB module is properly configured to give access to the objects only to those principals (users) that have legitimate rights to indeed GET or SET (change/create/delete) them.

10. IANA Considerations

Option #1:

The MIB module in this document uses the following IANA-assigned OBJECT IDENTIFIER values recorded in the SMI Numbers registry:

Descriptor -----	OBJECT IDENTIFIER value -----
sampleMIB	{ mib-2 XXX }

Option #2:

Editor's Note (to be removed prior to publication): the IANA is requested to assign a value for "XXX" under the 'mib-2' subtree and to record the assignment in the SMI Numbers registry. When the assignment has been made, the RFC Editor is asked to replace "XXX" (here and in the MIB module) with the assigned value and to remove this note.

Note well: prior to official assignment by the IANA, an internet draft MUST use placeholders (such as "XXX" above) rather than actual numbers. See RFC4181 Section 4.5 for an example of how this is done in an internet draft MIB module.

Option #3:

This memo includes no request to IANA.

11. Contributors

Arnold Mattheus
Deutsche Telekom
Darmstadt
Germany
email a.mattheus@telekom.de

Manuel Paul
Deutsche Telekom
Berlin
Germany
email Manuel.Paul@telekom.de

Frank Luennemann
Deutsche Telekom
Munster
Germany
email Frank.Luennemann@telekom.de

Scott Mansfield
Ericsson Inc.
email scott.mansfield@ericsson.com

Najam Saquib
Cisco
Ludwig-Erhard-Strasse 3
ESCHBORN, HESSEN 65760
GERMANY
email nasaquib@cisco.com

Walid Wakim
Cisco
9501 Technology Blvd
ROSEMONT, ILLINOIS 60018
UNITED STATES
email wwakim@cisco.com

Ori Gerstel
Cisco
32 HaMelacha St., (HaSharon Bldg)
SOUTH NETANYA, HAMERKAZ 42504
ISRAEL
email ogerstel@cisco.com

12. References

12.1. Normative References

- [RFC2863] McCloghrie, K. and F. Kastenholz, "The Interfaces Group MIB", RFC 2863, June 2000.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, April 1999.
- [RFC2580] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Conformance Statements for SMIv2", STD 58, RFC 2580, April 1999.
- [RFC3591] Lam, H-K., Stewart, M., and A. Huynh, "Definitions of Managed Objects for the Optical Interface Type", RFC 3591, September 2003.
- [ITU.G698.2] International Telecommunications Union, "Amplified multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.
- [ITU.G709] International Telecommunications Union, "Interface for the Optical Transport Network (OTN)", ITU-T Recommendation G.709, March 2003.
- [ITU.G872] International Telecommunications Union, "Architecture of optical transport networks", ITU-T Recommendation G.872, November 2001.
- [ITU.G798] International Telecommunications Union, "Characteristics of optical transport network hierarchy equipment functional blocks", ITU-T Recommendation G.798, October 2010.
- [ITU.G874] International Telecommunications Union, "Management

aspects of optical transport network elements", ITU-T Recommendation G.874, July 2010.

[ITU.G874.1]

International Telecommunications Union, "Optical transport network (OTN): Protocol-neutral management information model for the network element view", ITU-T Recommendation G.874.1, January 2002.

[ITU.G959.1]

International Telecommunications Union, "Optical transport network physical layer interfaces", ITU-T Recommendation G.959.1, November 2009.

[ITU.G826]

International Telecommunications Union, "End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections", ITU-T Recommendation G.826, November 2009.

[ITU.G8201]

International Telecommunications Union, "Error performance parameters and objectives for multi-operator international paths within the Optical Transport Network (OTN)", ITU-T Recommendation G.8201, April 2011.

[ITU.G694.1]

International Telecommunications Union, "Spectral grids for WDM applications: DWDM frequency grid", ITU-T Recommendation G.694.1, June 2002.

[ITU.G7710]

International Telecommunications Union, "Common equipment management function requirements", ITU-T Recommendation G.7710, May 2008.

12.2. Informative References

- [RFC3410] Case, J., Mundy, R., Partain, D., and B. Stewart, "Introduction and Applicability Statements for Internet-Standard Management Framework", RFC 3410, December 2002.
- [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.
- [RFC4181] Heard, C., "Guidelines for Authors and Reviewers of MIB Documents", BCP 111, RFC 4181, September 2005.

[I-D.kunze-g-698-2-management-control-framework]
Kunze, R., "A framework for Management and Control of
optical interfaces supporting G.698.2",
draft-kunze-g-698-2-management-control-framework-00 (work
in progress), July 2011.

[RFC4054] Strand, J. and A. Chiu, "Impairments and Other Constraints
on Optical Layer Routing", RFC 4054, May 2005.

Appendix A. Change Log

This optional section should be removed before the internet draft is
submitted to the IESG for publication as an RFC.

Note to RFC Editor: please remove this appendix before publication as
an RFC.

Appendix B. Open Issues

Note to RFC Editor: please remove this appendix before publication as
an RFC.

Authors' Addresses

Gabriele Galimberti (editor)
Cisco
Via Philips,12
20052 - Monza
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

Ruediger Kunze (editor)
Deutsche Telekom
Dddd, xx
Berlin
Germany

Phone: +49xxxxxxxxxxxx
Email: RKunze@telekom.de

Hing-Kam Lam (editor)
Alcatel-Lucent
600-700 Mountain Avenue, Murray Hill
New Jersey, 07974
USA

Phone: +1732331347
Email: kam.lam@alcatel-lucent.com

Dharini Hiremagalur (editor)
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 California
USA

Phone: +1408
Email: dharinih@juniper.net

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2013

Rakesh Gandhi
Zafar Ali
Gabriele Maria Galimberti
Cisco Systems, Inc.
October 15, 2012

RSVP-TE Extensions For Signaling GMPLS Restoration LSP
draft-gandhi-ccamp-gmpls-restoration-lsp-00

Abstract

Generalized Multi-Protocol Label Switching (GMPLS) RSVP-TE recovery signaling extensions are specified in [RFC4872] and [RFC4873]. In transport networks, there are requirements where GMPLS recovery scheme need to employ restoration LSP while keeping resources for the working and/ or protecting LSPs reserved in the network. Currently GMPLS recovery procedures do not address these requirements. This document proposes RSVP-TE extensions for GMPLS recovery with restoration LSP.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 18, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

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

Table of Contents

1. Introduction	3
2. Conventions used in this document	4
3. Signaling Extensions	5
3.1. Signaling Procedure	5
3.2. Resource Sharing	6
3.3. Protection Switchover	7
3.4. Compatibility	7
4. IANA Considerations	7
5. Security Considerations	8
6. Acknowledgement	8
7. References	8
7.1. Normative references	8
7.2. Informative References	9
Authors' Addresses	9

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) extends MPLS to include support for different switching technologies [RFC3471] [RFC3473]. These switching technologies provide several protection schemes [RFC4426][RFC4427] (e.g. 1+1, 1:N and M:N). GMPLS RSVP-TE signaling has been extended to support various recovery schemes to establish Label Switched Paths (LSPs) [RFC4872][RFC4873], typically working LSP and protecting LSP. [RFC4427] Section 7 specifies various schemes for GMPLS restoration.

In GMPLS recovery schemes currently considered, restoration LSP is signaled after the failure has been detected and notified on the working LSP. These schemes assume that working LSP is removed from the network before restoration LSP is signaled. In transport network, as working LSP are typically signaled over a nominal path, there are many scenarios where service providers would like to keep resources associated with the working LSP reserved. This is to make sure that the service (working LSP) can use the nominal path when the failure is repaired. Consequently, in transport networks one can employ a recovery scheme where a new restoration LSP is signaled while working LSP and/ or protecting LSP are not torn down in control plane due to a failure. Restoration LSP differs from a secondary LSP in the way that secondary LSP does not reserve resources in the data plane and is not able to carry any traffic until it is refreshed whereas restoration LSP does reserve resources and is able to carry traffic.

One example of the recovery scheme considered in this draft is 1:R recovery. 1:R recovery is exemplified in Figure 1. In this example, working LSP on path A-B-C-Z is pre-established. Typically after a failure detection and notification on the working LSP, a second LSP on path A-H-I-J-Z is established as a restoration LSP. Unlike protection LSP, restoration LSP is signaled on as needed basis.

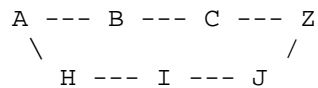


Figure 1: An example of 1:R recovery scheme

During failure, working LSP resources are not released and working and restoration LSPs coexist in the network. Nonetheless, working and restoration LSPs can share network resources. Typically when failure is recovered on the working LSP, restoration LSP is no longer required and torn down (e.g. revertive mode).

Another example of the recovery scheme considered in this draft is 1+1:R. In 1+1:R, a restoration LSP is signaled for the working LSP and/ or the protecting LSP after the failure has been detected and notified on the working LSP or the protecting LSP. The 1+1:R recovery is exemplified in Figure 2. In this example, working LSP on path A-B-C-Z and protecting LSP on path A-D-E-F-Z are pre-established. After a failure detection and notification on a working LSP or protecting LSP, a third LSP on path A-H-I-J-Z is established as a restoration LSP. The restoration LSP in this case provides protection against a second order failure. Restoration LSP is torn down when the failure on the working or protecting LSP is repaired.

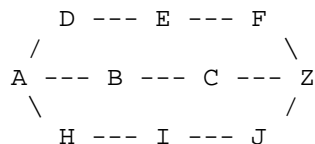


Figure 2: An example of 1+1:R recovery scheme

[RFC4872] Section 14 defines PROTECTION object for GMPLS recovery signaling. The PROTECTION object is used to identify primary and secondary LSPs using S bit and protecting and working LSPs using P bit. However, the PROTECTION object does not have a way to identify restoration LSP or signal protection type for the type of recovery considered by this document. This document defines a new flag in the RSVP PROTECTION object [RFC4872] [RFC4873] to identify the GMPLS restoration LSP and new LSP flags to signal LSP protection types addressed by this draft.

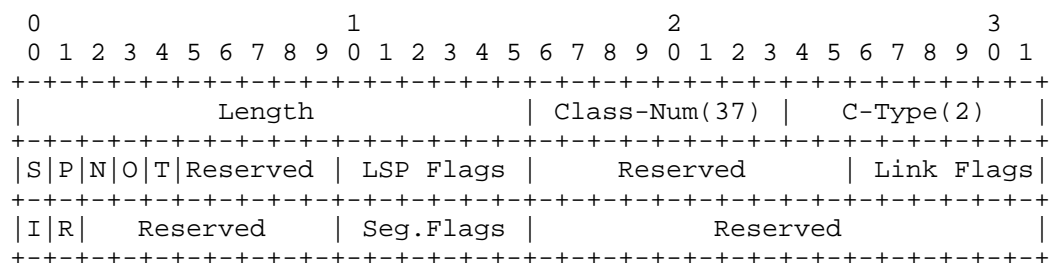
2. Conventions used in this document

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 [RFC2119].

3. Signaling Extensions

3.1. Signaling Procedure

The PROTECTION object [RFC4873] [RFC4872] format has been defined as follows. A new bit named T bit is defined in this object to indicate GMPLS restoration LSP.



Restoration (T) : 1 bit

When set to 1, this bit indicates that the request LSP is a restoration LSP. Working LSP is signaled with P bit set to 0 and T bit set to 0. Protecting LSP is signaled with P set to 1 and T bit set to 0.

When a restore LSP is signaled due to a failed protecting LSP, P bit and T bit MUST be set to 1. When a restore LSP is signaled due to failed working LSP, P bit is set to 0 and T bit is set to 1. When T bit is set to 1 and S bit set 1 is permitted but usage is outside the scope of this document.

LSP (Protection Type) Flags: 6 bits

Indicates the desired end-to-end LSP protection type. Following new values are defined in this document in addition to the ones defined in [RFC4872].

0x20 (TBD)	(Full) Rerouting with Restoration (1:R)
0x21 (TBD)	1:N:R Protection with Extra-Traffic with Restoration
0x22 (TBD)	1+1:R Unidirectional Protection with Restoration
0x23 (TBD)	1+1:R Bidirectional Protection with Restoration

The procedure for signaling all other fields in the PROTECTION object is specified in [RFC4872] and does not change other than specified in this section. As P bit and S bit are preserved when using restoration LSP, LSP protection types and LSP recovery procedures defined in [RFC4872] and [RFC4873] apply. Specifically, protection schemes defined in [RFC4872] namely 1+1 unidirectional protection, 1+1 bidirectional protection, (full) LSP rerouting, 1:N protection with extra-traffic do not change with the introduction of the restoration LSP.

A GMPLS recovery scheme "Rerouting without Extra-Traffic with Restoration LSP" is outside the scope of this document.

When using a GMPLS recovery mode, where working or protecting LSP are destroyed, and restoration LSP assumes the role of a working LSP or a protecting LSP, restoration LSP RSVP Path message MUST be refreshed by clearing the T-bit in the PROTECTION object.

3.2. Resource Sharing

Resource sharing may be desired between working LSP and restoration LSP or protecting LSP and restoration LSP. Resource sharing is typically achieved using the make-before-break procedures defined in [RFC3209]. It may not be desired to share resources between a working LSP and a protecting LSP. ASSOCIATION object with association type "resource sharing" defined in [RFC4873] is used to identify LSPs that can share resources by matching ASSOCIATION objects in the LSPs. When ASSOCIATION object is signaled with association type "resource sharing", reservation style present in the LSP is ignored and for all matching LSPs with identical ASSOCIATION objects are requested to share resources. For LSPs with non matching ASSOCIATION object or absence of the ASSOCIATION object, reservation style [FF/SE] present in the LSP is used for resource sharing as per [RFC3209].

3.3. Protection Switchover

Specific use case of the restoration LSP with protection switchover is outside the scope of this document. Also, protecting switching co-ordination (PSC) [RFC6378] mechanism is outside the scope of this document.

3.4. Compatibility

The PROTECTION object has already been defined with class numbers in the form 1lbbbbbb, which ensures compatibility with non-supporting nodes. Per [RFC2205], nodes not supporting this extension will ignore the object or the new T bit and LSP flag but forward it, unexamined and unmodified, in all messages resulting from this message.

4. IANA Considerations

IANA is requested to administer assignment of new values for namespace defined in this document and summarized in this section.

Within the current document, one new flag is defined in the PROTECTION object.

Value	Type
-----	-----
T bit (TBD)	Restoration LSP Type

Within the current document, following new LSP flags are defined in the PROTECTION object.

0x20 (TBD)	(Full) Rerouting with Restoration (1:R)
0x21 (TBD)	1:N:R Protection with Extra-Traffic with Restoration
0x22 (TBD)	1+1:R Unidirectional Protection with Restoration
0x23 (TBD)	1+1:R Bidirectional Protection with Restoration

5. Security Considerations

This document introduces no additional security considerations. For a general discussion on MPLS and GMPLS related security issues, see the MPLS/GMPLS security framework [RFC5920]. In addition, the considerations specified in [RFC4872] and [RFC4873] will apply.

6. Acknowledgement

The authors would like to thank George Swallow for the discussion on the GMPLS restoration.

7. References

7.1. Normative references

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.
- [RFC3471] Berger, L., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4872] Lang, J., Rekhter, Y., and D. Papadimitriou, "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.
- [RFC6378] Weingarten, Y., Bryant, S. Osborne, E. Sprecher, N., and A. Fulignoli, "MPLS Transport Profile (MPLS-TP) Linear Protection", RFC 6378, October 2011

7.2. Informative References

- [RFC4426] Lang, J., Rajagopalan B., and D.Papadimitriou, Editors,
 "Generalized Multiprotocol Label Switching (GMPLS)
 Recovery Functional Specification", RFC 4426, March 2006.
- [RFC4427] Mannie, E., Ed. and D. Papadimitriou, Ed., "Recovery
 (Protection and Restoration) Terminology for Generalized
 Multi-Protocol Label Switching, RFC 4427, March 2006.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS
 Networks", RFC 5920, July 2010.

Authors' Addresses

Rakesh Gandhi
Cisco Systems, Inc.

Email: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

Email: zali@cisco.com

Gabriele Maria Galimberti
Cisco Systems, Inc.

Email: ggalimbe@cisco.com

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2013

G. Galimberti, Ed.
G. Martinelli, Ed.
Cisco
D. Hiremagalur
G. Grammel
Juniper
July 6, 2012

A SNMP MIB to manage GMPLS TED with WSON specific support
draft-gmngm-ccamp-wson-snmp-mib-00

Abstract

This memo defines a portion of the Management Information Base (MIB) used by Simple Network Management Protocol (SNMP) for GMPLS based networks.

In particular in the context Wavelength Switching Optical Network (WSON) two sets of information were defined: a general constrains set (reusable by other technologies) and a WSON specific set. This document defines a MIB module for supporting GMPLS WSON specific information.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 7, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. The Internet-Standard Management Framework	3
3. Conventions	3
4. Overview	4
5. Structure of the MIB Module	4
5.1. tedWsonNodeTable	4
5.1.1. tedWsonResourceBlockInformation	4
5.1.2. tedWsonResourcePoolAccessibility	4
5.1.3. tedWsonResourceBlockWavelengthConstraints	4
5.1.4. tedWsonResourcePoolState	4
5.1.5. tedWsonBlockSharedAccessWavelengthAvailability	4
6. Relationship to Other MIB Modules	4
6.1. Relationship to the [TEMPLATE TODO] MIB	4
6.2. MIB modules required for IMPORTS	4
7. Definitions	4
8. Security Considerations	4
9. IANA Considerations	5
10. Contributors	6
11. References	6
11.1. Normative References	6
11.2. Informative References	7
Appendix A. Change Log	7
Appendix B. Open Issues	8
Authors' Addresses	8

1. Introduction

This memo defines a portion of the Management Information Base (MIB) used by Simple Network Management Protocol (SNMP) in GMPLS networks.

Extensions to current GMPLS to support Wavelength Switched Optical Networks (WSON) [RFC6163] include new objects with specific protocol extensions. Some information was selected as a generic constraint since they could be easily applied to other technologies than WSON. As such this [I-D.ietf-ccamp-gmpls-general-constraints-ospf-te] OSPF-TE was proposed and those information will be managed through a separated MIB [ref required].

In case of WSON some technology specific information are required and defined through [I-D.ietf-ccamp-rwa-info] and [I-D.ietf-ccamp-wson-signal-compatibility-ospf]. This MIB module will define objects related to WSON specific information.

[EDITOR NOTE] Very early draft to start MIB activity on GMPLS-WSON related extensions and collect feedback from working group.

2. The Internet-Standard Management Framework

For a detailed overview of the documents that describe the current Internet-Standard Management Framework, please refer to section 7 of RFC 3410 [RFC3410].

Managed objects are accessed via a virtual information store, termed the Management Information Base or MIB. MIB objects are generally accessed through the Simple Network Management Protocol (SNMP). Objects in the MIB are defined using the mechanisms defined in the Structure of Management Information (SMI). This memo specifies a MIB module that is compliant to the SMIv2, which is described in STD 58, RFC 2578 [RFC2578], STD 58, RFC 2579 [RFC2579] and STD 58, RFC 2580 [RFC2580].

3. Conventions

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]. In the description of OIDs the convention: Set (S) Get (G) and Trap (T) conventions will describe the action allowed by the parameter.

4. Overview

This MIB module should be used in conjunction with [I-D.ietf-ccamp-gmpls-ted-mib] since it only defines additional parameters to GMPLS TED MIB.

5. Structure of the MIB Module

5.1. tedWsonNodeTable

5.1.1. tedWsonResourceBlockInformation

5.1.2. tedWsonResourcePoolAccessibility

5.1.3. tedWsonResourceBlockWavelengthConstraints

5.1.4. tedWsonResourcePoolState

5.1.5. tedWsonBlockSharedAccessWavelengthAvailability

6. Relationship to Other MIB Modules

6.1. Relationship to the [TEMPLATE TODO] MIB

6.2. MIB modules required for IMPORTS

7. Definitions

[TEMPLATE TODO]: put your valid MIB module here.
A list of tools that can help automate the process of checking MIB definitions can be found at <http://www.ops.ietf.org/mib-review-tools.html>

8. Security Considerations

There are a number of management objects defined in this MIB module with a MAX-ACCESS clause of read-write and/or read-create. Such objects may be considered sensitive or vulnerable in some network environments. The support for SET operations in a non-secure environment without proper protection can have a negative effect on network operations. These are the tables and objects and their sensitivity/vulnerability:

o

Some of the readable objects in this MIB module (i.e., objects with a MAX-ACCESS other than not-accessible) may be considered sensitive or vulnerable in some network environments. It is thus important to control even GET and/or NOTIFY access to these objects and possibly to even encrypt the values of these objects when sending them over the network via SNMP.

SNMP versions prior to SNMPv3 did not include adequate security. Even if the network itself is secure (for example by using IPsec), even then, there is no control as to who on the secure network is allowed to access and GET/SET (read/change/create/delete) the objects in this MIB module.

It is RECOMMENDED that implementers consider the security features as provided by the SNMPv3 framework (see [RFC3410], section 8), including full support for the SNMPv3 cryptographic mechanisms (for authentication and privacy).

Further, deployment of SNMP versions prior to SNMPv3 is NOT RECOMMENDED. Instead, it is RECOMMENDED to deploy SNMPv3 and to enable cryptographic security. It is then a customer/operator responsibility to ensure that the SNMP entity giving access to an instance of this MIB module is properly configured to give access to the objects only to those principals (users) that have legitimate rights to indeed GET or SET (change/create/delete) them.

9. IANA Considerations

Option #1:

The MIB module in this document uses the following IANA-assigned OBJECT IDENTIFIER values recorded in the SMI Numbers registry:

Descriptor	OBJECT IDENTIFIER value
-----	-----
sampleMIB	{ mib-2 XXX }

Option #2:

Editor's Note (to be removed prior to publication): the IANA is requested to assign a value for "XXX" under the 'mib-2' subtree and to record the assignment in the SMI Numbers registry. When the assignment has been made, the RFC Editor is asked to replace "XXX"

(here and in the MIB module) with the assigned value and to remove this note.

Note well: prior to official assignment by the IANA, an internet draft MUST use placeholders (such as "XXX" above) rather than actual numbers. See RFC4181 Section 4.5 for an example of how this is done in an internet draft MIB module.

Option #3:

This memo includes no request to IANA.

10. Contributors

to be added.

11. References

11.1. Normative References

- [RFC2863] McCloghrie, K. and F. Kastenholz, "The Interfaces Group MIB", RFC 2863, June 2000.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, April 1999.
- [RFC2580] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Conformance Statements for SMIv2", STD 58, RFC 2580, April 1999.
- [I-D.ietf-ccamp-gmpls-general-constraints-ospf-te]
Zhang, F., Lee, Y., Han, J., Bernstein, G., and Y. Xu,
"OSPF-TE Extensions for General Network Element
Constraints",
draft-ietf-ccamp-gmpls-general-constraints-ospf-te-03
(work in progress), June 2012.
- [I-D.ietf-ccamp-wson-signal-compatibility-ospf]

Lee, Y. and G. Bernstein, "GMPLS OSPF Enhancement for Signal and Network Element Compatibility for Wavelength Switched Optical Networks", draft-ietf-ccamp-wson-signal-compatibility-ospf-08 (work in progress), April 2012.

[I-D.ietf-ccamp-gmpls-ted-mib]
Miyazawa, M., Otani, T., Kumaki, K., and T. Nadeau, "Traffic Engineering Database Management Information Base in support of MPLS-TE/GMPLS", draft-ietf-ccamp-gmpls-ted-mib-13 (work in progress), May 2012.

11.2. Informative References

[RFC3410] Case, J., Mundy, R., Partain, D., and B. Stewart, "Introduction and Applicability Statements for Internet-Standard Management Framework", RFC 3410, December 2002.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

[RFC4181] Heard, C., "Guidelines for Authors and Reviewers of MIB Documents", BCP 111, RFC 4181, September 2005.

[RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

[I-D.ietf-ccamp-rwa-info]
Lee, Y., Bernstein, G., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info-14 (work in progress), March 2012.

Appendix A. Change Log

This optional section should be removed before the internet draft is submitted to the IESG for publication as an RFC.

Note to RFC Editor: please remove this appendix before publication as an RFC.

Appendix B. Open Issues

Note to RFC Editor: please remove this appendix before publication as an RFC.

Authors' Addresses

Gabriele M. Galimberti (editor)
Cisco
Via Philips,12
20900 - Monza
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

Giovanni Martinelli (editor)
Cisco
Via Philips,12
20900 - Monza
Italy

Email: giomarti@cisco.com

Dharini Hiremagalur
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 CA
USA

Email: dharinih@juniper.net

Gert Grammel
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 CA
USA

Email: ggrammel@juniper.net

Network Working Group
Internet Draft
Category: Informational

Fatai Zhang, Ed.
Dan Li
Huawei
Han Li
CMCC
S. Belotti
Alcatel-Lucent
D. Ceccarelli
Ericsson
August 25, 2012

Expires: February 25, 2013

Framework for GMPLS and PCE Control of
G.709 Optical Transport Networks

draft-ietf-ccamp-gmpls-g709-framework-09.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on February 25, 2013.

Abstract

This document provides a framework to allow the development of protocol extensions to support Generalized Multi-Protocol Label Switching (GMPLS) and Path Computation Element (PCE) control of

Optical Transport Networks (OTN) as specified in ITU-T Recommendation G.709 as consented in October 2009.

Table of Contents

1. Introduction	2
2. Terminology	3
3. G.709 Optical Transport Network (OTN)	4
3.1. OTN Layer Network	4
3.1.1. Client signal mapping	5
3.1.2. Multiplexing ODUj onto Links	6
3.1.2.1. Structure of MSI information	8
4. Connection management in OTN	9
4.1. Connection management of the ODU	10
5. GMPLS/PCE Implications	12
5.1. Implications for LSP Hierarchy with GMPLS TE	12
5.2. Implications for GMPLS Signaling	13
5.3. Implications for GMPLS Routing	15
5.4. Implications for Link Management Protocol (LMP)	17
5.5. Implications for Control Plane Backward Compatibility ...	18
5.6. Implications for Path Computation Elements	19
6. Data Plane Backward Compatibility Considerations	20
7. Security Considerations	20
8. IANA Considerations	21
9. Acknowledgments	21
10. References	21
10.1. Normative References	21
10.2. Informative References	22
11. Authors' Addresses	23
12. Contributors	24
APPENDIX A: ODU connection examples	25

1. Introduction

OTN has become a mainstream layer 1 technology for the transport network. Operators want to introduce control plane capabilities based on Generalized Multi-Protocol Label Switching (GMPLS) to OTN networks, to realize the benefits associated with a high-function control plane (e.g., improved network resiliency, resource usage efficiency, etc.).

GMPLS extends MPLS to encompass time division multiplexing (TDM) networks (e.g., SONET/SDH, PDH, and G.709 sub-lambda), lambda switching optical networks, and spatial switching (e.g., incoming

port or fiber to outgoing port or fiber). The GMPLS architecture is provided in [RFC3945], signaling function and Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) extensions are described in [RFC3471] and [RFC3473], routing and OSPF extensions are described in [RFC4202] and [RFC4203], and the Link Management Protocol (LMP) is described in [RFC4204].

The GMPLS protocol suite including provision [RFC4328] provides the mechanisms for basic GMPLS control of OTN networks based on the 2001 revision of the G.709 specification [G709-V1]. Later revisions of the G.709 specification, including [G709-V3], have included some new features; for example, various multiplexing structures, two types of TSs (i.e., 1.25Gbps and 2.5Gbps), and extension of the Optical Data Unit (ODU) ODUj definition to include the ODUflex function.

This document reviews relevant aspects of OTN technology evolution that affect the GMPLS control plane protocols and examines why and how to update the mechanisms described in [RFC4328]. This document additionally provides a framework for the GMPLS control of OTN networks and includes a discussion of the implication for the use of the Path Computation Element (PCE) [RFC4655].

For the purposes of the control plane the OTN can be considered as being comprised of ODU and wavelength (OCh) layers. This document focuses on the control of the ODU layer, with control of the wavelength layer considered out of the scope. Please refer to [RFC6163] for further information about the wavelength layer.

2. Terminology

OTN: Optical Transport Network

ODU: Optical Channel Data Unit

OTU: Optical channel transport unit

OMS: Optical multiplex section

MSI: Multiplex Structure Identifier

TPN: Tributary Port Number

LO ODU: Lower Order ODU. The LO ODUj (j can be 0, 1, 2, 2e, 3, 4, flex.) represents the container transporting a client of the OTN that

is either directly mapped into an OTUk ($k = j$) or multiplexed into a server HO ODUk ($k > j$) container.

HO ODU: Higher Order ODU. The HO ODUk (k can be 1, 2, 2e, 3, 4.) represents the entity transporting a multiplex of LO ODUj tributary signals in its OPUk area.

ODUflex: Flexible ODU. A flexible ODUk can have any bit rate and a bit rate tolerance up to +/-100 ppm.

3. G.709 Optical Transport Network (OTN)

This section provides an informative overview of those aspects of the OTN impacting control plane protocols. This overview is based on the ITU-T Recommendations that contain the normative definition of the OTN. Technical details regarding OTN architecture and interfaces are provided in the relevant ITU-T Recommendations.

Specifically, [G872-2001] and [G872Am2] describe the functional architecture of optical transport networks providing optical signal transmission, multiplexing, routing, supervision, performance assessment, and network survivability. [G709-V1] defines the interfaces of the optical transport network to be used within and between subnetworks of the optical network. With the evolution and deployment of OTN technology many new features have been specified in ITU-T recommendations, including for example, new ODU0, ODU2e, ODU4 and ODUflex containers as described in [G709-V3].

3.1. OTN Layer Network

The simplified signal hierarchy of OTN is shown in Figure 1, which illustrates the layers that are of interest to the control plane. Other layers below OCh (e.g. Optical Transmission Section - OTS) are not included in this Figure. The full signal hierarchy is provided in [G709-V3].

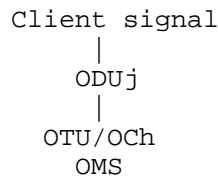


Figure 1 - Basic OTN signal hierarchy

Client signals are mapped into ODU_j containers. These ODU_j containers are multiplexed onto the OTU/OCh. The individual OTU/OCh signals are combined in the Optical Multiplex Section (OMS) using WDM multiplexing, and this aggregated signal provides the link between the nodes.

3.1.1. Client signal mapping

The client signals are mapped into a Low Order (LO) ODU_j. Appendix A gives more information about LO ODU.

The current values of *j* defined in [G709-V3] are: 0, 1, 2, 2e, 3, 4, Flex. The approximate bit rates of these signals are defined in [G709-V3] and are reproduced in Tables 1 and 2.

ODU Type	ODU nominal bit rate
ODU0	1 244 160 kbits/s
ODU1	239/238 x 2 488 320 kbit/s
ODU2	239/237 x 9 953 280 kbit/s
ODU3	239/236 x 39 813 120 kbit/s
ODU4	239/227 x 99 532 800 kbit/s
ODU2e	239/237 x 10 312 500 kbit/s
ODUflex for CBR Client signals	239/238 x client signal bit rate
ODUflex for GFP-F Mapped client signal	Configured bit rate

Table 1 - ODU types and bit rates

NOTE - The nominal ODU_k rates are approximately: 2 498 775.126 kbit/s (ODU1), 10 037 273.924 kbit/s (ODU2), 40 319 218.983 kbit/s (ODU3), 104 794 445.815 kbit/s (ODU4) and 10 399 525.316 kbit/s (ODU2e).

ODU Type	ODU bit-rate tolerance
ODU0	+ - 20 ppm
ODU1	+ - 20 ppm
ODU2	+ - 20 ppm
ODU3	+ - 20 ppm
ODU4	+ - 20 ppm
ODU2e	+ - 100 ppm
ODUflex for CBR Client signals	+ - 100 ppm
ODUflex for GFP-F Mapped client signal	+ - 100 ppm

Table 2 - ODU types and tolerance

One of two options is for mapping client signals into ODUflex depending on the client signal type:

- Circuit clients are proportionally wrapped. Thus the bit rate and tolerance are defined by the client signal.
- Packet clients are mapped using the Generic Framing Procedure (GFP). [G709-V3] recommends that the ODUflex(GFP) will fill an integral number of tributary slots of the smallest HO ODUk path over which the ODUflex(GFP) may be carried, and the tolerance should be +/-100ppm.

3.1.2. Multiplexing ODUj onto Links

The links between the switching nodes are provided by one or more wavelengths. Each wavelength carries one OCh, which carries one OTU, which carries one ODU. Since all of these signals have a 1:1:1 relationship, we only refer to the OTU for clarity. The ODUjs are mapped into the TS of the OPUk. Note that in the case where j=k the ODUj is mapped into the OTU/OCh without multiplexing.

The initial versions of G.709 [G709-V1] only provided a single TS granularity, nominally 2.5Gb/s. [G709-V3], approved in 2009, added an additional TS granularity, nominally 1.25Gb/s. The number and type of

TSs provided by each of the currently identified OTUk is provided below:

	2.5Gb/s	1.25Gb/s	Nominal Bit rate
OTU1	1	2	2.5Gb/s
OTU2	4	8	10Gb/s
OTU3	16	32	40Gb/s
OTU4	--	80	100Gb/s

To maintain backwards compatibility while providing the ability to interconnect nodes that support 1.25Gb/s TS at one end of a link and 2.5Gb/s TS at the other, the 'new' equipment will fall back to the use of a 2.5Gb/s TS if connected to legacy equipment. This information is carried in band by the payload type.

The actual bit rate of the TS in an OTUk depends on the value of k. Thus the number of TS occupied by an ODUj may vary depending on the values of j and k. For example an ODU2e uses 9 TS in an OTU3 but only 8 in an OTU4. Examples of the number of TS used for various cases are provided below:

- ODU0 into ODU1, ODU2, ODU3 or ODU4 multiplexing with 1,25Gbps TS granularity
 - o ODU0 occupies 1 of the 2, 8, 32 or 80 TS for ODU1, ODU2, ODU3 or ODU4
- ODU1 into ODU2, ODU3 or ODU4 multiplexing with 1,25Gbps TS granularity
 - o ODU1 occupies 2 of the 8, 32 or 80 TS for ODU2, ODU3 or ODU4
- ODU1 into ODU2, ODU3 multiplexing with 2.5Gbps TS granularity
 - o ODU1 occupies 1 of the 4 or 16 TS for ODU2 or ODU3
- ODU2 into ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU2 occupies 8 of the 32 or 80 TS for ODU3 or ODU4
- ODU2 into ODU3 multiplexing with 2.5Gbps TS granularity
 - o ODU2 occupies 4 of the 16 TS for ODU3
- ODU3 into ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU3 occupies 31 of the 80 TS for ODU4
- ODUflex into ODU2, ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity

- o ODUflex occupies n of the 8, 32 or 80 TS for ODU2, ODU3 or ODU4 ($n \leq \text{Total TS numbers of ODUk}$)
- ODU2e into ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU2e occupies 9 of the 32 TS for ODU3 or 8 of the 80 TS for ODU4

In general the mapping of an ODU j (including ODUflex) into the OTU k TSs is determined locally, and it can also be explicitly controlled by a specific entity (e.g., head end, NMS) through Explicit Label Control [RFC3473].

3.1.2.1. Structure of MSI information

When multiplexing an ODU j into a HO ODU k ($k > j$), G.709 specifies the information that has to be transported in-band in order to allow for correct demultiplexing. This information, known as Multiplex Structure Information (MSI), is transported in the OPU k overhead and is local to each link. In case of bidirectional paths the association between TPN and TS must be the same in both directions.

The MSI information is organized as a set of entries, with one entry for each HO ODU j TS. The information carried by each entry is:

- Payload Type: the type of the transported payload.
- Tributary Port Number (TPN): the port number of the ODU j transported by the HO ODU k . The TPN is the same for all the TSs assigned to the transport of the same ODU j instance.

For example, an ODU2 carried by a HO ODU3 is described by 4 entries in the OPU3 overhead when the TS size is 2.5 Gbit/s, and by 8 entries when the TS size is 1.25 Gbit/s.

On each node and on every link, two MSI values have to be provisioned:

- The TxMSI information inserted in OPU (e.g., OPU3) overhead by the source of the HO ODU k trail.
- The expectedMSI information that is used to check the acceptedMSI information. The acceptedMSI information is the MSI valued received in-band, after a 3 frames integration.

The sink of the HO ODU trail checks the complete content of the acceptedMSI information against the expectedMSI.

If the acceptedMSI is different from the expectedMSI, then the traffic is dropped and a payload mismatch alarm is generated.

Provisioning of TPN can be performed either by network management system or control plane. In the last case, control plane is also responsible for negotiating the provisioned values on a link by link base.

4. Connection management in OTN

OTN-based connection management is concerned with controlling the connectivity of ODU paths and optical channels (OCh). This document focuses on the connection management of ODU paths. The management of OCh paths is described in [RFC6163].

While [G872-2001] considered the ODU as a set of layers in the same way as SDH has been modeled, recent ITU-T OTN architecture progress [G872-Am2] includes an agreement to model the ODU as a single layer network with the bit rate as a parameter of links and connections. This allows the links and nodes to be viewed in a single topology as a common set of resources that are available to provide ODU_j connections independent of the value of j. Note that when the bit rate of ODU_j is less than the server bit rate, ODU_j connections are supported by HO-ODU (which has a one-to-one relationship with the OTU).

From an ITU-T perspective, the ODU connection topology is represented by that of the OTU link layer, which has the same topology as that of the OCh layer (independent of whether the OTU supports HO-ODU, where multiplexing is utilized, or LO-ODU in the case of direct mapping). Thus, the OTU and OCh layers should be visible in a single topological representation of the network, and from a logical perspective, the OTU and OCh may be considered as the same logical, switchable entity.

Note that the OTU link layer topology may be provided via various infrastructure alternatives, including point-to-point optical connections, flexible optical connections fully in the optical domain, flexible optical connections involving hybrid sub-lambda/lambda nodes involving 3R, etc.

The document will be updated to maintain consistency with G.872 progress when it is consented for publication.

4.1. Connection management of the ODU

LO ODU_j can be either mapped into the OTU_k signal ($j = k$), or multiplexed with other LO ODU_js into an OTU_k ($j < k$), and the OTU_k is mapped into an OCh. See Appendix A for more information.

From the perspective of control plane, there are two kinds of network topology to be considered.

(1) ODU layer

In this case, the ODU links are presented between adjacent OTN nodes, which is illustrated in Figure 2. In this layer there are ODU links with a variety of TSs available, and nodes that are ODXCs. Lo ODU connections can be setup based on the network topology.

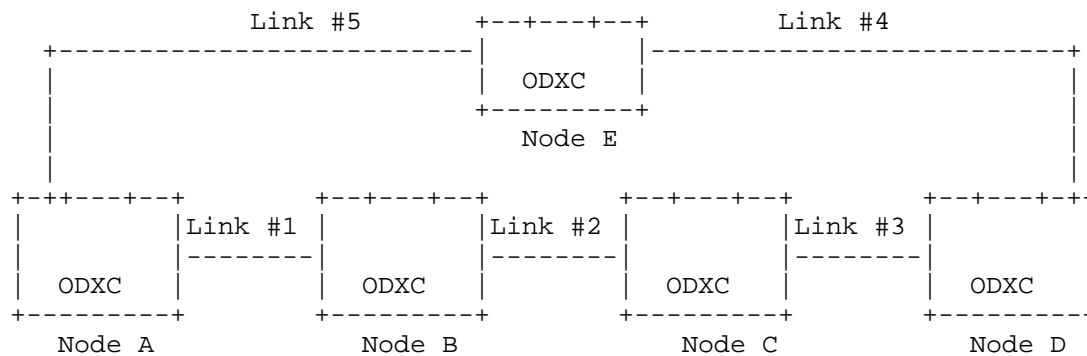


Figure 2 - Example Topology for LO ODU connection management

If an ODU_j connection is requested between Node C and Node E routing/path computation must select a path that has the required number of TS available and that offers the lowest cost. Signaling is then invoked to set up the path and to provide the information (e.g., selected TS) required by each transit node to allow the configuration of the ODU_j to OTU_k mapping ($j = k$) or multiplexing ($j < k$), and demapping ($j = k$) or demultiplexing ($j < k$).

(2) ODU layer with OCh switching capability

In this case, the OTN nodes interconnect with wavelength switched node (e.g., ROADM, OXC) that are capable of OCh switching, which is illustrated in Figure 3 and Figure 4. There are ODU layer and OCh layer, so it is simply a MLN. OCh connections may be created on demand, which is described in section 5.1.

In this case, an operator may choose to allow the underlined OCh layer to be visible to the ODU routing/path computation process in which case the topology would be as shown in Figure 4. In Figure 3 below, instead, a cloud representing OCH capable switching nodes is represented. In Figure 3, the operator choice is to hide the real RWA network topology.

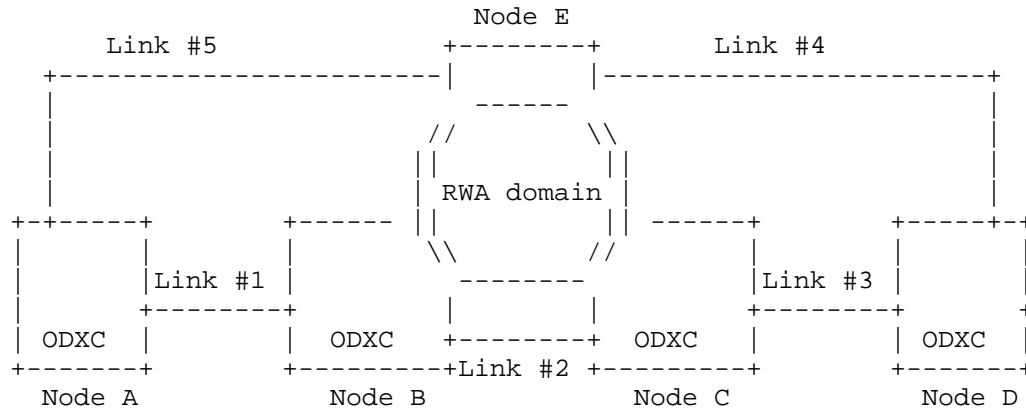


Figure 3 - RWA Hidden Topology for LO ODU connection management

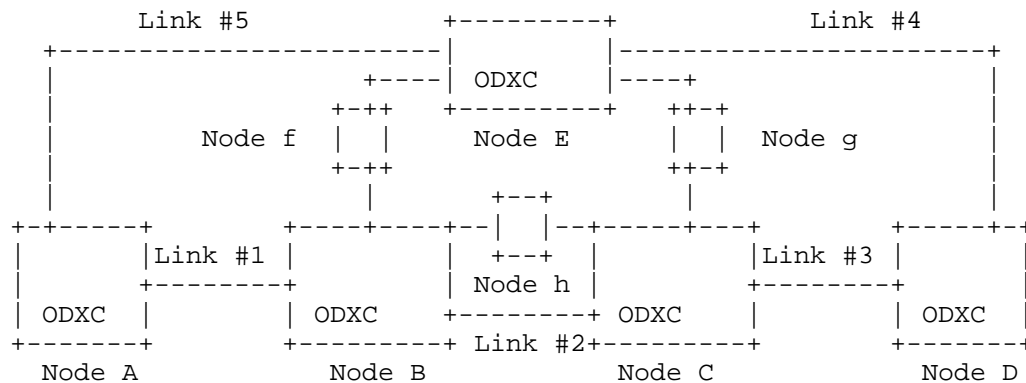


Figure 4 - RWA Visible Topology for LO ODUj connection management

In Figure 4, the cloud of previous figure is substitute by the real topology. The nodes f, g, h are nodes with OCH switching capability.

In the examples (i.e., Figure 3 and Figure 4), we have considered the case in which LO-ODUj connections are supported by OCh connection, and the case in which the supporting underlying connection can be also made by a combination of HO-ODU/OCh connections.

In this case, the ODU routing/path selection process will request an HO-ODU/OCh connection between node C and node E from the RWA domain. The connection will appear at ODU level as a Forwarding Adjacency, which will be used to create the ODU connection.

5. GMPLS/PCE Implications

The purpose of this section is to provide a set of requirements to be evaluated for extensions of the current GMPLS protocol suite and the PCE applications and protocols to encompass OTN enhancements and connection management.

5.1. Implications for LSP Hierarchy with GMPLS TE

The path computation for ODU connection request is based on the topology of ODU layer, including OCh layer visibility.

The OTN path computation can be divided into two layers. One layer is OCh/OTUk, the other is ODUj. [RFC4206] and [RFC6107] define the mechanisms to accomplish creating the hierarchy of LSPs. The LSP management of multiple layers in OTN can follow the procedures defined in [RFC4206], [RFC6107] and related MLN drafts.

As discussed in section 4, the route path computation for OCh is in the scope of WSON [RFC6163]. Therefore, this document only considers ODU layer for ODU connection request.

LSP hierarchy can also be applied within the ODU layers. One of the typical scenarios for ODU layer hierarchy is to maintain compatibility with introducing new [G709-V3] services (e.g., ODU0, ODUflex) into a legacy network configuration (containing [G709-V1] or [G709-V2] OTN equipment). In this scenario, it may be needed to consider introducing hierarchical multiplexing capability in specific network transition scenarios. One method for enabling multiplexing hierarchy is by introducing dedicated boards in a few specific places in the network and tunneling these new services through [G709-V1] or [G709-V2] containers (ODU1, ODU2, ODU3), thus postponing the need to upgrade every network element to [G709-V3] capabilities.

In such case, one ODUj connection can be nested into another ODUk (j<k) connection, which forms the LSP hierarchy in ODU layer. The creation of the outer ODUk connection can be triggered via network planning, or by the signaling of the inner ODUj connection. For the former case, the outer ODUk connection can be created in advance based on network planning. For the latter case, the multi-layer network signaling described in [RFC4206], [RFC6107] and [RFC6001] (including related modifications, if needed) are relevant to create the ODU connections with multiplexing hierarchy. In both cases, the outer ODUk connection is advertised as a Forwarding Adjacency (FA).

5.2. Implications for GMPLS Signaling

The signaling function and Resource reSerVation Protocol-Traffic Engineering (RSVP-TE) extensions are described in [RFC3471] and [RFC3473]. For OTN-specific control, [RFC4328] defines signaling extensions to support G.709 Optical Transport Networks Control as defined in [G709-V1].

As described in Section 3, [G709-V3] introduced some new features that include the ODU0, ODU2e, ODU4 and ODUFlex containers. The mechanisms defined in [RFC4328] do not support such new OTN features, and protocol extensions will be necessary to allow them to be controlled by a GMPLS control plane.

[RFC4328] defines the LSP Encoding Type, the Switching Type and the Generalized Protocol Identifier (Generalized-PID) constituting the common part of the Generalized Label Request. The G.709 Traffic Parameters are also defined in [RFC4328]. The following signaling aspects should be considered additionally since [RFC4328] was published:

- Support for specifying the new signal types and the related traffic information

The traffic parameters should be extended in signaling message to support the new optical Channel Data Unit (ODUj) including:

- ODU0
- ODU2e
- ODU4
- ODUFlex

For ODUFlex, since it has a variable bandwidth/bit rate BR and a bit rate tolerance T, the (node local) mapping process should be aware of the bit rate and tolerance of the ODUj being multiplexed in order to select the correct number of TS and the fixed/variable

stuffing bytes. Therefore, bit rate and bit rate tolerance should also be carried in the Traffic Parameter in the signaling of connection setup request.

For other ODU signal types, the bit rates and tolerances of them are fixed and can be deduced from the signal types.

- Support for LSP setup using different Tributary Slot Granularity (TSG)

The signaling protocol should be able to identify the type of TS (i.e., the 2.5 Gbps TS granularity and the new 1.25 Gbps TS granularity) to be used for establishing an H-LSP which will be used to carry service LSP(s) requiring specific TS type.

- Support for LSP setup of new ODUk/ODUflex containers with related mapping and multiplexing capabilities

New label must be defined to carry the exact TS allocation information related to the extended mapping and multiplexing hierarchy (For example, ODU0 into ODU2 multiplexing (with 1.25Gbps TS granularity)), in order to setting up the ODU connection.

- Support for Tributary Port Number allocation and negotiation

Tributary Port Number needs to be configured as part of the MSI information (See more information in Section 3.1.2.1). A new extension object has to be defined to carry TPN information if control plane is used to configure MSI information.

- Support for ODU Virtual Concatenation (VCAT) and Link Capacity Adjustment Scheme (LCAS)

GMPLS signaling should support the creation of Virtual Concatenation of ODUk signal with $k=1, 2, 3$. The signaling should also support the control of dynamic capacity changing of a VCAT container using LCAS ([G.7042]). [RFC6344] has a clear description of VCAT and LCAS control in SONET/SDH and OTN networks.

- Support for Control of Hitless Adjustment of ODUflex (GFP)

[G.7044] has been created in ITU-T to specify hitless adjustment of ODUflex (GFP) (HAO) that is used to increase or decrease the bandwidth of an ODUflex (GFP) that is transported in an OTN network.

The procedure of ODUflex (GFP) adjustment requires the participation of every node along the path. Therefore, it is recommended to use the control plane signaling to initiate the adjustment procedure in order to avoid the manual configuration at each node along the path.

From the perspective of control plane, the control of ODUflex resizing is similar to control of bandwidth increasing and decreasing described in [RFC3209]. Therefore, the SE style can be used for control of HAO.

All the extensions above should consider the extensibility to match future evolvement of OTN.

5.3. Implications for GMPLS Routing

The path computation process needs to select a suitable route for an ODUj connection request. In order to perform the path computation, it needs to evaluate the available bandwidth on each candidate link. The routing protocol should be extended to convey some information to represent ODU TE topology.

GMPLS Routing [RFC4202] defines Interface Switching Capability Descriptor of TDM which can be used for ODU. However, some issues discussed below, should also be considered.

Interface Switching Capability Descriptors present a new constraint for LSP path computation. [RFC4203] defines the switching capability and related Maximum LSP Bandwidth and the Switching Capability specific information. When the Switching Capability field is TDM the Switching Capability Specific Information field includes Minimum LSP Bandwidth, an indication whether the interface supports Standard or Arbitrary SONET/SDH, and padding. Hence a new Switching Capability value needs to be defined for [G709-V3] ODU switching in order to allow the definition of a new Switching Capability Specific Information field definition. The following requirements should be considered:

- Support for carrying the link multiplexing capability

As discussed in section 3.1.2, many different types of ODUj can be multiplexed into the same OTUk. For example, both ODU0 and ODU1 may be multiplexed into ODU2. An OTU link may support one or more types of ODUj signals. The routing protocol should be capable of carrying this multiplexing capability.

- Support any ODU and ODUflex

The bit rate (i.e., bandwidth) of TS is dependent on the TS granularity and the signal type of the link. For example, the bandwidth of a 1.25G TS in an OTU2 is about 1.249409620 Gbps, while the bandwidth of a 1.25G TS in an OTU3 is about 1.254703729 Gbps.

One LO ODU may need different number of TSs when multiplexed into different HO ODUs. For example, for ODU2e, 9 TSs are needed when multiplexed into an ODU3, while only 8 TSs are needed when multiplexed into an ODU4. For ODUflex, the total number of TSs to be reserved in a HO ODU equals the maximum of [bandwidth of ODUflex / bandwidth of TS of the HO ODU].

Therefore, the routing protocol should be capable of carrying the necessary and sufficient link bandwidth information for performing accurate route computation for any of the fixed rate ODUs as well as ODUflex.

- Support for differentiating between terminating and switching capability

Due to internal constraints and/or limitations, the type of signal being advertised by an interface could be just switched (i.e. forwarded to switching matrix without multiplexing/demultiplexing actions), just terminated (demuxed) or both of them. The capability advertised by an interface needs further distinction in order to separate termination and switching capabilities.

Therefore, to allow the required flexibility, the routing protocol should clearly distinguish the terminating and switching capability.

- Support for Tributary Slot Granularity advertisement

[G709-V3] defines two types of TS but each link can only support a single type at a given time. In order to perform a correct path computation (i.e. the LSP end points have matching Tributary Slot Granularity values) the Tributary Slot Granularity needs to be advertised.

- Support different priorities for resource reservation

How many priorities levels should be supported depends on the operator's policy. Therefore, the routing protocol should be capable of supporting either no priorities or up to 8 priority levels as defined in [RFC4202].

- Support link bundling

Link bundling can improve routing scalability by reducing the amount of TE links that has to be handled by routing protocol. The routing protocol should be capable of supporting bundling multiple OTU links, at the same line rate and muxing hierarchy, between a pair of nodes as a TE link. Note that link bundling is optional and is implementation dependent.

- Support for Control of Hitless Adjustment of ODUflex (GFP)

The control plane should support hitless adjustment of ODUflex, so the routing protocol should be capable of differentiating whether an ODU link can support hitless adjustment of ODUflex (GFP) or not, and how much resource can be used for resizing. This can be achieved by introducing a new signal type "ODUflex(GFP-F), resizable" that implies the support for hitless adjustment of ODUflex (GFP) by that link.

As mentioned in Section 5.1, one method of enabling multiplexing hierarchy is via usage of dedicated boards to allow tunneling of new services through legacy ODU1, ODU2, ODU3 containers. Such dedicated boards may have some constraints with respect to switching matrix access; detection and representation of such constraints is for further study.

5.4. Implications for Link Management Protocol (LMP)

As discussed in section 5.3, Path computation needs to know the interface switching capability of links. The switching capability of two ends of the link may be different, so the link capability of two ends should be correlated.

The Link Management Protocol (LMP) [RFC4204] provides a control plane protocol for exchanging and correlating link capabilities.

It is not necessary to use LMP to correlate link-end capabilities if the information is available from another source such as management configuration or automatic discovery/negotiation within the data plane.

Note that LO ODU type information can be, in principle, discovered by routing. Since in certain case, routing is not present (e.g. UNI case) we need to extend link management protocol capabilities to cover this aspect. In case of routing presence, the discovering procedure by LMP could also be optional.

- Correlating the granularity of the TS

As discussed in section 3.1.2, the two ends of a link may support different TS granularity. In order to allow interconnection the node with 1.25Gb/s granularity should fall back to 2.5Gb/s granularity.

Therefore, it is necessary for the two ends of a link to correlate the granularity of the TS. This ensures the correct use and of the TE link.

- Correlating the supported LO ODU signal types and multiplexing hierarchy capability

Many new ODU signal types have been introduced in [G709-V3], such as ODU0, ODU4, ODU2e and ODUflex. It is possible that equipment does not support all the LO ODU signal types introduced by those new standards or drafts. Furthermore, since multiplexing hierarchy is not allowed before [G709-V3], it is possible that only one end of an ODU link can support multiplexing hierarchy capability, or the two ends of the link support different multiplexing hierarchy capabilities (e.g., one end of the link supports ODU0 into ODU1 into ODU3 multiplexing while the other end supports ODU0 into ODU2 into ODU3 multiplexing).

For the control and management consideration, it is necessary for the two ends of an HO ODU link to correlate which types of LO ODU can be supported and what multiplexing hierarchy capabilities can be provided by the other end.

5.5. Implications for Control Plane Backward Compatibility

With the introduction of G709-v3, there may be OTN networks composed of a mixture of nodes, some of which support [G709-V1] and run control plane protocols defined in [RFC4328], while others support [G709-V3] and new OTN control plane characterized in this document. Note that a third case, for the sake of completeness, consists on G709-V1 nodes with a new OTN control plane, but such nodes can be considered as new nodes with limited capabilities.

This section discusses the compatibility of nodes implementing the control plane procedures defined [RFC4328], in support of [G709-V1], and the control plane procedures defined to support [G709-V3], as outlined by this document.

Compatibility needs to be considered only when controlling ODU1 or ODU2 or ODU3 connection, because [G709-V1] only support these three

ODU signal types. In such cases, there are several possible options including:

- A node supporting [G709-V3] could support only the [G709-V3] related control plane procedures, in which case both types of nodes would be unable to jointly control an LSP for an ODU type that both nodes support in the data plane. Note that this case is supported by the procedures defined in [RFC3473] as a different Switching Capability/Type value is used for the different control plane versions.
- A node supporting [G709-V3] could support both the [G709-V3] related control plane and the control plane defined in [RFC4328].
 - o Such a node could identify which set of procedure to follow when initiating an LSP based on the Switching Capability value advertised in routing.
 - o Such a node could follow the set of procedures based on the Switching Type received in signaling messages from an upstream node.
 - o Such a node, when processing a transit LSP, could select which signaling procedures to follow based on the Switching Capability value advertised in routing by the next hop node.

5.6. Implications for Path Computation Elements

[PCE-APS] describes the requirements for GMPLS applications of PCE in order to establish GMPLS LSP. PCE needs to consider the GMPLS TE attributes appropriately once a PCC or another PCE requests a path computation. The TE attributes which can be contained in the path calculation request message from the PCC or the PCE defined in [RFC5440] includes switching capability, encoding type, signal type, etc.

As described in section 5.2.1, new signal types and new signals with variable bandwidth information need to be carried in the extended signaling message of path setup. For the same consideration, PCECP also has a desire to be extended to carry the new signal type and related variable bandwidth information when a PCC requests a path computation.

6. Data Plane Backward Compatibility Considerations

If TS auto-negotiation is supported, a node supporting 1.25Gbps TS can interwork with the other nodes that supporting 2.5Gbps TS by combining Specific TSs together in data plane. The control plane must support this TS combination.

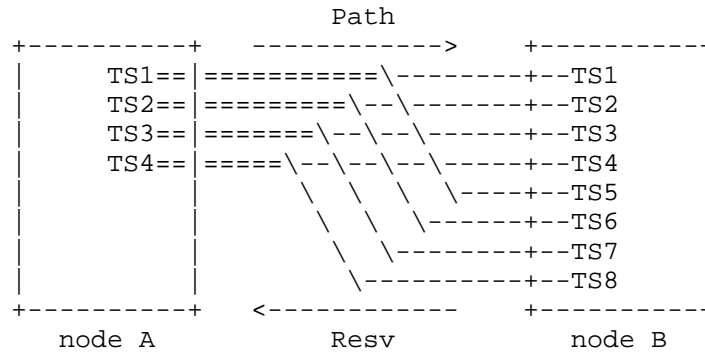


Figure 5 - Interworking between 1.25Gbps TS and 2.5Gbps TS

Take Figure 5 as an example. Assume that there is an ODU2 link between node A and B, where node A only supports the 2.5Gbps TS while node B supports the 1.25Gbps TS. In this case, the TS#i and TS#i+4 (where $i \leq 4$) of node B are combined together. When creating an ODU1 service in this ODU2 link, node B reserves the TS#i and TS#i+4 with the granularity of 1.25Gbps. But in the label sent from B to A, it is indicated that the TS#i with the granularity of 2.5Gbps is reserved.

In the contrary direction, when receiving a label from node A indicating that the TS#i with the granularity of 2.5Gbps is reserved, node B will reserved the TS#i and TS#i+4 with the granularity of 1.25Gbps in its data plane.

7. Security Considerations

The use of control plane protocols for signaling, routing, and path computation opens an OTN to security threats through attacks on those protocols. The data plane technology for an OTN does not introduce any specific vulnerabilities, and so the control plane may be secured using the mechanisms defined for the protocols discussed.

For further details of the specific security measures refer to the documents that define the protocols ([RFC3473], [RFC4203], [RFC4205],

[RFC4204], and [RFC5440]). [RFC5920] provides an overview of security vulnerabilities and protection mechanisms for the GMPLS control plane.

8. IANA Considerations

This document makes not requests for IANA action.

9. Acknowledgments

We would like to thank Maarten Visser and Lou Berger for their review and useful comments.

10. References

10.1. Normative References

- [RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol LabelSwitching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, Jan 2006.
- [RFC3471] Berger, L., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] L. Berger, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4201] K. Kompella, Y. Rekhter, Ed., "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.
- [RFC4202] K. Kompella, Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] K. Kompella, Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

- [RFC4205] K. Kompella, Y. Rekhter, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4205, October 2005.
- [RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [RFC4206] K. Kompella, Y. Rekhter, Ed., " Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC6107] K. Shiimoto, A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC6107, February 2011.
- [RFC6001] Dimitri Papadimitriou et al, "Generalized Multi-Protocol Label Switching (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC6001, February 21, 2010.
- [RFC5440] JP. Vasseur, JL. Le Roux, Ed., " Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [RFC6344] G. Bernstein et al, "Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS)", RFC6344, August, 2011.
- [G709-V3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709 Recommendation and Amendment2, April 2011.

10.2. Informative References

- [G709-V1] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation and Amendment1, November 2001.
- [G709-V2] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation, March 2003.
- [G7042] ITU-T, "Link capacity adjustment scheme (LCAS) for virtual concatenated signals", G.7042/Y.1305, March 2006.

- [G872-2001] ITU-T, "Architecture of optical transport networks", G.872 Recommendation, November 2001.
- [G872-Am2] ITU-T, "Architecture of optical transport networks", G.872 Recommendation and Amendment 2, July 2010.
- [G.7044] ITU-T, "Hitless adjustment of ODUflex", G.7044 (and Amendment 1), February 2012.
- [HZang00] H. Zang, J. Jue and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks", Optical Networks Magazine, January 2000.
- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks (WSN)", RFC6163, April 2011.
- [PCE-APS] Tomohiro Otani, Kenichi Ogaki, Diego Caviglia, and Fatai Zhang, "Requirements for GMPLS applications of PCE", draft-ietf-pce-gmpls-aps-req, Work in Progress.
- [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC5920, July 2010.

11. Authors' Addresses

Fatai Zhang (editor)
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Dan Li
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28973237
Email: huawei.danli@huawei.com

Han Li
China Mobile Communications Corporation
53 A Xibianmennei Ave. Xuanwu District
Beijing 100053 P.R. China

Phone: +86-10-66006688
Email: lihan@chinamobile.com

Sergio Belotti
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6863033

Email: sergio.belotti@alcatel-lucent.it

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

12. Contributors

Jianrui Han
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China

Phone: +86-755-28972913
Email: hanjianrui@huawei.com

Malcolm Betts
Huawei Technologies Co., Ltd.

Email: malcolm.betts@huawei.com

Pietro Grandi
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6864930

Email: pietro_vittorio.grandi@alcatel-lucent.it

Eve Varma
Alcatel-Lucent
1A-261, 600-700 Mountain Av
PO Box 636
Murray Hill, NJ 07974-0636
USA
Email: eve.varma@alcatel-lucent.com

APPENDIX A: ODU connection examples

This appendix provides a description of ODU terminology and connection examples. This section is not normative, and is just intended to facilitate understanding.

In order to transmit a client signal, an ODU connection needs to be created first. From the perspective of [G709-V3] and [G872-Am2], some types of ODUs (i.e., ODU1, ODU2, ODU3, ODU4) may assume either a client or server role within the context of a particular networking domain:

(1) An ODU_j client that is mapped into an OTU_k server. For example, if a STM-16 signal is encapsulated into ODU1, and then the ODU1 is mapped into OTU1, the ODU1 is a LO ODU (from a multiplexing perspective).

(2) An ODU_j client that is mapped into an ODU_k ($j < k$) server occupying several TSs. For example, if ODU1 is multiplexed into ODU2, and ODU2 is mapped into OTU2, the ODU1 is a LO ODU and the ODU2 is a HO ODU (from a multiplexing perspective).

Thus, a LO ODU_j represents the container transporting a client of the OTN that is either directly mapped into an OTU_k ($k = j$) or multiplexed into a server HO ODU_k ($k > j$) container. Consequently, the HO ODU_k represents the entity transporting a multiplex of LO ODU_j tributary signals in its OPU_k area.

In the case of LO ODU_j mapped into an OTU_k ($k = j$) directly, Figure 6 give an example of this kind of LO ODU connection.

In Figure 6, The LO ODU_j is switched at the intermediate ODXC node. OCh and OTU_k are associated with each other. From the viewpoint of connection management, the management of OTU_k is similar with OCh. LO ODU_j and OCh/OTU_k have client/server relationships.

For example, one LO ODU₁ connection can be setup between Node A and Node C. This LO ODU₁ connection is to be supported by OCh/OTU₁ connections, which are to be set up between Node A and Node B and between Node B and Node C. LO ODU₁ can be mapped into OTU₁ at Node A, demapped from it in Node B, switched at Node B, and then mapped into the next OTU₁ and demapped from this OTU₁ at Node C.

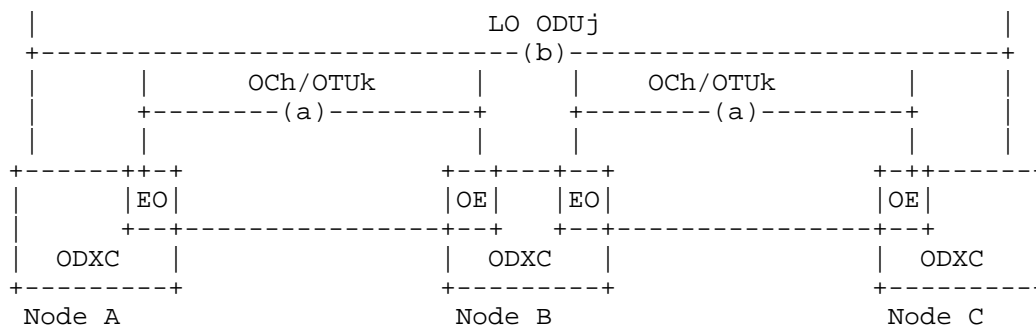


Figure 6 - Connection of LO ODU_j (1)

In the case of LO ODU_j multiplexing into HO ODU_k, Figure 7 gives an example of this kind of LO ODU connection.

In Figure 7, OCh, OTU_k, HO ODU_k are associated with each other. The LO ODU_j is multiplexed/de-multiplexed into/from the HO ODU at each ODXC node and switched at each ODXC node (i.e. trib port to line port, line card to line port, line port to trib port). From the viewpoint of connection management, the management of these HO ODU_k and OTU_k are similar to OCh. LO ODU_j and OCh/OTU_k/HO ODU_k have client/server relationships. When a LO ODU connection is setup, it will be using the existing HO ODU_k (/OTU_k/OCh) connections which have been set up. Those HO ODU_k connections provide LO ODU links, of which the LO ODU connection manager requests a link connection to support the LO ODU connection.

For example, one HO ODU₂ (/OTU₂/OCh) connection can be setup between Node A and Node B, another HO ODU₃ (/OTU₃/OCh) connection can be

setup between Node B and Node C. LO ODU1 can be generated at Node A, switched to one of the 10G line ports and multiplexed into a HO ODU2 at Node A, demultiplexed from the HO ODU2 at Node B, switched at Node B to one of the 40G line ports and multiplexed into HO ODU3 at Node B, demultiplexed from HO ODU3 at Node C and switched to its LO ODU1 terminating port at Node C.

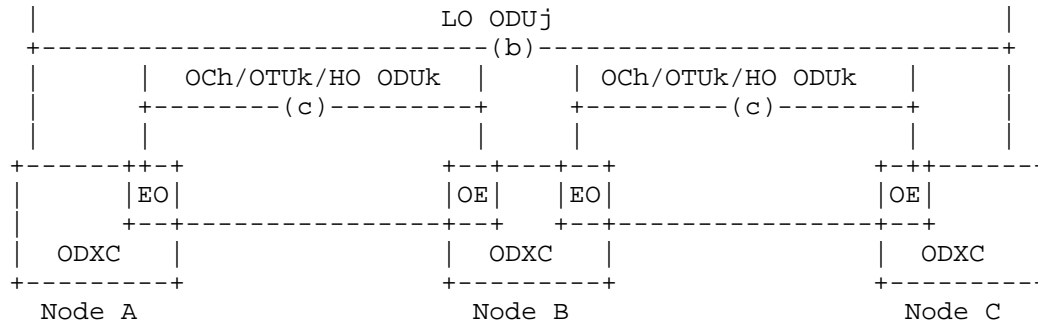


Figure 7 - Connection of LO ODUj (2)

Intellectual Property

The IETF Trust 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 any IETF 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.

Copies of Intellectual Property 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 any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of RFC 5378. No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under RFC 5378, shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

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

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 28, 2013

D. Ceccarelli, Ed.
D. Caviglia
Ericsson
F. Zhang
D. Li
Huawei Technologies
S. Belotti
P. Grandi
Alcatel-Lucent
R. Rao
K. Pithewan
Infinera Corporation
J. Drake
Juniper
August 27, 2012

Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS)
Control of Evolving G.709 OTN Networks
draft-ietf-ccamp-gmpls-ospf-g709v3-03

Abstract

The recent revision of ITU-T Recommendation G.709 [G709-V3] has introduced new fixed and flexible ODU containers, enabling optimized support for an increasingly abundant service mix.

This document describes OSPF routing protocol extensions to support Generalized MPLS (GMPLS) control of all currently defined ODU containers, in support of both sub-lambda and lambda level routing granularity.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on February 28, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
1.1. Terminology	3
2. OSPF-TE Extensions	3
3. TE-Link Representation	4
4. ISCD format extensions	5
4.1. Switch Capability Specific Information	7
5. Examples	13
5.1. MAX LSP Bandwidth fields in the ISCD	13
5.2. Example of T,S and TSG utilization	15
5.2.1. Example of different TSGs	16
5.3. Example of ODUflex advertisement	18
5.4. Example of single stage muxing	20
5.5. Example of multi stage muxing - Unbundled link	22
5.6. Example of multi stage muxing - Bundled links	23
5.7. Example of component links with non homogeneous hierarchies	25
6. Compatibility	27
7. Security Considerations	27
8. IANA Considerations	27
9. Contributors	28
10. Acknowledgements	29
11. References	29
11.1. Normative References	29
11.2. Informative References	30
Authors' Addresses	31

1. Introduction

G.709 OTN [G709-V3] includes new fixed and flexible ODU containers, two types of Tributary Slots (i.e., 1.25Gbps and 2.5Gbps), and supports various multiplexing relationships (e.g., ODUj multiplexed into ODUk ($j < k$)), two different tributary slots for ODUk ($K=1, 2, 3$) and ODUflex service type, which is being standardized in ITU-T. In order to present this information in the routing process, this document provides OTN technology specific encoding for OSPF-TE.

For a short overview of OTN evolution and implications of OTN requirements on GMPLS routing please refer to [OTN-FWK]. The information model and an evaluation against the current solution are provided in [OTN-INFO].

The routing information for Optical Channel Layer (OCh) (i.e., wavelength) is out of the scope of this document. Please refer to [WSO-Frame] for further information.

1.1. Terminology

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 [RFC2119].

2. OSPF-TE Extensions

In terms of GMPLS based OTN networks, each OTUk can be viewed as a component link, and each component link can carry one or more types of ODUj ($j < k$).

Each TE LSA can carry a top-level link TLV with several nested sub-TLVs to describe different attributes of a TE link. Two top-level TLVs are defined in [RFC 3630]. (1) The Router Address TLV (referred to as the Node TLV) and (2) the TE link TLV. One or more sub-TLVs can be nested into the two top-level TLVs. The sub-TLV set for the two top-level TLVs are also defined in [RFC 3630] and [RFC 4203].

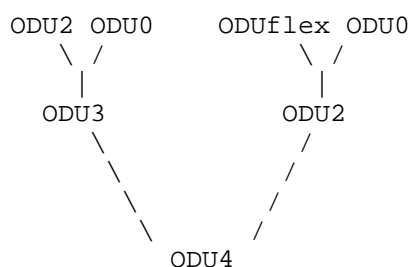
As discussed in [OTN-FWK] and [OTN-INFO], the OSPF-TE must be extended so to be able to advertise the termination and switching capabilities related to each different ODUj and ODUk/OTUk and the advertisement of related multiplexing capabilities. This leads to the need to define a new Switching Capability value and associated new Switching Capability for the ISCD.

In the following we will use ODUj to indicate a service type that is multiplexed into an higher order ODU, ODUk an higher order ODU

including an ODUj and ODUk/OTUk to indicate the layer mapped into the OTUk. Moreover ODUj(S) and ODUk(S) are used to indicate ODUj and ODUk supporting switching capability only, and the ODUj->ODUk format is used to indicate the ODUj into ODUk multiplexing capability.

This notation can be iterated as needed depending on the number of multiplexing levels. In the following the term "multiplexing tree" is used to identify a multiplexing hierarchy where the root is always a server ODUk/OTUk and any other supported multiplexed container is represented with increasing granularity until reaching the leaf of the tree. The tree can be structured with more than one branch if the server ODUk/OTUk supports more than one hierarchy.

If for example a multiplexing hierarchy like the following one is considered:



The ODU4 is the root of the muxing tree, ODU3 and ODU2 are containers directly multiplexed into the server and then ODU2, ODU0 are the leaves of the ODU3 branch, while ODUflex and ODU0 are the leaves of the ODU2 one. This means that on this traffic card it is possible to have the following multiplexing capabilities:

```

ODU2->ODU3->ODU4
ODU0->ODU3->ODU4
ODUflex->ODU2->ODU4
ODU0->ODU2->ODU4
  
```

3. TE-Link Representation

G.709 ODUk/OTUk Links are represented as TE-Links in GMPLS Traffic Engineering Topology for supporting ODUj layer switching. These TE-Links can be modeled in multiple ways. Some of the prominent

representations are captured below.

OTUk physical Link(s) can be modeled as a TE-Link(s). The TE-Link is termed as OTUk-TE-Link. The OTUk-TE-Link advertises ODUj switching capacity. The advertised capacity could include ODUk switching capacity. Figure-1 below provides an illustration of one hop OTUk TE-links.

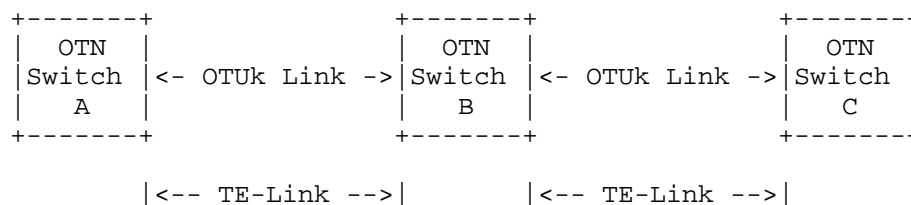


Figure 1: OTUk TE-Links

It is possible to create TE-Links that span more than one hop by creating FA between non-adjacent nodes. Such TE-Links are also termed OTUk-TE-Links. As in the one hop case, these types of OTUk-TE-Links also advertise ODUj switching capacity. The advertised capacity could include ODUk switching capacity.

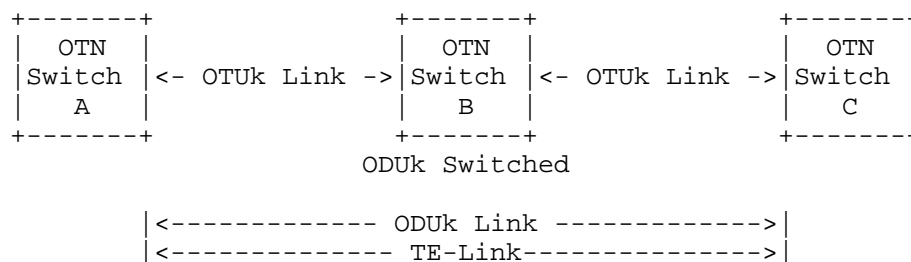


Figure 2: Multiple hop TE-Link

4. ISCD format extensions

The Interface Switching Capability Descriptor describes switching capability of an interface [RFC 4202]. This document defines a new Switching Capability value for OTN [G.709-v3] as follows:

Value	Type
-----	----
101 (TBA by IANA)	OTN-TDM capable (OTN-TDM)

Switching Capability and Encoding values MUST be used as follows:

Switching Capability = OTN-TDM
 Encoding Type = G.709 ODUk (Digital Path) [as defined in RFC4328]

Both fixed and flexible ODUs use the same switching type and encoding values. When Switching Capability and Encoding fields are set to values as stated above, the Interface Switching Capability Descriptor MUST be interpreted as follows:

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
+	+	+	+
Switching Cap	Encoding	Reserved	
+	+	+	+
	Max LSP Bandwidth at priority 0		
+	+	+	+
	Max LSP Bandwidth at priority 1		
+	+	+	+
	Max LSP Bandwidth at priority 2		
+	+	+	+
	Max LSP Bandwidth at priority 3		
+	+	+	+
	Max LSP Bandwidth at priority 4		
+	+	+	+
	Max LSP Bandwidth at priority 5		
+	+	+	+
	Max LSP Bandwidth at priority 6		
+	+	+	+
	Max LSP Bandwidth at priority 7		
+	+	+	+
	Switch Capability Specific Information		
+	(variable length)		
+	+	+	+

Maximum LSP Bandwidth

The MAX LSP bandwidth field is used according to [RFC4203]: i.e. 0 <= Max LSP Bandwidth <= ODUk/OTUk and intermediate values are those on

the branch of OTN switching hierarchy supported by the interface. E.g. in the OTU4 link it could be possible to have ODU4 as MAX LSP Bandwidth for some priorities, ODU3 for others, ODU2 for some others etc. The bandwidth unit MUST be in bytes per second and the encoding MUST be in IEEE floating point format. The discrete values for various ODUs is shown in the table below.

ODU Type	ODU nominal bit rate	Value in Byte/Sec
ODU0	1 244 160 kbits/s	0x4D1450C0
ODU1	239/238 x 2 488 320 kbit/s	0x4D94F048
ODU2	239/237 x 9 953 280 kbit/s	0x4E959129
ODU3	239/236 x 39 813 120 kbit/s	0X4F963367
ODU4	239/227 x 99 532 800 kbit/s	0x504331E3
ODU2e	239/237 x 10 312 500 kbit/s	0x4E9AF70A
ODUflex for CBR Client signals	239/238 x client signal bit rate	MAX LSP BANDWIDTH
ODUflex for GFP-F Mapped client signal	Configured bit rate	MAX LSP BANDWIDTH
ODU flex resizable	Configured bit rate	MAX LSP BANDWIDTH

A single ISCD MAY be used for the advertisement of unbundled or bundled links supporting homogeneous multiplexing hierarchies and the same Tributary Slot Granularity (TSG). A different ISCD MUST be used for each different muxing hierarchy (muxing tree in the following examples) and different TSG supported within the TE Link, if it includes component links with differing characteristics.

4.1. Switch Capability Specific Information

The technology specific part of the OTN ISCD may include a variable number of sub-TLVs called Bandwidth sub-TLVs. The muxing hierarchy tree MUST be encoded as an order independent list of them. Two types of Bandwidth TLV are defined (TBA by IANA):

- Type 1 - Unreserved Bandwidth for fixed containers
- Type 2 - Unreserved/MAX LSP Bandwidth for flexible containers

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 1 (Unres-fix)										Length																													
Signal type										Num of stages					T S		TSG		Res		Priority																		
Stage#1										...					Stage#N					Padding																			
Unres ODUj at Prio 0										Unres ODUj at Prio 1																													
Unres ODUj at Prio 2										Unres ODUj at Prio 3																													
Unres ODUj at Prio 4										Unres ODUj at Prio 5																													
Unres ODUj at Prio 6										Unres ODUj at Prio 7																													

Figure 4: Bandwidth TLV - Type 1 -

The values of the fields shown in figure 4 are explained after figure 6.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 2 (Unres/MAX-var)										Length																													
Signal type										Num of stages					T	S	TSG		Res		Priority																		
Stage#1										...					Stage#N					Padding																			
Unreserved Bandwidth at priority 0																																							
Unreserved Bandwidth at priority 1																																							
Unreserved Bandwidth at priority 2																																							
Unreserved Bandwidth at priority 3																																							
Unreserved Bandwidth at priority 4																																							
Unreserved Bandwidth at priority 5																																							
Unreserved Bandwidth at priority 6																																							
Unreserved Bandwidth at priority 7																																							
MAX LSP Bandwidth at priority 0																																							
MAX LSP Bandwidth at priority 1																																							
MAX LSP Bandwidth at priority 2																																							
MAX LSP Bandwidth at priority 3																																							
MAX LSP Bandwidth at priority 4																																							
MAX LSP Bandwidth at priority 5																																							
MAX LSP Bandwidth at priority 6																																							
MAX LSP Bandwidth at priority 7																																							

Figure 5: Bandwidth TLV - Type 2 -

- Signal Type: Indicates the ODU type being advertised

Value	Type
-----	-----
1	ODU1
2	ODU2
3	ODU3
4	ODU4
10	ODU0
11	ODU2e
20	ODUflex CBR
21	ODUflex GFP-F resizable
22	ODUflex GFP-F non resizable
230-256	Experimental

With respect to ODUflex, ODUflex CBR and ODUflex GFP-F MUST always be advertised separately as they use different adaptation functions. In the case both GFP-F resizable and non resizable (i.e. 21 and 22) are supported, Signal Type 21 implicitly supports also signal Signal Type 22, so only Signal Type 21 MUST be advertised. Signal Type 22 MUST be used only for non resizable resources.

- Number of stages: Indicates the number of multiplexing stages level. It MUST be equal to 0 when a server layer is being advertised, 1 in case of single stage muxing, 2 in case of dual stage muxing, etc.

- Flags:

- T Flag (bit 17): Indicates whether the advertised bandwidth can be terminated. When T=1, the signal type can be terminated, when T=0, the signal type cannot be terminated.

- S Flag (bit 18): Indicates whether the advertised bandwidth can be switched. When S=1, the signal type can be switched, when S=0, the signal type cannot be switched.

The value 00 in both T and S bits MUST NOT be used.

- TSG: Tributary Slot Granularity (3bit): Used for the advertisement of the supported Tributary Slot granularity

- 0 - Reserved
- 1 - 1.25 Gbps/2.5Gbps
- 2 - 2.5 Gbps only

- 3 - 1.25 Gbps only
- 4 - Don't care
- 5-7 - Reserved

Where value 1 is used on those interfaces where the fallback procedure is enabled and the default value of 1.25 Gbps can be falled back to 2.5 if needed. Values 2 and 3 are used where there is no chance to modify the TSG. In the former case the interface being advertised is a G.709v1 and in the latter the interface is a G.709v3 with fallback procedure disabled or unavailable. Value 4 is used for non multiplexed signal (i.e. non OTN client).

- Priority :8 bits field with 1 flag for each priority. Bit set indicates priority supported, bit cleared means priority not supported. The priority 0 is related to the most significant bit. When no priority is supported, priority 0 MUST be advertised.

- Stage#1 ... Stage#N : These fields are 8 bits long. Their number is variable and a field is present for each stage of the muxing hierarchy. The last one MUST always indicate the server ODU container (ODUk/OTUk). The values of the Stage fields MUST be the same ones defined for the Signal Type field. If the number of stages is 0, then no Stage fields MUST be included.

- Padding: Given that the number of Stages is variable, padding to 32 bits field MUST be used when needed.

- Unreserved Bandwidth/Max LSP BW : In case of fixed containers (Type=1) the Unreserved Bandwidth field MUST be 16 bits long and indicates the Unreserved Bandwidth in number of available containers. Only Unreserved/MAX LSP BW fields for supported priorities MUST be included, in order of increasing prioritiiy (0 to 7). In case the number of supported priorities is odd, a 16 bits all zeros padding field MUST be added. On the other hand, in case of variable containers (Type 2) the Unreserved/MAX LSP Bandwidth fields MUST be 32 bits long and expressed in IEEE floating point format. The advertisement of the MAX LSP bandwidth MUST take into account HO OPUk bit rate tolerance and be calculated according to the following formula:

$$\text{Max LSP BW} = (\# \text{ available TS}) * (\text{ODTUK.ts nominal bit rate}) * (1 - \text{HO OPUk bit rate tolerance})$$

Only Unreserved/MAX LSP bandwidth for supported priorities MUST be advertised.

5. Examples

The examples in the following pages are not normative and are not intended to infer or mandate any specific implementation.

5.1. MAX LSP Bandwidth fields in the ISCD

This example shows how the MAX LSP Bandwidth fields of the ISCD are filled accordingly to the evolving of the TE-link bandwidth occupancy. In the example an OTU4 link is considered, with supported priorities 0,2,4,7 and muxing hierarchy ODU1->ODU2->ODU3->ODU4.

At time T0, with the link completely free, the advertisement would be:

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Switching Cap								Encoding								Reserved															
Max LSP Bandwidth at priority 0 = 100Gbps																															
Max LSP Bandwidth at priority 1 = 0																															
Max LSP Bandwidth at priority 2 = 100Gbps																															
Max LSP Bandwidth at priority 3 = 0																															
Max LSP Bandwidth at priority 4 = 100Gbps																															
Max LSP Bandwidth at priority 5 = 0																															
Max LSP Bandwidth at priority 6 = 0																															
Max LSP Bandwidth at priority 7 = 100Gbps																															
Switch Capability Specific Information																															
(variable length)																															

Figure 6: Example 1 - MAX LSP Bandwidth fields in the ISCD @T0

At time T1 an ODU3 at priority 2 is set-up, so for priority 0 the MAX LSP Bandwidth is still equal to the ODU4 bandwidth, while for priorities from 2 to 7 (excluding the non supported ones) the MAX LSP Bandwidth is equal to ODU3, as no more ODU4s are available and the

next supported ODUj in the hierarchy is ODU3. The advertisement is updated as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Switching Cap										Encoding										Reserved																			
Max LSP Bandwidth at priority 0 = 100Gbps																																							
Max LSP Bandwidth at priority 1 = 0																																							
Max LSP Bandwidth at priority 2 = 40Gbps																																							
Max LSP Bandwidth at priority 3 = 0																																							
Max LSP Bandwidth at priority 4 = 40Gbps																																							
Max LSP Bandwidth at priority 5 = 0																																							
Max LSP Bandwidth at priority 6 = 0																																							
Max LSP Bandwidth at priority 7 = 40Gbps																																							
Switch Capability Specific Information																																							
(variable length)																																							

Figure 7: Example 1 - MAX LSP Bandwidth fields in the ISCD @T1

At time T2 an ODU2 at priority 4 is set-up. The first ODU3 is no longer available since T1 as it was kept by the ODU3 LSP, while the second is no more available and just 3 ODU2 are left in it. ODU2 is now the MAX LSP bandwidth for priorities higher than 4. The advertisement is updated as follows:

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Switching Cap								Encoding								Reserved															
								Max LSP Bandwidth at priority 0 = 100Gbps																							
								Max LSP Bandwidth at priority 1 = 0																							
								Max LSP Bandwidth at priority 2 = 40Gbps																							
								Max LSP Bandwidth at priority 3 = 0																							
								Max LSP Bandwidth at priority 4 = 10Gbps																							
								Max LSP Bandwidth at priority 5 = 0																							
								Max LSP Bandwidth at priority 6 = 0																							
								Max LSP Bandwidth at priority 7 = 10Gbps																							
								Switch Capability Specific Information																							
								(variable length)																							

Figure 8: Example 1 - MAX LSP Bandwidth fields in the ISCD @T2

5.2. Example of T,S and TSG utilization

In this example an interface with Tributary Slot Type 1.25 Gbps and fallback procedure enabled is considered (TSG=1). It supports the simple ODU1->ODU2->ODU3 hierarchy and priorities 0 and 3. Suppose that in this interface the ODU3 signal type can be both switched or terminated, the ODU2 can only be terminated and the ODU1 switched only. For the advertisement of the capabilities of such interface a single ISCD is used and its format is as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 1 (Unres-fix)										Length																													
Sig type=ODU1										#stages= 2										T0 S1 001 Res										1 0 0 1 0 0 0 0									
Stage#1=ODU2										Stage#2=ODU3										Padding																			
Unres ODU1 at Prio 0										Unres ODU1 at Prio 3																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU2										#stages= 1										T1 S0 001 Res										1 0 0 1 0 0 0 0									
Stage#1=ODU3										Padding																													
Unres ODU2 at Prio 0										Unres ODU2 at Prio 3																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU3										#stages= 0										T1 S1 001 Res										1 0 0 1 0 0 0 0									
Unres ODU3 at Prio 0										Unres ODU3 at Prio 3																													

Figure 9: Example 2 - TSG, T and S utilization

5.2.1. Example of different TSGs

In this example two interfaces with homogeneous hierarchies but different Tributary Slot Types are considered. The first one supports a G.709v1 interface (TSG=2) while the second one a G.709v3 interface with fallback procedure disabled (TSG=3). Both of them support ODU1->ODU2->ODU3 hierarchy and priorities 0 and 3. For the advertisement of the capabilities of such interfaces two different ISCDs are used and the format of their SCSIs is as follows:

```

SCSI of ISCD 1 - TSG=2
0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU1 | #stages= 2 |T|S| 2 | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU2 | Stage#2=ODU3 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Unres ODU1 at Prio 0          |          Unres ODU1 at Prio 3          |
+-----+-----+-----+-----+-----+-----+-----+-----+

SCSI of ISCD 2 - TSG=3
0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU1 | #stages= 2 |T|S| 3 | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU2 | Stage#2=ODU3 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Unres ODU1 at Prio 0          |          Unres ODU1 at Prio 3          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 10: Example 2.1 - Different TSGs utilization

A particular case in which hierarchies with the same muxing tree but with different exported TSG must be considered as non homogenous hierarchies is the case in which an H-LSP and the client LSP are terminated on the same egress node. What can happen is that a loose ero is used at the hop where the signaled LSP is nested into the H-LSP (penultimate hop of the LSP).

In the following figure, node C receives from A a loose ERO towards node E and must choose between the ODU2 H-LSP on if1 or the one on if2. In case the H-LSP on if1 exports a TS=1,25Gbps and if2 a TS=2,5Gbps and the service LSP being signaled needs a 1,25Gbps tributary slot, only the H-LSP on if1 can be used to reach node E. For further details please see section 4.1 of the [OTN-INFO].

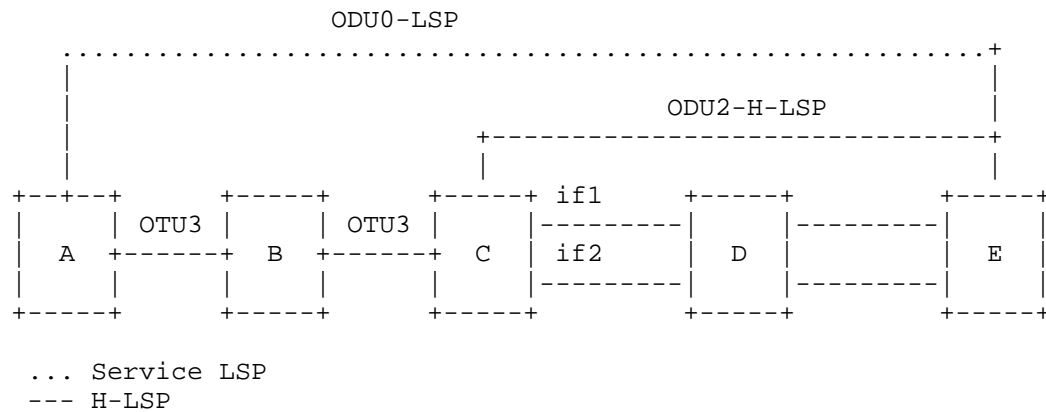


Figure 11: Example - Service LSP and H-LSP terminating on the same node

5.3. Example of ODUflex advertisement

In this example the advertisement of an ODUflex->ODU3 hierarchy is shown. In case of ODUflex advertisement the MAX LSP bandwidth needs to be advertised and in some cases also information about the Unreserved bandwidth could be useful. The amount of Unreserved bandwidth does not give a clear indication of how many ODUflex LSP can be set up either at the MAX LSP Bandwidth or at different rates, as it gives no information about the spatial allocation of the free TSs.

An indication of the amount of Unreserved bandwidth could be useful during the path computation process, as shown in the following example. Supposing there are two TE-links (A and B) with MAX LSP Bandwidth equal to 10 Gbps each. In case 50Gbps of Unreserved Bandwidth are available on Link A, 10Gbps on Link B and 3 ODUflex LSPs of 10 GBps each, have to be restored, for sure only one can be restored along Link B and it is probable (but not sure) that two of them can be restored along Link A.

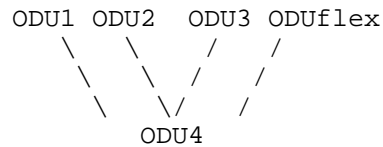
In the case of ODUflex advertisement the Type 2 Bandwidth TLV is used.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 2 (Unres/MAX-var)										Length																													
S. type=ODUflex										#stages= 1					T S TSG					Res					Priority														
Stage#1=ODU3										Padding																													
Unreserved Bandwidth at priority 0																																							
Unreserved Bandwidth at priority 1																																							
Unreserved Bandwidth at priority 2																																							
Unreserved Bandwidth at priority 3																																							
Unreserved Bandwidth at priority 4																																							
Unreserved Bandwidth at priority 5																																							
Unreserved Bandwidth at priority 6																																							
Unreserved Bandwidth at priority 7																																							
MAX LSP Bandwidth at priority 0																																							
MAX LSP Bandwidth at priority 1																																							
MAX LSP Bandwidth at priority 2																																							
MAX LSP Bandwidth at priority 3																																							
MAX LSP Bandwidth at priority 4																																							
MAX LSP Bandwidth at priority 5																																							
MAX LSP Bandwidth at priority 6																																							
MAX LSP Bandwidth at priority 7																																							

Figure 12: Example 3 - ODUflex advertisement

5.4. Example of single stage muxing

Supposing there is 1 OTU4 component link supporting single stage muxing of ODU1, ODU2, ODU3 and ODUFlex, the supported hierarchy can be summarized in a tree as in the following figure. For sake of simplicity we assume that also in this case only priorities 0 and 3 are supported.



and the related SCSIs as follows:

```

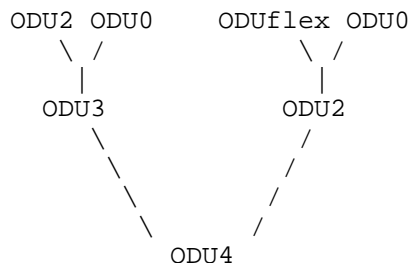
0                                     1                                     2                                     3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU4 | #stages= 0 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Unres ODU4 at Prio 0 =1          |  Unres ODU4 at Prio 3 =1          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU1 | #stages= 1 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Unres ODU1 at Prio 0 =40          |  Unres ODU1 at Prio 3 =40          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU2 | #stages= 1 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Unres ODU2 at Prio 0 =10          |  Unres ODU2 at Prio 3 =10          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU3 | #stages= 1 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Unres ODU3 at Prio 0 =2          |  Unres ODU3 at Prio 3 =2          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 2 (Unres/MAX-var)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|S. type=ODUflex| #stages= 1 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |          Padding          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Unreserved Bandwidth at priority 0 =100Gbps          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Unreserved Bandwidth at priority 3 =100Gbps          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          MAX LSP Bandwidth at priority 0 =100Gbps          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          MAX LSP Bandwidth at priority 3 =100Gbps          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 13: Example 4 - Single stage muxing

5.5. Example of multi stage muxing - Unbundled link

Supposing there is 1 OTU4 component link with muxing capabilities as shown in the following figure:



and supported priorities 0 and 3, the advertisement is composed by the following Bandwidth TLVs:

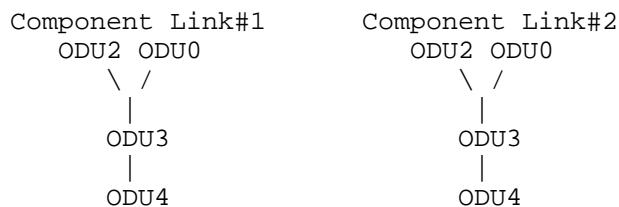
0										1										2										3																																																																					
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1																																																																				
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Type = 1 (Unres-fix)																				Length																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Sig type=ODU4																				#stages= 0																				T S TSG																				Res																				1 0 0 1 0 0 0 0									
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Unres ODU4 at Prio 0 =1																				Unres ODU4 at Prio 3 =1																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Type = 1 (Unres-fix)																				Length																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Sig type=ODU3																				#stages= 1																				T S TSG																				Res																				1 0 0 1 0 0 0 0									
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Stage#1=ODU4																				Padding																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Unres ODU3 at Prio 0 =2																				Unres ODU3 at Prio 3 =2																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Type = 1 (Unres-fix)																				Length																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Sig type=ODU2																				#stages= 1																				T S TSG																				Res																				1 0 0 1 0 0 0 0									
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Stage#1=ODU4																				Padding																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					
										Unres ODU2 at Prio 0 =10																				Unres ODU2 at Prio 3 =10																																																																					
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+																																																																					

Type = 1 (Unres-fix)				Length			
Sig type=ODU2	#stages= 2	T S	TSG	Res	1 0 0 1 0 0 0 0		
Stage#1=ODU3	Stage#2=ODU4	Padding					
Unres ODU2 at Prio 0 =8				Unres ODU2 at Prio 3 =8			
Type = 1 (Unres-fix)				Length			
Sig type=ODU0	#stages= 2	T S	TSG	Res	1 0 0 1 0 0 0 0		
Stage#1=ODU3	Stage#2=ODU4	Padding					
Unres ODU0 at Prio 0 =64				Unres ODU0 at Prio 3 =64			
Type = 1 (Unres-fix)				Length			
Sig type=ODU0	#stages= 2	T S	TSG	Res	1 0 0 1 0 0 0 0		
Stage#1=ODU2	Stage#2=ODU4	Padding					
Unres ODU0 at Prio 0 =80				Unres ODU0 at Prio 3 =80			
Type = 2 (Unres/MAX-var)				Length			
S.type=ODUflex	#stages= 2	T S	TSG	Res	1 0 0 1 0 0 0 0		
Stage#1=ODU2	Stage#2=ODU4	Padding					
Unreserved Bandwidth at priority 0 =100Gbps							
Unreserved Bandwidth at priority 3 =100Gbps							
MAX LSP Bandwidth at priority 0 =10Gbps							
MAX LSP Bandwidth at priority 3 =10Gbps							

Figure 14: Example 5 - Multi stage muxing - Unbundled link

5.6. Example of multi stage muxing - Bundled links

In this example 2 OTU4 component links with the same supported TSG and homogeneous muxing hierarchies are considered. The following muxing capabilities trees are supported:



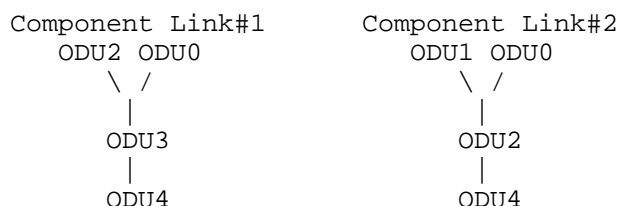
Considering only supported priorities 0 and 3, the advertisement is as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 1 (Unres-fix)										Length																													
Sig type=ODU4										#stages= 0										T S TSG Res 1 0 0 1 0 0 0 0																			
Unres ODU4 at Prio 0 =2										Unres ODU4 at Prio 3 =2																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU3										#stages= 1										T S TSG Res 1 0 0 1 0 0 0 0																			
Stage#1=ODU4										Padding																													
Unres ODU3 at Prio 0 =4										Unres ODU3 at Prio 3 =4																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU2										#stages= 2										T S TSG Res 1 0 0 1 0 0 0 0																			
Stage#1=ODU3										Stage#2=ODU4										Padding																			
Unres ODU2 at Prio 0 =16										Unres ODU2 at Prio 3 =16																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU0										#stages= 2										T S TSG Res 1 0 0 1 0 0 0 0																			
Stage#1=ODU3										Stage#2=ODU4										Padding																			
Unres ODU0 at Prio 0 =128										Unres ODU0 at Prio 3 =128																													

Figure 15: Example 6 - Multi stage muxing - Bundled links

5.7. Example of component links with non homogeneous hierarchies

In this example 2 OTU4 component links with the same supported TSG and non homogeneous muxing hierarchies are considered. The following muxing capabilities trees are supported:



Considering only supported priorities 0 and 3, the advertisement uses two different ISCDs, one for each hierarchy. In the following figure, the SCSI of each ISCD is shown:

SCSI of ISCD 1 - Component Link#1

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Type = 1 (Unres-fix)										Length																													
Sig type=ODU4										#stages= 0										T S TSG										Res 1 0 0 1 0 0 0 0									
Unres ODU4 at Prio 0 =1										Unres ODU4 at Prio 3 =1																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU3										#stages= 1										T S TSG										Res 1 0 0 1 0 0 0 0									
Stage#1=ODU4										Padding																													
Unres ODU3 at Prio 0 =2										Unres ODU3 at Prio 3 =2																													
Type = 1 (Unres-fix)										Length																													
Sig type=ODU2										#stages= 2										T S TSG										Res 1 0 0 1 0 0 0 0									
Stage#1=ODU3										Stage#2=ODU4										Padding																			

```

|   Unres ODU2 at Prio 0 =8   |   Unres ODU2 at Prio 3 =8   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Type = 1 (Unres-fix)   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU0 | #stages= 2 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU3 | Stage#2=ODU4 |   Padding   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Unres ODU0 at Prio 0 =64   |   Unres ODU0 at Prio 3 =64   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

SCSI of ISCD 2 - Component Link#2

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Type = 1 (Unres-fix)   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU4 | #stages= 0 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Unres ODU4 at Prio 0 =1   |   Unres ODU4 at Prio 3 =1   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Type = 1 (Unres-fix)   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU2 | #stages= 1 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |   Padding   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Unres ODU2 at Prio 0 =10   |   Unres ODU2 at Prio 3 =10   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Type = 1 (Unres-fix)   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU1 | #stages= 2 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU2 | Stage#2=ODU4 |   Padding   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Unres ODU1 at Prio 0 =40   |   Unres ODU1 at Prio 3 =40   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Type = 1 (Unres-fix)   |   Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Sig type=ODU0 | #stages= 2 |T|S| TSG | Res |1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU2 | Stage#2=ODU4 |   Padding   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Unres ODU0 at Prio 0 =80   |   Unres ODU0 at Prio 3 =80   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 16: Example 7 - Multi stage muxing - Non homogeneous hierarchies

6. Compatibility

All implementations of this document MAY support also advertisement as defined in [RFC4328]. When nodes support both advertisement methods, implementations MUST support the configuration of which advertisement method is followed. The choice of which is used is based on policy and is out of scope of the document. This enables nodes following each method to identify similar supporting nodes and compute paths using only the appropriate nodes.

7. Security Considerations

This document specifies the contents of Opaque LSAs in OSPFv2. As Opaque LSAs are not used for SPF computation or normal routing, the extensions specified here have no direct effect on IP routing. Tampering with GMPLS TE LSAs may have an effect on the underlying transport (optical and/or SONET-SDH) network. [RFC3630] suggests mechanisms such as [RFC2154] to protect the transmission of this information, and those or other mechanisms should be used to secure and/or authenticate the information carried in the Opaque LSAs.

8. IANA Considerations

Upon approval of this document, IANA will make the assignment of a new Switching Capability value for the existing ISCD located at <http://www.iana.org/assignments/ospf-traffic-eng-tlvs/ospf-traffic-eng-tlvs.xml>:

15 Interface Switching Capability Descriptor [RFC4203]

Switching capability	Description	Reference
101 (suggested)	OTN-TDM capable (OTN-TDM)	[This.I-D]

This document defines the following sub-TLVs of the ISCD TLV:

Value	Sub-TLV
1	Unreserved Bandwidth for fixed containers
2	Unreserved/MAX LSP bandwidth for flexible containers

9. Contributors

Xiaobing Zi, Huawei Technologies

Email: zixiaobing@huawei.com

Francesco Fondelli, Ericsson

Email: francesco.fondelli@ericsson.com

Marco Corsi

Email: corsi.marco@gmail.com

Eve Varma, Alcatel-Lucent

Email: eve.varma@alcatel-lucent.com

Jonathan Sadler, Tellabs

Email: jonathan.sadler@tellabs.com

Lyndon Ong, Ciena

Email: lyong@ciena.com

Ashok Kunjidhapatham

akunjidhapatham@infinera.com

Snigdho Bardalai

sbardalai@infinera.com

Steve Balls

Steve.Balls@metaswitch.com

Jonathan Hardwick

Jonathan.Hardwick@metaswitch.com

Xihua Fu

fu.xihua@zte.com.cn

Cyril Margaria

cyril.margaria@nsn.com

Malcolm Betts

Malcolm.betts@zte.com.cn

10. Acknowledgements

The authors would like to thank Fred Gruman and Lou Berger for the precious comments and suggestions.

11. References

11.1. Normative References

[OTN-FWK] F.Zhang, D.Li, H.Li, S.Belotti, D.Ceccarelli, "Framework for GMPLS and PCE Control of G.709 Optical Transport networks, work in progress draft-ietf-ccamp-gmpls-g709-framework-05", September 2011.

[OTN-INFO] S.Belotti, P.Grandi, D.Ceccarelli, D.Caviglia, F.Zhang,

D.Li, "Information model for G.709 Optical Transport Networks (OTN), work in progress draft-ietf-ccamp-otn-g709-info-model-02", October 2011.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2154] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", RFC 2154, June 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC2370] Coltun, R., "The OSPF Opaque LSA Option", RFC 2370, July 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [RFC6001] Papadimitriou, D., Vigoureux, M., Shiimoto, K., Brungard, D., and J.L. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, October 2010.

11.2. Informative References

- [G.709] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), February 2001.
- [G.709-v3]

ITU-T, "Draft revised G.709, version 3", consented
by ITU-T on Oct 2009.

Authors' Addresses

Daniele Ceccarelli (editor)
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28972912

Email: zhangfatai@huawei.com

Dan Li
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28973237

Email: danli@huawei.com

Sergio Belotti
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

Pietro Vittorio Grandi
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: pietro_vittorio.grandi@alcatel-lucent.com

Rajan Rao
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA

Email: rrao@infinera.com

Khuzema Pithewan
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA

Email: kpithewan@infinera.com

John E Drake
Juniper

Email: jdrake@juniper.net

Network Working Group
Internet Draft
Category: Standards Track

Fatai Zhang, Ed.
Huawei
Guoying Zhang
CATR
Sergio Belotti
Alcatel-Lucent
D. Ceccarelli
Ericsson
Khuzema Pithewan
Infinera
August 27, 2012

Expires: February 27, 2013

Generalized Multi-Protocol Label Switching (GMPLS) Signaling
Extensions for the evolving G.709 Optical Transport Networks Control

draft-ietf-ccamp-gmpls-signaling-g709v3-04.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on February 27, 2013.

Abstract

Recent progress in ITU-T Recommendation G.709 standardization has introduced new ODU containers (ODU0, ODU4, ODU2e and ODUflex) and

enhanced Optical Transport Networking (OTN) flexibility. Several recent documents have proposed ways to modify GMPLS signaling protocols to support these new OTN features.

It is important that a single solution is developed for use in GMPLS signaling and routing protocols. This solution must support ODUk multiplexing capabilities, address all of the new features, be acceptable to all equipment vendors, and be extensible considering continued OTN evolution.

This document describes the extensions to the Generalized Multi-Protocol Label Switching (GMPLS) signaling to control the evolving Optical Transport Networks (OTN) addressing ODUk multiplexing and new features including ODU0, ODU4, ODU2e and ODUFlex.

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 [RFC2119].

Table of Contents

1. Introduction	3
2. Terminology	4
3. GMPLS Extensions for the Evolving G.709 - Overview	4
4. Generalized Label Request	5
5. Extensions for Traffic Parameters for the Evolving G.709	7
5.1. Usage of ODUFlex(CBR) Traffic Parameters	8
5.2. Usage of ODUFlex(GFP) Traffic Parameters	10
6. Generalized Label	11
6.1. New definition of ODU Generalized Label	11
6.2. Examples	14
6.3. Label Distribution Procedure	15
6.3.1. Notification on Label Error	16
6.4. Supporting Virtual Concatenation and Multiplication	17
7. Supporting Hitless Adjustment of ODUFlex (GFP)	17
8. Control Plane Backward Compatibility Considerations.....	18
9. Security Considerations	19
10. IANA Considerations.....	19
11. References	20
11.1. Normative References	20
11.2. Informative References	21

12. Contributors	21
13. Authors' Addresses	22
14. Acknowledgment	24

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] extends MPLS to include Layer-2 Switching (L2SC), Time-Division Multiplex (e.g., SONET/SDH, PDH, and ODU), Wavelength (OCh, Lambdas) Switching, and Spatial Switching (e.g., incoming port or fiber to outgoing port or fiber). [RFC3471] presents a functional description of the extensions to Multi-Protocol Label Switching (MPLS) signaling required to support Generalized MPLS. RSVP-TE-specific formats and mechanisms and technology specific details are defined in [RFC3473].

With the evolution and deployment of G.709 technology, it is necessary that appropriate enhanced control technology support be provided for G.709. [RFC4328] describes the control technology details that are specific to foundation G.709 Optical Transport Networks (OTN), as specified in the ITU-T Recommendation G.709 [G709-V1], for ODUk deployments without multiplexing.

In addition to increasing need to support ODUk multiplexing, the evolution of OTN has introduced additional containers and new flexibility. For example, ODU0, ODU2e, ODU4 containers and ODUFlex are developed in [G709-V3].

In addition, the following issues require consideration:

- Support for Hitless Adjustment of ODUFlex (GFP) (HAO), which is defined in [G.7044].
- Support for Tributary Port Number. The Tributary Port Number has to be negotiated on each link for flexible assignment of tributary ports to tributary slots in case of LO-ODU over HO-ODU (e.g., ODU2 into ODU3).

Therefore, it is clear that [RFC4328] has to be updated or superseded in order to support ODUk multiplexing, as well as other ODU enhancements introduced by evolution of OTN standards.

This document updates [RFC4328] extending the G.709 ODUk traffic parameters and also presents a new OTN label format which is very flexible and scalable.

2. Terminology

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 [RFC2119].

3. GMPLS Extensions for the Evolving G.709 - Overview

New features for the evolving OTN, for example, new ODU0, ODU2e, ODU4 and ODUFlex containers are specified in [G709-V3]. The corresponding new signal types are summarized below:

- Optical Channel Transport Unit (OTUk):
 - . OTU4
- Optical Channel Data Unit (ODUk):
 - . ODU0
 - . ODU2e
 - . ODU4
 - . ODUFlex

A new Tributary Slot Granularity (TSG) (i.e., 1.25 Gbps) is also described in [G709-V3]. Thus, there are now two TS granularities for the foundation OTN ODU1, ODU2 and ODU3 containers. The TS granularity at 2.5 Gbps is used on legacy interfaces while the new 1.25 Gbps is used on the new interfaces.

In addition to the support of ODUk mapping into OTUk ($k = 1, 2, 3, 4$), the evolving OTN [G.709-V3] encompasses the multiplexing of ODUj ($j = 0, 1, 2, 2e, 3, flex$) into an ODUk ($k > j$), as described in Section 3.1.2 of [OTN-FWK].

Virtual Concatenation (VCAT) of OPUk (OPUk-Xv, $k = 1/2/3$, $X = 1...256$) is also supported by [OTN-V3]. Note that VCAT of OPU0 / OPU2e / OPU4 / OPUFlex is not supported per [OTN-V3].

[RFC4328] describes GMPLS signaling extensions to support the control for G.709 Optical Transport Networks (OTN) [G709-V1]. However, [RFC4328] needs to be updated because it does not provide the means to signal all the new signal types and related mapping and multiplexing functionalities. Moreover, it supports only the deprecated auto-MSI mode which assumes that the Tributary Port Number is automatically assigned in the transmit direction and not checked in the receive direction.

This document extends the G.709 traffic parameters described in [RFC4328] and presents a new flexible and scalable OTN label format.

Additionally, procedures about Tributary Port Number assignment through control plane are also provided in this document.

4. Generalized Label Request

The Generalized Label Request, as described in [RFC3471], carries the LSP Encoding Type, the Switching Type and the Generalized Protocol Identifier (G-PID).

[RFC4328] extends the Generalized Label Request, introducing two new code-points for the LSP Encoding Type (i.e., G.709 ODUk (Digital Path) and G.709 Optical Channel) and adding a list of G-PID values in order to accommodate [G709-v1].

This document follows these extensions and a new Switching Type is introduced to indicate the ODUk switching capability [G709-V3] in order to support backward compatibility with [RFC4328], as described in [OTN-FWK]. The new Switching Type (101, TBA by IANA) is defined in [OTN-OSPF].

This document also updates the G-PID values defined in [RFC4328]:

Value	G-PID Type
-----	-----
47	ODU-2.5G: transport of Digital Paths at 2.5, 10 and 40 Gbps via 2.5Gbps TSG
49	CBRa: asynchronous Constant Bit Rate (i.e., mapping of CBR2G5, CBR10G and CBR40G)
50	CBRb: bit synchronous Constant Bit Rate (i.e., mapping of CBR2G5, CBR10G, CBR40G, CBR10G3 and supra-2.488 CBR Gbit/s signal (carried by OPUflex))
32	ATM: mapping at 1.25, 2.5, 10 and 40 Gbps
51	BSOT: non-specific client Bit Stream with Octet Timing (i.e., Mapping of 1.25, 2.5, 10, 40 and 100 Gbps Bit Stream)
52	BSNT: non-specific client Bit Stream without Octet Timing (i.e., Mapping of 1.25, 2.5, 10, 40 and 100 Gbps Bit Stream)

Note: Values 32, 47, 49 and 50 include mapping of SDH.

In the case of ODU multiplexing, the LO ODU (i.e., the client signal) may be multiplexed into HO ODU via 1.25G TSG, 2.5G TSG or any one of them (i.e., TSG Auto_Negotiation is enabled). Since the G-PID type "ODUk" defined in [RFC4328] is only used for 2.5Gbps TSG, two new G-PID types are defined as follows:

- ODU-1.25G: transport of Digital Paths at 1.25, 2.5, 10, 40 and 100 Gbps via 1.25Gbps TSG
- ODU-any: transport of Digital Paths at 1.25, 2.5, 10, 40 and 100 Gbps via 1.25 or 2.5Gbps TSG (i.e., the fallback procedure is enabled and the default value of 1.25Gbps TSG can be fallen back to 2.5Gbps if needed)

In addition, some other new G-PID types are defined to support other new client signals described in [G709-V3]:

- CBRc: Mapping of constant bit-rate signals with justification into OPUk (k = 0, 1, 2, 3, 4) via GMP (i.e., mapping of sub-1.238, supra-1.238 to sub-2.488, close-to 9.995, close-to 40.149 and close-to 104.134 Gbit/s CBR client signal)
- 1000BASE-X: Mapping of a 1000BASE-X signal via timing transparent transcoding into OPU0
- FC-1200: Mapping of a FC-1200 signal via timing transparent transcoding into OPU2e

The following table summarizes the new G-PID values with respect to the LSP Encoding Type:

Value	G-PID Type	LSP Encoding Type
-----	-----	-----
59(TBA)	G.709 ODU-1.25G	G.709 ODUk
60(TBA)	G.709 ODU-any	G.709 ODUk
61(TBA)	CBRc	G.709 ODUk
62(TBA)	1000BASE-X	G.709 ODUk (k=0)
63(TBA)	FC-1200	G.709 ODUk (k=2e)

Note: Values 59 and 60 include mapping of SDH.

5. Extensions for Traffic Parameters for the Evolving G.709

The traffic parameters for G.709 are defined as follows:

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
Signal Type	Reserved	NMC/ Tolerance	
NVC		Multiplier (MT)	
	Bit_Rate		

The Signal Type MUST be extended in order to cover the new Signal Type introduced by the evolving OTN. The new Signal Type values are extended as follows:

Value	Type
0	Not significant
1	ODU1 (i.e., 2.5 Gbps)
2	ODU2 (i.e., 10 Gbps)
3	ODU3 (i.e., 40 Gbps)
4	ODU4 (i.e., 100 Gbps)
5	Reserved (for future use)
6	OCh at 2.5 Gbps
7	OCh at 10 Gbps
8	OCh at 40 Gbps
9	OCh at 100 Gbps
10	ODU0 (i.e., 1.25 Gbps)
11	ODU2e (i.e., 10Gbps for FC1200 and GE LAN)
12~19	Reserved (for future use)
20	ODUflex(CBR) (i.e., 1.25*N Gbps)
21	ODUflex(GFP-F), resizable (i.e., 1.25*N Gbps)
22	ODUflex(GFP-F), non resizable (i.e., 1.25*N Gbps)
23~255	Reserved (for future use)

NMC/Tolerance:

This field is redefined from the original definition in [RFC4328]. NMC field defined in [RFC4328] cannot be fixed value for an end-to-end circuit involving dissimilar OTN link types. For example, ODU2e requires 9 TS on ODU3 and 8 TS on ODU4. Usage of NMC field is

deprecated and SHOULD be used only with [RFC4328] generalized label format for backwards compatibility reasons. For the new generalized label format as defined in this document this field MUST be interpreted as Tolerance.

In case of ODUflex(CBR), the Bit_Rate and Tolerance fields MUST be used together to represent the actual bandwidth of ODUflex, where:

- The Bit_Rate field indicates the nominal bit rate of ODUflex(CBR) expressed in bytes per second, encoded as a 32-bit IEEE single-precision floating-point number (referring to [RFC4506] and [IEEE]). The value contained in the Bit Rate field has to keep into account both 239/238 factor and the Transcoding factor.
- The Tolerance field indicates the bit rate tolerance (part per million, ppm) of the ODUflex(CBR) encoded as an unsigned integer, which is bounded in 0~100ppm.

For example, for an ODUflex(CBR) service with Bit_Rate = 2.5Gbps and Tolerance = 100ppm, the actual bandwidth of the ODUflex is:

$$2.5\text{Gbps} * (1 \pm 100\text{ppm})$$

In case of ODUflex(GFP), the Bit_Rate field is used to indicate the nominal bit rate of the ODUflex(GFP), which implies the number of tributary slots requested for the ODUflex(GFP). Since the tolerance of ODUflex(GFP) makes no sense on tributary slot resource reservation, the Tolerance field for ODUflex(GFP) is not necessary and MUST be filled with 0.

In case of other ODUk signal types, the Bit_Rate and Tolerance fields are not necessary and MUST be set to 0.

The usage of the NVC and Multiplier (MT) fields are the same as [RFC4328].

5.1. Usage of ODUflex(CBR) Traffic Parameters

In case of ODUflex(CBR), the information of Bit_Rate and Tolerance in the ODUflex traffic parameters MUST be used to determine the total number of tributary slots N in the HO ODUk link to be reserved. Here:

N = Ceiling of

$$\frac{\text{ODUflex(CBR) nominal bit rate} * (1 + \text{ODUflex(CBR) bit rate tolerance})}{\text{ODTUK.ts nominal bit rate} * (1 - \text{HO OPUk bit rate tolerance})}$$

In this formula, the ODUflex(CBR) nominal bit rate is the bit rate of the ODUflex(CBR) on the line side, i.e., the client signal bit rate after applying the 239/238 factor (according to clause 7.3 table 7.2 of [G709-V3]) and the transcoding factor T (if needed) on the CBR client. According to clauses 17.7.3, 17.7.4 and 17.7.5 of [G709-V3]:

$\text{ODUflex(CBR) nominal bit rate} = \text{CBR client bit rate} * (239/238) / T$

The ODTUk.ts nominal bit rate is the nominal bit rate of the tributary slot of ODUk, as shown in Table 1 (referring to [G709-V3]).

Table 1 - Actual TS bit rate of ODUk (in Gbps)

ODUk.ts	Minimum	Nominal	Maximum
ODU2.ts	1.249 384 632	1.249 409 620	1.249 434 608
ODU3.ts	1.254 678 635	1.254 703 729	1.254 728 823
ODU4.ts	1.301 683 217	1.301 709 251	1.301 735 285

Note that:

Minimum bit rate of ODTUk.ts =
 $\text{ODTUk.ts nominal bit rate} * (1 - \text{HO OPUk bit rate tolerance})$

Maximum bit rate of ODTUk.ts =
 $\text{ODTUk.ts nominal bit rate} * (1 + \text{HO OPUk bit rate tolerance})$

Where: HO OPUk bit rate tolerance = 20ppm

Therefore, a node receiving a PATH message containing ODUflex(CBR) nominal bit rate and tolerance can allocate precise number of tributary slots and set up the cross-connection for the ODUflex service.

Note that for different ODUk, the bit rates of the tributary slots are different, and so the total number of tributary slots to be reserved for the ODUflex(CBR) MAY not be the same on different HO ODUk links.

An example is given below to illustrate the usage of ODUflex(CBR) traffic parameters.

As shown in Figure 1, assume there is an ODUflex(CBR) service requesting a bandwidth of (2.5Gbps, +/-100ppm) from node A to node C. In other words, the ODUflex traffic parameters indicate that Signal

Type is 20 (ODUflex(CBR)), Bit_Rate is 2.5Gbps and Tolerance is 100ppm.

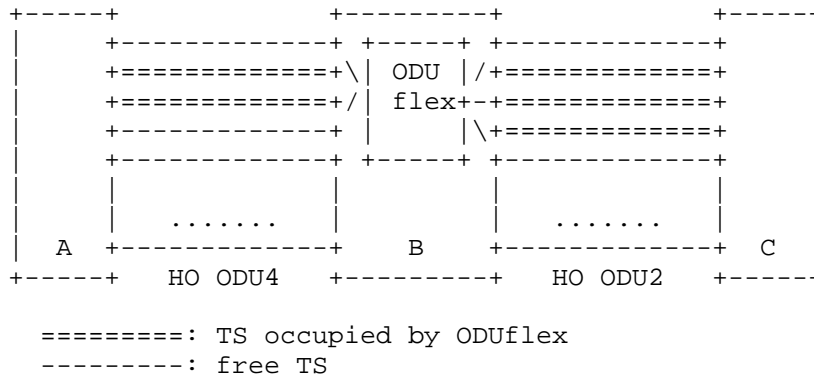


Figure 1 - Example of ODUflex(CBR) Traffic Parameters

- On the HO ODU4 link between node A and B:

The maximum bit rate of the ODUflex(CBR) equals $2.5\text{Gbps} * (1 + 100\text{ppm})$, and the minimum bit rate of the tributary slot of ODU4 equals $1.301\ 683\ 217\text{Gbps}$, so the total number of tributary slots $N1$ to be reserved on this link is:

$$N1 = \text{ceiling} (2.5\text{Gbps} * (1 + 100\text{ppm}) / 1.301\ 683\ 217\text{Gbps}) = 2$$

- On the HO ODU2 link between node B and C:

The maximum bit rate of the ODUflex equals $2.5\text{Gbps} * (1 + 100\text{ppm})$, and the minimum bit rate of the tributary slot of ODU2 equals $1.249\ 384\ 632\text{Gbps}$, so the total number of tributary slots $N2$ to be reserved on this link is:

$$N2 = \text{ceiling} (2.5\text{Gbps} * (1 + 100\text{ppm}) / 1.249\ 384\ 632\text{Gbps}) = 3$$

5.2. Usage of ODUflex(GFP) Traffic Parameters

[G709-V3-A2] recommends that the ODUflex(GFP) will fill an integral number of tributary slots of the smallest HO ODUk path over which the ODUflex(GFP) may be carried, as shown in Table 2.

Table 2 - Recommended ODUflex(GFP) bit rates and tolerance

ODU type	Nominal bit-rate	Tolerance
ODUflex(GFP) of n TS, 1<=n<=8	n * ODU2.ts	+/-100 ppm
ODUflex(GFP) of n TS, 9<=n<=32	n * ODU3.ts	+/-100 ppm
ODUflex(GFP) of n TS, 33<=n<=80	n * ODU4.ts	+/-100 ppm

According to this table, the Bit_Rate field for ODUflex(GFP) MUST equal to one of the 80 values listed below:

1 * ODU2.ts; 2 * ODU2.ts; ...; 8 * ODU2.ts;
 9 * ODU3.ts; 10 * ODU3.ts, ...; 32 * ODU3.ts;
 33 * ODU4.ts; 34 * ODU4.ts; ...; 80 * ODU4.ts.

In this way, the number of required tributary slots for the ODUflex(GFP) (i.e., the value of "n" in Table 2) can be deduced from the Bit_Rate field.

6. Generalized Label

[RFC3471] has defined the Generalized Label which extends the traditional label by allowing the representation of not only labels which are sent in-band with associated data packets, but also labels which identify time-slots, wavelengths, or space division multiplexed positions. The format of the corresponding RSVP-TE Generalized Label object is defined in the Section 2.3 of [RFC3473].

However, for different technologies, it usually needs to use specific label rather than the Generalized Label. For example, the label format described in [RFC4606] could be used for SDH/SONET, the label format in [RFC4328] for G.709.

6.1. New definition of ODU Generalized Label

In order to be compatible with new types of ODU signal and new types of tributary slot, the following new ODU label format MUST be used:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          TPN          |   Reserved   |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
~          Bit Map          .....~
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The ODU Generalized Label is used to indicate how the LO ODU_j signal is multiplexed into the HO ODU_k link. Note that the LO ODU_j signal type is indicated by traffic parameters, while the type of HO ODU_k link can be figured out locally according to the identifier of the selected interface carried in the IF_ID RSVP_HOP Object.

TPN (12 bits): indicates the Tributary Port Number (TPN) for the assigned Tributary Slot(s).

- In case of LO ODU_j multiplexed into HO ODU₁/ODU₂/ODU₃, only the lower 6 bits of TPN field are significant and the other bits of TPN MUST be set to 0.
- In case of LO ODU_j multiplexed into HO ODU₄, only the lower 7 bits of TPN field are significant and the other bits of TPN MUST be set to 0.
- In case of ODU_j mapped into OTU_k (j=k), the TPN is not needed and this field MUST be set to 0.

As per [G709-V3], The TPN is used to allow for correct demultiplexing in the data plane. When an LO ODU_j is multiplexed into HO ODU_k occupying one or more TSs, a new TPN value is configured at the two ends of the HO ODU_k link and is put into the related MSI byte(s) in the OPU_k overhead at the (traffic) ingress end of the link, so that the other end of the link can learn which TS(s) is/are used by the LO ODU_j in the data plane.

According to [G709-V3], the TPN field MUST be set as according to the following tables:

Table 3 - TPN Assignment Rules (2.5Gbps TS granularity)

HO ODU _k	LO ODU _j	TPN	TPN Assignment Rules
ODU ₂	ODU ₁	1~4	Fixed, = TS# occupied by ODU ₁
ODU ₃	ODU ₁	1~16	Fixed, = TS# occupied by ODU ₁
	ODU ₂	1~4	Flexible, != other existing LO ODU ₂ s' TPNs

Table 4 - TPN Assignment Rules (1.25Gbps TS granularity)

HO ODUk	LO ODUj	TPN	TPN Assignment Rules
ODU1	ODU0	1~2	Fixed, = TS# occupied by ODU0
ODU2	ODU1	1~4	Flexible, != other existing LO ODUs' TPNs
	ODU0 & ODUflex	1~8	Flexible, != other existing LO ODU0s and ODUflexes' TPNs
ODU3	ODU1	1~16	Flexible, != other existing LO ODUs' TPNs
	ODU2	1~4	Flexible, != other existing LO ODU2s' TPNs
	ODU0 & ODU2e & ODUflex	1~32	Flexible, != other existing LO ODU0s and ODU2es and ODUflexes' TPNs
ODU4	Any ODU	1~80	Flexible, != ANY other existing LO ODUs' TPNs

Note that in the case of "Flexible", the value of TPN MAY not be corresponding to the TS number as per [G709-V3].

Length (12 bits): indicates the number of bit of the Bit Map field, i.e., the total number of TS in the HO ODUk link.

In case of an ODUk mapped into OTUk, there is no need to indicate which tributary slots will be used, so the length field MUST be set to 0.

Bit Map (variable): indicates which tributary slots in HO ODUk that the LO ODUj will be multiplexed into. The sequence of the Bit Map is consistent with the sequence of the tributary slots in HO ODUk. Each bit in the bit map represents the corresponding tributary slot in HO ODUk with a value of 1 or 0 indicating whether the tributary slot will be used by LO ODUj or not.

Padded bits are added behind the Bit Map to make the whole label a multiple of four bytes if necessary. Padded bit MUST be set to 0 and MUST be ignored.

Note that the Length field in the label format MAY also be used to indicate the TS type of the HO ODUk (i.e., TS granularity at 1.25Gbps or 2.5Gbps) since the HO ODUk type can be known from IF_ID RSVP_HOP Object. In some cases when there is no LMP (Link Management Protocol)

or routing to make the two end points of the link to know the TSG, the TSG information used by another end can be deduced from the label format. For example, for HO ODU2 link, the value of the length field will be 4 or 8, which indicates the TS granularity is 2.5Gbps or 1.25Gbps, respectively.

6.2. Examples

The following examples are given in order to illustrate the label format described in the previous sections of this document.

(1) ODUk into OTUk mapping:

In such conditions, the downstream node along an LSP returns a label indicating that the ODUk (k=1, 2, 3, 4) is directly mapped into the corresponding OTUk. The following example label indicates an ODU1 mapped into OTU1.

```

      0               1               2               3
0 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          TPN = 0          |   Reserved   |   Length = 0   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

(2) ODUj into ODUk multiplexing:

In such conditions, this label indicates that an ODUj is multiplexed into several tributary slots of OPUk and then mapped into OTUk. Some instances are shown as follow:

- ODU0 into ODU2 Multiplexing:

```

      0               1               2               3
0 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          TPN = 2          |   Reserved   |   Length = 8   |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 1 0 0 0 0 0 0 |          Padded Bits (0)          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This above label indicates an ODU0 multiplexed into the second tributary slot of ODU2, wherein there are 8 TS in ODU2 (i.e., the type of the tributary slot is 1.25Gbps), and the TPN value is 2.

- ODU1 into ODU2 Multiplexing with 1.25Gbps TS granularity:


```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          TPN = 1          |   Reserved   |   Length = 8   |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 1 0 1 0 0 0 0 |          Padded Bits (0)          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This above label indicates an ODU1 multiplexed into the 2nd and the 4th tributary slot of ODU2, wherein there are 8 TS in ODU2 (i.e., the type of the tributary slot is 1.25Gbps), and the TPN value is 1.

- ODU2 into ODU3 Multiplexing with 2.5Gbps TS granularity:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          TPN = 1          |   Reserved   |   Length = 16   |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 1 1 0 1 0 1 0 0 0 0 0 0 0 0 0 |          Padded Bits (0)          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This above label indicates an ODU2 multiplexed into the 2nd, 3rd, 5th and 7th tributary slot of ODU3, wherein there are 16 TS in ODU3 (i.e., the type of the tributary slot is 2.5Gbps), and the TPN value is 1.

6.3. Label Distribution Procedure

This document does not change the existing label distribution procedures [RFC4328] for GMPLS except that the new ODUk label MUST be processed as follows.

When a node receives a generalized label request for setting up an ODUj LSP from its upstream neighbor node, the node MUST generate an ODU label according to the signal type of the requested LSP and the free resources (i.e., free tributary slots of ODUk) that will be reserved for the LSP, and send the label to its upstream neighbor node.

In case of ODUj to ODUk multiplexing, the node MUST firstly determine the size of the Bit Map field according to the signal type and the tributary slot type of ODUk, and then set the bits to 1 in the Bit Map field corresponding to the reserved tributary slots. The node MUST also assign a valid TPN, which MUST not collide with other TPN value used by existing LO ODU connections in the selected HO ODU link, and configure the expected multiplex structure identifier (ExMSI) using this TPN. Then, the assigned TPN MUST be filled into the label.

In case of ODUk to OTUk mapping, TPN field MUST be set to 0. Bit Map information is not REQUIRED and MUST not be included, so Length field MUST be set to 0 as well.

In order to process a received ODU label, the node MUST firstly learn which ODU signal type is multiplexed or mapped into which ODU signal type accordingly to the traffic parameters and the IF_ID RSVP_HOP Object in the received message.

In case of ODUj to ODUk multiplexing, the node MUST retrieve the reserved tributary slots in the ODUk by its downstream neighbor node according to the position of the bits that are set to 1 in the Bit Map field. The node determines the TS type (according to the total TS number of the ODUk, or pre-configured TS type), so that the node, based on the TS type, can multiplex the ODUj into the ODUk. The node MUST also retrieve the TPN value assigned by its downstream neighbor node from the label, and fill the TPN into the related MSI byte(s) in the OPUk overhead in the data plane, so that the downstream neighbor node can check whether the TPN received from the data plane is consistent with the ExMSI and determine whether there is any mismatch defect.

In case of ODUk to OTUk mapping, the size of Bit Map field MUST be 0 and no additional procedure is needed.

Note that the procedures of other label related objects (e.g., Upstream Label, Label Set) are similar to the one described above.

Note also that the TPN in the label_ERO MAY not be assigned (i.e., TPN field = 0) if the TPN is requested to be assigned locally.

6.3.1. Notification on Label Error

When receiving an ODUk label from the neighbor node, the node SHOULD check the integrity of the label. An error message containing an "Unacceptable label value" indication ([RFC3209]) SHOULD be sent if one of the following cases occurs:

- Invalid value in the length field.
- The selected link only supports 2.5Gbps TS granularity while the Length field in the label along with ODUk signal type indicates the 1.25Gbps TS granularity;
- The label includes an invalid TPN value that breaks the TPN assignment rules;

- The reserved resources (i.e., the number of "1" in the Bit Map field) do not match with the Traffic Parameters.

6.4. Supporting Virtual Concatenation and Multiplication

As per [RFC6344], the VCGs can be created using Co-Signaled style or Multiple LSPs style.

In case of Co-Signaled style, the explicit ordered list of all labels reflects the order of VCG members, which is similar to [RFC4328]. In case of multiplexed virtually concatenated signals (NVC > 1), the first label indicates the components of the first virtually concatenated signal; the second label indicates the components of the second virtually concatenated signal; and so on. In case of multiplication of multiplexed virtually concatenated signals (MT > 1), the first label indicates the components of the first multiplexed virtually concatenated signal; the second label indicates components of the second multiplexed virtually concatenated signal; and so on.

In case of Multiple LSPs style, multiple control plane LSPs are created with a single VCG and the VCAT Call can be used to associate the control plane LSPs. The procedures are similar to section 6 of [RFC6344].

7. Supporting Hitless Adjustment of ODUflex (GFP)

[G.7044] describes the procedure of ODUflex (GFP) hitless resizing using LCR (Link Connection Resize) and BWR (Bandwidth Resize) protocols in OTN data plane.

For the control plane, signaling messages are REQUIRED to initiate the adjustment procedure. Section 2.5 and Section 4.6.4 of [RFC3209] describe how the Share Explicit (SE) style is used in TE network for bandwidth increasing and decreasing, which SHOULD be still applicable for triggering the ODUflex (GFP) adjustment procedure in data plane.

Note that the SE style SHOULD be used at the beginning when creating a resizable ODUflex connection (Signal Type = 21). Otherwise an error with Error Code "Conflicting reservation style" SHOULD be generated when performing bandwidth adjustment.

- Bandwidth increasing

In order to increase the bandwidth of an ODUflex (GFP) connection, a Path message with SE style (keeping Tunnel ID unchanged and

assigning a new LSP ID) is sent along the path.

A downstream node compares the old Traffic Parameters (stored locally) with the new one carried in the Path message, to determine the number of TS to be added. After choosing and reserving new free TS, the downstream node sends back a Resv message carrying both the old and new LABEL Objects in the SE flow descriptor, so that its upstream neighbor can determine which TS are added. And the LCR protocol between each pair of neighbor nodes is triggered.

On the source node, the BWR protocol will be triggered by the successful completion of LCR protocols on every hop after Resv message is processed. On success of BWR, the source node SHOULD send a PathTear message to delete the old control state (i.e., the control state of the ODUflex (GFP) before resizing) on the control plane.

- Bandwidth decreasing

The SE style SHOULD also be used for ODUflex bandwidth decreasing. For each pair of neighbor nodes, the sending and receiving Resv message with old and new LABEL Objects will trigger the first step of LCR between them to perform LCR handshake. On the source node, the BWR protocol will be triggered by the successful completion of LCR handshake on every hop after Resv message is processed. On success of BWR, the second step of LCR, i.e., link connection decrease procedure will be started on every hop of the connection.

Similarly, after completion of bandwidth decreasing, a ResvErr message SHOULD be sent to tear down the old control state.

8. Control Plane Backward Compatibility Considerations

As described in [OTN-FWK], since the [RFC4328] has been deployed in the network for the nodes that support [G709-V1], control plane backward compatibility SHOULD be taken into consideration. More specifically:

- o Nodes supporting this document SHOULD support [OTN-OSPF].
- o Nodes supporting this document MAY support [RFC4328] signaling.

- o A node supporting both sets of procedures (i.e., [RFC4328] and this document) is NOT REQUIRED to signal an LSP using both procedures, i.e., to act as a signaling version translator.
- o Ingress nodes that support both sets of procedures MAY select which set of procedures to follow based on routing information or local policy.
- o Per [RFC3473], nodes that do not support this document will generate a PathErr message, with a "Routing problem/Switching Type" indication.

9. Security Considerations

This document introduces no new security considerations to the existing GMPLS signaling protocols. Referring to [RFC3473], further details of the specific security measures are provided. Additionally, [GMPLS-SEC] provides an overview of security vulnerabilities and protection mechanisms for the GMPLS control plane.

10. IANA Considerations

- G.709 SENDER_TSPEC and FLOWSPEC objects:

The traffic parameters, which are carried in the G.709 SENDER_TSPEC and FLOWSPEC objects, do not require any new object class and type based on [RFC4328]:

- o G.709 SENDER_TSPEC Object: Class = 12, C-Type = 5 [RFC4328]
- o G.709 FLOWSPEC Object: Class = 9, C-Type = 5 [RFC4328]

- Generalized Label Object:

The new defined ODU label (Section 6) is a kind of generalized label. Therefore, the Class-Num and C-Type of the ODU label is the same as that of generalized label described in [RFC3473], i.e., Class-Num = 16, C-Type = 2.

11. References

11.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, Jan 2006.
- [RFC3209] D. Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC3209, December 2001.
- [RFC3471] Berger, L., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] L. Berger, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC6344] G. Bernstein et al, "Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS)", RFC6344, August 2011.
- [OTN-FWK] Fatai Zhang et al, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", draft-ietf-ccamp-gmpls-g709-framework, Work in Progress, August 2012.
- [OTN-INFO] S. Belotti et al, "Information model for G.709 Optical Transport Networks (OTN)", draft-ietf-ccamp-otn-g709-info-model, Work in Progress, July 2012.
- [OTN-OSPF] D. Ceccarelli et al, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolving G.709 OTN Networks", draft-ietf-ccamp-gmpls-ospf-g709v3, Work in Progress, April 2012.
- [G709-V3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709/Y.1331, December 2009.

[G709-V3-A2] ITU-T, "Interfaces for the Optical Transport Network (OTN) Amendment 2", G.709/y.1331 Amendment 2, April 2011.

11.2. Informative References

[G709-V1] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation (and Amendment 1), February 2001 (November 2001).

[G709-V2] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation, March 2003.

[G798-V2] ITU-T, "Characteristics of optical transport network hierarchy equipment functional blocks", G.798, December 2006.

[G798-V3] ITU-T, "Characteristics of optical transport network hierarchy equipment functional blocks", G.798v3, consented June 2010.

[G.7044] ITU-T, "Hitless adjustment of ODUflex", G.7044 (and Amendment 1), February 2012.

[RFC4506] M. Eisler, Ed., "XDR: External Data Representation Standard", RFC 4506, May 2006.

[IEEE] "IEEE Standard for Binary Floating-Point Arithmetic", ANSI/IEEE Standard 754-1985, Institute of Electrical and Electronics Engineers, August 1985.

[GMPLS-SEC] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", Work in Progress, October 2009.

12. Contributors

Jonathan Sadler, Tellabs
Email: jonathan.sadler@tellabs.com

Kam LAM, Alcatel-Lucent
Email: kam.lam@alcatel-lucent.com

Xiaobing Zi, Huawei Technologies
Email: zixiaobing@huawei.com

Francesco Fondelli, Ericsson
Email: francesco.fondelli@ericsson.com

Lyndon Ong, Ciena
Email: lyong@ciena.com

Biao Lu, infinera
Email: blu@infinera.com

13. Authors' Addresses

Fatai Zhang (editor)
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Guoying Zhang
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.China
Phone: +86-10-68094272
Email: zhangguoying@mail.ritt.com.cn

Sergio Belotti
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6863033
Email: sergio.belotti@alcatel-lucent.it

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

Khuzema Pithewan
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089, USA
Email: kpithewan@infinera.com

Yi Lin
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972914
Email: yi.lin@huawei.com

Yunbin Xu
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.China
Phone: +86-10-68094134
Email: xuyunbin@mail.ritt.com.cn

Pietro Grandi
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6864930
Email: pietro_vittorio.grandi@alcatel-lucent.it

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: diego.caviglia@ericsson.com

Rajan Rao
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA
Email: rrao@infinera.com

John E Drake
Juniper
Email: jdrake@juniper.net

Igor Bryskin
Adva Optical
Email: IBryskin@advaoptical.com

14. Acknowledgment

The authors would like to thank Lou Berger and Deborah Brungard for their useful comments to the document.

Intellectual Property

The IETF Trust 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 any IETF 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.

Copies of Intellectual Property 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 any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of RFC 5378. No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under RFC 5378, shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

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

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: January 13, 2013

S. Belotti, Ed.
P. Grandi
Alcatel-Lucent
D. Ceccarelli, Ed.
D. Caviglia
Ericsson
F. Zhang
D. Li
Huawei Technologies
July 12, 2012

Information model for G.709 Optical Transport Networks (OTN)
draft-ietf-ccamp-otn-g709-info-model-04

Abstract

The recent revision of ITU-T recommendation G.709 [G.709-v3] has introduced new fixed and flexible ODU containers in Optical Transport Networks (OTNs), enabling optimized support for an increasingly abundant service mix.

This document provides a model of information needed by the routing and signaling process in OTNs to support Generalized Multiprotocol Label Switching (GMPLS) control of all currently defined ODU containers.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 13, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
1.1. Terminology	4
2. OSPF-TE requirements overview	4
3. RSVP-TE requirements overview	5
4. G.709 Digital Layer Info Model for Routing and Signaling	5
4.1. Tributary Slot Granularity	8
4.1.1. Data Plane Considerations	8
4.1.1.1. Payload Type and TSG relationship	8
4.1.1.2. Fall-back procedure	10
4.1.2. Control Plane considerations	10
4.2. Tributary Port Number	14
4.3. Signal type	14
4.4. Bit rate and tolerance	15
4.5. Unreserved Resources	16
4.6. Maximum LSP Bandwidth	16
4.7. Distinction between terminating and switching capability	16
4.8. Priority Support	19
4.9. Multi-stage multiplexing	19
4.10. Generalized Label	20
5. Security Considerations	20
6. IANA Considerations	20
7. Contributors	21
8. Acknowledgements	21
9. References	21
9.1. Normative References	21
9.2. Informative References	22
Authors' Addresses	23

1. Introduction

GMPLS[RFC3945] extends MPLS to include Layer-2 Switching (L2SC), Time-Division Multiplexing (e.g., SONET/SDH, PDH, and OTN), Wavelength (OCh, Lambdas) Switching and Spatial Switching (e.g., incoming port or fiber to outgoing port or fiber).

The establishment of LSPs that span only interfaces recognizing packet/cell boundaries is defined in [RFC3036, RFC3212, RFC3209]. [RFC3471] presents a functional description of the extensions to Multi-Protocol Label Switching (MPLS) signaling required to support GMPLS. ReSource reserVation Protocol-Traffic Engineering (RSVP-TE) -specific formats, mechanisms and technology specific details are defined in [RFC3473].

From a routing perspective, Open Shortest Path First-Traffic Engineering (OSPF-TE) generates Link State Advertisements (LSAs) carrying application-specific information and floods them to other nodes as defined in [RFC5250]. Three types of opaque LSA are defined, i.e. type 9 - link-local flooding scope, type 10 - area-local flooding scope, type 11 - AS flooding scope.

Type 10 LSAs are composed of a standard LSA header and a payload including one top-level TLV and possible several nested sub-TLVs. [RFC3630] defines two top-level TLVs: Router Address TLV and Link TLV; and nine possible sub-TLVs for the Link TLV, used to carry link related TE information. The Link type sub-TLVs are enhanced by [RFC4203] in order to support GMPLS networks and related specific link information. In GMPLS networks each node generates TE LSAs to advertise its TE information and capabilities (link-specific or node-specific) through the network. The TE information carried in the LSAs are collected by the other nodes of the network and stored into their local Traffic Engineering Databases (TED).

In a GMPLS enabled G.709 Optical Transport Networks (OTN), routing and signaling are fundamental in order to allow automatic calculation and establishment of routes for ODUk LSPs. The recent revision of ITU-T Recommendation G.709 [G709-V3] has introduced new fixed and flexible ODU containers that augment those specified in foundation OTN. As a result, it is necessary to provide OSPF-TE and RSVP-TE extensions to allow GMPLS control of all currently defined ODU containers.

This document provides the information model needed by the routing and signaling processes in OTNs to allow GMPLS control of all currently defined ODU containers.

OSPF-TE and RSVP-TE requirements are defined in [OTN-FWK], while

protocol extensions are defined in [OTN-OSPF] and [OTN-RSVP].

1.1. Terminology

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 [RFC2119].

2. OSPF-TE requirements overview

[OTN-FWK] provides a set of functional routing requirements summarized below :

- Support for link multiplexing capability advertisement: The routing protocol has to be able to carry information regarding the capability of an OTU link to support different type of ODUs
- Support of any ODUk and ODUFlex: The routing protocol must be capable of carrying the required link bandwidth information for performing accurate route computation for any of the fixed rate ODUs as well as ODUFlex.
- Support for differentiation between switching and terminating capacity
- Support for the client server mappings as required by [G.7715.1]. The list of different mappings methods is reported in [G.709-v3]. Since different methods exist for how the same client layer is mapped into a server layer, this needs to be captured in order to avoid the set-up of connections that fail due to incompatible mappings.
- Support different priorities for resource reservation. How many priorities levels should be supported depends on operator policies. Therefore, the routing protocol should be capable of supporting either no priorities or up to 8 priority levels as defined in [RFC4202].
- Support link bundling of component links at the same line rate and with same muxing hierarchy.
- Support for Tributary Slot Granularity (TSG) advertisement.

3. RSVP-TE requirements overview

[OTN-FWK] also provides a set of functional signaling requirements summarized below :

- Support for LSP setup of new ODUk/ODUflex containers with related mapping and multiplexing capabilities
- Support for LSP setup using different Tributary Slot granularity
- Support for Tributary Port Number allocation and negotiation
- Support for constraint signaling
- Support for TSG signaling

4. G.709 Digital Layer Info Model for Routing and Signaling

The digital OTN layered structure is comprised of digital path layer networks (ODU) and digital section layer networks (OTU). An OTU section layer supports one ODU path layer as client and provides monitoring capability for the OCh. An ODU path layer may transport a heterogeneous assembly of ODU clients. Some types of ODUs (i.e., ODU1, ODU2, ODU3, ODU4) may assume either a client or server role within the context of a particular networking domain. ITU-T G.872 recommendation provides two tables defining mapping and multiplexing capabilities of OTNs, which are reproduced below.

ODU client	OTU server
ODU 0	-
ODU 1	OTU 1
ODU 2	OTU 2
ODU 2e	-
ODU 3	OTU 3
ODU 4	OTU 4
ODU flex	-

Figure 1: OTN mapping capability

ODU client	ODU server
1,25 Gbps client	ODU 0
-	
2,5 Gbps client	ODU 1
ODU 0	
10 Gbps client	ODU 2
ODU0,ODU1,ODUflex	
10,3125 Gbps client	ODU 2e
-	
40 Gbps client	ODU 3
ODU0,ODU1,ODU2,ODU2e,ODUflex	
100 Gbps client	ODU 4
ODU0,ODU1,ODU2,ODU2e,ODU3,ODUflex	
CBR clients from greater than 2.5 Gbit/s to 100 Gbit/s: or GFP-F mapped packet clients from 1.25 Gbit/s to 100 Gbit/s.	ODUflex
-	

Figure 2: OTN multiplexing capability

How an ODU_k connection service is transported within an operator network is governed by operator policy. For example, the ODU_k connection service might be transported over an ODU_k path over an OTU_k section, with the path and section being at the same rate as that of the connection service (see Table 1). In this case, an entire lambda of capacity is consumed in transporting the ODU_k connection service. On the other hand, the operator might exploit different multiplexing capabilities in the network to improve infrastructure efficiencies within any given networking domain. In

this case, ODUk multiplexing may be performed prior to transport over various rate ODU servers (as per Table 2) over associated OTU sections.

From the perspective of multiplexing relationships, a given ODUk may play different roles as it traverses various networking domains.

As detailed in [OTN-FWK], client ODUk connection services can be transported over:

- o Case A) one or more wavelength sub-networks connected by optical links or
- o Case B) one or more ODU links (having sub-lambda and/or lambda bandwidth granularity)
- o Case C) a mix of ODU links and wavelength sub-networks.

This document considers the TE information needed for ODU path computation and parameters needed to be signaled for LSP setup.

The following sections list and analyze each type of data that needs to be advertised and signaled in order to support path computation and LSP setup.

4.1. Tributary Slot Granularity

ITU-T recommendation defines two type of TS granularity. This TS granularity is defined per layer, meaning that both ends of a link can select proper TS granularity differently for each supported layer, based on the rules below:

- If both ends of a link are new cards supporting both 1.25Gbps TS and 2.5Gbps TS, then the link will work with 1.25Gbps TS.
- If one end is a new card supporting both the 1.25Gbps and 2,5Gbps TS, and the other end is an old card supporting just the 2.5Gbps TS, the link will work with 2.5Gbps TS.

4.1.1. Data Plane Considerations

4.1.1.1. Payload Type and TSG relationship

As defined in G.709 an ODUk container consist of an OPUk (Optical Payload Unit) plus a specific ODUk Overhead (OH). OPUk OH information is added to the OPUk information payload to create an OPUk. It includes information to support the adaptation of client signals. Within the OPUk overhead there is the payload structure

identifier (PSI) that includes the payload type (PT). The payload type (PT) is used to indicate the composition of the OPUk signal. When an ODUj signal is multiplexed into an ODUk, the ODUj signal is first extended with frame alignment overhead and then mapped into an Optical channel Data Tributary Unit (ODTU). Two different types of ODTU are defined in G.709:

- ODTUjk ((j,k) = {(0,1), (1,2), (1,3), (2,3)}; ODTU01, ODTU12, ODTU13 and ODTU23) in which an ODUj signal is mapped via the asynchronous mapping procedure (AMP), defined in clause 19.5 of G.709.
- ODTUk.ts ((k,ts) = (2,1..8), (3,1..32), (4,1..80)) in which a lower order ODU (ODU0, ODU1, ODU2, ODU2e, ODU3, ODUflex) signal is mapped via the generic mapping procedure (GMP), defined in clause 19.6 of G.709.

G.709 introduces also a logical entity, called ODTUGk, characterizing the multiplexing of the various ODTU. The ODTUGk is then mapped into OPUK. ODTUjk and ODTUk.ts signals are directly time-division multiplexed into the tributary slots of an HO OPUk.

When PT is assuming value 20 or 21, together with OPUk type (K=1,2,3,4), it is used to discriminate two different ODU multiplex structure ODTUGx :

- Value 20: supporting ODTUjk only,
- Value 21: supporting ODTUk.ts or ODTUk.ts and ODTUjk.

The discrimination is needed for OPUk with K =2 or 3, since OPU2 and OPU3 are able to support both the different ODU multiplex structures. For OPU4 and OPU1, only one type of ODTUG is supported: ODTUG4 with PT=21 and ODTUG1 with PT=20. (see table Figure 6). The relationship between PT and TS granularity, is in the fact that the two different ODTUGk discriminated by PT and OPUk are characterized by two different TS granularities of the related OPUk, the former at 2.5 Gbps, the latter at 1.25Gbps.

In order to complete the picture, in the PSI OH there is also the Multiplex Structure Identifier (MSI) that provides the information on which tributary slots the different ODTUjk or ODTUk.ts are mapped into the related OPUk. The following figure shows how the client traffic is multiplexed till the OPUk layer.

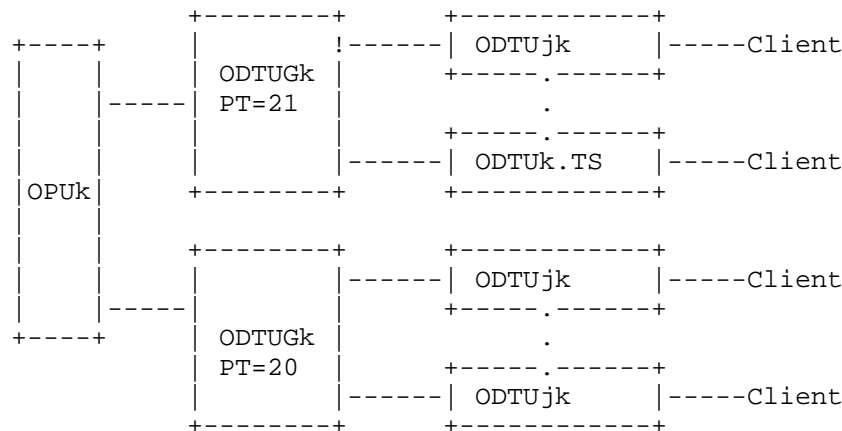


Figure 3: OTN client multiplexing

4.1.1.2. Fall-back procedure

SG15 ITU-T G.798 recommendation describes the so called PT=21-to-PT=20 interworking process that explains how two equipments with interfaces with different PayloadType, and hence different TS granularity (1.25Gbps vs. 2.5Gbps), can be coordinated so to permit the equipment with 1.25 TS granularity to adapt his TS allocation accordingly to the different TS granularity (2.5Gbps) of a neighbor.

Therefore, in order to let the NE change TS granularity accordingly to the neighbour requirements, the AUTOpayloadtype needs to be set. When both the neighbors (link or trail) have been configured as structured, the payload type received in the overhead is compared to the transmitted PT. If they are different and the transmitted PT=21, the node must fallback to PT=20. In this case the fall-back process makes the system self consistent and the only reason for signaling the TS granularity is to provide the correct label (i.e. label for PT=21 has twice the TS number of PT=20). On the other side, if the AUTOpayloadtype is not configured, the RSVP-TE consequent actions in case of TS mismatch need to be defined.

4.1.2. Control Plane considerations

When setting up an ODUj over an ODUk, it is possible to identify two types of TSG, the server and the client one. The server TSG is used to map an end to end ODUj onto a server ODUk LSP or links. This parameter can not be influenced in any way from the ODUj LSP: ODUj LSP will be mapped on tributary slots available on the different links/ODUk LSPs. When setting up an ODUj at a given rate, the fact

that it is carried over a path composed by links/FAs structured with 1.25Gbps or 2.5Gbps TS size is completely transparent to the end to end ODUj.

On the other side the client TSG is the tributary slot size that is exported towards the client layer. The client TSG information is one of the parameters needed to correctly select the adaptation towards the client layers at the end nodes and this is the only thing that the ODUj has to guarantee. When setting up an HO-ODUk/OTUk LSP or an H-LSP/FA, in the case where the egress interface cannot be identified from the ERO, it is necessary for the penultimate node to select an interface on the egress node that supports the TSG and ODU client hierarchy specified in signaling. It must then select an interface on itself that can be paired with the interface it selected.

In figure 4 an example of client and server TSG utilization in a scenario with mixed G.709 v2 and G.709 v3 interfaces is shown.

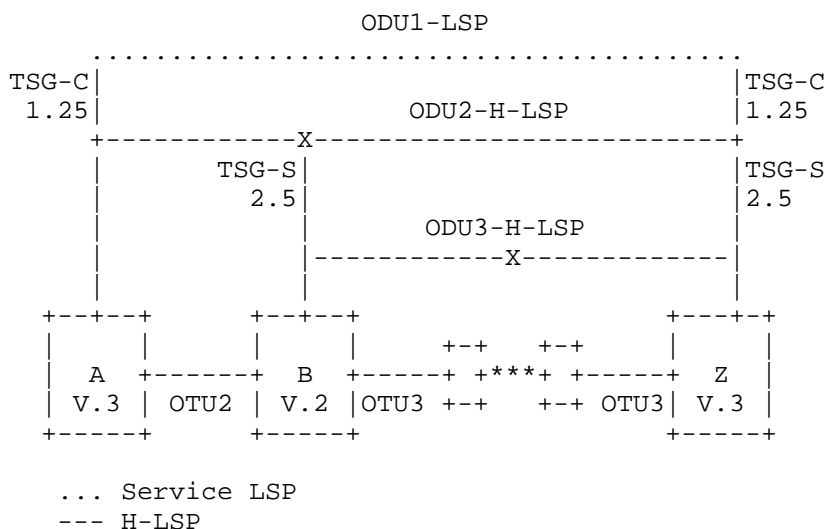


Figure 4: Client-Server TSG example

In this scenario, an ODU3 LSP is setup from node B to Z. Node B has an old interface able to support 2.5 TSG granularity, hence only client TSG equal to 2.5Gbps can be exported to ODU3 H-LSP possible clients. An ODU2 LSP is setup from node A to node Z with client TSG 1.25 signaled and exported towards clients. The ODU2 LSP is carried by ODU3 H-LSP from B to Z. Due to the limitations of old node B interface, the ODU2 LSP is mapped with 2.5Gbps TSG over the ODU3

H-LSP. Then an ODU1 LSP is setup from A to Z, carried by the ODU2 H-LSP and mapped over it using a 1.25Gbps TSG.

What is shown in the example is that the TSG processing is a per layer issue: even if the ODU3 H-LSP is created with TSG client at 2.5Gbps, the ODU2 H-LSP must guarantee a 1.25Gbps TSG client. ODU3 H-LSP is eligible from ODU2 LSP perspective since from the routing it is known that this ODU3 interface at node Z, supports an ODU2 termination exporting a TSG 1.25/2.5.

Moreover, with respect to the penultimate hop implications let's consider a further example in which the setup of an ODU3 path that is going to carry an ODU0 is considered. In this case it is needed the support of 1,25 GBps TS. The information related to the TSG is carried in the signaling and node C, having two different interfaces toward D with different TSGs, can choose the right one as depicted in the following figure. In case the full ERO is provided in the signaling with explicit interface declaration, there is no need for C to choose the right interface as it has been already decided by the ingress node or the PCE.

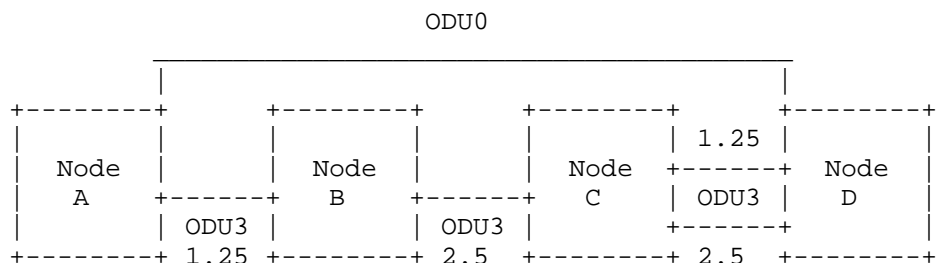


Figure 5: TSG in signaling

The TSG information is needed also in the routing protocol as the ingress node (A in the previous example) needs to know if the interfaces between C and D can support the required TSG. In case they cannot, A will compute an alternate path from itself to D.

In a multi-stage multiplexing environment any layer can have a different TSG structure, e.g. in a multiplexing hierarchy like ODU0->ODU2->ODU3, the ODU3 can be structured at TSG=2.5 in order to support an ODU2 connection, but this ODU2 connection can be a tunnel for ODU0, and hence structured with 1.25 TSG. Therefore any multiplexing level has to advertise his TSG capabilities in order to allow a correct path computation by the end nodes (both of the ODUk trail and of the H-LSP/FA).

The following table shows the different mapping possibilities depending on the TSG types. The client types are shown in the left column, while the different OPUk server and related TSGs are listed in the top row. The table also shows the relationship between the TSG and the payload type.

	2.5G TS			1.25G TS			
	OPU2	OPU3		OPU1	OPU2	OPU3	OPU4
ODU0	-	-		AMP PT=20	GMP PT=21	GMP PT=21	GMP PT=21
ODU1	AMP PT=20	AMP PT=20		-	AMP PT=21	AMP PT=21	GMP PT=21
ODU2	-	AMP PT=20		-	-	AMP PT=21	GMP PT=21
ODU2e	-	-		-	-	GMP PT=21	GMP PT=21
ODU3	-	-		-	-	-	GMP PT=21
ODUf1	-	-		-	GMP PT=21	GMP PT=21	GMP PT=21

Figure 6: ODUj into OPUk mapping types

The signaled TSGs information is not enough to have a complete choice since the penultimate hop node has to distinguish between interfaces with the same TSG (e.g. 1.25Gbps) whether the interface is able to support the right hierarchy, i.e. it is possible to have two interfaces both at 1.25 TSG but only one is supporting ODU0.

A dedicated optional object could be defined in order to carry the multiplexing hierarchy and adaptation information (i.e. TSG/PT, AMP/GMP) so to have a more precise choice capability. In this way, when the penultimate node receives such object, together with the Traffic Parameters Object, is allowed to choose the correct interface towards the egress node.

In conclusion both routing and signaling will need to be extended to appropriately represent the TSG/PT information. Routing will need to

represent a link's TSG and PT capabilities as well as the supported multiplexing hierarchy. Signaling will need to represent the TSG/PT and multiplexing hierarchy encoding.

4.2. Tributary Port Number

[RFC4328] supports only the deprecated auto-MSI mode which assumes that the Tributary Port Number is automatically assigned in the transmit direction and not checked in the receive direction.

As described in [G709-V3] and [G798-V3], the OPUk overhead in an OTUk frame contains n (n = the total number of TSs of the ODUk) MSI (Multiplex Structure Identifier) bytes (in the form of multi-frame), each of which is used to indicate the association between tributary port number and tributary slot of the ODUk.

The association between TPN and TS has to be configured by the control plane and checked by the data plane on each side of the link. (Please refer to [OTN-FWK] for further details). As a consequence, the RSVP-TE signaling needs to be extended to support the TPN assignment function.

4.3. Signal type

From a routing perspective, [RFC 4203] allows advertising foundation G.709 (single TS type) without the capability of providing precise information about bandwidth specific allocation. For example, in case of link bundling, dividing the unreserved bandwidth by the MAX LSP bandwidth it is not possible to know the exact number of LSPs at MAX LSP bandwidth size that can be set up. (see example fig. 3)

The lack of spatial allocation heavily impacts the restoration process, because the lack of information of free resources highly increases the number of crank-backs affecting network convergence time.

Moreover actual tools provided by OSPF-TE only allow advertising signal types with fixed bandwidth and implicit hierarchy (e.g. SDH/SONET networks) or variable bandwidth with no hierarchy (e.g. packet switching networks) but do not provide the means for advertising networks with mixed approach (e.g. ODUFlex CBR and ODUFlex packet).

For example, advertising ODU0 as MIN LSP bandwidth and ODU4 as MAX LSP bandwidth it is not possible to state whether the advertised link supports ODU4 and ODUFlex or ODU4, ODU3, ODU2, ODU1, ODU0 and ODUFlex. Such ambiguity is not present in SDH networks where the hierarchy is implicit and flexible containers like ODUFlex do not exist. The issue could be resolved by declaring 1 ISCD for each

signal type actually supported by the link.

Supposing for example to have an equivalent ODU2 unreserved bandwidth in a TE-link (with bundling capability) distributed on 4 ODU1, it would be advertised via the ISCD in this way:

MAX LSP Bw: ODU1

MIN LSP Bw: ODU1

- Maximum Reservable Bandwidth (of the bundle) set to ODU2
- Unreserved Bandwidth (of the bundle) set to ODU2

Moreover with the current IETF solutions, ([RFC4202], [RFC4203]) as soon as no bandwidth is available for a certain signal type it is not advertised into the related ISCD, losing also the related capability until bandwidth is freed.

In conclusion, the OSPF-TE extensions defined in [RFC4203] require a different ISCD per signal type in order to advertise each supported container. This motivates attempting to look for a more optimized solution, without proliferations of the number of ISCD advertised. The OSPF LSA is required to stay within a single IP PDU; fragmentation is not allowed. In a conforming Ethernet environment, this limits the LSA to 1432 bytes (Packet_MTU (1500 Bytes) - IP_Header (20 bytes) - OSPF_Header (28 bytes) - LSA_Header (20 bytes)).

With respect to link bundling, the utilization of the ISCD as it is, would not allow precise advertising of spatial bandwidth allocation information unless using only one component link per TE link.

On the other hand, from a singaling point of view, [RFC4328] describes GMPLS signaling extensions to support the control for G.709 OTNs [G709-V1]. However, [RFC4328] needs to be updated because it does not provide the means to signal all the new signal types and related mapping and multiplexing functionalities.

4.4. Bit rate and tolerance

In the current traffic parameters signaling, bit rate and tolerance are implicitly defined by the signal type. ODUflex CBR and Packet can have variable bit rates and tolerances (please refer to [OTN-FWK] table 2); it is thus needed to upgrade the signaling traffic parameters so to specify requested bit rates and tolerance values during LSP setup.

4.5. Unreserved Resources

Unreserved resources need to be advertised per priority and per signal type in order to allow the correct functioning of the restoration process. [RFC4203] only allows advertising unreserved resources per priority, this leads not to know how many LSPs of a specific signal type can be restored. As example it is possible to consider the scenario depicted in the following figure.

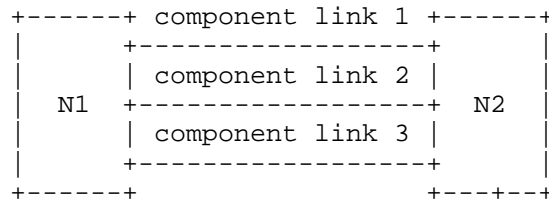


Figure 7: Concurrent path computation

Suppose to have a TE link comprising 3 ODU3 component links with 32TSs available on the first one, 24TSs on the second, 24TSs on the third and supporting ODU2 and ODU3 signal types. The node would advertise a TE link unreserved bandwidth equal to 80 TSs and a MAX LSP bandwidth equal to 32 TSs. In case of restoration the network could try to restore 2 ODU3 (64TSs) in such TE-link while only a single ODU3 can be set up and a crank-back would be originated. In more complex network scenarios the number of crank-backs can be much higher.

4.6. Maximum LSP Bandwidth

Maximum LSP bandwidth is currently advertised in the common part of the ISCD and advertised per priority, while in OTN networks it is only required for ODUflex advertising. This leads to a significant waste of bits inside each LSA.

4.7. Distinction between terminating and switching capability

The capability advertised by an interface needs further distinction in order to separate termination and switching capabilities. Due to internal constraints and/or limitations, the type of signal being advertised by an interface could be just switched (i.e. forwarded to switching matrix without multiplexing/demultiplexing actions), just terminated (demuxed) or both of them. The following figures help explaining the switching and terminating capabilities.

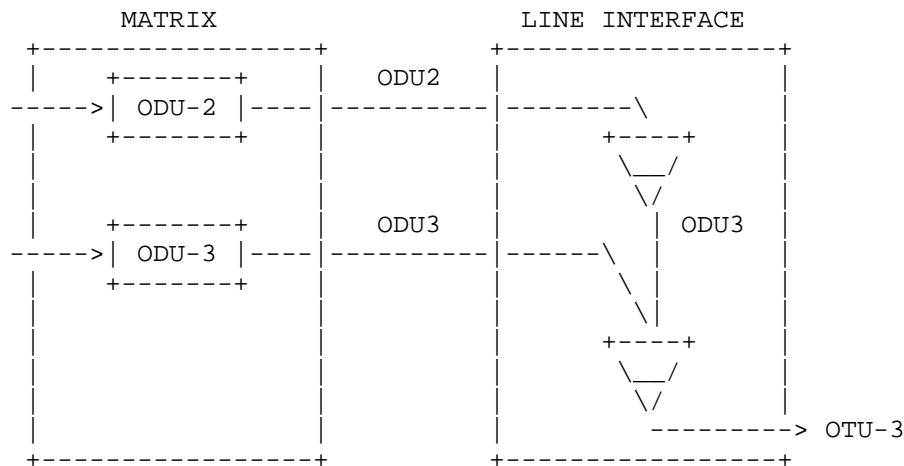


Figure 8: Switching and Terminating capabilities

The figure in the example shows a line interface able to:

- Multiplex an ODU2 coming from the switching matrix into and ODU3 and map it into an OTU3
- Map an ODU3 coming from the switching matrix into an OTU3

In this case the interface bandwidth advertised is ODU2 with switching capability and ODU3 with both switching and terminating capabilities.

This piece of information needs to be advertised together with the related unreserved bandwidth and signal type. As a consequence signaling must have the possibility to setup an LSP allowing the local selection of resources consistent with the limitations considered during the path computation.

In figures 6 and 7 there are two examples of the need of termination/switching capability differentiation. In both examples all nodes are supposed to support single-stage capability. The figure 6 addresses a scenario in which a failure on link B-C forces node A to calculate another ODU2 LSP path carrying ODU0 service along the nodes B-E-D. Being D a single stage capable node, it is able to extract ODU0 service only from ODU2 interface. Node A has to know that from E to D exists an available OTU2 link from which node D can extract the ODU0 service. This information is required in order to avoid that the OTU3 link is considered in the path computation.

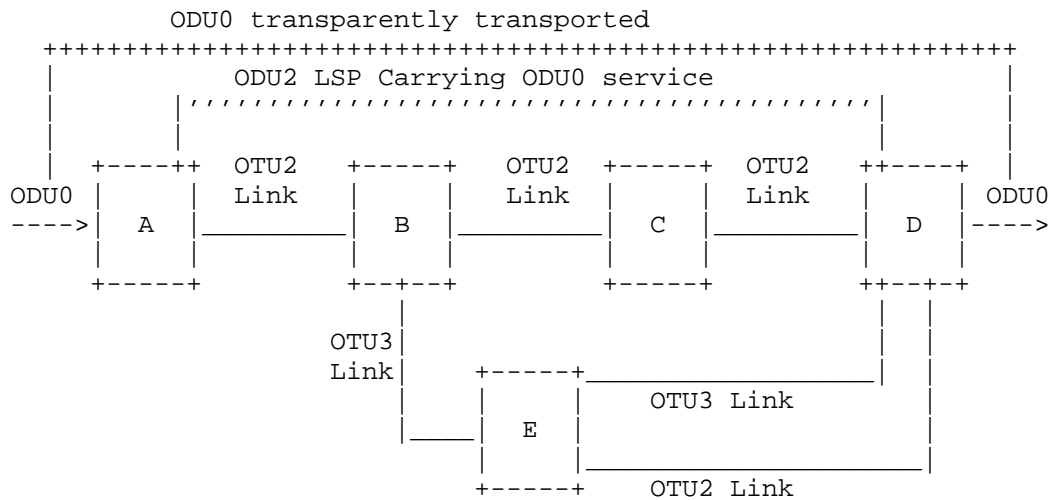


Figure 9: Switching and Terminating capabilities - Example 1

Figure 7 addresses the scenario in which the restoration of the ODU2 LSP (ABCD) is required. The two bundled component links between B and E could be used, but the ODU2 over the OTU2 component link can only be terminated and not switched. This implies that it cannot be used to restore the ODU2 LSP (ABCD). However such ODU2 unreserved bandwidth must be advertised since it can be used for a different ODU2 LSP terminating on E, e.g. (FBE). Node A has to know that the ODU2 capability on the OTU2 link can only be terminated and that the restoration of (ABCD) can only be performed using the ODU2 bandwidth available on the OTU3 link.

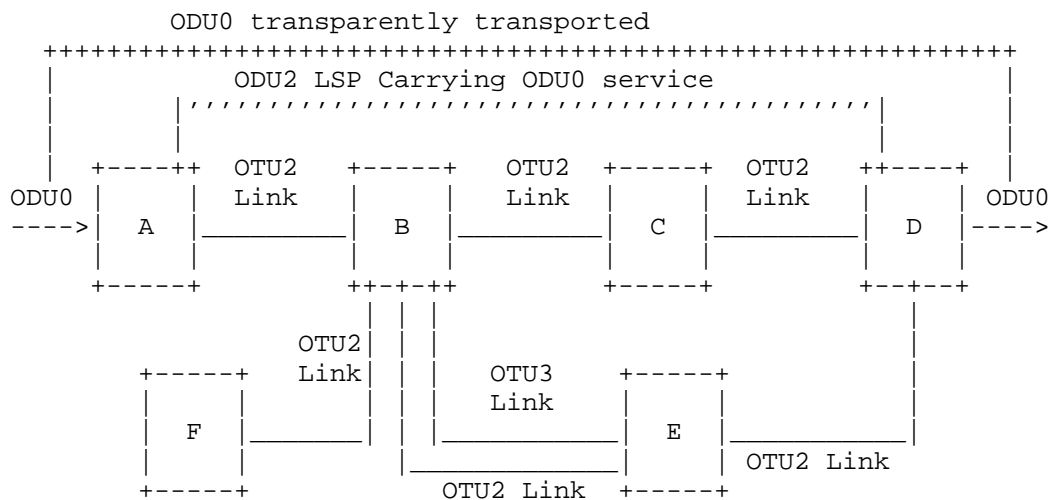


Figure 10: Switching and Terminating capabilities - Example 2

4.8. Priority Support

The IETF foresees that up to eight priorities must be supported and that all of them have to be advertised independently on the number of priorities supported by the implementation. Considering that the advertisement of all the different supported signal types will originate large LSAs, it is advised to advertise only the information related to the really supported priorities.

4.9. Multi-stage multiplexing

With reference to the [OTN-FWK], introduction of multi-stage multiplexing implies the advertisement of cascaded adaptation capabilities together with the matrix access constraints. The structure defined by IETF for the advertisement of adaptation capabilities is ISCD/IACD as in [RFC4202] and [RFC5339]. Modifications to ISCD/IACD, if needed, have to be addressed in the related encoding documents.

With respect to the routing, please note that in case of multi stage muxing hierarchy (e.g. ODU1->ODU2->ODU3), not only the ODUk/OTUk bandwidth (ODU3) and service layer bandwidth (ODU1) are needed, but also the intermediate one (ODU2). This is a typical case of spatial allocation problem.

Suppose in this scenario to have the following advertisement:

Hierarchy: ODU1->ODU2->ODU3

Number of ODU1==5

The number of ODU1 suggests that it is possible to have an ODU2 FA, but it depends on the spatial allocation of such ODU1s.

It is possible that 2 links are bundled together and 3 ODU1->ODU2->ODU3 are available on a component link and 2 on the other one, in such a case no ODU2 FA could be set up. The advertisement of the ODU2 is needed because in case of ODU1 spatial allocation (3+2), the ODU2 available bandwidth would be 0 (no ODU2 FA can be created), while in case of ODU1 spatial allocation (4+1) the ODU2 available bandwidth would be 1 (1 ODU2 FA can be created).

4.10. Generalized Label

The ODUk label format defined in [RFC4328] could be updated to support new signal types defined in [G709-V3] but would hardly be further enhanced to support possible new signal types.

Furthermore such label format may have scalability issues due to the high number of labels needed when signaling large LSPs. For example, when an ODU3 is mapped into an ODU4 with 1.25G tributary slots, it would require the utilization of thirty-one labels ($31 \times 4 \times 8 = 992$ bits) to be allocated while an ODUflex into an ODU4 may need up to eighty labels ($80 \times 4 \times 8 = 2560$ bits).

A new flexible and scalable ODUk label format needs to be defined.

5. Security Considerations

This document provides a model of information needed by the routing and signaling process in OTN networks. Such a model is very similar from a security standpoint of the information that can be currently conveyed via GMPLS routing protocols. For a general discussion on MPLS- and GMPLS-related security issues, see the MPLS/GMPLS security framework [RFC5920]

6. IANA Considerations

This informational document does not make any requests for IANA action.

7. Contributors

Jonathan Sadler, Tellabs

EMail: jonathan.sadler@tellabs.com

John Drake, Juniper

EMail: jdrake@juniper.net

Francesco Fondelli

Ericsson

Via Moruzzi 1

Pisa - 56100

Email: francesco.fondelli@ericsson.com

8. Acknowledgements

The authors would like to thank Eve Varma and Sergio Lanzone for their precious collaboration and review.

9. References

9.1. Normative References

[OTN-OSPF]

D.Ceccarelli,D.Caviglia,F.Zhang,D.Li,Y.Xu,P.Grandi,S.Belotti, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolutive G.709 OTN Networks", work in progress draft-ietf-ccamp-gmpls-ospf-g709-00, October 2011.

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.
- [RFC5339] Le Roux, J.L. and D. Papadimitriou, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 5339, September 2008.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.
- [RFC6107] Shiomoto, K. and A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC 6107, February 2011.

9.2. Informative References

- [G.709-v1] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), February 2001.
- [G.709-v2] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), March 2003.
- [G.709-v3] ITU-T, "Rec G.709, version 3", approved by ITU-T on December 2009.
- [G.872-am2] ITU-T, "Amendment 2 of G.872 Architecture of optical

transport networks for consent", consented by ITU-T on June 2010.

[OTN-FWK] F.Zhang, D.Li, H.Li, S.Belotti, D.Ceccarelli, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", work in progress draft-ietf-ccamp-gmpls-g709-framework-05, September 2011.

Authors' Addresses

Sergio Belotti (editor)
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

Pietro Vittorio Grandi
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: pietro_vittorio.grandi@alcatel-lucent.com

Daniele Ceccarelli (editor)
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28972912

Email: zhangfatai@huawei.com

Dan Li
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28973237

Email: danli@huawei.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

F. Zhang, Ed.
D. Li
Huawei
O. Gonzalez de Dios, Ed.
Telefonica I+D
C. Margaria
Nokia Siemens Networks
M. Hartley
Cisco
October 22, 2012

RSVP-TE Extensions for Collecting SRLG Information
draft-ietf-ccamp-rsvp-te-srlg-collect-01

Abstract

This document provides extensions for the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) to support automatic collection of Shared Risk Link Group (SRLG) Information for the TE link formed by a LSP.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. RSVP-TE Requirements	3
2.1. SRLG Collection Indication	3
2.2. SRLG Collection	3
2.3. SRLG Update	3
3. RSVP-TE Extensions (Encoding)	3
3.1. SRLG Collection Flag	4
3.2. SRLG sub-object	4
4. Signaling Procedures	5
4.1. SRLG Collection	5
4.2. SRLG Update	6
5. Manageability Considerations	7
5.1. Policy Configuration	7
5.2. Coherent SRLG IDs	7
6. Security Considerations	7
7. IANA Considerations	7
7.1. RSVP Attribute Bit Flags	8
7.2. ROUTE_RECORD Object	8
8. Contributing Authors	8
9. Acknowledgements	8
10. Normative References	9
Authors' Addresses	9

1. Introduction

It is important to understand which TE links in the network might be at risk from the same failures. In this sense, a set of links may constitute a 'shared risk link group' (SRLG) if they share a resource whose failure may affect all links in the set [RFC4202].

On the other hand, as described in [RFC4206] and [RFC6107], H-LSP (Hierarchical LSP) or S-LSP (stitched LSP) can be used for carrying one or more other LSPs. Both of the H-LSP and S-LSP can be formed as a TE link. In such cases, it is important to know the SRLG information of the LSPs that will be used to carry further LSPs.

This document provides an automatic mechanism to collect the SRLG for the TE link formed by a LSP. Note that how to use the collected SRLG information is out of scope of this document

2. RSVP-TE Requirements

2.1. SRLG Collection Indication

The head nodes of the LSP must be capable of indicating whether the SRLG information of the LSP should be collected during the signaling procedure of setting up an LSP. SRLG information should not be collected without an explicit request for it being made by the head node.

2.2. SRLG Collection

If requested, the SRLG information should be collected during the setup of an LSP. The endpoints of the LSP may use the collected SRLG information and use it for routing, sharing and TE link configuration purposes.

2.3. SRLG Update

When the SRLG information of an existing LSP for which SRLG information was collected during signaling changes, the relevant nodes of the LSP must be capable of updating the SRLG information of the LSP. This means that the signaling procedure must be capable of updating the new SRLG information.

3. RSVP-TE Extensions (Encoding)

3.1. SRLG Collection Flag

In order to indicate nodes that SRLG collection is desired, this document defines a new flag in the Attribute Flags TLV, which is carried in an LSP_REQUIRED_ATTRIBUTES Object: [Editor's note: LSP_ATTRIBUTES Object is also under consideration]

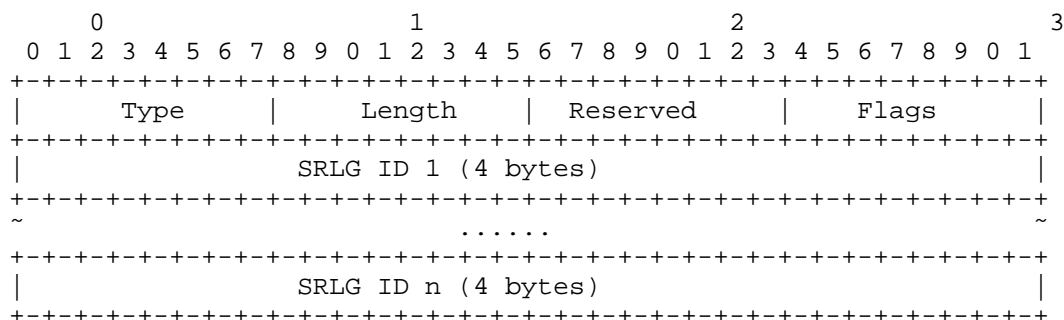
- o Bit Number (to be assigned by IANA, recommended bit zero): SRLG Collection flag

The SRLG Collection flag is meaningful on a Path message. If the SRLG Collection flag is set to 1, it means that the SRLG information should be reported to the head and tail node along the setup of the LSP.

The rules of the processing of the Attribute Flags TLV are not changed.

3.2. SRLG sub-object

This document defines a new RRO sub-object (ROUTE_RECORD sub-object) to record the SRLG information of the LSP. Its format is modeled on the RRO sub-objects defined in RFC 3209 [RFC3209].



Type

The type of the sub-object, to be assigned by IANA, which is recommended 34.

Length

The Length contains the total length of the sub-object in bytes, including the Type and Length fields. The Length depends on the number of SRLG IDs.

Flags

The Flags are used to indicate properties of the SRLG-list contained in the sub-object.

0x01 = SRLG-list edited

If set, this flag indicates that the SRLG-list contained in the RRO sub-object has been edited in some way by a node during signaling in accordance with that node's policy.

0x02 = Partial SRLG-list

If set, this flag indicates that the SRLG-list contained in this RRO sub-object is known to be incomplete.

SRLG Id

The 32-bit identifier of the SRLG.

Reserved

This field is reserved. It SHOULD be set to zero on transmission and MUST be ignored on receipt.

The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and ROUTE_RECORD Object are not changed.[Editor's note: The rules of processing LSP_ATTRIBUTES Object (which is under consideration) are also not changed]

4. Signaling Procedures

4.1. SRLG Collection

Typically, the head node gets the route information of an LSP by adding a RRO which contains the sender's IP addresses in the Path message. If a head node also desires SRLG recording, it sets the SRLG Collection Flag in the Attribute Flags TLV which can be carried in an LSP_REQUIRED_ATTRIBUTES Object.

When a node receives a Path message which carries an LSP_REQUIRED_ATTRIBUTES Object and the SRLG Collection Flag is set, if local policy determines that the SRLG information should not be provided to the endpoints, it must return a PathErr message to reject the Path message. Otherwise, it must add an SRLG sub-object to the RRO to carry the local SRLG information. Then it forwards the Path message to the next node in the downstream direction.

[Editor's note: It is under consideration that with the Path message

carries an LSP_REQUIRED_ATTRIBUTES Object and the SRLG Collection Flag is set, if local policy determines that the SRLG information should not be provided to the endpoints, the Path message should not be rejected and the SRLG sub-object must not be added]

Following the steps described above, the intermediate nodes of the LSP can collect the SRLG information in the RRO during the forwarding of the Path message hop by hop. When the Path message arrives at the tail node, the tail node can get the SRLG information from the RRO.

Before the Resv message is sent to the upstream node, the tail node adds an SRLG sub-object to the RRO. The collected SRLG information can be carried in the SRLG sub-object. Therefore, during the forwarding of the Resv message in the upstream direction, the SRLG information is not needed to be collected hop by hop.

Based on the above procedure, the endpoints can get the SRLG information automatically. Then the endpoints can for instance advertise it as a TE link to the routing instance based on the procedure described in [RFC6107] and configure the SRLG information of the FA automatically.

It is noted that a node (e.g. the edge node of a domain) may edit the RRO to remove the route information (e.g. node, interface identifier information) before forwarding it due to some reasons (e.g. confidentiality or reduce the size of RRO). A node MAY edit SRLG information within the RRO of a Path or Resv message if dictated by its local policy. If a node makes such an alteration to an existing RRO object, it SHOULD set the "SRLG-list edited" flag in the edited RRO sub-object to indicate to other nodes that this has been done.

[Editor's note: Two behaviors are under consideration: using LSP_REQUIRED_ATTRIBUTES the collection is mandatory, while using LSP_ATTRIBUTES the collection is desired, but not mandatory]

4.2. SRLG Update

When the SRLG information of a link is changed, the LSPs using that link should be aware of the changes. The procedures defined in Section 4.4.3 of RFC 3209 [RFC3209] MUST be used to refresh the SRLG information if the SRLG change is to be communicated to other nodes according to the local node's policy. If local policy is that the SRLG change should be suppressed or would result in no change to the previously signaled SRLG-list, the node need not send an update

5. Manageability Considerations

5.1. Policy Configuration

In a border node of inter-domain or inter-layer network, the following SRLG processing policy should be capable of being configured:

- o Whether the SRLG IDs of the domain or specific layer network can be exposed to the nodes outside the domain or layer network, or whether they should be summarized or removed entirely.
- o If SRLGs are summarized or removed, whether the "SRLG-list edited" flag is set in affected SRLG RRO-sub-objects.
- o If the SRLG IDs must not be exposed to the nodes outside of the domain or specific layer network by policy, the border node must reject the Path message desiring SRLG recording and send a PathErr message with the defined error code 'Policy Control Failure'/'Inter-domain policy failure'. [Editor's note: This last statement may be removed in next versions and do not impose such rejection.]

5.2. Coherent SRLG IDs

In a multi-layer multi-domain scenario, SRLG ids may be configured by different management entities in each layer/domain. In such scenarios, maintaining a coherent set of SRLG IDs is a key requirement in order to be able to use the SRLG information properly. Thus, SRLG IDs must be unique. Note that current procedure is targeted towards a scenario where the different layers and domains belong to the same operator, or to several coordinated administrative groups.

Further scenarios, where coherence in the SRLG IDs cannot be guaranteed are out of the scope of the present document and are left for further study.

6. Security Considerations

TBD.

7. IANA Considerations

7.1. RSVP Attribute Bit Flags

The IANA has created a registry and manages the space of attributes bit flags of Attribute Flags TLV as described in section 11.3 of [RFC5420]. It is requested that the IANA makes assignments from the Attribute Bit Flags.

This document introduces a new Attribute Bit Flag:

- o Bit number: TBD (10)
- o Defining RFC: this I-D
- o Name of bit: SRLG Collection Flag
- o The meaning of the Attribute Flags TLV on a Path is defined in this I-D

7.2. ROUTE_RECORD Object

IANA has made the following assignments in the "Class Names, Class Numbers, and Class Types" section of the "RSVP PARAMETERS" registry located at <http://www.iana.org/assignments/rsvp-parameters>. We request that IANA make assignments from the ROUTE_RECORD RFC 3209 [RFC3209] portions of this registry.

This document introduces a new RRO sub-object:

Type	Name	Reference
-----	-----	-----
TBD (34)	SRLG sub-object	This I-D

8. Contributing Authors

Zafar Ali
Cisco Systems
zali@cisco.com

9. Acknowledgements

The authors would like to thank Igor Bryskin, Ramon Casellas and Lou Berger for their useful comments to the document.

10. 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.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC5150] Ayyangar, A., Kompella, K., Vasseur, JP., and A. Farrel, "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.
- [RFC6107] Shiimoto, K. and A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC 6107, February 2011.

Authors' Addresses

Fatai Zhang (editor)
Huawei
F3-5-B RD Center
Bantian, Longgang District, Shenzhen 518129
P.R.China

Phone:
Email: zhangfatai@huawei.com

Dan Li
Huawei
F3-5-B RD Center
Bantian, Longgang District, Shenzhen 518129
P.R.China

Phone:
Email: danli@huawei.com

Oscar Gonzalez de Dios (editor)
Telefonica I+D
Don Ramon de la Cruz
Madrid, 28006
Spain

Phone: +34 913328832
Email: ogondio@tid.es

Cyril Margaria
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com

Matt Hartley
Cisco

Phone:
Email: mhartley@cisco.com

Network Working Group
Internet Draft
Intended status: Standards Track

Expires: September 2012

G. Bernstein
Grotto Networking
Sugang Xu
NICT
Y.Lee
Huawei
G. Martinelli
Cisco
Hiroaki Harai
NICT

March 7, 2012

Signaling Extensions for Wavelength Switched Optical Networks
draft-ietf-ccamp-wson-signaling-03.txt

Abstract

This memo provides extensions to Generalized Multi-Protocol Label Switching (GMPLS) signaling for control of wavelength switched optical networks (WSON). Such extensions are necessary in WSONs under a number of conditions including: (a) when optional processing, such as regeneration, must be configured to occur at specific nodes along a path, (b) where equipment must be configured to accept an optical signal with specific attributes, or (c) where equipment must be configured to output an optical signal with specific attributes. In addition this memo provides mechanisms to support distributed wavelength assignment with bidirectional LSPs, and choice in distributed wavelength assignment algorithms. These extensions build on previous work for the control of lambda and G.709 based networks.

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September 7, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

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 [RFC2119].

Table of Contents

1. Introduction.....	3
2. Terminology.....	3
3. Requirements for WSON Signaling.....	4
3.1. WSON Signal Characterization.....	4
3.2. Per LSP Network Element Processing Configuration.....	5
3.3. Bi-Directional WSON LSPs.....	5
3.4. Distributed Wavelength Assignment Support.....	6
3.5. Out of Scope.....	6
4. WSON Signal Traffic Parameters, Attributes and Processing.....	6
4.1. Traffic Parameters for Optical Tributary Signals.....	6
4.2. Signal Attributes and Processing.....	7
4.2.1. WSON Processing Object Encoding.....	8
5. Bidirectional Lightpath Setup.....	8
6. RWA Related.....	9
6.1. Wavelength Assignment Method Selection.....	9
7. Security Considerations.....	10

8. IANA Considerations.....	11
9. Acknowledgments.....	11
10. References.....	12
10.1. Normative References.....	12
10.2. Informative References.....	12
Author's Addresses.....	14
Intellectual Property Statement.....	15
Disclaimer of Validity.....	16

1. Introduction

This memo provides extensions to Generalized Multi-Protocol Label Switching (GMPLS) signaling for control of wavelength switched optical networks (WSON). Fundamental extensions are given to permit simultaneous bi-directional wavelength assignment while more advanced extensions are given to support the networks described in [RFC6163] which feature connections requiring configuration of input, output, and general signal processing capabilities at a node along a LSP

These extensions build on previous work for the control of lambda and G.709 based networks.

2. Terminology

CWDM: Coarse Wavelength Division Multiplexing.

DWDM: Dense Wavelength Division Multiplexing.

FOADM: Fixed Optical Add/Drop Multiplexer.

ROADM: Reconfigurable Optical Add/Drop Multiplexer. A reduced port count wavelength selective switching element featuring ingress and egress line side ports as well as add/drop side ports.

RWA: Routing and Wavelength Assignment.

Wavelength Conversion/Converters: The process of converting an information bearing optical signal centered at a given wavelength to one with "equivalent" content centered at a different wavelength. Wavelength conversion can be implemented via an optical-electronic-optical (OEO) process or via a strictly optical process.

WDM: Wavelength Division Multiplexing.

Wavelength Switched Optical Networks (WSON): WDM based optical networks in which switching is performed selectively based on the center wavelength of an optical signal.

AWG: Arrayed Waveguide Grating.

OXC: Optical Cross Connect.

Optical Transmitter: A device that has both a laser tuned on certain wavelength and electronic components, which converts electronic signals into optical signals.

Optical Responder: A device that has both optical and electronic components. It detects optical signals and converts optical signals into electronic signals.

Optical Transponder: A device that has both an optical transmitter and an optical responder.

Optical End Node: The end of a wavelength (optical lambdas) lightpath in the data plane. It may be equipped with some optical/electronic devices such as wavelength multiplexers/demultiplexer (e.g. AWG), optical transponder, etc., which are employed to transmit/terminate the optical signals for data transmission.

3. Requirements for WSON Signaling

The following requirements for GMPLS based WSON signaling are in addition to the functionality already provided by existing GMPLS signaling mechanisms.

3.1. WSON Signal Characterization

WSON signaling MUST convey sufficient information characterizing the signal to allow systems along the path to determine compatibility and perform any required local configuration. Examples of such systems include intermediate nodes (ROADMs, OXCs, Wavelength converters, Regenerators, OEO Switches, etc...), links (WDM systems) and end systems (detectors, demodulators, etc...). The details of any local configuration processes are out of the scope of this document.

From [RFC6163] we have the following list of WSON signal characteristic information:

List 1. WSON Signal Characteristics

1. Optical tributary signal class (modulation format).
2. FEC: whether forward error correction is used in the digital stream and what type of error correcting code is used
3. Center frequency (wavelength)
4. Bit rate
5. G-PID: General Protocol Identifier for the information format

The first three items on this list can change as a WSON signal traverses a network with regenerators, OEO switches, or wavelength converters. An ability to control wavelength conversion already exists in GMPLS signaling along with the ability to share client signal type information (G-PID). In addition, bit rate is a standard GMPLS signaling traffic parameter. It is referred to as Bandwidth Encoding in [RFC3471]. This leaves two new parameters: modulation format and FEC type, needed to fully characterize the optical signal.

3.2. Per LSP Network Element Processing Configuration

In addition to configuring a network element (NE) along an LSP to input or output a signal with specific attributes, we may need to signal the NE to perform specific processing, such as 3R regeneration, on the signal at a particular NE. In [RFC6163] we discussed three types of processing not currently covered by GMPLS:

- (A) Regeneration (possibly different types)
- (B) Fault and Performance Monitoring
- (C) Attribute Conversion

The extensions here MUST provide for the configuration of these types of processing at nodes along an LSP.

3.3. Bi-Directional WSON LSPs

WSON signaling MAY support LSP setup consistent with the wavelength continuity constraint for bi-directional connections. The following cases MAY be separately supported:

- (a) Where the same wavelength is used for both upstream and downstream directions
- (b) Where different wavelengths can be used for both upstream and downstream directions.

(Editor's Note: an evaluation of current GMPLS bidirectional solutions should be evaluated if they would fit to the current WSON needs.)

3.4. Distributed Wavelength Assignment Support

WSON signaling MAY support the selection of a specific distributed wavelength assignment method.

This method is beneficial in cases of equipment failure, etc., where fast provisioning used in quick recovery is critical to protect carriers/users against system loss. This requires efficient signaling which supports distributed wavelength assignment, in particular when the centralized wavelength assignment capability is not available.

As discussed in the [RFC6163] different computational approaches for wavelength assignment are available. One method is the use of distributed wavelength assignment. This feature would allow the specification of a particular approach when more than one is implemented in the systems along the path.

3.5. Out of Scope

This draft does not address signaling information related to optical impairments.

4. WSON Signal Traffic Parameters, Attributes and Processing

As discussed in [RFC6163] single channel optical signals used in WSONs are called "optical tributary signals" and come in a number of classes characterized by modulation format and bit rate. Although WSONs are fairly transparent to the signals they carry, to ensure compatibility amongst various networks devices and end systems it can be important to include key lightpath characteristics as traffic parameters in signaling [RFC6163].

4.1. Traffic Parameters for Optical Tributary Signals

In [RFC3471] we see that the G-PID (client signal type) and bit rate (byte rate) of the signals are defined as parameters and in

[RFC3473] they are conveyed Generalized Label Request object and the RSVP SENDER_TSPEC/FLOWSPEC objects respectively.

4.2. Signal Attributes and Processing

Section 3.2. gave the requirements for signaling to indicate to a particular NE along an LSP what type of processing to perform on an optical signal or how to configure that NE to accept or transmit an optical signal with particular attributes.

One way of accomplishing this is via a new EXPLICIT_ROUTE subobject. Reference [RFC3209] defines the EXPLICIT_ROUTE object (ERO) and a number of subobjects, while reference [RFC5420] defines general mechanisms for dealing with additional LSP attributes. Although reference [RFC5420] defines a RECORD_ROUTE object (RRO) attributes subobject, it does not define an ERO subobject for LSP attributes.

Regardless of the exact coding for the ERO subobject conveying the input, output, or processing instructions. This new "processing" subobject would follow a subobject containing the IP address, or the interface identifier [RFC3477], associated with the link on which it is to be used along with any label subobjects [RFC3473].

The WSON Signal Processing object is defined as an LSP_ATTRIBUTES and extends the PATH message. It is defined as the following:

```
<WSN Processing> ::= <hop information> <Transmitter Capabilities>  
<Receiver Capabilities> [<RegenerationCapabilities>]
```

```
<Receiver Capabilities> ::= <ModulationTypeList> <FECTypeList>  
<BitRateRange>
```

```
<Transmitter Capabilities> ::=
```

```
(ModulationTypeList> <FECTypeList> <BitRateRange>
```

Where:

<hop information>: Ipv4,Ipv6 address. Note: this not required if WSON Processing object become part of the ERO

<Transmitter Capabilities> is defined in [WSN-Encode].

<ReceiverCapabilities> is defined in [WSN-Encode].

<ModulationTypeList> is defined in [WSON-Encode]

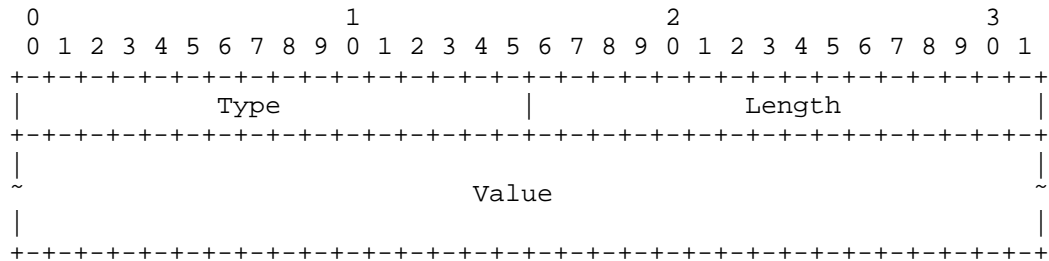
<FECTypeList> is defined in [WSON-Encode]

<BitRateRange> is defined in [WSON-Encode]

<RegenerationCapabilities> is defined in [WSON-Encode]

<RegenerationCapabilities> are only applied in the intermediate nodes of the LSP. The head and tail nodes will ignore regeneration capability processing.

4.2.1. WSON Processing Object Encoding



Type: to be defined by IANA

Value: sub-TLVs according to section 4.1.

5. Bidirectional Lightpath Setup

With the wavelength continuity constraint in CI-incapable [RFC3471] WSONs, where the nodes in the networks cannot support wavelength conversion, the same wavelength on each link along a unidirectional lightpath should be reserved. In addition to the wavelength

continuity constraint, requirement 3.2 gives us another constraint on wavelength usage in data plane, in particular, it requires the same wavelength to be used in both directions. [RFC6163] in section 6.1 reports on the implication to GMPLS signaling related to both bi-directionality and Distributed Wavelengths Assignment.

Current GMPLS solution defines a bidirectional LSP (as defined by [RFC3471]). The label distribution is based on Label_Set and Upstream_Label objects. In case of specific constraints such as the same wavelengths in both directions, it may require several signaling attempts using information from the Acceptable_Label_Set received from path error messages.

Some implementations may prefer using two unidirectional LSPs. This solution has been always available as per [RFC3209] however recent work introduces the association concept [RFC4872] and [ASSOC-Info]. Recent transport evolutions [ASSOC-ext] provide a way to associate two unidirectional LSPs as a bidirectional LSP. In line with this, a small extension can make this approach work for the WSON case.

6. RWA Related

6.1. Wavelength Assignment Method Selection

Routing + Distributed wavelength assignment (R+DWA) is one of the options defined by the [RFC6163]. The output from the routing function will be a path but the wavelength will be selected on a hop-by-hop basis.

Under this hypothesis the node initiating the signaling process needs to declare its own wavelength availability (through a label_set object). Each intermediate node may delete some labels due to connectivity constraints or its own assignment policy. At the end, the destination node has to make the final decision on the wavelength assignment among the ones received through the signaling process.

As discussed in [HZang00] a number of different wavelength assignment algorithms maybe employed. In addition as discussed in

[RFC6163] the wavelength assignment can be either for a unidirectional lightpath or for a bidirectional lightpath constrained to use the same lambda in both directions.

A simple TLV could be used to indication wavelength assignment directionality and wavelength assignment method. This would be placed in an LSP_REQUIRED_ATTRIBUTES object per [RFC5420]. The use of a TLV in the LSP required attributes object was pointed out in [Xu].

[TO DO: The directionality stuff needs to be reconciled with the earlier material]

Unique Wavelength: 0 same wavelength in both directions, 1 may use different wavelengths [TBD: shall we use only 1 bit]

Wavelength Assignment Method: 0 unspecified (any), 1 First-Fit, 2 Random, 3 Least-Loaded (multi-fiber). Others TBD.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Unique WL										WA Method										Reserved																			

7. Security Considerations

This document has no requirement for a change to the security models within GMPLS and associated protocols. That is the OSPF-TE, RSVP-TE, and PCEP security models could be operated unchanged.

However satisfying the requirements for RWA using the existing protocols may significantly affect the loading of those protocols. This makes the operation of the network more vulnerable to denial of service attacks. Therefore additional care maybe required to ensure that the protocols are secure in the WSON environment.

Furthermore the additional information distributed in order to address the RWA problem represents a disclosure of network capabilities that an operator may wish to keep private. Consideration should be given to securing this information.

8. IANA Considerations

TBD. Once finalized in our approach we will need identifiers for such things and modulation types, modulation parameters, wavelength assignment methods, etc...

9. Acknowledgments

Anyone who provide comments and helpful inputs

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [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.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.
- [RFC5420] Farrel, A., Ed., Papadimitriou, D., Vasseur, J.-P., and A. Ayyangar, " Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2006.

10.2. Informative References

- [WSON-CompOSPF] Y. Lee, G. Bernstein, "OSPF Enhancement for Signal and Network Element Compatibility for Wavelength Switched Optical Networks", work in progress: draft-lee-ccamp-wson-signal-compatibility-OSPF.
- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", work in progress: draft-bernstein-ccamp-wavelength-switched-03.txt, February 2008.

- [HZang00] H. Zang, J. Jue and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks", Optical Networks Magazine, January 2000.
- [Xu] S. Xu, H. Harai, and D. King, "Extensions to GMPLS RSVP-TE for Bidirectional Lightpath the Same Wavelength", work in progress: draft-xu-rsvpte-bidir-wave-01, November 2007.
- [Winzer06] Peter J. Winzer and Rene-Jean Essiambre, "Advanced Optical Modulation Formats", Proceedings of the IEEE, vol. 94, no. 5, pp. 952-985, May 2006.
- [G.959.1] ITU-T Recommendation G.959.1, Optical Transport Network Physical Layer Interfaces, March 2006.
- [G.694.1] ITU-T Recommendation G.694.1, Spectral grids for WDM applications: DWDM frequency grid, June 2002.
- [G.694.2] ITU-T Recommendation G.694.2, Spectral grids for WDM applications: CWDM wavelength grid, December 2003.
- [G.Sup43] ITU-T Series G Supplement 43, Transport of IEEE 10G base-R in optical transport networks (OTN), November 2006.
- [RFC4427] Mannie, E., Ed., and D. Papadimitriou, Ed., "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4427, March 2006.
- [RFC4872] Lang, J., Rekhter, Y., and Papadimitriou, D., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872,
- [ASSOC-Info] Berger, L., Faucheur, F., and A. Narayanan, "Usage of The RSVP Association Object", draft-ietf-ccamp-assoc-info-00 (work in progress), October 2010.
- [ASSOC-Ext] Zhang, F., Jing, R., "RSVP-TE Extension to Establish Associated Bidirectional LSP", draft-zhang-mppls-tp-rsvp-te-ext-associated-lsp-03 (work in progress), February 2011.

Author's Addresses

Greg M. Bernstein (editor)
Grotto Networking
Fremont California, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Nicola Andriolli
Scuola Superiore Sant'Anna, Pisa, Italy
Email: nick@sssup.it

Alessio Giorgetti
Scuola Superiore Sant'Anna, Pisa, Italy
Email: a.giorgetti@sssup.it

Lin Guo
Key Laboratory of Optical Communication and Lightwave Technologies
Ministry of Education
P.O. Box 128, Beijing University of Posts and Telecommunications,
P.R.China
Email: guolintom@gmail.com

Hiroaki Harai
National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi, Koganei,
Tokyo, 184-8795 Japan

Phone: +81 42-327-5418
Email: harai@nict.go.jp

Yuefeng Ji
Key Laboratory of Optical Communication and Lightwave Technologies
Ministry of Education
P.O. Box 128, Beijing University of Posts and Telecommunications,
P.R.China
Email: jyf@bupt.edu.cn

Daniel King
Old Dog Consulting

Email: daniel@olddog.co.uk

Young Lee (editor)
Huawei Technologies

1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

Sugang Xu
National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi, Koganei,
Tokyo, 184-8795 Japan

Phone: +81 42-327-6927
Email: xsg@nict.go.jp

Giovanni Martinelli
Cisco
Via Philips 12
20052 Monza, IT

Phone: +39 039-209-2044
Email: giomarti@cisco.com

Intellectual Property Statement

The IETF Trust 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 any IETF 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.

Copies of Intellectual Property 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 any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: April 25, 2013

D. Ceccarelli, Ed.
Ericsson
I. Bryskin
ADVA Optical Networking
V. Beeram
J. Drake
Juniper Networks
C. Margaria
Nokia Siemens Networks
October 22, 2012

Multi-domain network integration framework in the context of overlay
model
draft-many-ccamp-gmpls-overlay-model-01

Abstract

This document provides a framework for the integration of GMPLS based multi domain networks in the context of the overlay model. It defines terminology, requirements, interfaces and use cases characterizing the overlay model.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

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

Table of Contents

1. Introduction	3
1.1. Terminology	3
2. Definitions	3
3. Overlay interfaces	6
3.1. GMPLS UNI	6
3.2. GMPLS E-NNI	7
3.2.1. Pre-Provisioned Server Paths	8
3.2.2. Virtual Links	10
3.2.3. Virtual Nodes	11
4. Signaling: Info passed from Edge Network to Core network . . .	13
5. routing: Info passed from Core Network to Edge network . . .	14
6. Use cases	14
7. Compatibility	14
8. Security Considerations	14
9. IANA Considerations	15
10. Contributors	15
11. Acknowledgements	15
12. References	15
12.1. Normative References	15
12.2. Informative References	15
Authors' Addresses	15

1. Introduction

Generalized Multiprotocol Label Switching (GMPLS) defines both routing and signaling protocols for the creation of Label Switched Paths (LSPs) in various transport technologies and scenarios.

In multi domain scenarios, the GMPLS enables two different architectures: the peer model and the overlay model[RFC4208]. In the peer model edge nodes support both a routing and a signaling protocol and their interaction with the core nodes is basically the same that occurs among core nodes.

On the other side, in the overlay case, the interaction between edge nodes and core nodes is limited to signaling messages exchange and a very limited, if any, routing information exchange between core nodes and edge nodes regarding other edge nodes reachability. The edge nodes do not participate in the routing instance running in the core network domain and do not have any information about its topology.

This memo is focused on the overlay model frameworks and in particular addresses: definitions, methods (i.e. UNI interface, E-NNI interface) and use cases.

1.1. Terminology

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 [RFC2119].

2. Definitions

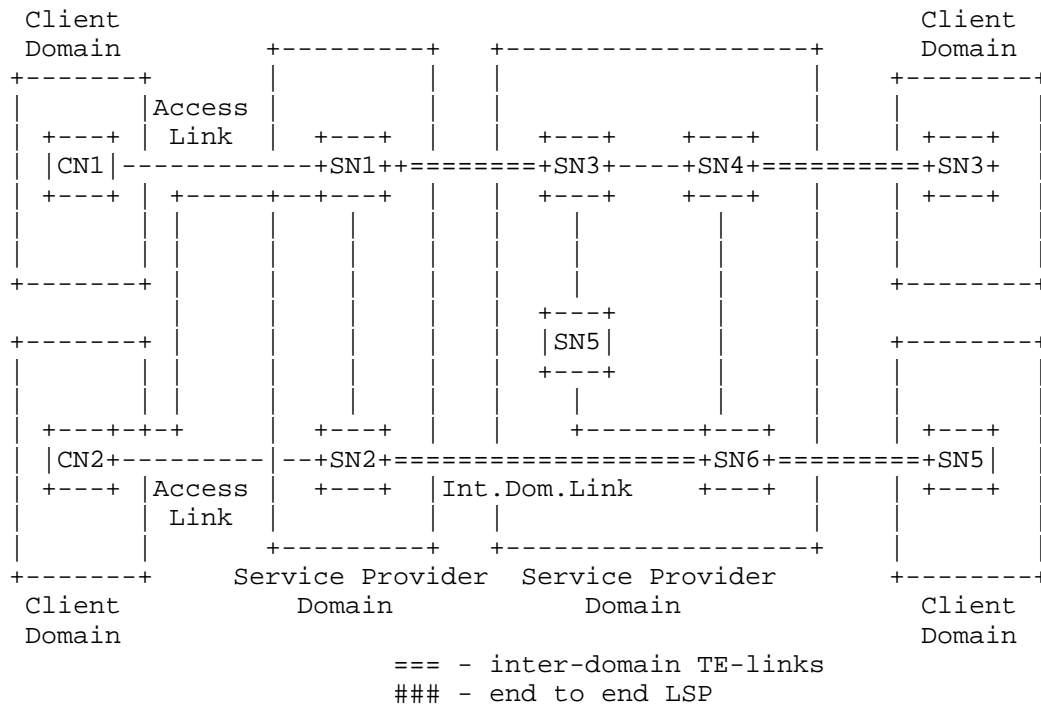


Figure 1: Overlay reference model

+ Client Domain: it is also called overlay network [RFC4208]. It is composed by the network elements directly connected to the server domain and has no knowledge of the server network topology but can retrieve routing information on how to reach other areas of the client domain via the server domain. The interaction with the server domain occurs via the overlay interfaces.

+ Client node: Node of the client domain directly connected to an access link. Client nodes can be ingress/egress nodes for overlay services and ingress/egress/transit nodes for overlay end-to-end LSPs.

+ Service Provider Domain: Network acting as server layer for the edge network. Multiple server domains can be interconnected.

+ Service provider node: Node of the server network without interfaces on access TE-Links.

+ Border Service Provider node: Node of the server network connected to access TE-links. It can receive signaling messages

from client nodes and provide them with routing information regarding the other client networks connected to the server network.

+ Access TE-link: TE-links between client nodes and server nodes. It only can be an interface supporting a transport technology (e.g. MPLS-TP, OTN, WDM).

+ Inter Domain TE-link: TE-links between server border nodes belonging to different server domains. A server network can be composed by multiple server domains. If an overlay interface is supposed to forward server domain topology information to the client network it must include topology information of all the server domain. As a consequence topology information of each server domain must be forwarded on the inter domain TE-links.

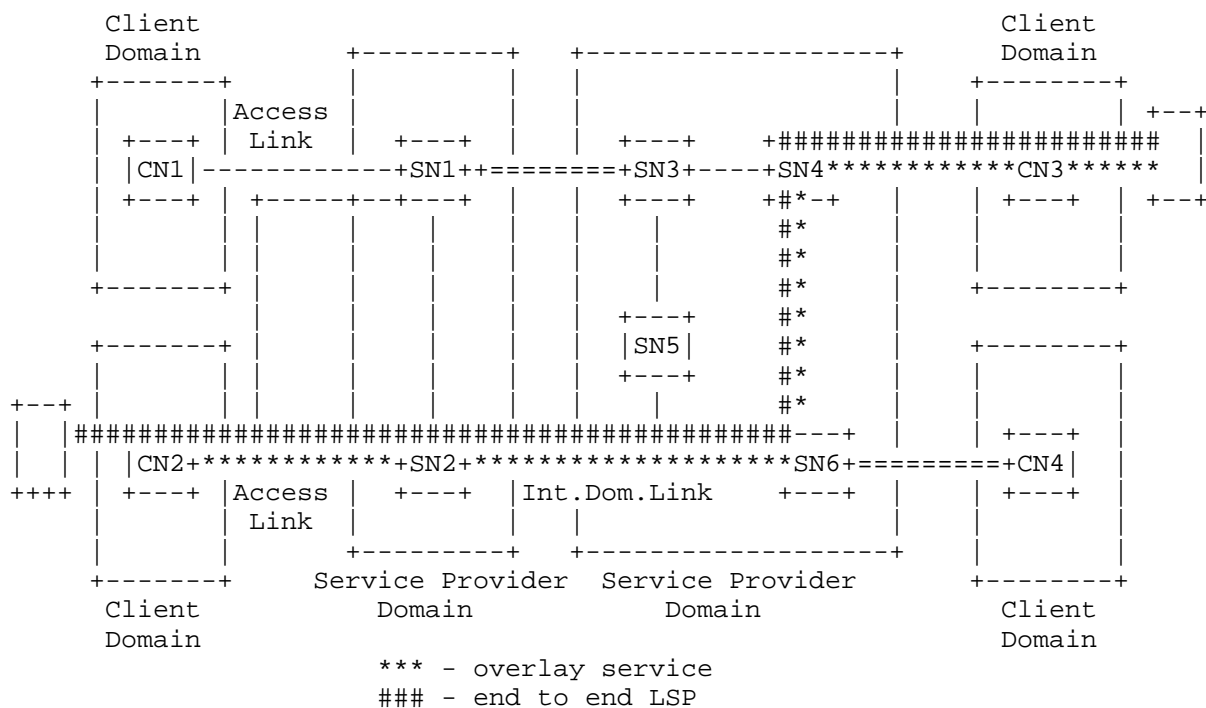


Figure 2: Overlay services and interfaces

+ Overlay interface: Overlay interfaces are used for setting up overlay services and allow the exchange of signaling and routing information. Two types of overlay interfaces are defined in the following, namely the GMPLS UNI and GMPLS E-NNI. Both of them allows inter-connecting devices in the client domain which are directly connected to the server domain and in either cases the service provides domain real topology is not known to the client domain but only a virtualized version.

++ GMPLS UNI: defined in [RFC4208] and augmented by [INC-ROUTE], [TE-REC], [XRO-SUB] and [OBJ-FUNC].

++ GMPLS E-NNI: defined in [E-NNI]. The E-NNI interface allows signaling and routing information exchange via the access TE-link.

+ Overlay service: It can only be established from a client node to a client node and have the service provider nodes as transit LSRs. Overlay services are injected by the client nodes into the domains where its interfaces (other than the overlay ones) belong to so to allow the establishment of end to end LSP crossing the client and service provider domains.

3. Overlay interfaces

3.1. GMPLS UNI

[statements]:

- the GMPLS UNI consider the service provider domains as opaque, no inter-domain routing IGP is used and paths (TE link) are not exported. Some (limited) properties of TE links may be requested by the client and may be provided by the server layer (based on policy), e.g., SRLG, metrics, etc.
- As there is no service provider domain IGP visibility the path computation constraints are to be considered by the border service provider path computation element. Those parameters can be provided to the border service provider PCE using PCEP or using signaling.
- The lack of IGP visibility do also require the routing information described in section 5 to be provided by signaling protocol (RRO)
- once a path is computed (and set up), it can be exported to the client layer as a TE-link with given TE parameters and SRLG

recording

[requirements]:

- U1. Must support signaling from client domain to server domain
- U2. Must support a path computation request in the server domain with given constraints
- U3. No server domain topology is exported to the client domain a priori (i.e. before the setup of a path in the server layer)
- U4. May support routing information towards the client server a posteriori (i.e. once a path is computed and set up, it may be exported to the client layer as a TE-link with given TE parameters and SRLG recording)
- U5. The path computation in the server domain can be either distributed (i.e. performed by the service provider border node) or centralized by a PCE.
- U6. ...

3.2. GMPLS E-NNI

[statements]:

[x] In comparison to GMPLS UNI the GMPLS E-NNI offer virtual topology visibility of the service provider network to the client. This offer the advantage of reusing of IGP mechanisms and information but may require more coordination from the border service provider nodes to offer such topology.

[x] paths in the service provider domain are precomputed. They can be pre-sigaled (i.e. fully committed FA-LSPs) or just pre-computed and sharable (virtual TE-links)

- FA-LSPs and virtual TE-links are exported to the client domain via the E-NNI interface with given properties (i.e. TE parameters). Virtual TE-links are activated via an external trigger (e.g. NMS)

[] TE-links may then injected in the domain the client node belongs to and used for end to end LSPs

[requirements]:

- E1. Must support signaling from client domain to server domain
- E2. Must support routing information towards the client server a priori (i.e. a set of pre-computed or pre-provisioned paths in the server domain must be exported to the client domain as provisioned or virtual TE-links with given TE parameters and SRLG recording)
- E3. No server domain topology is exported to the client domain
- E4. Server layer domain may be exported to the client domain as a virtual node
- E5. The path computation in the server domain can be either distributed (i.e. performed by the service provider border node) or centralized by a PCE.
- E6. ...

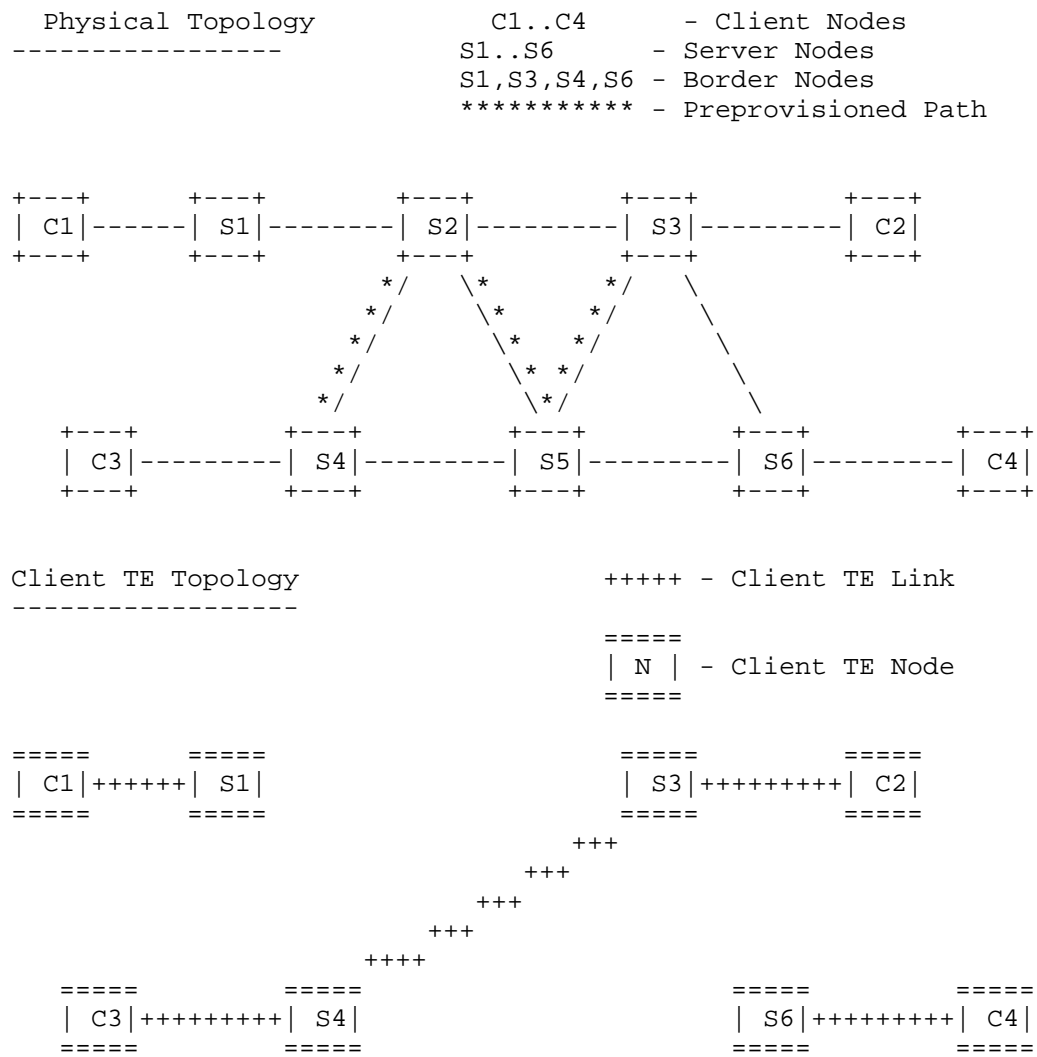
[description]

In order to be able to compute an end-to-end path between a pair of client nodes, the client TE topology must be sufficiently augmented to indicate via some fashion the paths through the server topology which can provide connectivity between nodes in the client topology. In other words, the feasible paths across each server network domain should be made available in terms of TE links/nodes to all the clients. GMPLS-ENNI provides the tools required to do this augmentation.

There are different ways of augmenting the client topology. This document discusses 3 approaches in the following subsections.

3.2.1. Pre-Provisioned Server Paths

In this approach, the paths in the server domain are provisioned up front and advertised as TE links in the client network domain. This means that relevant server network resources are committed before the service is setup in the client network.



3.2.2. Virtual Links

In this approach, potential paths in the server domain are advertised as Virtual TE links in the client network domain. The Virtual TE link is advertised just like a real TE link and does not require the allocation of resources in the server domain.

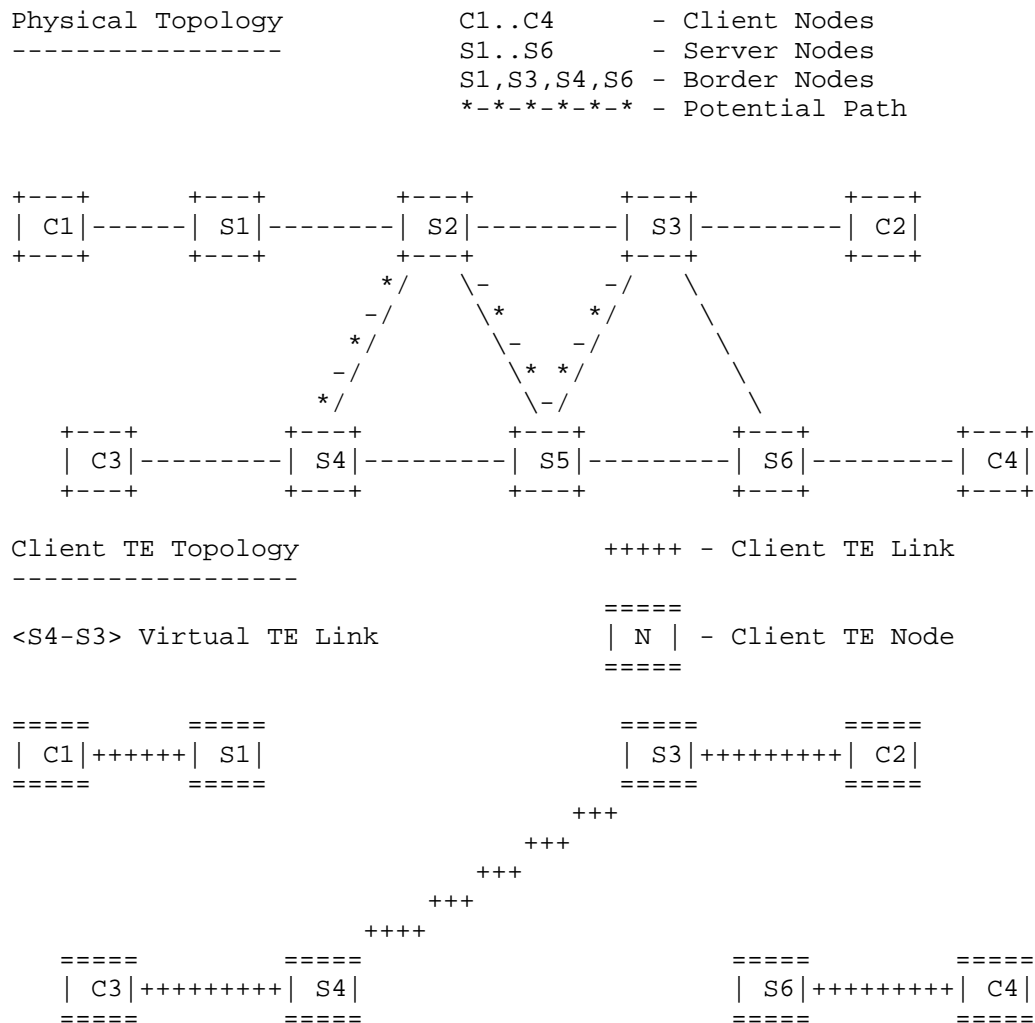


Figure 4: Virtual TE Links

In Figure 4 , the Virtual TE-Link [S4-S3] is advertised into the client domain based on the presense of the potential path [S4-S2-S5-S3] in the server domain. This makes the computation of an end-to-end Client Domain Path from C3 to C2 possible. The resources in the server domain get provisioned only when the corresponding client service gets signaled.

3.2.3. Virtual Nodes

In this approach, the server domain topology is presented to the clients as a Virtual TE Node.

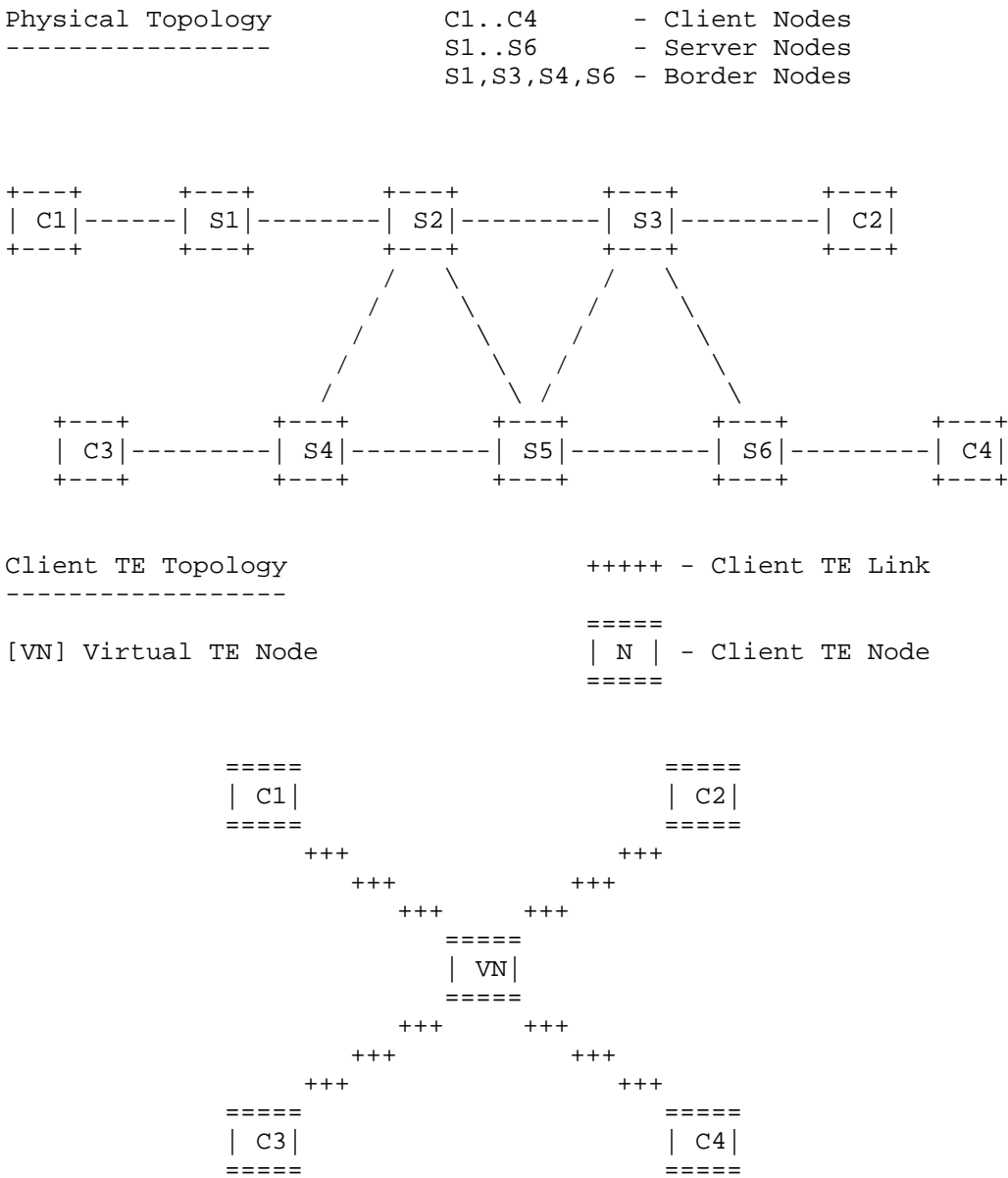


Figure 5: Virtual node

In Figure 5 , the Server Domain Topology is represented as a single

Virtual Node in the Client network.

4. Signaling: Info passed from Edge Network to Core network

The client node may need to constraint the path used in the service provider domains. Those constraints may be :

- Objective Functions
- TE Metric for a loose segment
- Node/Link/SRLG inclusion
- Node/Link/SRLG exclusion
- Path used by another LSP(from: draft-ali-xro-lsp-subobject)

(from: draft-ali-objective-functions) - the ingress node may need to request remote node to perform path computation or expansion. In such cases, ingress node needs to convey the required objective function to the remote node, to enable it to perform the desired path computation. Similarly, there are cases the ingress needs to indicate a TE metric bound for a loose segment that is expanded by a remote node.

(from: draft-ali-xro-lsp-subobject) - Moreover there are cases in which a remote node is requested to compute a path excluding LSPs currently existing or expected to exist within the network.

Hence what needs to be available to the remote node (ingress core node) is:

- + Objective function(draft-ali-objective-functions): A set of one or more specific optimization criteria that must be followed in expanding route of a TE-LSP in MPLS and GMPLS networks: Minimum TE Metric Cost, Minimum IGP Metric Cost, Minimum Latency, Minimum Latency Variation
- + Metric(draft-ali-objective-functions): the bound on the path metric that must not be exceeded for the loose segment to be considered as acceptable by the ingress node
- + Include Route(draft-ali-include-route): There are scenarios that require two or more LSPs or segments of LSPs to follow same route in the network: Explicit Inclusion Route

- + Exclude Route(draft-ali-xro-lsp0subobject):
- + Exclude Route(RFC4874)

5. routing: Info passed from Core Network to Edge network

(draft-beeram) - It is important to note that topology information is layer-specific; e.g. path computation and qualification operations occur within a given layer, and hence information about topology and resource availability are required for the specific layer to which the connection belongs. The topology and resource availability information required by a path computation element in the client layer is quite distinct from that required by a path computation element in the server layer network. Hence, the actual server layer traffic engineering links are of no importance for the client layer network. In fact, it can be desirable to block their advertisements into the client TE domain by the border nodes.

(draft-ali-te-metric-recording): Moreover there are many scenarios in which Traffic Engineering (TE) metrics such as cost, latency and latency variation associated with a Forwarding Adjacency (FA) or Routing Adjacency (RA) Label Switched Path (LSP) are not available to the ingress (edge node) and egress nodes.

6. Use cases

- + L1VPN (from draft-fedyk)
- + IP/MPLS Offloading with ENNI automation (from draft-enni)
- + Service Optimization and Restoration in Multi-layer networks (from draft enni)
- + Use of PCE and VNTM in Multi-layer Network Operation (from draft enni)

7. Compatibility

8. Security Considerations

9. IANA Considerations

10. Contributors

George Swallow, Cisco System

Zafar Ali, Cisco System

11. Acknowledgements

12. References

12.1. Normative References

- [E-NNI] Beeram V., Bryskin I., Doonan W., Drake J., Grammel G., Paul M., Kunze R., Armbruster F., de Dios O., "Generalized Multiprotocol Label Switching (GMPLS) External Network Network Interface (E-NNI): Virtual Link Enhancements for the Overlay Model", July 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.

12.2. Informative References

Authors' Addresses

Daniele Ceccarelli (editor)
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Igor Bryskin
ADVA Optical Networking

Email: ibryskin@advaoptical.com

Vishnu Pavan Beeram
Juniper Networks

Email: vbeeram@juniper.net

John Drake
Juniper Networks

Email: jdrake@juniper.net

Cyril Margaria
Nokia Siemens Networks

Email: cyril.margaria@nsn.com

CCAMP
Internet-Draft
Intended status: Standards Track
Expires: January 16, 2013

C. Margaria, Ed.
Nokia Siemens Networks
D. Schroetter
NUTS:ix GmbH
G. Martinelli
Cisco
S. Balls
B. Wright
Metaswitch
July 15, 2012

LSP Attribute in ERO
draft-margaria-ccamp-lsp-attribute-ero-01

Abstract

LSP attributes can be specified or recorded for whole path, but they cannot be targeted to a specific hop. This document proposes alternative ways to extend the semantic for RSVP ERO object to target LSP attributes to a specific hop.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 16, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Contributing Authors	3
1.2. Requirements Language	3
2. Requirements	4
3. Solutions	5
3.1. ERO LSP Attribute Subobject	5
3.1.1. ERO LSP_ATTRIBUTE subobject	5
3.1.2. Procedures	6
3.1.3. Pros and Cons	6
3.2. Information carried in the LSP Attribute object	7
3.2.1. Solution overview	7
3.2.2. ERO Hop Info Index Subobject	7
3.2.3. RRO Hop Info Index Subobject	8
3.2.4. LSP Attribute Hop Info TLV	8
3.2.5. Procedures	9
3.2.6. Pros and Cons	10
4. IANA Considerations	11
5. Security Considerations	12
6. Acknowledgments	13
7. References	14
7.1. Normative References	14
7.2. Informative References	14
Authors' Addresses	15

1. Introduction

Generalized MPLS (GMPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) can be route-constrained by making use of the Explicit Route (ERO) object and related sub-objects as defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], [RFC5520] and [RFC5553]. This document proposes mechanisms to target LSP attributes at a specific hop. This document present several solutions for discussion, final document will contains only one document after WG consensus.

1.1. Contributing Authors

1.2. Requirements Language

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].

2. Requirements

The requirement is to provide a generic mechanism to carry information related to specific nodes when signaling an LSP. This document does not restrict what that information can be used for. LSP attribute defined [RFC5420] should be expressed in ERO and SERO objects.

3. Solutions

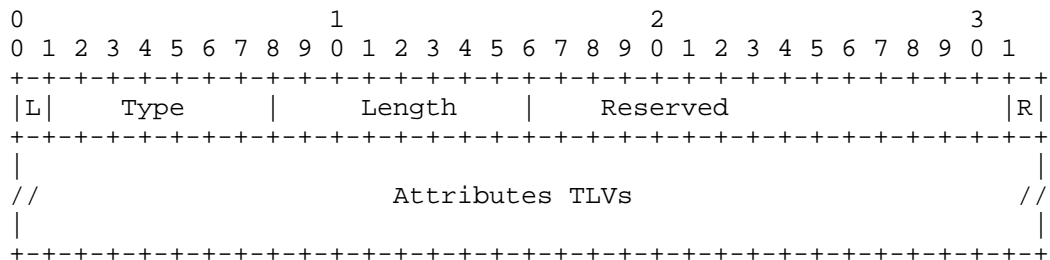
3.1. ERO LSP Attribute Subobject

The ERO LSP Attributes subobject may be carried in the ERO or SERO object if they are present. The subobject uses the standard format of an ERO subobject.

3.1.1. ERO LSP_ATTRIBUTE subobject

The length is variable and content MUST be the same as for the LSP_ATTRIBUTE object with Attributes TLVs.

The ERO LSP attribute subobject is defined as follows:



See [RFC3209] for a description of L parameters. The attributes TLV are encoded as defined in [RFC5420] section 3.

Type x TBD by IANA.

Length The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length MUST be always divisible by 4.

Reserved Reserved, must be set to 0 when the subobject is inserted in the ERO, MUST NOT be changed when a node process the ERO and must be ignored on the node addressed by the preceding ERO subobjects

R This bit reflects the LSP_REQUIRED_ATTRIBUTE and LSP_ATTRIBUTE semantic. When set indicates required LSP attributes to be processed by the node, when cleared the LSP attributes are not required as described in Section 3.1.2.

Attributes TLVs as defined in [RFC5420] section 3.

3.1.2. Procedures

As described in [RFC3209] and [RFC3473] the ERO is managed as a list where each hop information starts with a subobject identifying an abstract node or link. The LSP attribute subobject must be appended after the existing subobjects defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], [RFC5520] and [RFC5553]. Several LSP attribute subobject MAY be present.

If a node is processing an LSP attribute subobject and does not support handling of the subobject it will behave as described in [RFC3209] when an unrecognized ERO subobject is encountered. This node will return a PathErr with error code "Routing Error" and error value "Bad EXPLICIT_ROUTE object" with the EXPLICIT_ROUTE object included, truncated (on the left) to the offending unrecognized subobject.

When the R bit is set a node MUST examine the attribute TLV present in the subobject following the rules described in [RFC5420] section 5.2. When the R bit is not set a node MUST examine the attribute TLV present in the subobject following the rules described in [RFC5420] section 4.2. If more than one ERO LSP attribute subobject having the R bit set is present, the first one MUST be processed and the others SHOULD be ignored. If more than one ERO LSP attribute subject having the R bit cleared is present, the first one MUST be processed and the others SHOULD be ignored. [[anchor8: This need to be revised due to object length Pb --Ed.]]

3.1.3. Pros and Cons

This solution minimize the changes to the ERO object and so implementations can access all per-hop information when processing the ERO.

However, per hop ERO sub-objects are limited to 255 bytes in length which may limit its extensibility. Subsequent uses of this mechanism may wish to carry large amounts of contiguous information targeted at a single hop, which would need to split across multiple sub-objects.

It also requires the sub-object to be duplicated multiple times in the ERO if the same information needs to be targeted at multiple nodes.

3.2. Information carried in the LSP Attribute object

3.2.1. Solution overview

A new ERO/RRO sub-object (Hop info index) is defined to be an index/pointer to a new TLV (Hop info) carried in the LSP_ATTRIBUTES object. This TLV is formed of sub-TLVs carrying information targeted at a specific node.

3.2.2. ERO Hop Info Index Subobject

The ERO Hop Info Index subobject may be carried in the ERO or SERO object if they are present. The subobject uses the standard format of an ERO subobject.

The ERO Hop info index subobject is defined as follows:

										1										2										3																																																											
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1																																																										
L										Type																				Length										R										Res																				Index																			

See [RFC3209] for a description of L parameters.

Type Type x TBD by IANA.

Length The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length MUST be always divisible by 4.

Res Reserved, must be set to 0 when the subobject is inserted in the ERO, MUST NOT be changed when a node process the ERO and must be ignored on the node addressed by the preceding ERO subobjects.

R If set, the corresponding Hop Info TLV should be handled as required and according to the rules of the LSP_REQUIRED_ATTRIBUTE object. If clear, the corresponding Hop Info TLV should be handled as optional and according to the rules of the LSP_ATTRIBUTE object. This bit is present in the ERO to allow attributes mandatory on some node and optional on others.

Index A value used to refer to an LSP Attribute Hop Info TLV containing information targeted at the node processing this ERO.

Each hop on an LSP may have at most two ERO Hop Info Index subobjects associated with it. One for optional attributes, and one for

required attributes. Note that both these attributes are carried as separate Hop Info TLVs within the LSP_ATTRIBUTE object as they are not Required on the LSP as a whole.

3.2.3. RRO Hop Info Index Subobject

The RRO Hop Info Index subobject may be carried in the RRO object if it is present. The subobject uses the standard format of an RRO subobject.

The RRO Hop info index subobject is defined as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Type										Length										Res										Index									

Type x TBD by IANA.

Length The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length MUST be always divisible by 4.

Res Reserved, must be set to 0 when the subobject is inserted in the ERO, MUST NOT be changed when a node process the ERO and must be ignored on the node addressed by the preceding ERO subobjects.

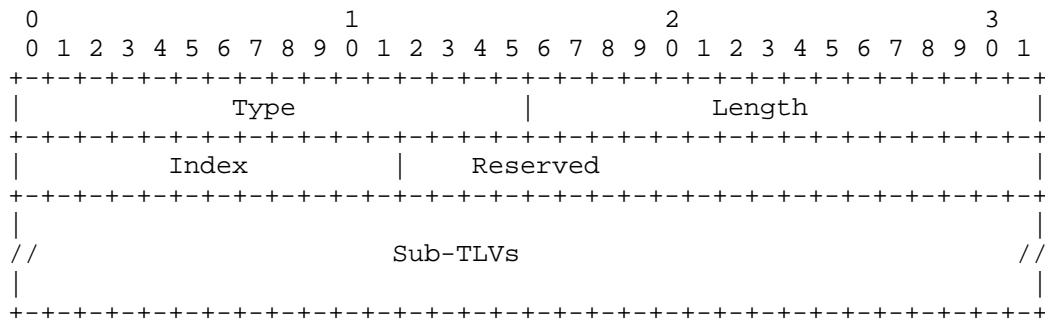
Index A value used to refer to an LSP Attribute Hop Info TLV containing information targeted at the node processing this ERO.

Each hop on an LSP may have at most one RRO Hop Info Index subobjects associated with it.

3.2.4. LSP Attribute Hop Info TLV

The LSP Attribute Hop Info TLV may be carried in the LSP Attribute object if present. It MUST be carried if an ERO Hop Info Index subobject is present in an ERO or SERO.

The LSP Attribute Hop Info TLV is defined as follows:



Type x TBD by IANA.

Length The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length MUST be always divisible by 4.

Index A value referred to by the Index field in the ERO Hop Info Index Subobject.

Reserved Reserved, must be set to 0 when the subobject is inserted in the LSP Attributes, MUST NOT be changed when a node process the LSP Attributes and must be ignored on the node processing the Hop Info TLV.

Sub-TLVs The information that is targeted at the specific hop or hops identified by the Index field.

This document defines 1 sub-TLV type as below.

3.2.4.1. Per Hop Attribute sub-TLV

The Per Hop Attribute sub-TLV is defined to be identical to the Attributes TLV in [RFC5420]. Thus using this sub-TLV means any Attribute TLV can now be targeted at specific nodes using the LSP Attribute Hop Info TLV.

Note that this means the number space for the Type value of Attributes for the whole LSP and those that can only ever be targeted at specific hops is shared.

3.2.5. Procedures

As described in [RFC3209] and [RFC3473] the ERO is managed as a list where each hop information starts with a subobject identifying an abstract node or link. The Hop Info Index subobject must be appended

after the existing subobjects defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], [RFC5520] and [RFC5553]. Only one Hop Info Index subobject may be added per node or link entry.

If a node is processing an ERO Hop Info Index subobject and does not support handling of the subobject it will behave as described in [RFC3209] when an unrecognized ERO subobject is encountered. This node will return a PathErr with error code "Routing Error" and error value "Bad EXPLICIT_ROUTE object" with the EXPLICIT_ROUTE object included, truncated (on the left) to the offending unrecognized subobject.

If the node does supports the Hop Info Index subobject it will look for a corresponding (Both having the same Index field value) LSP Attribute Hop Info TLV in the LSP Attribute object. If one is not present it will return a PathErr with error code "Routing Error" and error value "Bad EXPLICIT_ROUTE object".

A node processing the LSP Attribute Hop Info TLV should not alter it. It is valid for multiple ERO entries to refer to the same Hop Info TLV, thus targeting the same information at multiple nodes.

The RRO Hop Info Index subobject should be processed according to the rules of section 7.3.1 of [RFC5420]. A node inserting an RRO Hop Info Index subobject should not also insert an RRO Attributes subobject.

3.2.6. Pros and Cons

This solution is more complex in term of processing, but addresses some of the restrictions in the first solution. LSP Attribute TLVs allow a length of up to 65535 bytes and the indexing system allows multiple nodes to target the same information. The LSP Attribute Hop Info TLV may be extended by further sub-TLV types

Other objects may be candidate to contain the Indexed ERO attribute, for instance the ERO object with a new C-Type.

4. IANA Considerations

TBD once a final approach has been chosen.

5. Security Considerations

None.

6. Acknowledgments

The authors would like to thanks Lou Berger for his directions and Attila Takacs for inspiring this [I-D.kern-ccamp-rsvpte-hop-attributes].

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.
- [RFC4874] Lee, CY., Farrel, A., and S. De Cnodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4874, April 2007.
- [RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.
- [RFC5520] Bradford, R., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, April 2009.
- [RFC5553] Farrel, A., Bradford, R., and JP. Vasseur, "Resource Reservation Protocol (RSVP) Extensions for Path Key Support", RFC 5553, May 2009.

7.2. Informative References

- [I-D.kern-ccamp-rsvpte-hop-attributes]
Kern, A. and A. Takacs, "Encoding of Attributes of LSP intermediate hops using RSVP-TE",
draft-kern-ccamp-rsvpte-hop-attributes-00 (work in progress), October 2009.

Authors' Addresses

Cyril Margaria (editor)
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com

Dirk Schroetter
NUTS:iX GmbH
Langer Weg 88
Eschborn, 65185
DE

Phone: +49 6173 39198411
Email: dirk.schroetter@nutsix.de

Giovanni Martinelli
Cisco
via Philips 12
Monza 20900
IT

Phone: +39 039 209 2044
Email: giomarti@cisco.com

Steve Balls
Metaswitch
100 Church Street
Enfield EN2 6BQ
UJ

Phone: +44 208 366 1177
Email: steve.balls@metaswitch.com

Ben Wright
Metaswitch
100 Church Street
Enfield EN2 6BQ
UJ

Phone: +44 208 366 1177
Email: Ben.Wright@metaswitch.com

CCAMP
Internet-Draft
Intended status: Standards Track
Expires: January 31, 2013

G. Martinelli
M. Kattan
G. Galimberti
Cisco
A. Zanardi
CREATE-NET
July 30, 2012

Encoding for WSON Information Model with Impairments Validation.
draft-martinelli-ccamp-wson-iv-encode-00

Abstract

This document defines proper encoding for the Information Model to support Impairment-Aware (IA) Routing and Wavelength Assignment (RWA) function. This operation might be required in Wavelength Switched Optical Networks (WSON) that already support RWA, encoding defined here goes in addition to available WSON encoding and it is fully compatible with it.

As the information model, the encoding is independent from control plane architectures and protocol implementations. Its definitions must be reused in related protocol extensions.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 31, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal

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

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Encoding	3
2.1. Optical Parameter	3
2.2. Impairment Matrix	4
3. Acknowledgements	5
4. IANA Considerations	5
5. Security Considerations	5
6. References	6
6.1. Normative References	6
6.2. Informative References	6
Authors' Addresses	7

1. Introduction

In case of WSON where optical impairments plays a significant role, the framework document [RFC6566] defines related control plane architectural options for an Impairment Aware routing and wavelength assignment (IA-RWA). This document provides a suitable encoding for the related WSON Impairment Information Model defined [ID.martinelli-ccamp-wson-iv-info].

This document directly refers to ITU recommendations [ITU.G680] and [ITU.G697] as already detailed in the information model.

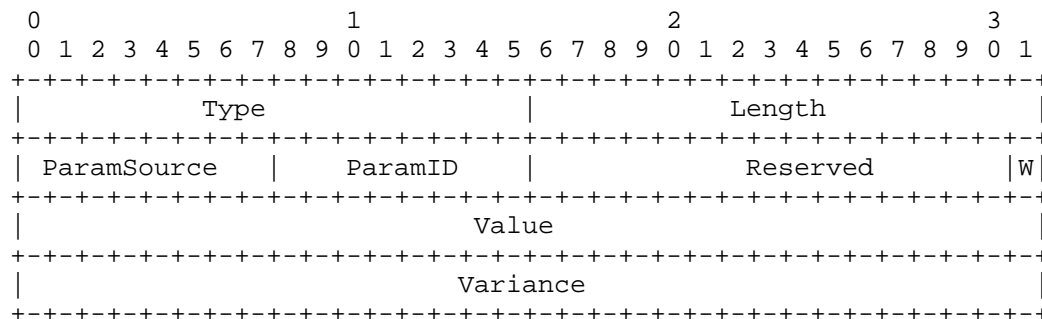
1.1. Requirements Language

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].

2. Encoding

2.1. Optical Parameter

The OPTICAL_PARAM is defined as a TLV object.



Where the following parameters are defined:

Type. IANA defined for Optical Parameter TLV.

Length. The length (bytes) of this TLV including the header.

ParamSource. Where this parameter was defined. Currently only [ITU.G697] has defined this with value 1.

ParamID. Parameter identifier according to the source. [ITU.G697] has defined the following identifiers:

1. Reserved
2. Reserved
3. Reserved
4. Reserved
5. OSNR
6. Q Factor
7. PMD
8. Residual Chromatic Dispersion

Value. Value for the parameter. As defined by [ITU.G697] is a 32 bit IEEE floating point number.

Variance. Variance for the parameter, a 32 bit IEEE floating point number.

In addition the following flags is defined:

W. The parameter has a wavelength dependency

[EDITOR NOTE: encoding to be refined. In case of wavelength dependency the label set can be added within parameter definition. To evaluate vs. with impairment matrix flags]

2.2. Impairment Matrix

As defined by the [ID.martinelli-ccamp-wson-iv-info] the Impairment Matrix follow the same structure as the Connectivity Matrix with some variations.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Connectivity										MatrixID										Reserved										W N									
Link Set A #1																																							
:										:																													
Link Set B #1																																							
:										:																													
Optical Parameter TLV(s)																																							
:										:																													

Where:

Connectivity: 2, Impairment Matrix.

MatrixID: matrix identifier, the scope of this integer number is shared with [I-D.ietf-ccamp-rwa-info].

N: Node scope flags. With this flag set there's no Link Set information but only a list of optical parameters TLVs that apply to the whole optical node.

W: Wavelength Dependency Matrix [EDITOR NOTE: to evaluate vs a flag at parameter level.]

3. Acknowledgements

TBD

4. IANA Considerations

This document reuse the ConnectivityMatrix object defined in [I-D.ietf-ccamp-general-constraint-encode].

This document defines the Optical Parameter Object.

5. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

6. References

6.1. Normative References

- [ITU.G680] International Telecommunications Union, "Physical transfer functions of optical network elements", ITU-T Recommendation G.680, July 2007.
- [ITU.G697] International Telecommunications Union, "Optical monitoring for dense wavelength division multiplexing systems", ITU-T Recommendation G.697, February 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

6.2. Informative References

- [I-D.ietf-ccamp-general-constraint-encode] Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", draft-ietf-ccamp-general-constraint-encode-08 (work in progress), July 2012.
- [I-D.ietf-ccamp-rwa-info] Lee, Y., Bernstein, G., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info-14 (work in progress), March 2012.
- [I-D.narten-iana-considerations-rfc2434bis] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", draft-narten-iana-considerations-rfc2434bis-09 (work in progress), March 2008.
- [ID-martinelli-ccamp-wson-iv-info] G. Martinelli, M. Kattan, G. Galimberti, A. Zanardi, "Information Model for Wavelength Switched Optical Networks (WSON) with Optical Impairments Validation", draft-martinelli-ccamp-wson-iv-info-00 (work in progress), 2012.
- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.
- [RFC6566] Lee, Y., Bernstein, G., Li, D., and G. Martinelli, "A Framework for the Control of Wavelength Switched Optical Networks (WSONs) with Impairments", RFC 6566, March 2012.

Authors' Addresses

Giovanni Martinelli
Cisco
via Philips 12
Monza, 20900
Italy

Phone: +39 039 2092044
Email: giomarti@cisco.com

Moustafa Kattan
Cisco
DUBAI, 500321
UNITED ARAB EMIRATES

Phone:
Email: mkattan@cisco.com

Gabriele M. Galimberti
Cisco
Via Philips,12
Monza 20900
Italy

Phone: +39 039 2091462
Email: ggalimbe@cisco.com

Andrea Zanardi
CREATE-NET
via alla Cascata 56 C, Povo
Trento 38100
Italy

Email: andrea.zanardi@create-net.org

CCAMP
Internet-Draft
Intended status: Informational
Expires: January 31, 2013

G. Martinelli
M. Kattan
G. Galimberti
Cisco
A. Zanardi
CREATE-NET
July 30, 2012

Information Model for Wavelength Switched Optical Networks (WSON) with
Optical Impairments Validation.
draft-martinelli-ccamp-wson-iv-info-00

Abstract

This document defines the Information Model to support Impairment-Aware (IA) Routing and Wavelength Assignment (RWA) function. This operation might be required in Wavelength Switched Optical Networks (WSON) that already support RWA and the Information model defined here goes in addition and it is fully compatible with the already defined information model for WSON.

This information model shall support all control plane architectural options defined for WSON with impairment validation.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 31, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal

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

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Properties of an Impairment Information Model	3
3. Background from WSOON Information Model	5
4. Optical Impairment Information Model	5
4.1. Node Information	6
4.2. Link Information	7
5. Encoding Considerations	7
6. Acknowledgements	8
7. IANA Considerations	8
8. Security Considerations	8
9. References	8
9.1. Normative References	8
9.2. Informative References	8
Appendix A. G.680 Essential information	9
Authors' Addresses	9

1. Introduction

In the context of Wavelength Switched Optical Network (WSON), [RFC6163] defines the basic framework for a GMPLS control plane. The associated info model [I-D.ietf-ccamp-rwa-info] defines all parameters required for the related RWA process. These references are the foundation but they do not consider the Optical Impairment case.

In case of WSON where optical impairments plays a significant role, the framework document [RFC6566] defines related control plane architectural options for an Impairment Aware routing and wavelength assignment (IA-RWA). Options include different combinations of Impairment Validation (IV) and RWA functions through control plane elements and operations (PCE, Routing, Signaling).

This document intent to provide the information model for the impairment aware case to allow the impairment validation function. It goes in addition with [I-D.ietf-ccamp-rwa-info] and the model itself is independ of any architectural option described by the framework and shall support all of them.

Models for the optical impairments are defined by ITU and the only available models are reported in [ITU.G680] to cover only the linear impairment case while non-linear case is left for further study. The information model defined here however provide a generic enough mechanism that could be easily extended to additional impairments models.

1.1. Requirements Language

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].

2. Properties of an Impairment Information Model

An information model may have several attributes or properties that need to be defined for each optical parameter made available to the control plane. The properties will help to determine how the control plane can deal with it depending on architectural options chosen within the overall impairment framework [RFC6566]. In some case properties value will help to identify the level of approximation supported by the IV process.

- o Time Dependency.
This will identify how the impairment may vary with time. There could be cases where there is no time dependency, while in other cases there may be need of impairment re- evaluation after a certain time. In this category, variations in impairments due to environmental factors such as those discussed in [G.sup47] are considered. In some cases a level of approximation will consider an impairment that has time dependency as constant. In this Information Model we do neglect this property.
- o Wavelength Dependency.
This property will identify if an impairment value can be considered as constant over all the wavelength spectrum of interest or if it has different values. Also in this case a detailed impairment evaluation might lead to consider the exact value while an approximation IV might take a constant value for all wavelengths. In this Information Model we consider both case: dependency / not dependency from a specific wavelengths. This property may appear directly in the Information model definitions or in the related encoding.
- o Linearity.
As impairments are representation of physical effects there are some that have a linear behavior while other are non-linear. Linear approximation is in scope of scenario C of [RFC6566]. The linearity implies the additivity optical quantities considered during an Impairment Validation process. As an additional approximation level, non-linear impairments as contribution into linear ones. This Information Model deals with the linear properties of optical impairments.
- o Multi-Channel.
There are cases where a channel's impairments take different values depending on the aside wavelengths already in place. In this case a dependency among different LSP is introduced and is typically a result of linear effects. This Information Model neglect this effects on neighbor LSPs.

The following table summarize the above considerations where in the first column reports the list of properties to be considered for each optical parameters, while second column state if this property is taken into account or not by this Information Model.

Property	Info Model Awareness
Time Dependency	no
Wavelength Dependency	yes
Linearity	yes
Multi-channel	no

Table 1: Optical Impairment Properties

3. Background from WSON Information Model

In this section we report terms already defined for the WSON-RWA (not impairment aware) as in [I-D.ietf-ccamp-rwa-info] and [I-D.ietf-ccamp-general-constraint-encode]. The purpose is to provide essential information that will be reused or extended for the impairment case.

In particular [I-D.ietf-ccamp-rwa-info] defines the connectivity matrix as the follow:

```
ConnectivityMatrix ::= <MatrixID> <ConnType> <Matrix>
```

However according to [I-D.ietf-ccamp-general-constraint-encode] this definitions can be further detailed as:

```
ConnectivityMatrix ::=
  <MatrixID> <ConnType> ((<LinkSet> <LinkSet>) ...)
```

This second definition highlights how the connectivity matrix is built by pairs of LinkSet objects identifying the internal node connectivity capability.

As a additional note, Connectivity Matrix belong to Node Information.

4. Optical Impairment Information Model

The idea behind this Information Model is to reuse the concept of the Connectivity Matrix and defines an Impairment Matrix that summarize optical impairments provided by the Node.

The goal of the information model is not to rephrase content from [ITU.G680] but only provide necessary building blocks that allow the IW-RWA process to apply the computational model defined by such recommendation. Then the [ITU.G680] model defined in section 9 provide information to calculate the following parameters:

- o OSNR. Section 9.1
- o Chromatic Dispersion (CD). Section 9.2
- o Polarization Mode Dispersion (PMD). Section 9.3
- o Polarization Dependent Loss (PDL). Section 9.3

It should be noted that [ITU.G697] already defines an encoding for all these parameters and in Section 5 we report some encoding consideration. The [ITU.G697] is mainly oriented for monitoring so the purpose is only reuse parameter definitions for those parameters required by Impairment Validation process.

The information model defined here make the assumption that the Optical Node is able to provide it's own contribution to such parameters. To this extend the information model intentionally ignore all internal detailed parameters that are used to by the formulas. As an additional note, as reported in in [ITU.G680] Section 10, each parameter can be reported as an OSNR contribution, in such way the Optical Node not necessarily embed optical computational capability but can provide an approximated contribution to optical impairments.

With the above considerations this Information Model is able provide an abstract view for an optical node to enable WSOON protocol extension with optical impairments validation.

4.1. Node Information

This model defines the Impairment Matrix as the following:

```
ImpairmentMatrix ::= <MatrixID> <ConnType>
  ((<LinkSet> <LinkSet> <ImparimentVector>) ...)
```

Where:

MatrixID. Is a unique identifier for the Matrix. This ID shall be unique in scope among all connectivity matrix defined in [I-D.ietf-ccamp-rwa-info] and all impairment matrix defined here.

ConnType. The type of matrix. Since values 0 and 1 are already defined. This document defines the value 2.

LinkSet. Same object definition and usage as [I-D.ietf-ccamp-general-constraint-encode].

ImpairmentVector is defined as list of optical parameters associated to the internal node connection.

ImpairmentVector ::= <OPTICAL_PARAM> ...

The set of OPTICAL_PARAM is identified by [ITU.G697] since they match with parameters required by the linear impairments evaluation provided by [ITU.G680]. This info model does not preclude any of such parameters and eventually new parameters can be added to the list.

4.2. Link Information

Currently not evaluated yet any information is required at Link Level however the same approach can be used as in case of Node Information section. The Link information defined in [I-D.ietf-ccamp-rwa-info] is extend in the following way:

```
<DynamicLinkInfo> ::= <LinkID> <AvailableLabels>
                        [<SharedBackupLabels>]
                        <ImpairmentVector>
```

With ImpairmentVector defined as previous section.

5. Encoding Considerations

Details about encoding will be defined in a separate document however worth remembering that, within [ITU.G697] Appending V, ITU already provides a guideline for encoding some optical parameters.

In particular [ITU.G697] indicates that each parameters shall be represented by a 32 bit floating point number.

As an additional consideration, actual values for parameters defined in the information models are provided by the Optical Node and it could provide by direct measurement or from some internal computation starting from indirect measurement. In any case the encoding shall

provide an the possibility to associate a variance with the parameter. This information will enable the function implementing IV-RWA process to make some additional considerations on wavelength feasibility.

6. Acknowledgements

TBD

7. IANA Considerations

This document does not have ant IANA requirement.

8. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

9. References

9.1. Normative References

[ITU.G680]

International Telecommunications Union, "Physical transfer functions of optical network elements", ITU-T Recommendation G.680, July 2007.

[ITU.G697]

International Telecommunications Union, "Optical monitoring for dense wavelength division multiplexing systems", ITU-T Recommendation G.697, February 2012.

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[I-D.ietf-ccamp-general-constraint-encode]

Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", draft-ietf-ccamp-general-constraint-encode-08 (work in progress), July 2012.

[I-D.ietf-ccamp-rwa-info]

Lee, Y., Bernstein, G., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info-14 (work in progress), March 2012.

[I-D.narten-iana-considerations-rfc2434bis]

Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", draft-narten-iana-considerations-rfc2434bis-09 (work in progress), March 2008.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

[RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

[RFC6566] Lee, Y., Bernstein, G., Li, D., and G. Martinelli, "A Framework for the Control of Wavelength Switched Optical Networks (WSONs) with Impairments", RFC 6566, March 2012.

Appendix A. G.680 Essential information

TBD if we need some info instead of reading [ITU.G680]

Authors' Addresses

Giovanni Martinelli
Cisco
via Philips 12
Monza, 20900
Italy

Phone: +39 039 2092044
Email: giomarti@cisco.com

Moustafa Kattan
Cisco
DUBAI, 500321
UNITED ARAB EMIRATES

Phone:
Email: mkattan@cisco.com

Gabriele M. Galimberti
Cisco
Via Philips,12
Monza 20900
Italy

Phone: +39 039 2091462
Email: ggalimbe@cisco.com

Andrea Zanardi
CREATE-NET
via alla Cascata 56 C, Povo
Trento 38100
Italy

Email: andrea.zanardi@create-net.org

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2013

O. Gonzalez de Dios, Ed.
Telefonica I+D
R. Casellas, Ed.
CTTC
F. Zhang
Huawei
X. Fu
ZTE
D. Ceccarelli
Ericsson
I. Hussain
Infinera
July 6, 2012

Framework for GMPLS based control of Flexi-grid DWDM networks
draft-ogrcetal-ccamp-flexi-grid-fwk-00

Abstract

This document defines a framework and the associated control plane requirements for the GMPLS based control of flexi-grid DWDM networks. To allow efficient allocation of optical spectral bandwidth for high bit-rate systems, the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) is extending the standard [G.694.1] to include the concept of flexible grid: a new DWDM grid has been developed within the ITU-T Study Group 15, by defining a set of nominal central frequencies, smaller channel spacings and the concept of "frequency slot". In such environment, a data plane connection is switched based on the allocated, variable-width optical spectrum frequency slot.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on January 7, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Requirements Language	4
2. Introduction	4
3. Acronyms	5
4. Terminology	6
5. DWDM flexi-grid enabled network element models	11
5.1. Switched Resources and Labels	11
5.2. Physical links	12
5.3. Transceivers	12
5.4. ROADMs	13
6. Layered Network Model	14
7. Topology view in Control Plane	15
8. Control Plane Requirements	15
8.1. Neighbor Discovery and Link Property Correlation	16
8.2. Path Computation / Routing and Spectrum Assignment (RSA)	16
8.2.1. Architectural Approaches to RSA	17
8.3. Routing / Topology dissemination	17
8.3.1. Available Frequency Ranges/slots of DWDM Links	18
8.3.2. Available Slot Width Ranges of DWDM Links	18
8.3.3. Tunable Optical Transmitters and Receivers	18
8.3.4. Hierarchical Spectrum Management	18
8.3.5. Information Model	19
8.4. Signaling requirements	20
8.4.1. Slot Width Requirement	20
8.4.2. Frequency Slot Representation	20
8.4.3. Relationship with MRN/MLN	20
9. Control Plane Procedures	20
10. Backwards (fixed-grid) compatibility, and WSON interworking	21
11. Misc & Summary of open Issues [To be removed at later versions]	22
12. Security Considerations	23
13. Contributing Authors	23
14. Acknowledgments	25
15. References	25
15.1. Normative References	25
15.2. Informative References	26
Authors' Addresses	26

1. Requirements Language

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 [RFC2119].

2. Introduction

The term "Flexible grid" (flexi-grid for short) as defined by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) study group 15 in the latest version of [G.694.1], refers to the updated set of nominal central frequencies (a frequency grid), channel spacings and optical spectrum management/allocation considerations that have been defined in order to allow an efficient and flexible allocation and configuration of optical spectral bandwidth for high bit-rate systems.

A key concept of flexi-grid is the "frequency slot"; a variable-sized optical frequency range that can be allocated to a data connection. As detailed later in the document, a frequency slot is characterized by its nominal central frequency, selected from the set of reference frequencies, and its slot width which, as per [G.694.1], is constrained to be a multiple of a given slot width granularity.

Compared to a traditional fixed grid network, which uses fixed size optical spectrum frequency ranges or "frequency slots" with typical channel separations of 100 or 50 GHz, a flexible grid network can select its data channels with with a more flexible choice of slot widths, allocating as much optical spectrum as required, and allowing higher bitrates (e.g., 100G or 400G or higher).

From a networking perspective, a flexible grid network is assumed to be a layered network [G.872][G.805], extending the OTN architecture and interfaces [G.709], in which the flexi-grid layer (also referred to as the media layer) is the server layer and the OCh Layer (also referred to as the signal layer) is the client layer. In the media layer, switching is based on a frequency slot, and the size of a media channel is given by the properties of the associated frequency slot. In this layered network, the media channel itself can be dimensioned to contain one or more Optical Channels.

As described in [RFC3945], GMPLS extends MPLS from supporting only Packet Switching Capable (PSC) interfaces and switching to also support four new classes of interfaces and switching that include Lambda Switch Capable (LSC).

A Wavelength Switched Optical Network (WSO), addressed in [RFC6163],

is a term commonly used to refer to the application/deployment of a Generalized Multi-Protocol Label Switching (GMPLS)-based control plane for the control (provisioning/recovery, etc) of a fixed grid WDM network. [editors' note: we need to think of the relationship of WSON and OCh switching. Are they equivalent? WSON includes regeneration, OCh does not? decoupling of lambda/OCh/OCC]

This document defines the framework for a GMPLS-based control of flexi-grid enabled DWDM networks (in the scope defined by ITU-T layered Optical Transport Networks [G.872], as well as a set of associated control plane requirements. An important design consideration relates to the decoupling of the management of the optical spectrum resource and the client signals to be transported. [Editor's note: a point was raised during the meeting that WSON has not made separation between OCh and Lambda (spectrum and signal are bundled). This needs to be confirmed.] A direct consequence of this "separation of concerns" is that, although not in scope of the present document, single-carrier / multi-carrier and related modulation formats, etc. could be supported. [Editor's note: the concept of frequency slot channel supporting multiple OCHs is defined in an ITU contribution. It is not a standard document yet.]

[Editors' note: this document will track changes and evolutions of [G.694.1] [G.872] documents until their final publication. This document is not expected to become RFC until then. Likewise, as agreed during IETF83, the consideration of the concepts of Super-channel (a collection of one or more frequency slots to be treated as unified entity for management and control plane) and consequently Contiguous Spectrum Super-channel (a super-channel with a single frequency slot) and Split-Spectrum super-channel (a super-channel with multiple frequency slots) is postponed until the ITU-T data plane includes such physical layer entities, e.g., an ITU-T contribution exists]

3. Acronyms

FS: Frequency Slot

FSCh: Frequency Slot Channel

NCF: Nominal Central Frequency

OCG: Optical Carrier Group

OCh: Optical Channel

OCC: Optical Channel Carrier

OTUk: Optical channel Transport Unit level k

ODUk: Optical channel Data Unit Level k

ODUj: Optical channel Data Unit Level j

SWG: Slot Width Granularity

4. Terminology

The following is a list of terms (see [G.694.1] and [G.872]) reproduced here for completeness. [Editors' note: regarding wavebands, we agreed NOT to use the term in flexigrid. The term has been used inconsistently in fixed-grid networks and overlaps with the definition of frequency slot. If need be, a question will be sent to ITU-T asking for clarification regarding wavebands.]

[Editors' note: *important* these terms are not yet final and they may change / be replaced or obsoleted at any time.]

- o Optical Channel Slot (definition in the scope of a fixed grid DWDM network, to be adapted to a flexi-grid). The optical spectrum frequency range (portion of optical spectrum) allocated / occupied by a single optical channel. Each optical channel signal has a defined carrier central frequency and required frequency slot width (the supported optical channel signal bandwidth plus source stability). Optical Channel slots within an optical multiplex section may be allocated (in-service) or may be unallocated (out-of-service). An in-service Optical Channel Slot may be carrying an Optical Channel Signal or not. Optical Channel Slots are switched in an Optical Channel Matrix.
- o Nominal Central Frequency Granularity: 6.25 GHz (note: sometimes referred to as 0.00625 THz).
- o Nominal Central Frequency: each of the allowed frequencies as per the definition of flexible DWDM grid in [G.694.1]. The set of nominal central frequencies can be built using the following expression $f = 193.1 \text{ THz} + n \times 0.00625 \text{ THz}$, where 193.1 THz is ITU-T "anchor frequency" for transmission over the C band, n is a positive or negative integer including 0.

```

-5 -4 -3 -2 -1 0 1 2 3 4 5      <- values of n
...+---+---+---+---+---+---+---+---+---+---+
      ^
      193.1 THz <- anchor frequency

```


or multiple Optical Channels may be transported.

- o Fiber Frequency Slot: the total allocable spectrum on a fiber ($n=0$ and $m= \text{infinity?}$). [Editors' note/CM: is this useful? is the spectrum bounded/symmetric w.r.t anchor frequency?]
- o Frequency Slot Channel: a topological construct that represents a piece of spectrum supported by a concatenation of media elements (fiber, amplifiers, filters...). This term is used to identify the end-to-end physical layer entity with its corresponding (one or more) frequency slots local at each link. [Editors' note:
 - * MB: a frequency slot is a local (i.e., to the link) concept, while a frequency slot channel has an end to end meaning.
 - * IH: the FSCh is the CTP layer that is defining the frequency slot connection matrix.
 - * CM: the CTP is the Frequency Slot and the Frequency Slot Channel the trail, the OCh being on top of the Channel.
 - * ITU-T mailing list defines Common Frequency Slot which may replace Frequency Slot Channel (?).]
- o Common Frequency Slot: the optical frequency range that is common to all of the devices in a particular path through the optical network. It is a logical construct derived from the frequency slots allocated to each device in the path (intersection). As an example, if there are two devices having slots with the same n but different m , then the common frequency slot has the smaller of the two m values. [Editors' note: this definition overlaps with Effective Frequency Slot] [Editors' note: clarify what happens when the resulting slot cannot be characterized with n and m , see Figure. Are we assuming that the same " n " applies?].

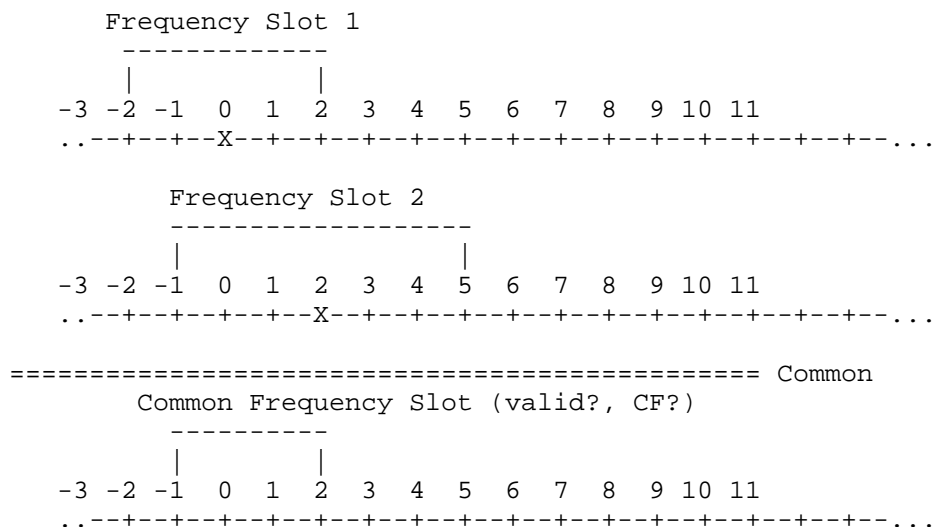


Figure 4. Common Frequency Slot

- o [Note: Following terminology is copied from ITU-T WP3 Q12 interim meeting [WD12R2]].
- o [Editors' note: if we accept that a frequency slot can support one or more optical channel signals do we need the following two definitions?).
- o Single-Channel Frequency Slot: a frequency slot associated with a single optical channel signal ((that carries a single OCh payload)).
- o Multi-Channel Frequency Slot: a frequency slot associated with multiple optical channel signals (i.e. multiple OChs). Note that if there are multiple optical signals within frequency slot, then each signal still has its own central frequency. That is, the term "central frequency" applies to an Optical signal and the term "nominal central frequency" applies to a frequency slot. In other words, the Frequency Slot central frequency is independent of the signals central frequencies.

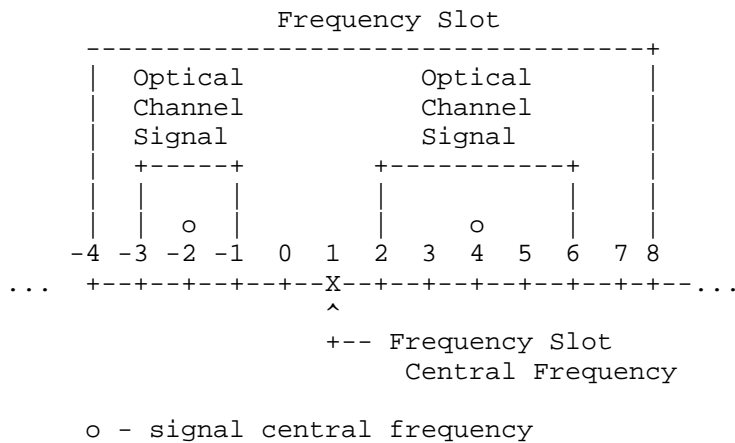


Figure 3. Frequency slot with 2 Optical channel signals

- o Network Channel (NCh): An end-to-end path through an optical network from a port on an OCh termination source to a port on an OCh termination sink (i.e. one OEO to another OEO). It is constructed from a concatenation of link channels and subnetwork channels.
- o Link Channel (LCh): A partial path through an optical network that provides a fixed relationship between the ports on a "subnetwork" or "access group" and the ports on another "subnetwork" or "access group". [Note: the terms subnetwork and access group are defined in G.805].
- o Subnetwork Channel (SNCh): A path through an optical subnetwork that provides a relationship across a subnetwork. It is formed by the association of "ports" on the boundary of the subnetwork.
- o Matrix Channel (MCh): A path through an optical matrix that provides a relationship across a matrix. It is formed by the association of "ports" on the boundary of the matrix.
- o Effective Frequency Slot: An attribute of a channel which identifies that part of the frequency slots allocated to the devices along the channel that is common to all

The following terms are defined in the scope of a GMPLS control plane. [Editors' note: the following ones were *not* agreed during IETF83 but are put here to be discussed.]

- o SSON: Spectrum-Switched Optical Network. An optical network in which a data plane connection is switched based on an optical spectrum frequency slot of a variable slot width, rather than based on a fixed grid and fixed slot width. Please note that a Wavelength Switched Optical Network (WSO) can be seen as a particular case of SSON in which all slot widths are equal and depend on the used channel spacing.
- o Flexi-LSP: a control plane construct that represents a data plane connection in which the switching involves a frequency slot. Different Flexi-LSPs may have different slot widths. The term flexi-LSP is used when needed to differentiate from regular WSO LSP in which switching is based on a nominal wavelength.
- o RSA: Routing and Spectrum Assignment. As opposed to the typical Routing and Wavelength Assignment (RWA) problem of traditional WDM networks, the flexibility in SSON leads to spectral contiguous constraint, which means that when assigning the spectral resources to single connections, the resources assigned to them must be contiguous over the entire connections in the spectrum domain.

5. DWDM flexi-grid enabled network element models

Similar to fixed grid networks, a flexible grid network is also constructed from subsystems that include Wavelength Division Multiplexing (WDM) links, tunable transmitters and receivers, Reconfigurable Optical Add/Drop Multiplexers (ROADMs), wavelength converters, and electro-optical network elements, all of them with flexible grid characteristics.

As stated in [G.694.1] the flexible DWDM grid defined in Clause 7 has a nominal central frequency granularity of 6.25 GHz and a slot width granularity of 12.5 GHz. However, devices or applications that make use of the flexible grid may not be capable of supporting every possible slot width or position. In other words, applications may be defined where only a subset of the possible slot widths and positions are required to be supported. For example, an application could be defined where the nominal central frequency granularity is 12.5 GHz (by only requiring values of n that are even) and that only requires slot widths as a multiple of 25 GHz (by only requiring values of m that are even).

5.1. Switched Resources and Labels

As per [G.872] [Editor's note/CM: we need to better distinguish between G.872 and contributions, it would help to see what is agreed and what is still open, the list below contains items as per MB/XF

slides]:

- o OCh Slots are switched in an Optical Channel Matrix.
- o The (link) physical layer entity, as defined by ITU-T is the Frequency Slot.
- o A frequency slot channel may be switched in a Frequency Slot Matrix [ITU-T contribution draft].
- o The frequency slot matrix connection cannot modify the center frequency or increase the bandwidth of the frequency slots present at its ports [Editors' note: this text comes from G.872 updated. This seems to constrain / limit to only a transparent segment? the "m" must be the same end to end while "n" can be change by the equivalent of a wavelength converter, but WC are not defined. Currently, we only consider the case that the frequency slot matrix connection cannot modify the center frequency or the bandwidth of the frequency slots present at its ports. The use cases of dynamically modifying the center frequency or the bandwidth of the frequency slots are for further study after the clear definition by ITU-T].
- o [Editors' note: we are not discarding O/E/O. If defined in a ITU-T network reference model with trail/terminations, considering optical channels i.e. with well-defined interfaces, reference points, and architectures. The implications of O/E/O will be also addressed once we have another context that includes them. In OTN from an OCh point of view end to end means from transponder to transponder, so if there is a 3R from ingress to egress there are 2 OCh which can have different 'n' and 'm'].

5.2. Physical links

5.3. Transceivers

Optical transmitters/receivers may have different restrictions on the following properties:

- o Available central frequencies: The set of central frequencies which can be used by an optical transmitter/receiver.
- o Slot width: The slot width needed by a transmitter/receiver. The slot width is dependent on bit rate and modulation format. For one specific transmitter, the bit rate and modulation format may be tunable, so slot width would be determined by the modulation format used at a given bit rate.

- o The minimum and maximum slot width.
- o The step granularity: the optical hardware may not be able to select parameters with the lowest granularity (e.g. 6.25 GHz for nominal central frequencies or 12.5 GHz for slot width granularity).

5.4. ROADMs

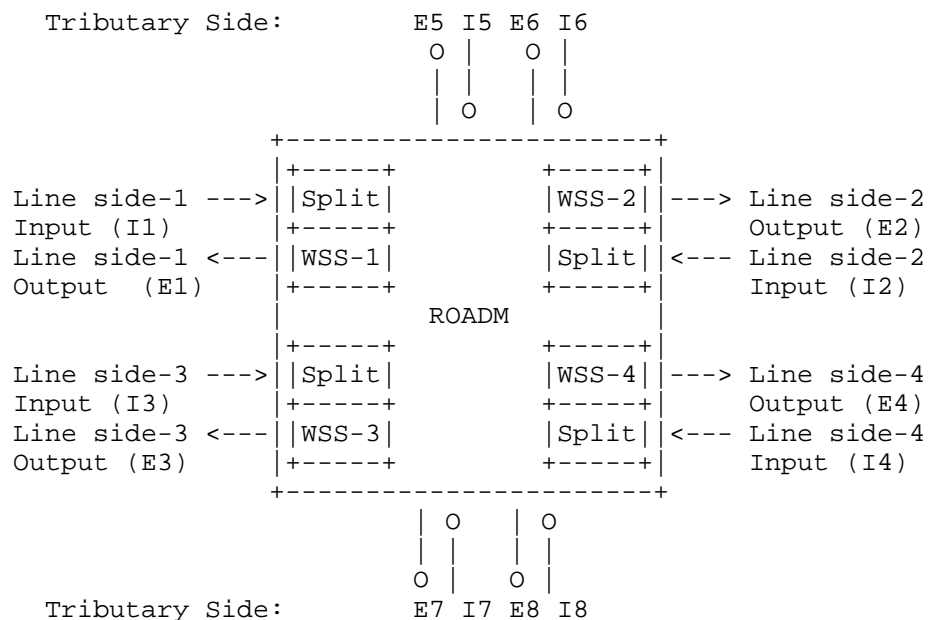


Figure 5. Simplified ROADM model with Line Sides and Tributaries

[Editor's note: different ROADM configuration such as C/CD/CDC will be added later.]

A Frequency slot matrix may have switching restrictions, for example, when it is realized using flexi-grid enabled ROADMs. A key feature of ROADMs is their highly asymmetric switching capability which is described in [RFC6163] in detail. The ports on ROADM include line side ports which are connected to DWDM links and tributary side input/output ports which can be connected to transmitters/receivers. The capability of ports on ROADM, which are characterized as follows:

- o Available frequency ranges: the set or union of frequency ranges that are not allocated (i.e. available). The relative grouping and distribution of available frequency ranges in a fiber is

usually referred to as ''fragmentation''.

- o Available slot width ranges: the set or union of slot width ranges supported by ROADM. It includes the following information.
 - * Slot width threshold: the minimum and maximum Slot Width supported by ROADM. For example, the slot width can be from 50GHz to 200GHz.
 - * Step granularity: the minimum step by which the optical filter bandwidth of ROADM can be increased or decreased. This parameter is typically equal to slot width granularity (i.e. 12.5GHz) or integer multiples of 12.5GHz.

6. Layered Network Model

[Editors' note: OTN hierarchy is not fully covered. It is important to understand, where the FSC sits in the OTN hierarchy. This is also important from control plane perspective as this layer becomes the connection end points of optical layer service]. OCh / flexi-grid layered model.

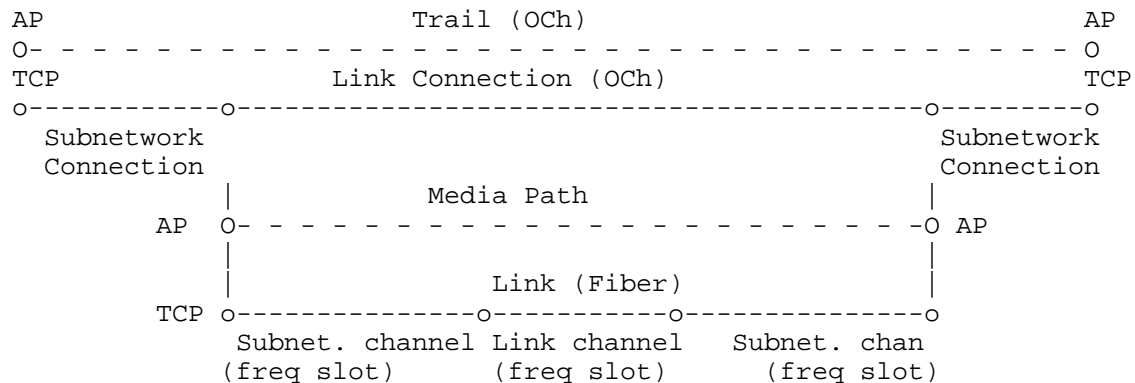


Figure 6. Layered Network Model G.805

[Editors' note: we are replicating the figure here for reference, until the ITU-T document is official.

The media path is a piece of spectrum that has been allocated to a path between two ports of a media device. [Editors'note/CM/IH: it seems the media path is equivalent to the FSC (freq.slot channel is between the AP?. Why use a new term media path?]

7. Topology view in Control Plane

[Note: the frequency slot matrix connection may interconnect one or more frequency slot channels which in turn may carry one or more OCh signals.]

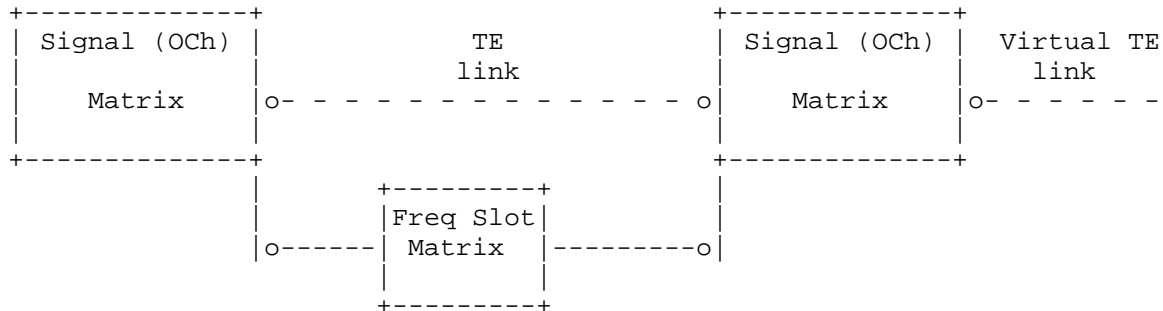


Figure 7. MRN/MLN topology view with TE link / FA

8. Control Plane Requirements

[Editor's note: The considered topology view is a layered network, in which the media layer corresponds to the server layer (flexigrid) and the signal layer corresponds to the client layer (OCh). This data plane modeling considers the flexigrid and the OCh as separate layers, especially considering both the single and multi-channel frequency slots. However, this has implications on the interop/interworking with WSON and OCh switching. We need to manage a MRN for OCh and stitching for WSON? In other words, a key part of the fwk is to define how can we have MRN/MLN hierarchical relationship with OCh/FS and yet stitching 1:1 between WSON and SSON? In this line: how does OCh switching and WSON relate, actually?]

[Editor's note: formal requirements such as noted in the comments will be added in a later version of the document].

Hierarchy spectrum management decouples media and signal, but from the point of view of the control plane, such separation of concerns implies the management of a MRN/MLN network. So Control Plane needs to differentiate signal LSP and media LSP. It should also need to support Hierarchy-LSP [RFC4206] The central frequency of each hop should be same along end-to-end media or signal LSP because of Spectrum Continuity Constraint. Otherwise some nodes need to convert the central frequency along media or signal LSP.

8.1. Neighbor Discovery and Link Property Correlation

[Editors' note: text from draft-li-ccamp-grid-property-lmp-01]

During the practical deployment procedure, fixed-grid optical nodes will be gradually replaced by flexible nodes. This will lead to an interworking problem between fixed-grid DWDM and flexible-grid DWDM nodes. Additionally, even two flexible-grid optical nodes may have different grid properties, leading to link property conflict.

Devices or applications that make use of the flexible-grid may not be able to support every possible slot width. In other words, applications may be defined where different grid granularity can be supported. Taking node F as an example, an application could be defined where the nominal central frequency granularity is 12.5 GHz requiring slot widths being multiple of 25 GHz. Therefore the link between two optical nodes with different grid granularity must be configured to align with the larger of both granularities. Besides, different nodes may have different slot width tuning ranges.

In summary, in a DWDM Link between two nodes, at least the following properties should be negotiated:

- Grid capability (channel spacing) - Between fixed-grid and flexible-grid nodes.

- Grid granularity - Between two flexible-grid nodes.

- Slot width tuning range - Between two flexible-grid nodes.

8.2. Path Computation / Routing and Spectrum Assignment (RSA)

Much like in WSON, in which if there is no (available) wavelength converters in an optical network, an LSP is subject to the ''wavelength continuity constraint'' (see section 4 of [RFC6163]), if the capability of shifting or converting an allocated frequency slot, the LSP is subject to the Optical ''Spectrum Continuity Constraint''.

Because of the limited availability of wavelength/spectrum converters (sparse translucent optical network) the wavelength/spectrum continuity constraint should always be considered. When available, information regarding spectrum conversion capabilities at the optical nodes may be used by RSA mechanisms.

The RSA process determines a route and frequency slot for a LSP. Hence, when a route is computed the spectrum assignment process (SA) should determine the central frequency and slot width based on the slot width and available central frequencies information of the

transmitter and receiver, and the available frequency ranges information and available slot width ranges of the links that the route traverses.

8.2.1. Architectural Approaches to RSA

Similar to RWA for fixed grids, different ways of performing RSA in conjunction with the control plane can be considered. The approaches included in this document are provided for reference purposes only; other possible options could also be deployed.

8.2.1.1. Combined RSA (R&SA)

In this case, a computation entity performs both routing and frequency slot assignment. The computation entity should have the detailed network information, e.g. connectivity topology constructed by nodes/links information, available frequency ranges on each link, node capabilities, etc.

The computation entity could reside either on a PCE or the ingress node.

8.2.1.2. Separated RSA (R+SA)

In this case, routing computation and frequency slot assignment are performed by different entities. The first entity computes the routes and provides them to the second entity; the second entity assigns the frequency slot.

The first entity should get the connectivity topology to compute the proper routes; the second entity should get the available frequency ranges of the links and nodes' capabilities information to assign the spectrum.

8.2.1.3. Routing and Distributed SA (R+DSA)

In this case, one entity computes the route but the frequency slot assignment is performed hop-by-hop in a distributed way along the route. The available central frequencies which meet the spectrum continuity constraint should be collected hop by hop along the route. This procedure can be implemented by the GMPLS signaling protocol.

8.3. Routing / Topology dissemination

In the case of combined RSA architecture, the computation entity needs to get the detailed network information, i.e. connectivity topology, node capabilities and available frequency ranges of the links. Route computation is performed based on the connectivity

topology and node capabilities; spectrum assignment is performed based on the available frequency ranges of the links. The computation entity may get the detailed network information by the GMPLS routing protocol. Compared with [RFC6163], except wavelength-specific availability information, the connectivity topology and node capabilities are the same as WSON, which can be advertised by GMPLS routing protocol (refer to section 6.2 of [RFC6163]). This section analyses the necessary changes on link information brought by flexible grids.

8.3.1. Available Frequency Ranges/slots of DWDM Links

In the case of flexible grids, channel central frequencies span from 193.1 THz towards both ends of the C band spectrum with 6.25 GHz granularity. Different LSPs could make use of different slot widths on the same link. Hence, the available frequency ranges should be advertised.

8.3.2. Available Slot Width Ranges of DWDM Links

The available slot width ranges needs to be advertised, in combination with the Available frequency ranges, in order to verify whether a LSP with a given slot width can be set up or not; this is is constrained by the available slot width ranges of the flexi-grid enabled ROADMs (the flexi-grid Frequency slot matrix). Depending on the availability of the slot width ranges, it is possible to allocate more spectrum than strictly needed by the LSP.

8.3.3. Tunable Optical Transmitters and Receivers

The slot width of a LSP is determined by the transmitter and receiver that could be mapped to ADD/DROP interfaces in WSON. Moreover their central frequency could be fixed or tunable, hence, both the slot width of an ADD/DROP interface and the available central frequencies should be advertised.

8.3.4. Hierarchical Spectrum Management

[Editors' note: the part on the hierarchy of the optical spectrum could be confusing, we can discuss it]. The total available spectrum on a fiber could be described as a resource that can be divided by a media device into a set of Frequency Slots. In terms of managing spectrum, it is necessary to be able to speak about different granularities of managed spectrum. For example, a part of the spectrum could be assigned to a third party to manage. This need to partition creates the impression that spectrum is a hierarchy in view of Management and Control Plane. The hierarchy is created within a management system, and it is an access right hierarchy only. It is a

management hierarchy without any actual resource hierarchy within fiber. The end of fiber is a link end and presents a fiber port which represents all of spectrum available on the fiber. Each spectrum allocation appears as Link Channel Port (i.e., frequency slot port) within fiber.

8.3.5. Information Model

Fixed DM grids can also be described via suitable choices of slots in a flexible DWDM grid. However, devices or applications that make use of the flexible grid may not be capable of supporting every possible slot width or central frequency position. Following is the definition of information model, not intended to limit any IGP encoding implementation. For example, information required for routing/path selection may be the set of available nominal central frequencies from which a frequency slot of the required width can be allocated. A convenient encoding for this information (may be as a frequency slot or sets of contiguous slices) is further study in IGP encoding document.

[Editor's note: to be discussed]

<Available Spectrum in Fiber for frequency slot> ::=

 <Available Frequency Range-List>
 <Available Central Frequency Granularity >
 <Available Slot Width Granularity>
 <Minimal Slot Width>
 <Maximal Slot Width>

<Available Frequency Range-List> ::=

 <Available Frequency Range >[< Available Frequency Range-List>]

<Available Frequency Range >::=

 <Start Spectrum Position><End Spectrum Position> |
 <Sets of contiguous slices>

<Available Central Frequency Granularity> ::= n x 6.25GHz,

 where n is positive integer, such as 6.25GHz, 12.5GHz, 25GHz, 50GHz
 or 100GHz

<Available Slot Width Granularity> ::= m x 12.5GHz,

 where m is positive integer

<Minimal Slot Width> ::= j x 12.5GHz,

 j is a positive integer

<Maximal Slot Width> ::= k x 12.5GHz,

 k is a positive integer (k >= j)

Figure 8. Routing Information model

8.4. Signaling requirements

Note on explicit label control

Compared with [RFC6163], except identifying the resource (i.e., fixed wavelength for WSON and frequency resource for flexible grids), the other signaling requirements (e.g., unidirectional or bidirectional, with or without converters) are the same as WSON described in the section 6.1 of [RFC6163]. In the case of routing and distributed SA, GMPLS signaling can be used to allocate the frequency slot to a LSP.

For R+DSA, the GMPLS signaling procedure is similar to the one described in section 4.1.3 of [RFC6163] except that the label set should specify the available nominal central frequencies that meet the slot width requirement of the LSP.

8.4.1. Slot Width Requirement

[Editors' note: the signaling requirements need to be discussed. This is just preliminary text].

In order to allocate a proper frequency slot for a LSP, the signaling should specify its slot width requirement. The intermediate nodes can collect the acceptable central frequencies that meet the slot width requirement hop by hop. The tail-end node also needs to know the slot width of a LSP to assign the proper frequency resource. Hence, the slot width requirement should be specified in the signaling message when a LSP is being set up. [Note: other methods may not need to collect availability]

8.4.2. Frequency Slot Representation

The frequency slot can be determined by the central frequency (n value) and slot width (m value). Such parameters should be able to be specified by the signaling protocol.

8.4.3. Relationship with MRN/MLN

8.4.3.1. OCh Layer

8.4.3.2. Media (frequency slot) layer

9. Control Plane Procedures

Resizing existing LSP(s) without deletion: refers to increase or

decrease of slot width value 'm' without changing the value of 'n'

[Editor's note: Restoration / Resizing a single LSP without deletion as well as timing constraints. As per ITU-T clarification on service affecting or non-service affecting (i.e., hitless) restoration, at present no hitless resizing protocol has been defined for OCh. Hitless resizing is defined for an ODU entity only.]

10. Backwards (fixed-grid) compatibility, and WSON interworking

- o SSON as evolution of WSON, same LSC, different Swcap?
- o Potential problems with having the same swcap but the label format changes w.r.t. wson
- o A new SwCap may need to be defined, LSC swcap already defined ISCD which can not be modified
- o Role of LSP encoding type?
- o Notion of hierarchy? There is no notion of hierarchy between WSON and flexi-grid / SSON - only interop / interwork.

Arguments for LSC switching capability

[QW] A LSP for an optical signal which has a bandwidth of 50GHz passes through both a fixed grid network and a flexible grid network. We assume that no OEOs exist in the LSP, so both the fixed grid path and flexible grid path occupy 50GHz. From the perspective of data plane, there is no change of the signal and no multiplexing when the fixed grid path interconnects with flexible grid path. From this scenario we can conclude that both fixed grid network path and flexible grid network path belong to the same layer. No notion of hierarchy exists between them.

[QW] stitching LSP which is described in [RFC5150] can be applied in one layer. LSP hierarchy allows more than one LSP to be mapped to an H-LSP, but in case of S-LSP, at most one LSP may be associated with an S-LSP. This is similar to the scenario of interconnection between fixed grid LSP and flexible grid LSP. Similar to an H-LSP, an S-LSP could be managed and advertised, although it is not required, as a TE link, either in the same TE domain as it was provisioned or a different one. Path setup procedure of stitching LSP can be applied in the scenario of interconnection between fixed grid path and flexible grid path.

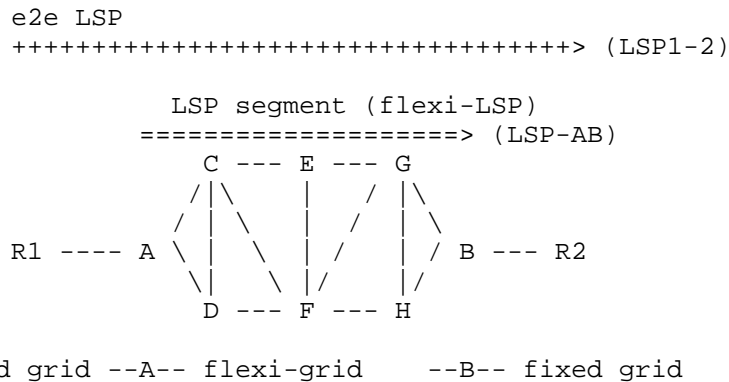


Figure 9. LSP Stitching [RFC5150] and relationship with fixed-flexi

11. Misc & Summary of open Issues [To be removed at later versions]

- o Will reuse a lot of work / procedures / encodings defined in the context of WSON
- o At data rates of GBps / TBps, encoding bandwidths with bytes per second unit and IEEE 32-bit floating may be problematic / non scalable.
- o Bandwidth fields not relevant since there is not a 1-to-1 mapping between bps and Hz, since it depends on the modulation format, fec, either there is an agreement on assuming best / worst case modulations and spectral efficiency.
- o Label I: "m" is inherent part of the label, part of the switching, allows encode the "lightpath" in a ERO using Explicit Label Control, Still maintains that feature a cross-connect is defined by the tuple (port-in, label-in, port-out, label-out), allows a kind-of "best effort LSP"
- o Label II: "m" is not part of the label but of the TSPEC, needs to be in the TSPEC to decouple client signal traffic specification and management of the optical spectrum, having in both places is redundant and open to incoherences, extra error checking.
- o Label III: both, It reflects both the concept of resource request allocation / reservation and the concept of being inherent part of the switching.

12. Security Considerations

TBD

13. Contributing Authors

Qilei Wang
ZTE
Ruanjian Avenue, Nanjing, China
wang.qilei@zte.com.cn

Malcolm Betts
ZTE
malcolm.betts@zte.com.cn

Sergio Belotti
Alcatel Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6863033
sergio.belotti@alcatel-lucent.com

Cyril Margaria
Nokia Siemens Networks
St Martin Strasse 76, Munich, 81541, Germany
+49 89 5159 16934
cyril.margaria@nsn.com

Xian Zhang
Huawei
zhang.xian@huawei.com

Yao Li
ZTE
Zijinghua Road, Nanjing, China
li.yao3@zte.com.cn

Fei Zhang
ZTE
Zijinghua Road, Nanjing, China
zhang.fei3@zte.com.cn

Lei Wang
ZTE
East Huayuan Road, Haidian district, Beijing, China
wang.lei131@zte.com.cn

Guoying Zhang
China Academy of Telecom Research
No.52 Huayuan Bei Road, Beijing, China
zhangguoying@rittt.cn

Takehiro Tsuritani
KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino, Saitama, Japan
tsuri@kddilabs.jp

Lei Liu
KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino, Saitama, Japan
le-liu@kddilabs.jp

Eve Varma
Alcatel-Lucent
+1 732 239 7656
eve.varma@alcatel-lucent.com

Young Lee
Huawei

Jianrui Han
Huawei

Sharfuddin Syed
Infinera

Rajan Rao
Infinera

Marco Sosa
Infinera

Biao Lu
Infinera

Abinder Dhillon
Infinera

Felipe Jimenez Arribas
Telefonica I+D

Andrew G. Malis
Verizon

Adrian Farrel
Old Dog Consulting

Daniel King
Old Dog Consulting

14. Acknowledgments

The authors would like to thank Pete Anslow for his insights and clarifications.

15. References

15.1. Normative References

- [G.709] International Telecommunications Union, "ITU-T Recommendation G.709: Interfaces for the Optical Transport Network (OTN).", March 2009.
- [G.805] International Telecommunications Union, "ITU-T Recommendation G.805: Generic functional architecture of transport networks.", March 2000.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC5150] Ayyangar, A., Kompella, K., Vasseur, JP., and A. Farrel, "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.
- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

15.2. Informative References

- [G.694.1] International Telecommunications Union, "ITU-T Recommendation G.694.1, Spectral grids for WDM applications: DWDM frequency grid, draft v1.6 2011/12", 2011.
- [G.872] International Telecommunications Union, "ITU-T Recommendation G.872, Architecture of optical transport networks, draft v0.12 2012/03 (for discussion)", 2012.
- [WD12R2] International Telecommunications Union, WD12R2, Q12-SG15, ZTE, Ciena WP3, "Proposed media layer terminology for G.872", 05 2012.

Authors' Addresses

Oscar Gonzalez de Dios (editor)
Telefonica I+D
Don Ramon de la Cruz 82-84
Madrid, 28045
Spain

Phone: +34913128832
Email: ogondio@tid.es

Ramon Casellas (editor)
CTTC
Av. Carl Friedrich Gauss n.7
Castelldefels, Barcelona
Spain

Phone: +34 93 645 29 00
Email: ramon.casellas@cttc.es

Fatai Zhang
Huawei
Huawei Base, Bantian, Longgang District
Shenzhen, 518129
China

Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Xihua Fu
ZTE
Ruanjian Avenue
Nanjing,
China

Email: fu.xihua@zte.com.cn

Daniele Ceccarelli
Ericsson
Via Calda 5
Genova,
Italy

Phone: +39 010 600 2512
Email: daniele.ceccarelli@ericsson.com

Iftekhhar Hussain
Infinera
140 Caspian Ct.
Sunnyvale, 94089
USA

Phone: 408-572-5233
Email: ihussain@infinera.com

INTERNET-DRAFT
Intended Status: Standard Track
Expires: February 2, 2013

Khuzema Pithewan
Rajan Rao
Infinera
August 1, 2012

OSPF-TE extensions for MLNMRN based on OTN
draft-rao-ccamp-otn-mlnmrn-ospfte-ext-00.txt

Abstract

This document specifies OSPF extensions for multi-layer/multi-region where one of the regions is OTN.

Status of this Memo

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

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

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

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	3
2.	OTN Layer Identification	3
3.	Interface Adjustment Capability Extensions for OTN	4
4	Procedure	5
5	IANA Considerations	5
6	Security Considerations	5
7	References	5
8	Authors' Addresses	5

1 Introduction

This document specifies the OSPF extensions required to work in multi-region networks involving OTN. The specification is based on the requirement as specified in RFC 5212. As per the said RFC, ISCD characterizes the information associated to one or more network layers. Same RFC also says that the information about the adjustment capabilities of the nodes in the network allow the path computation process to select an end-to-end multi-layer or multi-region path that includes links with different switching capabilities joined by LSRs that can adapt (i.e., adjust) the signal between the links. By inference, information about the adjustment capabilities should be able to identify a layer in ISCD, if ISCD specifies more than one layer.

RFC6001 specifies how to advertise adjustment capabilities between two switching regions. IACD definition has provision to extend it for a specific technology through Adjustment Capability Specific information (ACSI) field, if required. ACSI field can be used to identify a layer in the multi-layer ISCD. OTN being defined as multi-layer ISCD, the corresponding IACD needs to be extended to be able to carry layer identification so as to enable multi-layer/multi-region path computation.

2. OTN Layer Identification

[GMPLS-OTN-OSPF] defines attributes that identifies a layer in multi-layer OTN ISCD. These attributes are part of Bandwidth sub-TLV in Switch capability specific information of ISCD. These attributes are reproduced here for completeness sake.

- * Signal Type: Layer for which bandwidth is being advertised.
- * Hierarchy : also called as multiplexing branch that specifies all the layers between server layer and signal type.
- * TSG : Time Slot Granularity

3. Interface Adjustment Capability Extensions for OTN

RFC6001 defines IACD sub-TLV as follows. Please refer to the RFC for definition of individual fields of the sub-TLV.

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Lower SC								Lower Encoding								Upper SC								Upper Encoding							
								Max LSP Bandwidth at priority 0																							
								Max LSP Bandwidth at priority 1																							
								Max LSP Bandwidth at priority 2																							
								Max LSP Bandwidth at priority 3																							
								Max LSP Bandwidth at priority 4																							
								Max LSP Bandwidth at priority 5																							
								Max LSP Bandwidth at priority 6																							
								Max LSP Bandwidth at priority 7																							
								Adjustment Capability-specific information																							
								(variable)																							

Adjustment Capability-specific information abbreviated as ACSI henceforth for OTN G.709v3 carries LayerID Sub-TLV which is defined as follows

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Type = 1 (LayerID TLV)																Length															
								Signal type								Num of stages								TSG							
								Stage#2								...								Stage#N							
																Res								Stage#1							
																								Padding							

The definition & meaning of fields used in the above sub-TLV is same as in bandwidth sub-TLV of ISCD as defined in [GMPLS-OTN-OSPF]. This LayerID sub-TLV is applicable only when one of the regions is OTN, which means either lower or upper SC and Encoding type MUST have Switch Cap as OTN-TDM and encoding type as G.709 ODUk. The 8

priorities of the BW as defined in main IACD structure, is adjustment capability between the two regions where one of the region is identifies by LayerID sub-TLV.

4 Procedure

A node advertising IACD for the bandwidth between a non-OTN interface and an OTN interface MUST include the Layer ID TLV as part of ACSI as defined above. For multi-region path computation, the path computing node MUST look at the LayerID sub-TLV (in ACSI part of IACD) if lower/upper SC and encoding type is OTN-TDM and G.709 ODUk respectively to identify the layer in OTN ISCD.

5 IANA Considerations

TBD

6 Security Considerations

TBD

7 References

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

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

[GMPLS-OTN-OSPF] Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS)

8 Authors' Addresses

Khuzema Pithewan
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: kpithewan@infinera.com

Rajan Rao
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: rrao@infinera.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

A. Takacs
F. Fondelli
B. Tremblay
Ericsson
Z. Ali
Cisco Systems
October 22, 2012

RSVP-TE Recovery Extension for data plane initiated reversion,
protection timer and SNC options signalling
draft-takacs-ccamp-revertive-ps-07

Abstract

RSVP-TE recovery extensions are specified in [RFC4872] and [RFC4873]. Currently recovery signalling does not support the request for revertive protection, recovery timers values, and sub-network connection (SNC) protection options. This document extends the PROTECTION Object format allowing sub-TLVs, and defines three sub-TLVs to carry wait-to-restore and hold-off intervals and SNC protection options.

Requirements Language

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

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	4
2. Updated PROTECTION Object format and sub-TLVs	6
3. Error handling	9
4. IANA Considerations	10
5. Security Considerations	11
6. References	12
Authors' Addresses	13

1. Introduction

Generalised MPLS (GMPLS) extends MPLS to include support for different switching technologies [RFC3471]. These switching technologies provide several protection schemes [RFC4426][RFC4427] (e.g. 1+1, 1:N, M:N). Many characteristics of those protection schemes are common regardless of the switching technology (e.g. TDM, LSC, etc). GMPLS RSVP-TE signalling has been extended to support the various protection schemes and establish connections (Label Switched Paths (LSPs)) configuring its specific protection characteristics [RFC4426][RFC4872].

Currently RSVP-TE extensions do not address the configuration of protection switching timers. It also does not provide information on the protection switching operation mode (i.e., revertive or non-revertive) and sub-network connection (SNC) protection options.

The Hold-off time (HOFF) is defined as the time between the reporting of signal fail or degrade, and the initialization of the recovery switching operation [RFC4427]. This timer is useful to limit the number of switch actions when multiple layers of recovery are being used, or in case of 1+1 unidirectional protection scheme [G.808.1] to prevent too early switching due to the differential delay between the short and long path.

The Wait-to-Restore time (WTR) is defined as a period of time that must elapse after a recovered fault before an LSP can be used again to transport the normal traffic and/or to select the normal traffic from the LSP [RFC4427]. The WTR time is fundamental in revertive mode of operation, to prevent frequent operation of the protection switch due to an intermittent defect [G.808.1].

Reversion refers to the process of moving normal traffic back to the original working LSP after the failure is cleared and the path is repaired [RFC4426][RFC4427][RFC4872]. In transport networks reversion is desirable since the protection path may not be optimal from a routing and resource consumption point of view, additionally, moving traffic back to the working LSP allows the protection resources to be used to protect other LSPs. On the other hand, reversion requires that the working resources remain allocated during failure. The operator needs to have the choice between revertive and non-revertive protection to balance the pros and cons in a given situation.

WTR and HOFF timers and SNC protection options must be accurately configured at both ends of the LSP. Operators may need to tune WTR and HOFF timers on a per LSP basis to ensure best protection switching performance (e.g., account for differential delays between

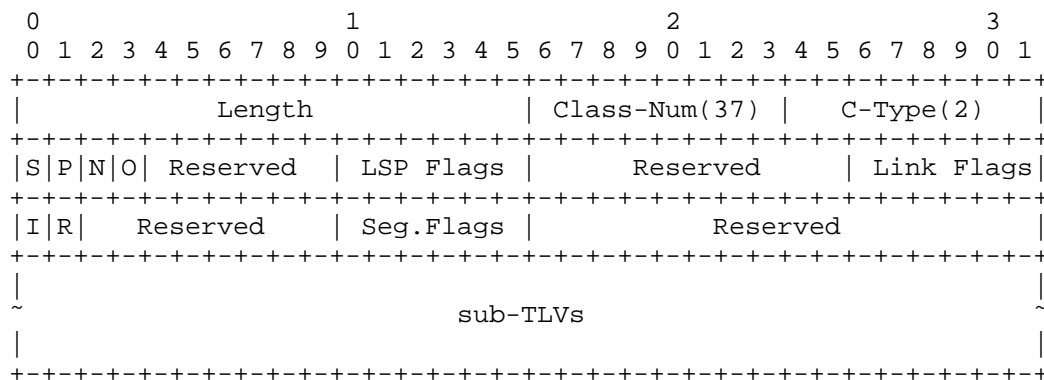
worker and protection paths). Currently these values are either pre-configured to a default value (and so may be suboptimal for some of the LSPs) or need to be manually set/tuned after the connections have been established. Since these parameters are important for recovery in transport networks, it is desirable that GMPLS RSVP-TE protection signalling carries the necessary information.

This document extends the PROTECTION Object format allowing sub-TLVs, and defines three sub-TLVs to carry WTR and HOFF timer values and SNC protection options.

2. Updated PROTECTION Object format and sub-TLVs

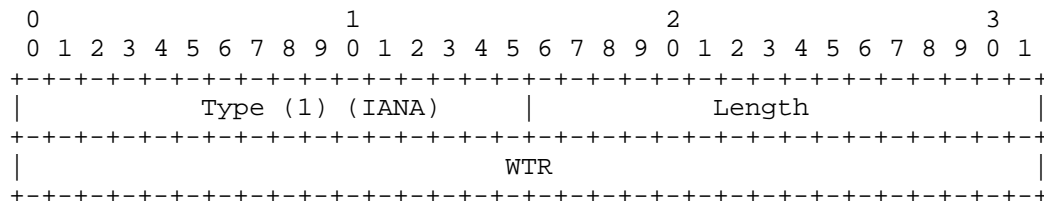
In [RFC4872] and [RFC4873] the PROTECTION object is specified to support end-to-end and segment recovery. In order to ease addition of protection attributes the PROTECTION Object is extended to carry sub-TLVs. The new format updates the PROTECTION Object format of C-Type 2. The updated format is depicted below. IANA is requested to maintain the TLV space for the PROTECTION Object.

We retained C-Type to ensure that nodes not capable of interpreting the new format (sub-TLVs) will still be able to process the object without being required to generate an error; while nodes recognising the new format will process the TLVs accordingly. The processed sub-TLV MUST be included in the PROTECTION Object sent in the Resv message upstream, to ensure that the sender can maintain a consistent view of the actual protection configuration of the LSP.



This document specifies three new sub-TLVs.

WTR - Wait-to-Restore time sub-TLV specifies the WTR time. If the WTR field is 0 the protection switching operation mode is non-revertive, otherwise revertive operation with the signalled timer (in milliseconds) is requested. The value 0xffffffff is reserved, and refers to a locally pre-configured WTR value.



HOFF - Hold-off time sub-TLV specifies the HOFF time. The values are in milliseconds. The value 0xffffffff is reserved, and refers to a locally pre-configured HOFF value.

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Type (2) (IANA)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Length                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     HOFF                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

SNC protection options sub-TLV specifies attributes of the SNC protection.

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Type (3) (IANA)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
| mode |                                     Reserved                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Mode field (3 bits): The mode values are defined as follows:

SNC Protection Mode	Value
-----	-----
Reserved	0x0
SNC/N (Sub Network Connection protection with Non-intrusive monitoring)	0x1
SNC/I (Sub Network Connection protection with Inherent monitoring)	0x2
SNC/S (Sub Network Connection protection with Sub-layer monitoring)	0x3

Reserved field (29 bits): This field is reserved for future use. It MUST be set to 0 when sent and MUST be ignored when received.

In the case of end-to-end protection the PROTECTION Object is inserted at the top level in the Path message, the WTR, HOFF and SNC

options sub-TLVs correspond to the end-to-end protection. In the case when a segment of the LSP is to be protected and the WTR and HOFF timers and SNC protection options for the protection segment are to be set by signalling, explicit segment recovery control has to be used, i.e., the PROTECTION Object with the desired timers set must be inserted in the appropriate Secondary Explicit Route Object (SERO).

3. Error handling

In the case a specific configuration of the timers or SNC option is not supported the corresponding error should be generated and sent in the PathErr message: "Routing Problem/Unsupported WTR value" or "Routing Problem/Unsupported HOFF value" or "Routing Problem/Unsupported SNC protection mode".

4. IANA Considerations

4.1. New TLV space for the PROTECTION object

A new TLV space needs to be opened and maintained for the PROTECTION Object in the "Class Names, Class Numbers, and Class Types " Registry.

4.3. New RSVP error sub-code

For Error Code = 24 "Routing Problem" (see [RFC3209]) the following sub-codes are defined.

Sub-code -----	Value -----
Unsupported WTR value	To be assigned by IANA (suggest value: 80)
Unsupported HOFF value	To be assigned by IANA (suggest value: 81)
Unsupported SNC protection mode	To be assigned by IANA (suggest value: 82)

5. Security Considerations

This document introduces no new security issues. The considerations in [RFC4872] and [RFC4873] apply.

6. References

- [G.808.1] "Generic protection switching -- Linear trail and subnetwork protection", ITU-T Recommendation G.808.1, March 2006.
- [IEEE-PBBTE]
"IEEE 802.1Qay Draft Standard for Provider Backbone Bridging Traffic Engineering", work in progress.
- [RFC3471] "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC4426] "Generalized Multi-Protocol Label Switching (GMPLS) Recovery Functional Specification", RFC 4426, March 2006.
- [RFC4427] "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4427, March 2006.
- [RFC4872] "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.
- [RFC4873] "GMPLS Segment Recovery", RFC 4873, May 2007.

Authors' Addresses

Attila Takacs
Ericsson
Laborc u. 1.
Budapest, 1037
Hungary

Email: attila.takacs@ericsson.com

Francesco Fondelli
Ericsson
Via Negrone
Genova, 16153
Italy

Email: francesco.fondelli@ericsson.com

Benoit Tremblay
Ericsson
8400 Decarie.
Montreal, Quebec H4P 2N2
Canada

Email: benoit.c.tremblay@ericsson.com

Zafar Ali
Cisco Systems

Email: zali@cisco.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 29, 2013

Qilei Wang
Xihua Fu
ZTE Corporation
Feb 25, 2013

OSPF extensions for support spectrum sub-band allocation
draft-wang-ccamp-flexigrid-wavelength-range-ospf-02.txt

Abstract

This document addresses the requirements and routing protocol extension of spectrum sub-band allocation in order to help reduce non-linear effect and raise spectrum utilization rate in the scenario of indiscriminately positioning of various channels with different bit rates.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on August 29, 2013.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. Conventions used in this document	4
3. Overview of the Solution	5
4. Extension of routing protocol	6
4.1. Relationship with WSON	6
4.2. Extensions of OSPF Protocol to Support Spectrum Group Allocation	7
4.2.1. Spectrum sub-band Allocation by Bitrates	7
5. Security Considerations	9
6. IANA considerations	9
7. References	9
7.1. Normative References	9
7.2. Informative References	9
Authors' Addresses	10

1. Introduction

In current DWDM systems, completely freedom and indiscriminate positioning of various channels with different bit rates is likely to lead to dramatically impaired system performance due to XPM (Cross-phased Modulation) effect and low spectrum utilization rate.

Cross phase modulation (XPM) is known as the phenomenon that variations of intensity of one optical signal can change the refractive index of the fiber, and modulate the phase of the other optical signals co-propagating in the same fiber.

When DWDM was first introduced, the typical wavelength data rate was 2.5Gb/s. 10Gb/s wavelength was enabled because of the development of higher performance optical modulators. However, the same simple modulation format was used at both 2.5Gb/s and 10Gb/s. The modulation technique is Intensity Modulation with Direct Detection (IM-DD). Barrier appeared if we want a wavelength to transport data with 40 Gb/s bitrate or more. After years of development, coherent technology is introduced to broken the limit for 40Gb/s and soon for 100Gb/s transmission.

Intensity modulation direct detection (IM-DD) systems are less sensitive to the variation of phase of the signal and is going to bring the variation of the intensity, then the changes of refractive index of the fiber. Just the reverse, optical coherent systems is a phase sensitive system. The modulated phase due to XPM has significant influence on the system performance. This is due to the fact that the phase modulation due to XPM can be transformed into intensity modulation through the chromatic dispersion of the fiber and will result in the distortion of signal. For example, if we mix 10Gbit/s NRZ modulated channels with 100Gbit/s xPSK modulated channels indiscriminately, XPM would have a detrimental effect on the 100Gbit/s signal if this are caused by 10Gbit/s signal; Complex modulation formats (e.g., 16QAM) would be used to modulate signal beyond 100Gbit/s (e.g., 400Gbit/s, 1Tbit/s), while QAM modulation format experiences both intensity and phase modulation, a QAM signal may affect another QAM signal due to XPM effect.

In current DWDM system with different bit rates, general advice is to group the channels with the same bit rates into the same spectrum sub-band to avoid the detrimental XPM effect.

Except the advantage described above, grouping of channels with the same bitrates will help reduce fragment in flexible grid network. Two kinds of DWDM application is described in the newest version of [G.694.1], one is fixed grid, and the other is flexible grid. Fixed grid is the traditional DWDM application with fixed channel slot

width (e.g., 50GHz) which has been deployed in large scale in the network and flexible grid is a new kind of DWDM application with different channel slot widths which can be used to transport large bandwidth data. The flexible grid technology is also a DWDM application which is different from fixed grid application. Flexible grid is a kind of new DWDM application with different channel slot width (e.g., 50GHz, 87.5GHz etc). Signals with different bit rates may occupy different slot widths and use different modulation format. Grouping of signals with the same bit rates into the same spectrum may also be a better choice than completely freedom positioning.

Frequently setup and release of flexible grid optical channels which occupy different slot widths would result in the spectrum fragments that can't be used anymore, and this will cause the decline of spectrum utilization rate. Global Concurrent Optimization (GCO), which combines with the implementation of the stateful PCE, can be used for defragmentation and raise spectrum utilization rate. But it's complicated to apply the GCO in current network, because this will involve a large scale of resources and computation. Another method that can be used is to group the channels with the same bit rates into the same spectrum sub-band and this is also a feasible solution to help raise spectrum utilization rate.

According to the previous description, grouping of channels with the same bit rates maybe a good choice to reduce the XPM effect and raise spectrum utilization rate, especially in network with different slot widths. So in order to take the grouping into consideration during the path computation, a sub-band spectrum of the available spectrum SHOULD be allocated in advance and this spectrum allocation information SHOULD be known by path computation element when compute the path. This document mainly addresses the routing protocol extension to support the advertisement of the spectrum allocation information. Policy may be used by operator when split spectrum supported on a link into several sub-bands. One spectrum sub-band can only be used for path (optical channel) setup with specific attributes, for example, with specific bitrates and/or modulation format.

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 [RFC2119].

3. Overview of the Solution

In current DWDM system with different bit rates, general advice is to group the channels with the same bit rates and modulation format into the same spectrum sub-band to avoid the detrimental XPM effect. As signals of the same bitrates usually use the same modulation format on a specific link, this document mainly pays attention to grouping of channels with the same bitrates. Figure 1 shows an example, in which part of the spectrum frequency slots (i.e., the left sub-band) are allocated to 10 Gbit/s channels while another part of the frequency slots (i.e., the right sub-band) are allocated to 100 Gbit/s channels. Currently an NRZ modulation format is employed for 10 Gbit/s channels, while 40 Gbit/s and/or 100 Gbit/s channels are mostly phase-modulated (e.g. xPSK) signals. As noted above it is common practice to keep an appropriate guard band between channels with different bit rates and/or modulation format to minimize nonlinear effects induced signal degradations and to group channels with the same bit rate and modulation format.

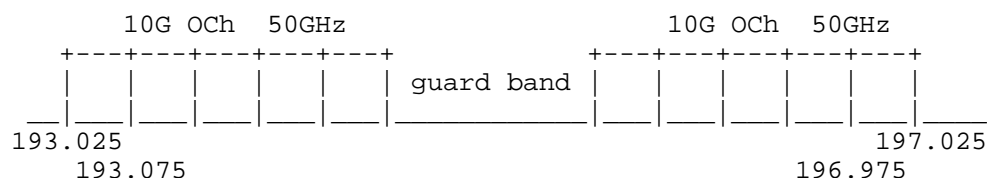


Figure 1: Group of Channels

[Notes: According to the simulation results about grouped spectrum allocation and ungrouped spectrum allocation for 100Gbit/s and 400Gbit/s transmissions are carried. A grouped spectrum allocation only resulted in a small improvement (about 0.2 and 0.5 dB) compared to ungrouped spectrum allocation for 100G and 400G data transmissions respectively. The penalty induced by mixed QPSK and 16QAM signals seems to be much smaller than in a QPSK and NRZ hybrid system. This may be explained by the similar spectrum of the QPSK and 16QAM formats for the same baud rate. These results seem to indicate that systems operated with only phase modulated signals may show significantly lower impairments compared to systems operated with a mix of NRZ and phase modulated signals.]

[Note: we are not sure if grouping of channels with the same bit rate is needed in future. According to the discussion in ITU-T, a part of

the experts express that: we should use grouping of channels with the same bit rate in the short term, however, in the long term, it may be forced to use indiscriminate positioning of the spectrum.]

Some other advantages are also brought if operator split available spectrum into several sub-bands and one spectrum sub-band can only be used for optical channel setup with the same bit rate. A spectrum sub-band which can only be used for optical channel setup with the same bit rate will help reduce fragments. Channels in the same spectrum group sub-band with the same bitrates looks almost like fixed grid technology with the same slot width, and they won't generate fragments in the process of path setup and release. It may also be convenient for operator to manage the network if the operator groups the optical channel with the same bit rate.

When control plane uses path computation element to setup an end-to-end path through the DWDM network, spectrum available information and restriction information (e.g., spectrum partition information) should be taken into consideration in order to compute a suitable end-to-end path. The spectrum sub-band information needs to be advertised by routing protocol in order to help the path computation. Section 4 describes the extension of OSPF routing protocol to advertise these spectrum sub-band information in order to help path computation.

4. Extension of routing protocol

4.1. Relationship with WSON

WSON related work can be re-used in this document to advertise the spectrum group information. This Section addresses the routing extension work of the features which is describe in the previous section base on the current WSON work in IETF CCAMP Group. In the document [draft-ietf-ccamp-general-constraint-encode], a new link sub-TLV called Port Label Restrictions sub-TLV is defined. Descriptions about Port Label Restrictions sub-TLV in the draft [draft-ietf-ccamp-gmpls-general-constraints-ospf-te] are introduced here: "Port label restrictions describe the label restrictions that the network element (node) and link may impose on a port. These restrictions represent what labels may or may not be used on a link and are intended to be relatively static. More dynamic information is contained in the information on available labels. Port label restrictions are specified relative to the port in general or to a specific connectivity matrix for increased modeling flexibility" and "For example, Port Label Restrictions describes the wavelength restrictions that the link and various optical devices such as OXCs, ROADMs, and waveband multiplexers may impose on a port in WSON".

According to the previous description, the restrictions information carried in the port label restriction sub-TLV are used to represent what wavelength/spectrum may or may not be used on a link and are relatively static. The spectrum group allocation information described in this document can be regarded as label restrictions which are imposed by network element (node) on a port, and the network element include various optical devices such as OXCs, ROADMs and waveband multiplexers and so on. These spectrum sub-bands restrictions represent the spectrum allocation information and a spectrum sub-band can only be used for channels setup with specific bitrates and/or modulation format. This restriction is relative static and can be carried in Port Label Restrictions sub-TLV.

4.2. Extensions of OSPF Protocol to Support Spectrum Group Allocation

Spectrum sub-band allocation information should be known by path computation element if operators want to compute an end-to-end optical channel path. As described in the previous section, Port Label Restrictions sub-TLV can be used to carry this spectrum sub-band allocation restriction information. Figure 1 is the format of Port Label Restrictions sub-TLV which is described in [draft-ietf-ccamp-general-constraint-encode] and definition of the parameters included in this sub-TLV can be found in this document.

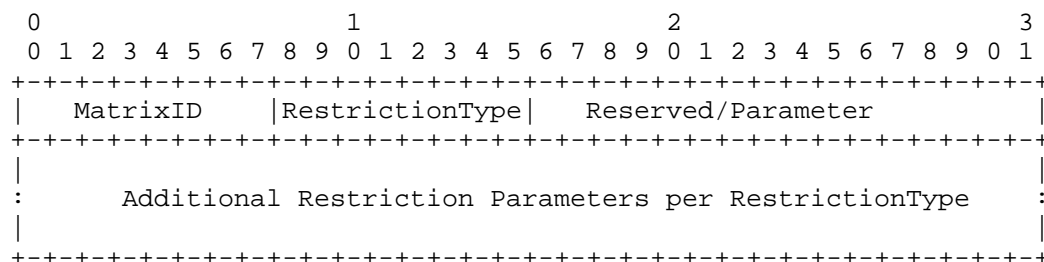


Figure 2: Port Label Restrictions sub-TLV

4.2.1. Spectrum sub-band Allocation by Bitrates

As described in section 3, channels on a single fiber with the same bitrates usually use the same modulation format, especially when the equipments come from one vendor. So here operator can split spectrum into several spectrum sub-bands by bitrates. In this section, OSPF protocol is extended to cover the spectrum sub-bands allocation information by bitrates.

The spectrum sub-bands allocation information by bitrates is needed in the process of path computation if an end-to-end path needs to be computed by path computation element and this information SHOULD be advertised by routing protocols. Figure 3 gives a new kinds of Port Label Restrictions sub-TLV which mainly extent the Additional Restriction Parameters field to cover the spectrum sub-bands allocation information. The parameters in the Additional Restriction Parameters field include Bit Rate which indicates the bitrates of the specific spectrum sub-bands and spectrum boundaries information of the sub-band.

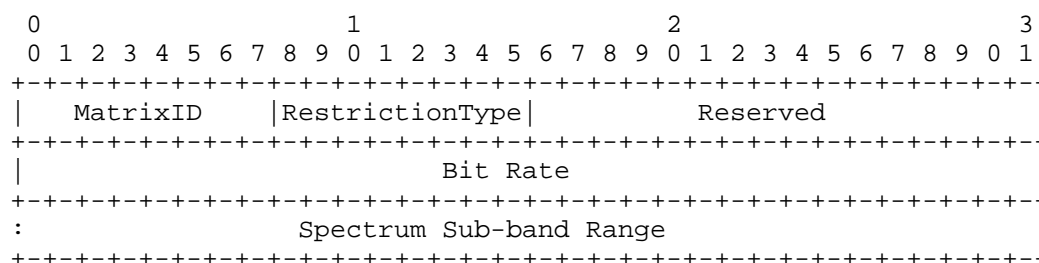


Figure 3

Definition of MatrixID and RestrictionType field can be found in the document [draft-ietf-ccamp-general-constraint-encode]. Value of RestrictionType needs to be assigned by IANA.

[Note: As several routing documents exist in the CCAMP and the label_set object encoding is not determined, this document use "spectrum sub-band range" to represent spectrum allocation information temporarily.]

"Bit Rate" field indicates the bitrates of the specific spectrum sub-band.

In some situation, modulation format information may also be needed to help allocation wavelength range, as signals with the same bitrates on a single fiber can use different modulation format. In this case, modulation formats information is needed to be carried in Port Label Restrictions sub-TLV. Wavelength that is supported by subsystems can be partitioned to service traditional fixed grid technology.

5. Security Considerations

TBD

6. IANA considerations

TBD.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSOs)", RFC 6163, April 2011.

7.2. Informative References

- [Cross Phase Modulation]
International Journal of Electronics and Computer Science Engineering, "Effect on WDM due to Cross Phase Modulation".
- [G.694.1 v1]
International Telecommunications Union, "Draft revised G.694.1 version 1.3".
- [WD6-12] International Telecommunications Union, "Discussion on progressing Q.6/15's work on DWDM applications using flexible grid".
- [flexible-grid-ospf-ext]
Fatai Zhang, Xiaobing Zi, Ramon Casellas, O. Gonzalez de Dios, and D. Ceccarelli, "GMPLS OSPF-TE Extensions in support of Flexible-Grid in DWDM Networks", draft-zhang-ccamp-flexible-grid-ospf-ext-00.txt .
- [flexigrid-lambda-label]
D. King, A. Farrel, Y. Li, F. Zhang, and R. Casellas, "Generalized Labels for the Flexi-Grid in Lambda-Switch-Capable (LSC) Label Switching Routers", draft-farrkingel-ccamp-flexigrid-lambda-label-01.txt .

[super-channel-label]

Iftekhar Hussain, Abinder Dhillon, Zhong Pan, Marco Sosa
and Bert Basch, Steve Liu, Andrew G. Malis, "Generalized
Label for Super-Channel Assignment on Flexible Grid",
draft-hussain-ccamp-super-channel-label-02.txt .

Authors' Addresses

Qilei Wang
ZTE Corporation

Email: wang.qilei@zte.com.cn

Xihua Fu
ZTE Corporation
ZTE Plaza, No.10, Tangyan South Road, Gaoxin District
Xi'an 210012
P.R.China

Email: fu.xihua@zte.com.cn

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

Qilei Wang
Xihua Fu
ZTE Corporation
Oct 22, 2012

RSVP-TE Extensions for GMPLS control of Spectrum Switched Optical
Networks (SSONs)
draft-wang-ccamp-gmpls-sson-rsvpte-02.txt

Abstract

A new architecture of optical transport networks which is addressed in the newest version of G.872 is being developed in ITU-T SG15. Compared with previous G.872 technology, this new technology allows the switch of large chunk of contiguous spectrum which may be wider than the spectrum occupied by a single optical channel signal.

Since current control plane technology isn't able to control this kind of application, this document describes the signaling extension to support the control of this kind of new spectrum utilization and implementation way. This document also addresses the interworking between WSON optical channel and SSON (Spectrum Switched Optical Network) optical channel.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal

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

Table of Contents

1. Introduction	3
1.1. Conventions used in this document	4
2. Terminology	4
3. Requirements and Modeling of SSON	5
3.1. Hierarchy between Optical Channel and Media Channel	5
3.2. Switching Type	6
3.3. Media Channel	6
3.3.1. Label Format	6
3.3.2. Traffic Parameters	6
3.3.3. Grid Attributes of Forwarding Adjacency	7
3.4. Optical Channel	7
3.4.1. Overview of Flexible Grid and Fixed Grid	7
3.4.2. Interwork between WSON OCh signal and SSON OCh signal	7
4. Signaling Protocol Extensions to Support Control of Media Layer	8
4.1. Switching Type	9
4.2. Label Format Extensions of Media Channel Layer	9
4.3. Traffic Parameters of Media Channel Layer	10
5. Signaling Procedures	10
5.1. RSVP-TE Signaling Procedures to Support the Setup of Frequency Slot Channel	10
5.1.1. Centralized Spectrum Assignment	10
5.1.2. Distributed Spectrum Assignment	10
5.2. Interwork between WSON signal and SSON signal	11
6. Security Considerations	13
7. References	13
7.1. Normative References	13
7.2. Informative References	13
Authors' Addresses	14

1. Introduction

In the newest version of G.872, a new kind of spectrum utilization and implementation way is introduced to current optical network. A new kind of entity called media channel which is similar to LSP is introduced. According to the description in [G.872v16], the media channel is a topological construct that represents both the path through the media and the resource (frequency slot) that it occupies. A media channel is bounded by ports on media elements. A media channel may be dimensioned to carry more than one OCh-P signal. That is to say, a chunk of contiguous spectrum which can be occupied by a group of optical channels can be forwarded via wide-band filters as a whole without filtering and switching everything down to the individual OCh (Optical Channel) level in long-haul systems. Compared with narrowband, wideband filters and switching has many advantages, for example, building OCh Signals for management convenience, maximizing the reach and traversing more nodes.

Following is the description taken from [G.872v16], "below the OCh, the entities that provide for configuration of the media channels are described separately from the entities that provide management of the collections of the OCh-P signals that traverse the media". According to the description, we can conclude that the containment relationship exists between media channels and OCh signals, and the containment relationship should be announced to the source and end nodes of the media channels. Intermediate nodes are unnecessary to know the containment relationship because the media channel can be switched as a whole without filtering and switching everything down to individual OCh level.

Two kinds of entities which are spectrum configuration entity and signal management entity should be provide by nodes creating media channels from the perspective of control plane.

Note: Optical channels switched in a media channel may have different spectrum bandwidth.

From the perspective of management plane and control plane, containment relationship indicates that hierarchy exists between the media channel and optical channel. GMPLS protocol are needed to extent to help manage this kind of spectrum utilization and implementation way. This document first describes the coexistence of media channel and optical channel, then the layer model base on the hierarchy between optical channels (OCh) and media channel from management plane or control plane perspective and defines signaling protocol extension to support the control of media channel. As the flexible grid framework document describes both the media channel and optical channel, this document also give a detail description about

the interworking between WSON optical channel and SSON (Spectrum Switched Optical Network) optical channel.

1.1. 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 [RFC2119].

2. Terminology

- o Frequency slot: As defined by Q6 in clause 3.1.2 of G.694.1, a frequency slot is a frequency range which is allocated to a slot and unavailable to other slots within a flexible grid. A frequency slot is defined by its nominal central frequency and its slot width. Detailed description can be found in the framework document.
- o Media channel: as defined in [G.872v16], media channel is used to indicate a media association that represents both the topology (i.e., the path through the media) and the resource (frequency slot) that it occupies. A media channel is bounded by ports on media elements and can span any combination of network elements and fibers.
- o Effective frequency slot: The effective frequency slot of a media channel is that part of the frequency slots of the filters along the media channel that is common to all of the filters' frequency slots. It is described by its nominal central frequency and its slot width.
- o Nominal central frequency (of a frequency slot) - as used by Q6/SG15 in G.694.1. This parameter is associated with a grid position on the fixed grid and a slot in the flexible grid.
- o Network media channel: The end-to-end channel allocated to transport a single OCh payload signal is called a network media channel and supports a single OCh payload network connection.
- o SSON: Spectrum-Switched Optical Network. This concept and definition is introduced from the framework document. An optical network in which a data plane connection is switched based on an optical spectrum frequency slot of a variable slot width, rather than based on a fixed grid and fixed slot width.

3. Requirements and Modeling of SSON

3.1. Hierarchy between Optical Channel and Media Channel

Spectrum may be allocated in larger and contiguous piece than spectrum occupied by a single optical channel which is called Frequency Slot. The frequency slot is described by its nominal central frequency and its slot width [ITU-T G.694.1]. The media Channel is a topological construct that represents both the path through the media and the resource (frequency slot) that it occupies. A media channel is bounded by ports on media elements and can span any combination of network elements and fibers, while the frequency slot is a local concept, while the media channel has an end-to-end meaning. A media channel may be dimensioned to carry more than one OCh P signal. Network media channel is a specific type of media channel which can only be used to transport a single OCh signal and support a single OCh connection.

From the perspective of control plane or management plane, hierarchy exists between media channel and optical channel, as media channel can be used to transport optical channel signals. During the process of path setup, containment relationship between optical channel signal and media channel should be conveyed through signaling and announced to source node and end node in order to help group and detach optical channel signals from one another. Dependency relationship needs to be explicitly told.

Notes: no hierarchy exists in either media channels or optical channels.

Media channels can be switched in a media Matrix. The media channel matrix provides flexible connectivity for the media channels. Media ports at the edge of a media channel matrix may be created and broken to help route media channel path. Media channel connection will limit the connectivity of optical channel signals over the network element within the frequency slot. Media channel matrixes are not mandatory to have the function of optical channel matrix as signals in the media channel can be switched as a whole.

In the case where the switching granularity of the media Matrix allows for independent switching of each OCh, it can be decided as a matter of policy that a request to establish an OCh connection will, internal to the NE, establish a network media channel Matrix Connection of the same spectrum slot width, and the network media Matrix Connection can be released when the OCh connection is released.

As more than one optical channel signal can be carried in a media

channel, notion of hierarchy exists between them. current optical transport network can be modeled into two layers and managed independently from the perspective of management, one is optical channel layer and the other is media layer. Optical channel layer can be modeled as higher layer and media channel layer can be modeled as lower layer. Media layer LSP created in high layer appear as a data link in optical channel layer. One or more optical channel LSPs can be nested into media channel LSP. That is to say, media channel LSP appears as a H-LSP in the higher layer optical channel LSP.

As a kind of resource, spectrum allows for the utilization by both optical channel layer and media layer. It's not necessary to setup media LSP first before the setup of OCh LSP. Coexistence of OCh and media channel in the same link is permitted.

3.2. Switching Type

Switching type can be used to indicate the type of switching that should be performed on a particular link. According to the modeling in the previous section, a new switching type should be defined to indicate the switching capability of media channel layer.

3.3. Media Channel

3.3.1. Label Format

Section 3.3 of [RFC3471] defines waveband switching: "A waveband represents a set of contiguous wavelengths which can be switched together to a new waveband". This is similar to the media channel switching, because they both switch multiple wavelengths or spectrum as a unit.

But the wavelength label defined in [RFC3471] only has significance between neighbors, in order to control the setup and release of media channel with RSVP-TE signaling, a new media channel label which has definite information of nominal central frequency and slot width of the spectrum is needed. This chunk of spectrum can be used for subsequent setup of optical channel path.

3.3.2. Traffic Parameters

In current network, like MPLS network, OTN network, signaling can be used to reserve bandwidth (i.e., bitrates) at each node along the path when set up LSPs. The bandwidth information describes the end-to-end traffic characteristic of a LSP, so the signaling SHOULD be able to carry bandwidth information that a LSP need to occupy.

In the process of the setup of media channel, the most critical

traffic characteristic of a media channel LSP is spectrum, i.e., the spectrum width that a LSP can occupy. For example, if a third party wants to manage and operate a chunk of spectrum by itself, carrier could use the signaling to set up a media channel with a specific spectrum width to satisfy the requirement. Carrier doesn't care how this spectrum can be used by the party and how many data this chunk of spectrum can bear. So when we use signaling to set up a media channel, spectrum resource information (i.e., spectrum width) should be carried in the signaling to reserve the spectrum resource along the path.

3.3.3. Grid Attributes of Forwarding Adjacency

Media matrix connection may interconnect one or more media channels, which in turn may carry one or more OCh signals. In the case the media matrix just allow the switching of spectrum as a whole, internal flexible grid or fixed grid attributes are unnecessary to be known by the forwarding adjacency end points.

3.4. Optical Channel

[Notes: This section mainly addresses the current status of optical channel interconnection, including interconnection between WSON optical channel signal and SSON optical channel signal.]

3.4.1. Overview of Flexible Grid and Fixed Grid

Fixed grid signals have fixed slot width (e.g., 50GHz), while flexible grid signals allow different slot widths (e.g., 50GHz, 87.5GHz). GMPLS and PCE control of fixed grid network (i.e., WSON, Wavelength Switched Optical Network) is close to mature in IETF CCAMP, while flexible grid control plane technology is still being developed in IETF. This section mainly focuses on the interconnection between WSON optical channel signal and SSON optical channel signal.

3.4.2. Interwork between WSON OCh signal and SSON OCh signal

Some open issues are listed in the recent flexible grid framework document and still need to be resolved if we want to push the framework document forward. Part of these issues which may have relation to the interwork between SSON and WSON are listed here:

- 1). If a new switching capability is needed to represent SSON optical channel layer?
- 2). Potential problems with having the same switching capability but the label format changes compared with WSON optical channel layer.

3). Role of LSP encoding type? I think the issue listed here intends to say if a new LSP encoding type is needed for flexible grid optical channel layer.

4). Notion of hierarchy? There is no notion of hierarchy between flexible grid OCh and fixed grid OCh.

Just from my perspective, I think SSON optical channel layer should use the same switching capability as WSON optical channel layer. Some words are given here to describe my opinion. A LSP which has a bandwidth of 50GHz pass through both WSON network and SSON network. We assume that no OEOs exist in the LSP, so both the WSON optical channel path and SSON optical channel path occupy 50GHz. From the perspective of data plane, there is no change of the signal and no multiplexing when the WSON optical channel path interconnects with SSON optical channel path. From this scenario we can conclude that both WSON optical channel layer and SSON optical channel layer belong to the same layer. No notion of hierarchy exists between them. Based on these words, I think both WSON optical channel layer and SSON optical channel layer should use the same switching capability.

The previous words mention the issues 1) and 4). Another two issues are to be discussed in the following description in the process of path setup.

Because there is no notion of hierarchy exists between WSON optical channel layer and SSON optical channel layer, hierarchy LSP which is addressed in [RFC4206] and [RFC6107] can't be applied. But stitching LSP which is described in [RFC5150] can be applied in one layer. LSP hierarchy allows more than one LSP to be mapped to an H-LSP, but in case of S-LSP, at most one LSP may be associated with an S-LSP. This is similar to the scenario of interconnection between WSON OCh LSP and SSON OCh LSP. Similar to an H-LSP, an S-LSP could be managed and advertised, although it is not required, as a TE link, either in the same TE domain as it was provisioned or a different one. Path setup procedure of stitching LSP can be applied in the scenario of interconnection between WSON optical channel path and SSON optical channel path.

4. Signaling Protocol Extensions to Support Control of Media Layer

This section mainly addresses the signaling protocol extension in order to support the control of spectrum-switched optical network media layer and the facilitating of the setup of forwarding adjacency in G.872 optical transport network.

4.1. Switching Type

A new switching type is defined here for media channel layer.

Value	Type
-----	-----
XX(IANA)	Media Channel Switched Capable (MCSC)

Figure 1: Switching Capability

4.2. Label Format Extensions of Media Channel Layer

According to the description in the section 3.2, label should be able to describe the frequency slot characteristic in order to facilitate the switch of this large piece of spectrum. Label format of flexible grid can be introduced here to depict the label of media channel.

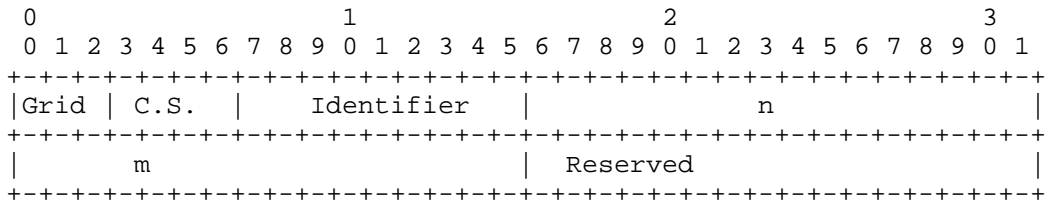


Figure 2: label

Grid Type: 3. A new grid value to support flexible grid.

The meaning of C.S. and identifier is maintained from [RFC6205] and [draft-farrkingel-ccamp-flexigrid-lambda-label].

Similar to the definition in [draft-farrkingel-ccamp-flexigrid-lambda-label], n is used to identify the nominal central frequency, and m (16bits) is used to identify slot width of the media channel.

[Notes: here we use 16 bits to represent the "m" value, because 8 bits maybe not enough for the setup of media channel with large chunk of spectrum. This document is different from current flexible grid document in CCAMP because of different model way.]

4.3. Traffic Parameters of Media Channel Layer

Similar to the original signaling which carry the information of Bandwidth (i.e., bitrates) that a LSP may reserve at each node along the path, signaling that is used to set up media channel SHOULD be able to carry the information of spectrum width. The spectrum width traffic parameters can be organized as follow, and this information is carried in the Sender_Tspec object within a path message.

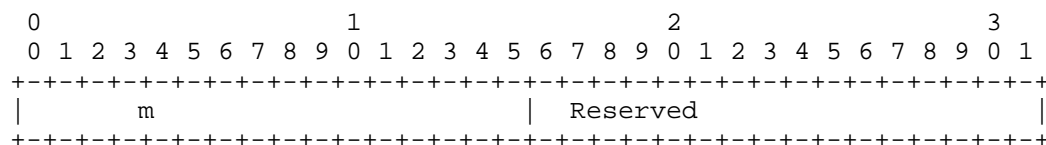


Figure 3: traffic parameter

m (16 bits): the spectrum width is specified by $m \times 12.5$ GHz.

5. Signaling Procedures

5.1. RSVP-TE Signaling Procedures to Support the Setup of Frequency Slot Channel

5.1.1. Centralized Spectrum Assignment

In this case, both of the route and the frequency slot information (i.e., central frequency and spectrum width) are provided by the PCE or ingress node. When signaling a LSP, the assigned label information is carried in the ERO label sub-object which is addressed in [RFC3473]. When the nodes along the LSP receive the path message carrying the ERO and ERO label sub-object, the procedure of path setup is the same as the procedure which is described in [RFC3473] and [RFC4003]. RRO and RRO label sub-object are used to record the label information of the egress.

5.1.2. Distributed Spectrum Assignment

In this case, only the route is provided by a PCE or ingress node before the signaling procedure. The available spectrum SHALL be collected hop by hop and the egress node SHOULD select a proper label for the LSP. After the route is computed, the ingress node SHOULD find out the available spectrum for the LSP on the next link of the route.

Then a path message is sent to the next node along the path according to the route information. The path message MUST contain a Sender_Tspec object to specify the spectrum width of the media channel. A Label_set object SHALL be added to the path message, which contains the candidate available spectrum for the LSP on the next link.

When an intermediate node receives the path message, it can derive the spectrum width information from the Sender_Tspec object. Then it SHOULD find the available spectrum for the LSP on the next link of the route similar to the ingress node. The common part of the two available spectrum sets. If the new set is null, the path message SHALL be rejected by a patherr message. Otherwise, the Label_set object in the path message SHALL be updated according to the new set and the path message is forwarded to the next node according to the route.

When an egress node receives a path message, it SHOULD select an available spectrum from the Label_set object based on local policy and determine the media channel base on the spectrum width and the available spectrum. Then a Resv message is responded so that the nodes along the LSP can establish the optical cross-connect based on the Label object which is determined by the spectrum width in the traffic parameters and the available spectrum in the Label_set object.

5.2. Interwork between WSON signal and SSON signal

The path setup procedure of WSON OCh signal's interworking with SSON OCh signal is described as follows:

Let's take the following network into consideration.

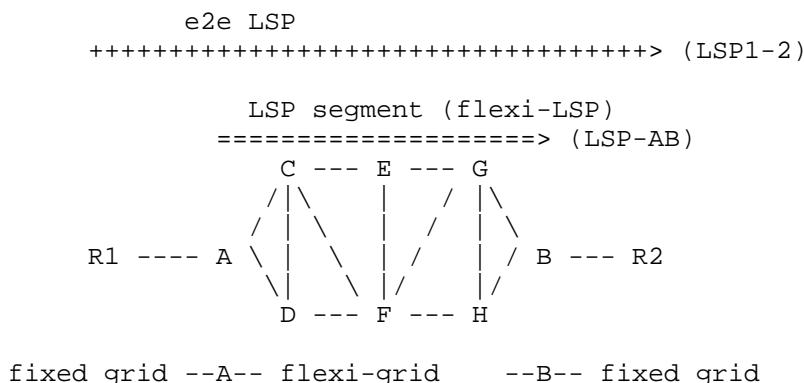


Figure 4: Interworking between WSON OCh and SSON OCh

In this scenario, R1 and R2 are traditional WSON signal capable nodes, A and B are both WSON optical channel signal and SSON optical channel signal capable nodes, the other nodes are SSON optical channel capable nodes. We assume that a 40Gbit/s LSP from R1 to R2 needs to be set up.

Node R1 prepares signaling path message for the end-to-end path setup from R1 to the destination node R2. Before R1 sends path message, R1 should first send a path computation request to the path computation element in order to compute an end-to-end path from R1 to R2. After path computation, PCRep message which contains ERO and label information is send back to R1 from PCE.

R1 encapsulates the path message which contains ERO to explicitly indicate the path and label used and RRO to record the path traversed and label used by node traversed. Then R1 sends the path message to the next hop node A. Here we assume path computation element is capable of fixed grid and flexible grid path computation, and the ERO contain the path information (R1, A, B, R2). When the path message arrives at node A, node A verifies the path message and finds incomplete ERO information, then send another path computation request message to the PCE in order to obtain the whole path information. PCE sends path computation response message which contain ERO (A, D, F, H, B) and label information. Here the label is flexible label information which is addressed in [draft-farrkingell].

To facilitate the control of stitching LSP boundaries, we may use a different encoding type for flexible grid to help control. Encoding type can be used to help stitching LSP boundaries control. Stitching LSP boundaries control looks like FA-LSP boundaries control, but has many differences.

After matching the switching type and encoding type of the interface, Node A blocks the signaling process and decides to set up a stitching LSP according to the flexible grid LSP setup procedure using another signaling process. Procedure for set up stitching LSP can be found in RFC5150. The stitching LSP can be seen as a TE link in the fixed grid network. After the setup of stitching LSP between A and B, A then continues the blocking signaling procedure and sends the path message to the next hop B directly and finishes the end-to-end LSP.

In this scenario of interconnection between WSON OCh and SSON OCh, dynamic stitching LSP setup is frequent, static stitching LSP configuration may not be needed here.

6. Security Considerations

TBD

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.

7.2. Informative References

- [G.694.1 v6]
International Telecommunications Union, "Draft revised G.694.1 version 1.6".
- [G.872 v16]
International Telecommunications Union, "Draft revised Recommendation ITU-T G.872".
- [flexible-grid-ospf-ext]
Fatai Zhang, Xiaobing Zi, Ramon Casellas, O. Gonzalez de Dios, and D. Ceccarelli, "GMPLS OSPF-TE Extensions in support of Flexible-Grid in DWDM Networks", draft-zhang-ccamp-flexible-grid-ospf-ext-00.txt .
- [flexible-grid-requirements]
Fatai Zhang, Xiaobing Zi, O. Gonzalez de Dios, and Ramon Casellas, "Requirements for GMPLS Control of Flexible Grids", draft-zhang-ccamp-flexible-grid-requirements-01.txt .
- [flexigrid-lambda-label]
D. King, A. Farrel, Y. Li, F. Zhang, and R. Casellas, "Generalized Labels for the Flexi-Grid in Lambda-Switch-Capable (LSC) Label Switching Routers", draft-farrkingel-ccamp-flexigrid-lambda-label-01.txt .
- [ospf-ext-constraint-flexi-grid]
L Wang, Y Li, "OSPF Extensions for Routing Constraint Encoding in Flexible-Grid Networks", draft-wangl-ccamp-ospf-ext-constraint-flexi-grid-00.txt .

Authors' Addresses

Qilei Wang

ZTE Corporation

Email: wang.qilei@zte.com.cn

Xihua Fu

ZTE Corporation

ZTE Plaza, No.10, Tangyan South Road, Gaoxin District

Xi'an

P.R.China

Email: fu.xihua@zte.com.cn

