

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

Zafar Ali
George Swallow
Clarence Filsfils
Luyuan Fang
Cisco Systems
Kenji Kumaki
KDDI Corporation
Ruediger Kunze
Deutsche Telekom AG
July 16, 2012

Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)
extension for signaling Objective Function and Metric Bound
draft-ali-ccamp-rc-objective-function-metric-bound-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 January 15, 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.

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

In particular networks such as those used by financial institutions, network performance criteria such as latency are becoming as critical to data path selection. However cost is still an important consideration. This leads to a situation where path calculation involves multiple metrics and more complex objective functions.

When using GMPLS control plane, 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. This document defines extensions to the RSVP-TE Protocol to allow an ingress node to request the required objective function for the path computation, as well as a metric bound to influence route computation decisions at a remote node(s).

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 signaling extensions.....	4
2.1. Objective Function (OF) Subobject.....	4
2.1.1. Minimum TE Metric Cost Path Objective Function....	6
2.1.2. Minimum IGP Metric Cost Path Objective Function...	6
2.1.3. Minimum Latency Path Objective Function.....	6
2.1.4. Minimum Latency Variation Path Objective Function.	7
2.2. Metric subobject.....	7
2.3. Processing Rules for the OF Subobjects.....	8
2.4. Processing Rules for the Metric subobject.....	9
3. Security Considerations.....	10
4. IANA Considerations.....	10
5. Acknowledgments.....	11
6. References.....	11
6.1. Normative References.....	11
6.2. Informative References.....	11

1. Introduction

As noted in [OSPF-TE-METRIC] and [ISIS-TE-METRIC], in certain networks such as financial information networks (e.g. stock market data providers), performance criteria (e.g. latency) are becoming as critical to data path selection along with other metrics. Such networks may require selection of a path that minimizes end-to-end latency. Or a path may need to be found that minimizes some other TE metric, but subject a latency bound. Thus there is a requirement to be able to find end-to-end paths with different optimization criteria.

When the entire route for an LSP is computed at the ingress node, this requirement can be met by a local decision at that node. However, there are scenarios where partial or full route computations are performed by remote nodes. The scenarios 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 UNI-Network (server) node [RFC 4208];

In these scenarios, there is a need for the ingress node to convey the optimization criteria including the TE metrics (e.g., IGP metric, TE metric, hop counts, latency, etc.) to be used for the path computation to the node performing route computation or expansion. Similarly, there is a need for the ingress node to

indicate a TE metric bound for the loose segment being expanded by a remote node.

[RFC5541] defines extensions to the Path Computation Element communication Protocol (PCEP) to allow a Path Computation Client (PCC) indicate in a path computation request the desired objective function. [RFC5440] defines extension to the PCEP to allow a PCC indicate in a path computation request a bound on given TE metric(s). This draft defines similar mechanisms for the RSVP-TE protocol allowing an ingress node to indicate in a Path request the desired objective function along with any associated TE metric bound(s). This information is used by the nodes performing route expansion to find the "best" candidate route.

2. RSVP-TE signaling extensions

This section defines RSVP-TE signaling extensions required to address the above-mentioned requirements. Two new ERO subobject types, Objective Function (OF) and Metric, are defined for this purpose. Their purpose is as follows.

- . OF subobject conveys a set of one or more specific optimization criteria that MUST be followed in expanding route of a TE-LSP in MultiProtocol Label Switching (MPLS) and GMPLS networks.
- . Metric subobject indicates the bound on the path metric that MUST NOT be exceeded for the loose segment to be considered as acceptable by the ingress node.

The scope of the Metric and OF subobjects is the node performing the expansion for loose ERO and the subsequent ERO subobject that identifies an abstract node. The following subsection provides the details.

2.1. Objective Function (OF) Subobject

A new ERO subobject type Objective Function (OF) is defined in order for the ingress node to indicate the required objective function on a loose hop. The ERO subobject type OF is optional. It MAY be carried within an ERO object of RSVP-TE Path message. The OF subobject has the following format:

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	0	1	2	3	4	5	6	7	8	9																														
L																Type																Length																OF Code																Reserved															

```

+-----+
|                                     |
|//                               Optional TLV(s)                               //|
|                                     |
+-----+

```

The fields of OF subobject are defined as follows:

L bit: The L bit SHOULD be set, so that the subobject represents a loose hop in the explicit route.

Type: The Type is to be assigned by IANA (suggested value: 66).

Length: The Length contains the total length of the subobject in bytes, including the Type field, the Length field and the length of the optional TLV(s). When there is no optional TLV, the Length is 4.

OF Code (1 byte): The identifier of the objective function. The following OF code values are suggested. These values are to be assigned by IANA.

- * OF code value 0 is reserved.

- * OF code value 1 (to be assigned by IANA) is for Minimum TE Metric Cost Path (MTMCP) OF defined in this document. See definition of MTCP OF in the following.

- * OF code value 2 (to be assigned by IANA) is for Minimum Interior Gateway Protocol (IGP) Metric Cost Path (MIMCP) OF defined in the following.

- * OF code value 3 (to be assigned by IANA) is for Minimum Load Path (MLP) OF as defined in RFC5541.

- * OF code value 4 (to be assigned by IANA) is for Maximum Residual Bandwidth Path (MBP) OF as defined in RFC5541.

- * OF code value 5 (to be assigned by IANA) is for Minimize Aggregate Bandwidth Consumption (MBC) OF as defined in RFC5541.

- * OF code value 6 (to be assigned by IANA) is for Minimize the Load of the most loaded Link (MLL) OF as defined in RFC5541.

- * OF code value 7 is skipped (to keep the objective function code values consistent between [RFC5541] and this draft.

* OF code value 8 (to be assigned by IANA) is for Minimum Latency Path (MLP) OF defined in this document. See definition of MLP OF in the following.

* OF code value 9 (to be assigned by IANA) is for Minimum Latency Variation Path (MLVP) OF defined in this document. See definition of MLVP OF in the following.

Other objective functions may be defined in future.

Reserved (1 byte): This field MUST be set to zero on transmission and MUST be ignored on receipt.

Optional TLVs may be defined in the future to encode objective function parameters.

2.1.1. Minimum TE Metric Cost Path Objective Function

Minimum TE Metric Cost Path (MTMCP) OF is defined as an Objective Function where a path is computed such that the sum of the TE metric of the links along the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that the sum of the TE metric of the links along the route is minimized.

2.1.2. Minimum IGP Metric Cost Path Objective Function

Minimum IGP Metric Cost Path (MIMCP) OF is defined as an Objective Function where a path is computed such that the sum of the IGP metric of the links along the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that the sum of the IGP metric of the links along the route is minimized.

2.1.3. Minimum Latency Path Objective Function

Minimum Latency Path (MLP) OF is defined as an Objective Function where a path is computed such that latency of the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that overall latency of the loose hop is minimized.

2.1.4. Minimum Latency Variation Path Objective Function

Minimum Latency Variation Path (MLVP) OF is defined as an Objective Function where a path is computed such that latency variation in the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that overall latency variation of the loose hop is minimized.

2.2. Metric subobject

The ERO subobject type Metric is optional. It MAY be carried within an ERO object of RSVP-TE Path message. This subobject has the following format:

```

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   | metric-type |   Reserved   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     metric-bound                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

The fields of the Metric subobject are defined as follows:

L bit: The L bit SHOULD be set, so that the subobject represents a loose hop in the explicit route.

Type: The Type is to be assigned by IANA (suggested value: 67).

Length: The Length is 8.

Metric-type (8 bits): Specifies the metric type associated with the partial route expended by the node processing the loose ERO. The following values are currently defined:

- * T=1: cumulative IGP cost
- * T=2: cumulative TE cost
- * T=3: Hop Counts
- * T=4: Cumulative Latency
- * T=5: Cumulative Latency Variation

Reserved: This field MUST be set to zero on transmission and MUST be ignored on receipt.

Metric-bound (32 bits): The metric-bound indicates an upper bound for the path metric that MUST NOT be exceeded for the ERO expending node to consider the computed path as acceptable. The metric bound is encoded in 32 bits using IEEE floating point format as defined in [IEEE.754.1985]).

2.3. Processing Rules for the OF Subobjects

The basic processing rules of an ERO are not altered. Please refer to [RFC3209] for details.

The scope of the OF subobject is the previous ERO subobject that identifies an abstract node, and the subsequent ERO subobject that identifies an abstract node. Multiple OF subobjects may be present between any pair of abstract nodes.

The following conditions SHOULD result in Path Error with error code "Routing Problem" and error subcode "Bad EXPLICIT_ROUTE object":

- . If the first OF subobject is not preceded by a subobject identifying the next hop.
- . If the OF subobject follows a subobject that does not have the L-bit set.

If the processing node does not understand the OF subobject, it SHOULD send a PathErr with the error code "Routing Error" and error value of "Bad Explicit Route Object" toward the sender [RFC3209].

If the processing node understands the OF subobject and the ERO passes the above mentioned sanity check and any other sanity checks associated with other ERO subobjects local to the node, the node takes the following actions:

- . If the node supports the requested OF(s), the node expands the loose hop using the requested Objective Functions(s) as minimization criterion (criteria) for computing the route to the next abstract node. After processing, the OF subobjects

are removed from the ERO. The rest of the steps for the loose ERO processing follow procedures outlined in [RFC3209].

- . If the node understands the OF subobject but does not support any or all of the requested OF(s), it SHOULD send a Path Error with error code "Routing Problem" and a new error subcode "Unsupported Objective Function". The error subcode "Unsupported Objective Function" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 107).
- . If the node understands the OF subobject and supports all of the requested OF(s) but cannot perform route computation with all objective functions considered together as optimization criteria for the path computation, it SHOULD send a Path Error with error code "Routing Problem" and a new error subcode "Objective Function too complex". The error subcode "Objective Function too complex" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 108).
- . If the objective function is supported but policy does not permit applying it, the processing node SHOULD send a Path Error with error code "Policy control failure" (value 2) and subcode "objective function not allowed". The error subcode "objective function not allowed" for Path Error code "Policy control failure" is to be assigned by IANA (Suggested Value: 105).

2.4. Processing Rules for the Metric subobject

The basic processing rules of an ERO are not altered. Please refer to [RFC3209] for details.

The scope of the Metric subobject is between the previous ERO subobject that identifies an abstract node, and the subsequent ERO subobject that identifies an abstract node. Multiple Metric subobjects may be present between any pair of abstract nodes.

The following conditions SHOULD result in Path Error with error code "Routing Problem" and error subcode "Bad EXPLICIT_ROUTE object":

- . If the first Metric subobject is not preceded by a subobject identifying the next hop.

- . If the Metric subobject follows a subobject that does not have the L-bit set.

If the processing node does not understand the Metric subobject, it SHOULD send a PathErr with the error code "Routing Error" and error value of "Bad Explicit Route Object" toward the sender [RFC3209].

If the processing node understands the Metric subobject and the ERO passes the above mentioned sanity check and any other sanity checks associated with other ERO subobjects local to the node, the node takes the following actions:

- . For all the Metric subobject(s), the node expands the loose hop such that the requested metric bound(s) are met for the route between the two abstract nodes in the ERO. After processing, the Metric subobjects are removed from the ERO. The rest of the steps for the loose ERO processing follow procedure outlined in [RFC3209].
- . If the node understands the Metric subobject but cannot find a route to the next abstract node such that the requested metric bound(s) can be satisfied, it SHOULD send a Path Error with error code "Routing Problem" and a new error subcode "No route available toward destination with the requested metric bounds". The error subcode "No route available toward destination with the requested metric bounds" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 109).

3. Security Considerations

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

4. IANA Considerations

This document adds the following two new subobject of the existing entry for ERO (20, EXPLICIT_ROUTE):

Value	Description
-----	-----

TBA (suggest value: 66) Objective Function (OF) subobject

TBA (suggest value: 67) Metric subobject

These subobject may be present in the Explicit Route Object, but not in the Route Record Object.

OF Code values carried in OF subobject requires an IANA entry with suggested values as defined in section 2.1.

5. Acknowledgments

Authors would like to thank Matt Hartley, Ori Gerstel, Gabriele Maria Galimberti, 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.
- [IEEE.754.1985] IEEE Standard 754, "Standard for Binary Floating-Point Arithmetic", August 1985.

6.2. Informative References

- [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

ID draft-ali-ccamp-rc-objective-function-metric-bound-02.txt

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

George Swallow
Cisco Systems.
swallow@cisco.com

Clarence Filsfils
Cisco Systems.
cfilsfil@cisco.com

Luyuan Fang
Cisco Systems.
lufang@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
Ori Gerstel
Cisco Systems
Kenji Kumaki
KDDI Corporation
Ruediger Kunze
Deutsche Telekom AG
July 16, 2012

Include Routes - Extension to
Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)
draft-ali-ccamp-rsvp-te-include-route-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 January 15, 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 scenarios that require two or more LSPs or segments of LSPs to follow same route in the network. This document specifies methods to communicate route inclusions along the loose hops during path setup using the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) protocol.

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 signaling extensions.....	5
2.1. Explicit Inclusion Route Subobject (EIRS).....	5
2.2. EIRS Subobject Processing Rule.....	8
2.3. Processing of EIRS with XRO and EXRS.....	9
3. Security Considerations.....	9
4. IANA Considerations.....	9
5. Acknowledgments.....	10
6. References.....	10
6.1. Normative References.....	10
6.2. Informative References.....	11

1. Introduction

There are scenarios that require two or more LSPs to follow same route in the network. E.g., many deployments require member LSPs of a bundle/ aggregated link (or Forwarding Adjacency (FA)) follow the same route. Possible reasons for two or more LSPs to follow the same end-to-end or partial route include, but are not limited to:

- . Fate sharing: an application may require that two or more LSP fail together. In the example of bundle link this would mean that if one component goes down, the entire bundle goes down.
- . Homogeneous Attributes: it is often required that two or more LSPs have the same TE metrics like latency, delay variation, etc. In the example of a bundle/ aggregated link this would meet the requirement that all component links (FAs) of a bundle should have same latency and delay variation. As noted in [OSPF-TE-METRIC] and [ISIS-TE-METRIC], in certain networks, such as financial information networks, network performance (e.g. latency and latency variation) is becoming critical and hence having bundles with component links (FAs) with homogeneous delay and delay variation is important.

The RSVP-TE specification, "RSVP-TE: Extensions to RSVP for LSP Tunnels" [RFC3209] and GMPLS extensions to RSVP-TE, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions" [RFC3473] allow abstract nodes and resources to be explicitly included in a path setup. However, such inclusion may not be possible when a loose hop is expanded. It is obviously possible to divide the loose hop into multiple loose hops and construct an inclusion in that fashion. However, there are scenarios where division of a loose hop into multiple explicit loose hops is not possible. Included (but not limited to) are the following:

- . When the destination is in another area, AS, or across a UNI, the ingress node may not have full visibility of the topology. In cases where the ingress node lacks sufficient topological knowledge around the loose hop, it is not able to divide a loose hop into a proper sequence of strict or a sequence of finer-grained loose hops.
- . The ingress node requires two Label Switched Paths (LSPs) to follow the same route but has no knowledge of how a loose hop of a reference LSP was expanded. The ingress node requires certain SLRGs to be explicitly "included" when the loose hop is expanded. This document defines inclusion use of the SRLG subobject defined in [RFC4874].

When the entire route of LSPs that need to follow the same path is computed by the ingress node, the aforementioned requirements can be met by a local decision at the ingress node. However, there are scenarios where a full route computation is not performed at the ingress but instead is performed by remote nodes. This case creates a need for relevant affinity requirements to be communicated to the route expanding 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 UNI-Network (server) node;

This document addresses these requirements and defines procedures that may be used to signal LSPs such that the entire LSP or LSP segments follow the same route.

2. RSVP-TE signaling extensions

A new ERO subobject type the Explicit Inclusion Route Subobject (EIRS) is introduced to indicate an inclusion between a pair of included nodes or abstract nodes. The ERO subobject encoding and processing rules are similar to Explicit Exclusion Route Subobject (EXRS) subobject of ERO defined in [RFC4874], with the exception of include vs. exclude usage.

2.1. Explicit Inclusion Route Subobject (EIRS)

The Explicit Inclusion Route Subobject (EIRS) defines abstract nodes or resources (such as links, SRLG, Circuit IDs (see [DRAFT-LSP-XRO]), unnumbered interfaces, or labels, etc.) that must or should be used on the path between two inclusive abstract nodes or resources in the explicit route. An EIRS is an ERO subobject that contains one or more subobjects of its own, called EIRS subobjects. Each EIRS may carry multiple inclusions. The inclusion is encoded exactly as for XRO subobjects and prefixed by an additional Type and Length.

The format of the EIRS 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|L|   Type   |   Length   |   Reserved   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
//               one or more EIRS subobjects               //
|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

An example of EIRS for SLRG inclusion (SRLG Id 1 and SRLG Id 2) is provided in the following. This example is referenced in the following description.

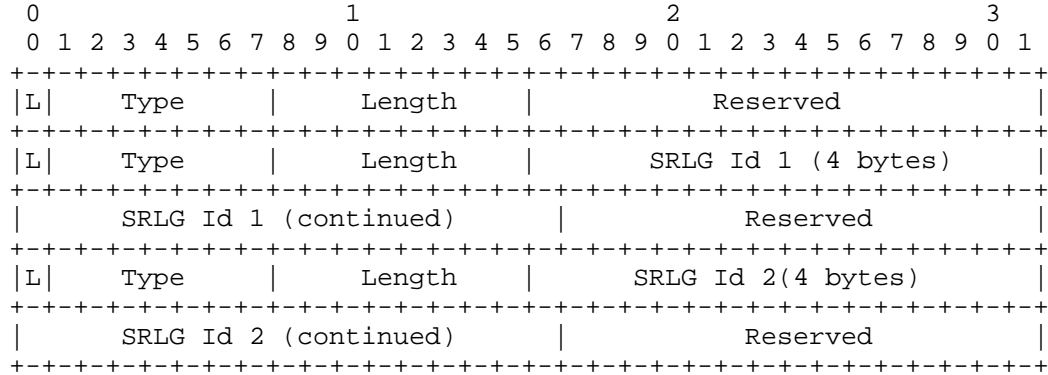


Figure 1: Example of EIRS with SRLG subobjects

Please note that there are two or more "L bits" in an EIRS. The following convention is used to reference the individual "L bits".

EIRS.L: The L bit of the header of the EIRS subobject.
E.g., EIRS.L refers to the first L bit in EIRS example in Figure 1.

EIRS.SubobjectN.L: The L bit of the nth subobject of EIRS.
E.g., EIRS.Subobject2.L refers to the third L bit in EIRS example in Figure 1 (i.e., the L bit to define the expected treatment of SRLG ID2 value).

The fields of the EIRS subobject are defined as follows:

EIRS.L bit: The L bit is an attribute of the EIRS subobject. The L bit SHOULD be set, so that the subobject represents a loose hop in the explicit route.

EIRS.Type: The type of the subobject is to be defined by IANA (Suggested Value: 68).

EIRS.Reserved: This field is reserved. It SHOULD be set to zero on transmission and MUST be ignored on receipt.

EIRS subobjects: An EIRS subobject indicates the abstract node or resource to be included in the path. The format of an EIRS subobject is exactly the same as the format of a subobject in the eXclude Route Object (XRO) (See [RFC4874] and [DRAFT-LSP-XRO-SUB]). This is with the exception of the interpretation of the "EIRS.SubobjectN.L bit" of the subobjects, as detailed in the following.

EIRS.SubobjectN.L bit: For all supported subobjects of EIRS, the EIRS.SubobjectN.L bit has the following interpretation.

- EIRS.SubobjectN.L = 0 indicates that the attribute specified MUST be included.
- EIRS.SubobjectN.L = 1 indicates that the attribute specified SHOULD be included.

An EIRS may include all subobjects defined in this document for the XRO (See [RFC4874] and [DRAFT-LSP-XRO-SUB]). Specifically, an EIRS may include the following subobjects:

EIRS.SubobjectN.Type = 1: IPv4 address [RFC3209].

EIRS.SubobjectN.Type = 2: IPv6 address [RFC3209].

EIRS.SubobjectN.Type = 3: Label [RFC6001].

EIRS.SubobjectN.Type = 4: Unnumbered Interface ID [RFC3477].

EIRS.SubobjectN.Type = 32: Autonomous system number [RFC3209].

EIRS.SubobjectN.Type = 34: SRLG [RFC4874].

EIRS.SubobjectN.Type = 35: Switching Capability (SC) [RFC6001].

EIRS.SubobjectN.Type = TBD (suggested value 37): LSP [DRAFT-LSP-XRO-SUB].

Please note that EIRS.SubobjectN.Type = 33: Explicit Exclusion Route subobject (EXRS) [RFC4874] is not supported.

2.2. EIRS Subobject Processing Rule

The scope of the inclusion is the previous ERO subobject that identifies a node or an abstract node, and the subsequent ERO subobject that identifies a node or an abstract node. The processing rules of the EIRS are the same as the processing rule of the EXRS, with the exception that EIRS subobjects request resource inclusion, whereas EXRS subobjects request resource exclusion.

Multiple inclusions may be present between any pair of nodes or abstract nodes. An EIRS may be present when an EXRS is also present in the ERO and/ or an XRO is also present in the path message. Section 2.3 discusses details of processing of the EIRS with the XRO object and the EXRS subobject of ERO.

If the processing node does not understand the EIRS subobject, it behaves as described in [RFC3209] when an unrecognized ERO subobject is encountered. This means that 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 EIRS subobject.

If the EIRS.L bit is not set, the processing node SHOULD generate a Path Error with error code "Routing Problem" and error subcode "Bad EXPLICIT_ROUTE object".

If the processing node understands the EIRS subobject and all the subobjects contained in the EIRS, it takes the following steps:

- . For all subobjects contained in the EIRS such that EIRS.SubobjectN.L = 0, the processing node finds a path that MUST include the resource attribute identified by the EIRS.SubobjectN.
- . For all subobjects contained in the EIRS such that EIRS.SubobjectN.L = 1, the processing node finds a path that MUST include the resource attribute identified by the EIRS.SubobjectN.
- . If the processing node fails to find a route such that the all resources identified in the EIRS.SubobjectN for all N can be included in the route (depending on EIRS.SubobjectN.L bit setting), the node SHOULD return a PathErr with the error code "Routing Problem" and error value "Route Blocked by Include

Route". The error subcode "Route Blocked by Include Route" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 110).

If the processing node understands the EIRS subobject but does not understand or support a subobject contained in the EIRS (say EIRS.SubobjectN), it SHOULD 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 EIRS subobject containing the unsupported EIRS.subobjectN.

A node MAY reject a Path message if the EIRS is too large or complicated for the local implementation or as governed by local policy. In this case, the node SHOULD send a PathErr message with the error code "Routing Error" and error subcode "EIRS Too Complex". An ingress node receiving this error code/subcode combination MAY reduce the complexity of the EIRS. The error subcode "EIRS Too Complex" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 111).

2.3. Processing of EIRS with XRO and EXRS

A node performing ERO expansion MAY find an XRO in the Path message and both EIRS and EXRS subobjects in ERO. In this case, the processing node MUST include all resources identified in the EIRS and exclude all resources identified in the EXRS and XRO.

If the constraints identified by the EIRS, EXRS and XRO conflict each other, the processing node SHOULD send a PathErr message with the error code "Routing Error" and error subcode "inconsistent include/ exclude constraints". The error subcode "inconsistent include/ exclude constraints" for Path Error code "Routing Problem" is to be assigned by IANA (Suggested Value: 112).

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

This document adds the following new subobject of the existing entry for ERO (20, EXPLICIT_ROUTE):

Value	Description
-----	-----
TBA (suggest value: 68)	Explicit Inclusion Route Subobject (EIRS)

These subobject may be present in the Explicit Route Object, but not in the Route Record Object.

5. Acknowledgments

Authors would like to thank Matt Hartley, Gabriele Maria Galimberti, 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.
- [RFC 3473] Berger, L., Ed., "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.
- [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.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[DRAFT-LSP-XRO-SUB] Ali, Z., Swallow, G., Filsfils, C., et al,
"Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) LS

P

Route Diversity using Exclude Routes", draft-ali-ccamp-xro-lsp-subobject, work in progress.

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, Inc.
Email: zali@cisco.com

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

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

Ori Gerstel
Cisco Systems
ogerstel@cisco.com

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

Internet-Draft

draft-ali-ccamp-rsvp-te-include-route-02.txt

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								
+-----+																																							

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
Internet-Draft
Updates: 3209 (if approved)
Intended status: Informational
Expires: November 24, 2012

S. Balls, Ed.
J. Hardwick
Metaswitch
C. Margaria
Nokia Siemens Networks
May 23, 2012

Relaxing RSVP Loop Checking
draft-balls-ccamp-relax-loop-check-01

Abstract

This specification relaxes the rules governing loop checking within RSVP. These were originally defined in [RFC3209] and are too strict for the requirements of today's data planes.

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 November 24, 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.

1. Introduction

Generalized MPLS (GMPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) are prohibited from passing through a single node more than once. Today's data planes are such that allowing certain loops is required for some LSPs to be possible.

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. General Overview

With today's data planes it is acceptable for a single data flow (LSP) to pass through a single control plane node on more than one occasion on the path from source to destination. Currently control plane protocols will prevent such a path being managed in the control plane as they explicitly prevent routing loops. It is desirable for such LSPs to be able to be managed in the same way as non-looping LSPs.

2.1. Example in WDM networks

In WDM networks it can be necessary to route the data via an additional box in order to fulfil regeneration or wavelength conversion requirements. For example, consider the following simple example.

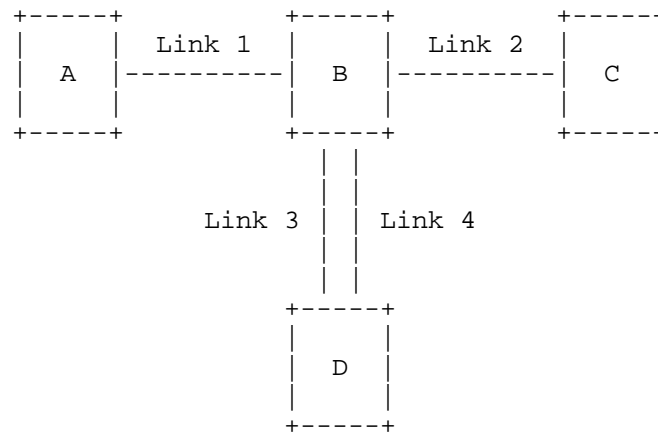


Figure 1

If node B cannot perform wavelength conversion but Link 1 and Link 2 do not have a common free wavelength then the only way to set up a path from node A to node C will be via node D. This requires two passes through node B which to RSVP looks like a loop.

2.2. Example using Connectivity Matrices

In any type of network a specific node may have connectivity restrictions that limit the output ports available given the input ports. For example, given the above network, where node B has the following connectivity restrictions.

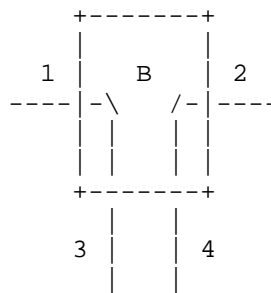


Figure 2

As in the above example, the only way to set up a path from node A to

node C will be via node D. This requires two passes through node B which to RSVP looks like a loop.

2.3. Example with additional label restrictions

Connections between ports on a node may be restricted based on labels. Consider the following network.

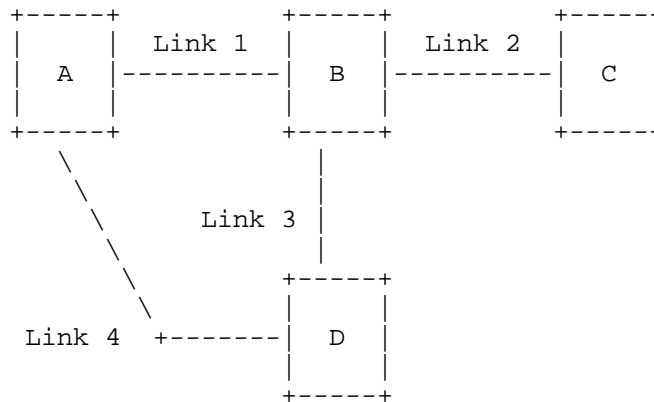


Figure 3

This network has the following properties.

- o Node A is electro-optical outputting Lambda 1 and can switch Lambda 2.
- o Node D can convert between Lambda 1 and Lambda 2.
- o Link 1 and Link 3 have Lambda 1 available.
- o All links have Lambda 2 available.

To setup a path from A to C in this network, the LSP must pass through Link 1 twice: once using Lambda 1 and once using Lambda 2. This results in the path A-B-D-A-B-C being taken which requires two passes through node A.

3. Existing workaround

In current networks it is possible to support such paths either

through management configuration at each node, or splitting the path into two or more signalling sessions. In the above examples this can be achieved with one session from A to D, and a second session from D to C. It would also require management on node D to join the data paths together. It is desirable that a single signalling session can be used to set up such paths, thus only requiring management input at the ingress.

4. Solution

4.1. Overview

To support such networks, the rules governing RSVP loop checking are relaxed. No changes to protocol messages are made.

4.2. Assumptions and limitations

These changes are only applicable to GMPLS out of band signalling when using point to point data links.

4.3. General Rules

The following rules govern the changes in behaviour that allow RSVP loop checking to be relaxed while still setting up non-looping data paths in RSVP.

- o For each pass through the node, the pair of inbound and outbound data interfaces and labels must be different.

Are there any further rules when are loops OK and when not OK?

4.4. RRO handling

Section 4.4.4 of [RFC3209] states that RSVP must reject a Path message if the receiving router is already in the RRO. This is now relaxed to allow such a condition provided a different interface-label pair is used in each case. If the router has existing session state for a received Path message, and it MUST verify that the newly requested data path (input and output interface and label) is different from the existing data path(s) for that session, and the existing data path(s) is (are) present earlier in the RRO. If this is not the case, the router MUST return a "Routing problem" PathErr message with the error value "loop detected".

In order to carry out this checking correctly, specific interfaces and labels SHOULD be recorded in the RRO. If this is not the case, each node can only verify the path is acceptable against local state

and should not reject the RRO if the local state is valid.

It is allowable for local policy to exist to limit the number of different paths through a router in a single LSP instance. If this limit is exceeded the router SHOULD return a "Routing problem" PathErr message with the error value "loop detected". This local policy is not intended to be advertised in routing. It is present as a backstop to protect against malicious Path messages consuming all resources on the router.

4.5. ERO handling

Sections 4.3.4.1 and 4.3.5 of [RFC3209] also state that RSVP must detect and avoid loops. This checking is also relaxed to allow loops in the cases stated above. Again, local policy can limit the number of different paths through a router in a single LSP instance. A router may "look ahead" in the ERO to determine such local policy will be exceeded in advance of it happening and SHOULD return a "Routing problem" PathErr message with the error value "loop detected" in such a case.

When calculating or expanding an ERO a router may include multiple entries through a single router. If the ERO contains loose hops that form a loop, and a node determines a non-looping route is available, it MAY remove the loop from the ERO.

4.6. Interface handling

As stated in the general rules, an implementation supporting multiple passes through a node must ensure that for each pass the input and output interfaces and labels are different.

Internally, this means that if a Path message is received using a different input interface this may no longer mean the LSP has been rerouted upstream. Implementations must check the RRO to determine the correct behaviour when processing such a Path message. Care must be taken to handle valid cases where the incoming label can change.

4.7. Signalling

For the avoidance of doubt, no new signalling is being defined in this draft.

The behaviour of refresh or error messages is unchanged and should therefore be sent along the looped path (if present). Nodes SHOULD NOT shortcut the loop.

4.8. Error Handling

How to behave when receiving a PathErr with error value "loop detected" is out of scope of this draft and is a local implementation decision. For example, it may choose to try and recalculate the path mandating that the error node is avoided, or does not support looping.

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

In principle these changes to RSVP pose no security exposures over and above [RFC3209]. However, by allowing loops a single LSP can now consume multiple resources. As suggested local policy can limit the number of paths and thus the resource a single LSP can consume.

7. 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.

Authors' Addresses

Steve Balls (editor)
Metaswitch
100 Church Street
Enfield, EN2 6BQ
UK

Phone: +44 208 366 1177
Email: steve.balls@metaswitch.com

Jonathan Hardwick
Metaswitch
100 Church Street
Enfield, EN2 6BQ
UK

Phone: +44 208 366 1177
Email: jonathan.hardwick@metaswitch.com

Cyril Margaria
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: January 17, 2013

D. Ceccarelli, Ed.
F. Fondelli
Ericsson
S. Belotti
D. Papadimitriou, Ed.
Alcatel-Lucent
July 16, 2012

Multi layer implications in GMPLS controlled networks
draft-bcg-ccamp-gmpls-ml-implications-01

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 January 17, 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. Applicability Scenarios	3
3. Requirements	6
4. Evaluation	7
5. Missing information	12
6. IANA Considerations	13
7. Contributors	13
8. Acknowledgements	13
9. References	13
9.1. Normative References	13
9.2. Informative References	14
Authors' Addresses	14

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. 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.

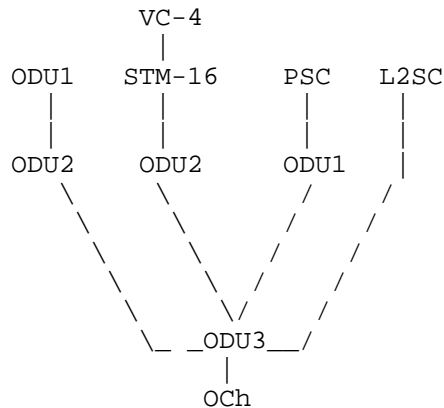


Figure 1: Muxing tree

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.

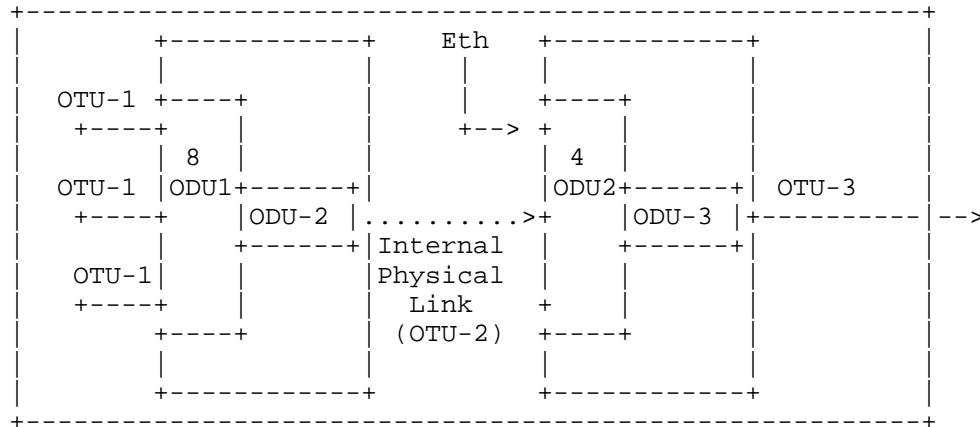


Figure 2: Cascaded muxponder

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 ODU2->ODU3 muxponder (through interfaces Y4 and Y5), which in turn is client to an OCh WSS.

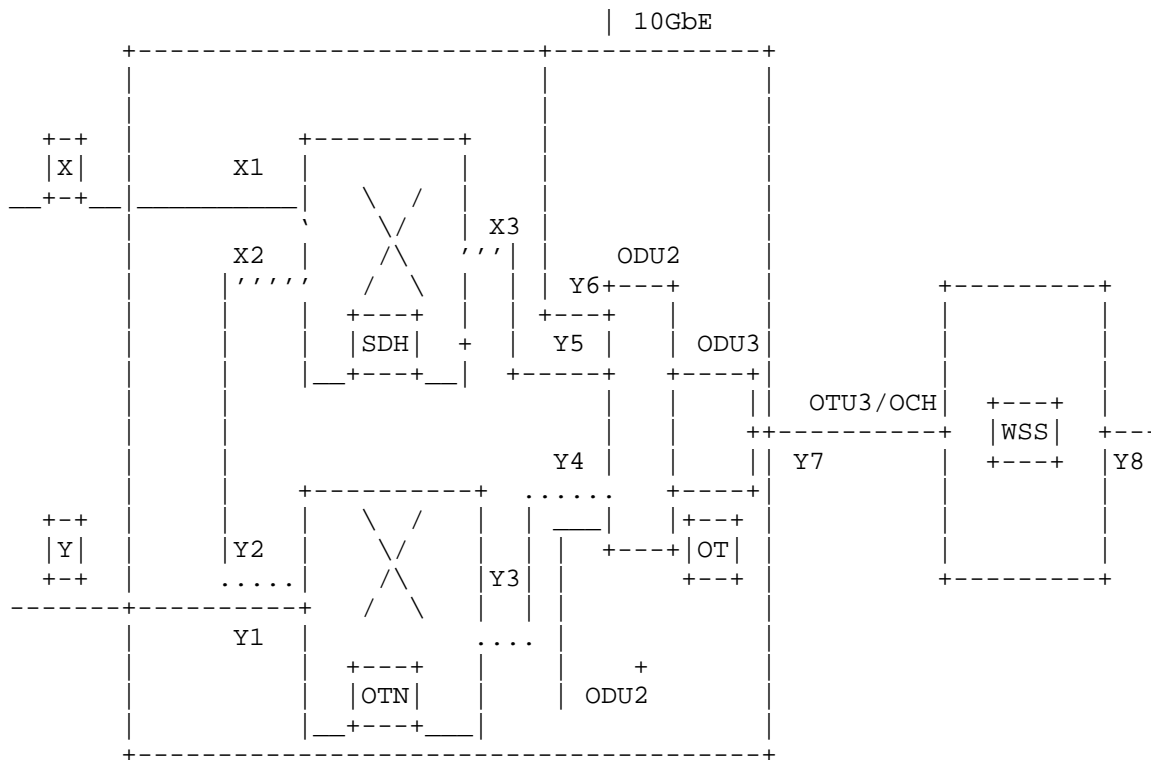


Figure 3: Hybrid node with optical muxponder and different switching matrices

3. Requirements

In order to deal with all the scenarios depicted in the previous sections, protocol extensions need to take into account the following set of requirements.

1. It must be possible to identify from which branch of X to which branch of Y the mapping is performed. Due to a restricted connectivity to a given switching layer, not all the indicated branches are really available. An example of such limitations can be seen in figure Figure 3, where for example the SDH client can be mapped only on interface Y5 of the muxponder board or the 10GbE 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.

4. 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 (associated to the adjustment pool) is set to the SC value of ISCD sub-TLV of the interface that gets access to the adjustment pool and the "upper" SC value of the IACD (associated to the adjustment pool) determines the SC capability of the resource pool. 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 (associated to the adjustment pool) is set to the SC value of ISCD sub-TLV of the interface(s) that gets access to the adjustment pool and the "upper" SC value of the IACD (associated to the adjustment pool) determines the SC capability of the adjustment pool itself, however, this "upper SC" value is not associated to any ISCD sub-TLV (compared to the single pool model where the "upper SC" value corresponds to at least one of the SC value associated to the ISCD sub-TLVs). In other terms, the (lower) SC value associated to each adjustment pool shall be covered by at least one of the SC associated to the ISCD sub-TLVs. 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.

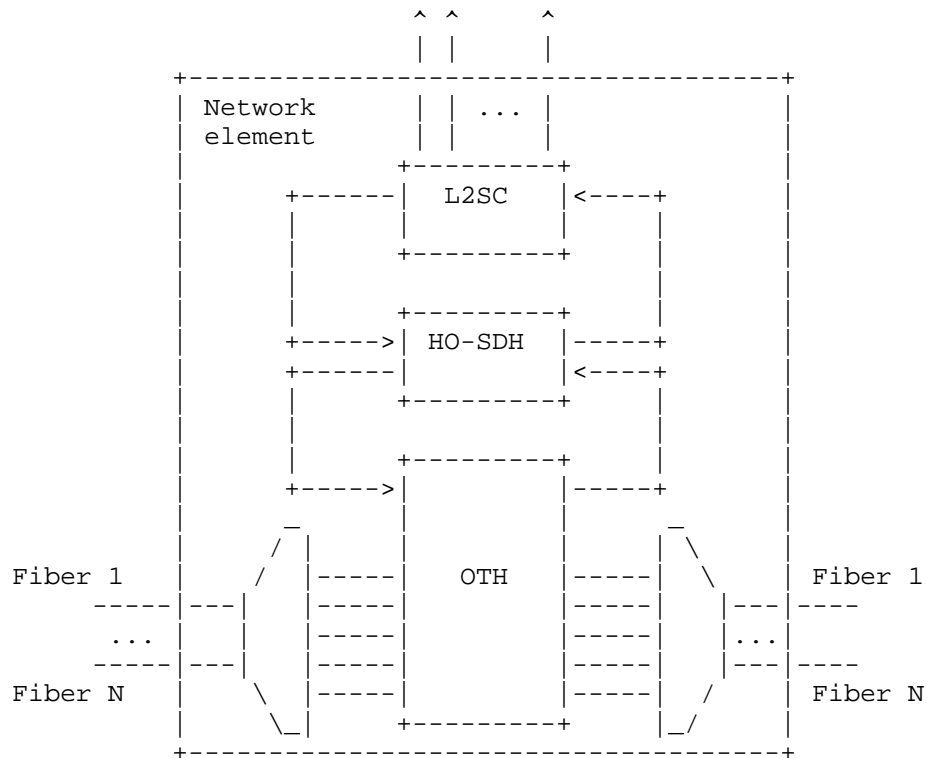


Figure 4: Example of single pool model

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.

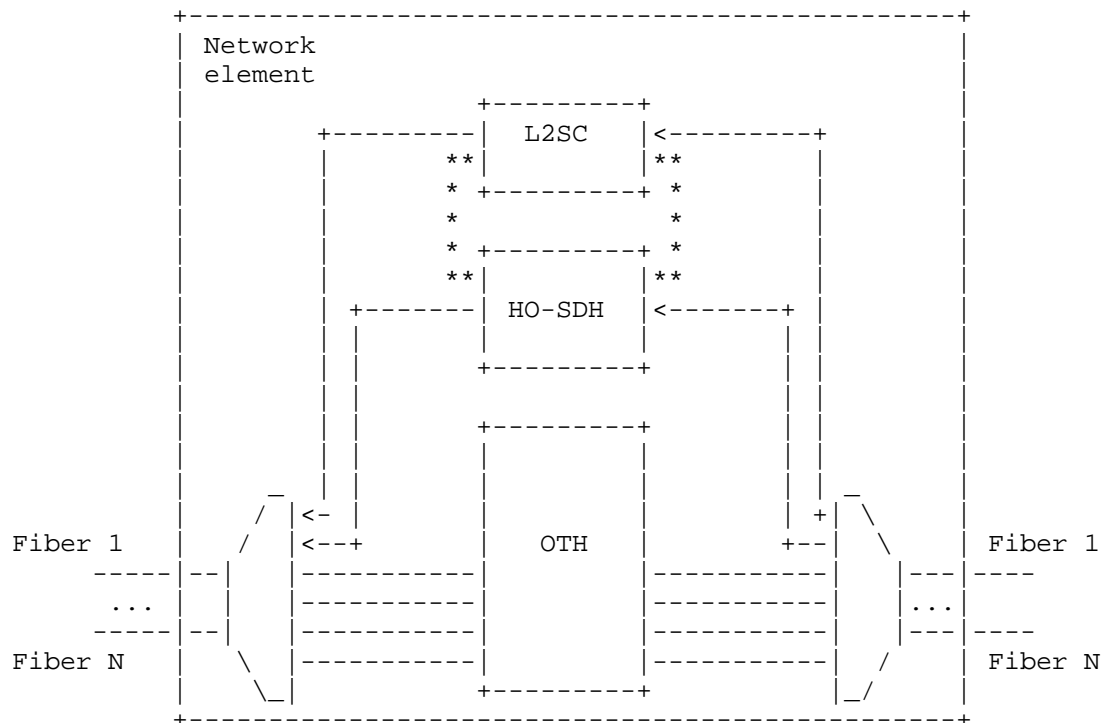


Figure 5: Example of multiple pool model - No interconnection between OTH and HO-SDH

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.

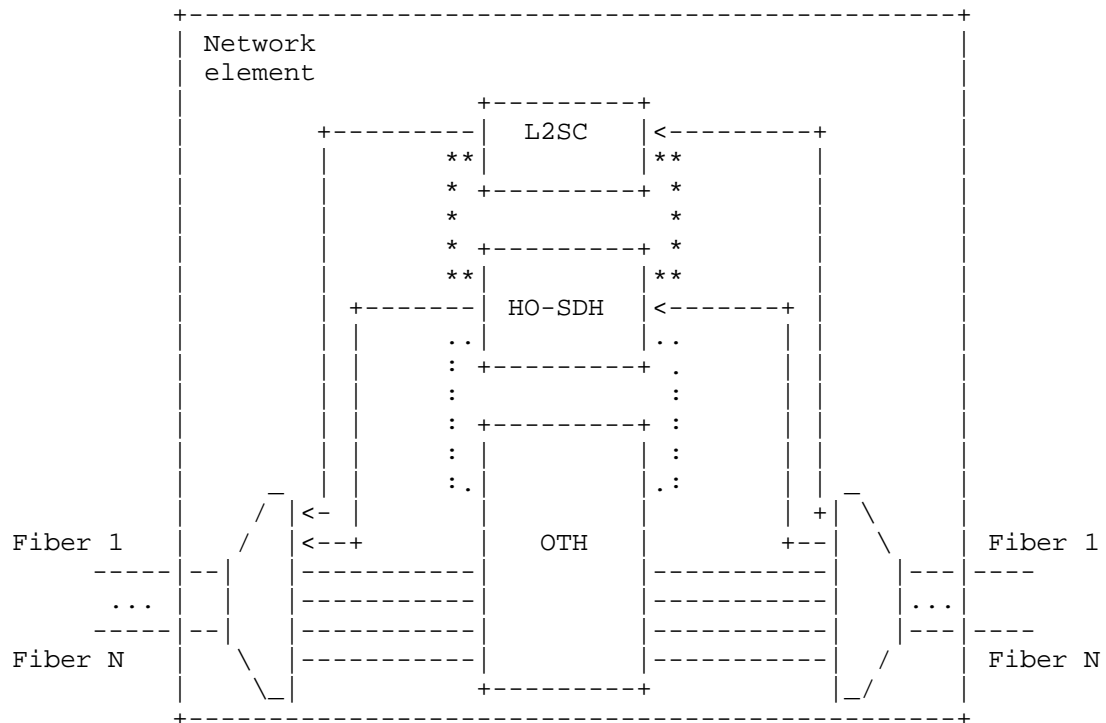


Figure 6: Example of multiple pool model - With interconnection between OTH and HO-SDH

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.

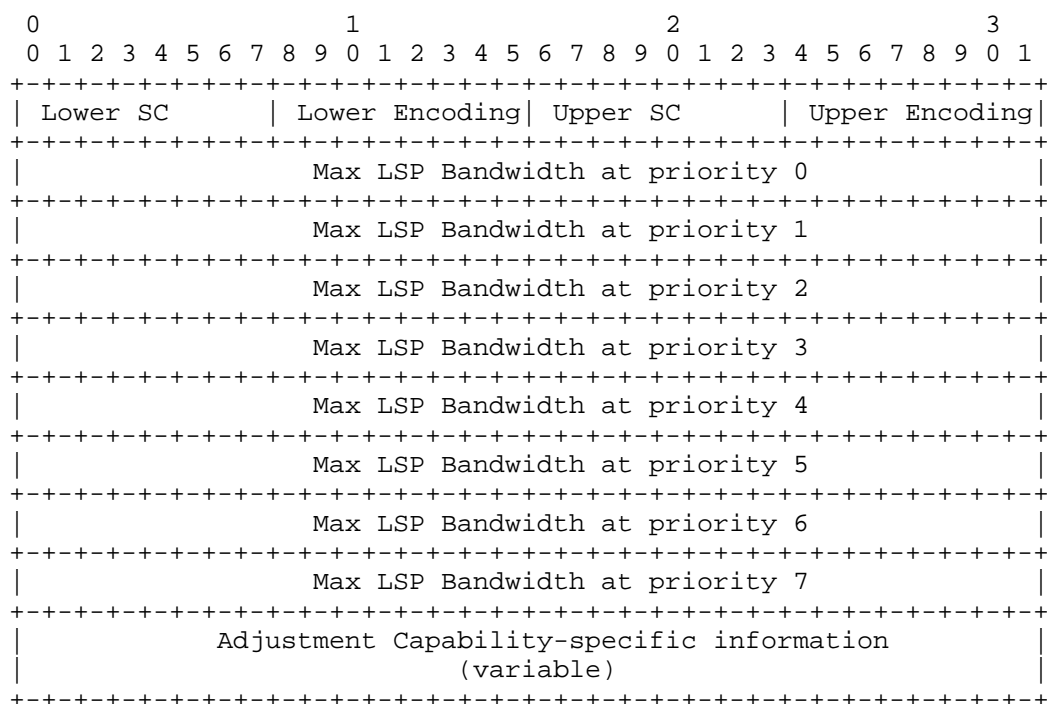


Figure 7: IACD format

5. Missing information

The pieces of information needed for addressing the requirements listed in Section 3 are:

- Adaptation information from a client to a server layer
- 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).

6. IANA Considerations

TBD

7. Contributors

TBD

8. Acknowledgements

TBD

9. References

9.1. 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 JL. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, October 2010.

9.2. 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.

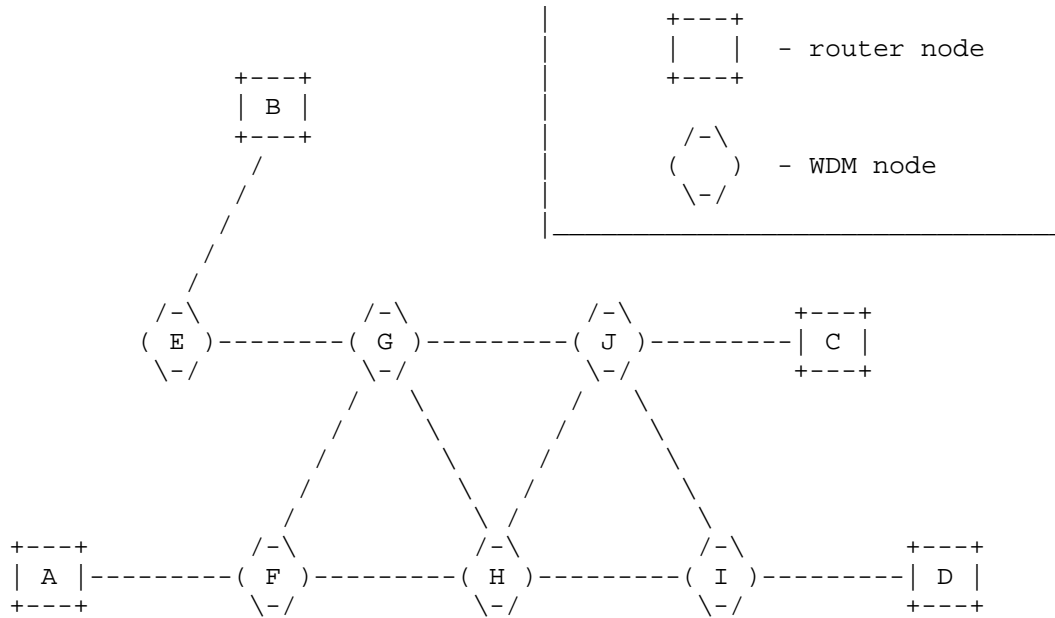


Figure 1: Sample hybrid topology

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.

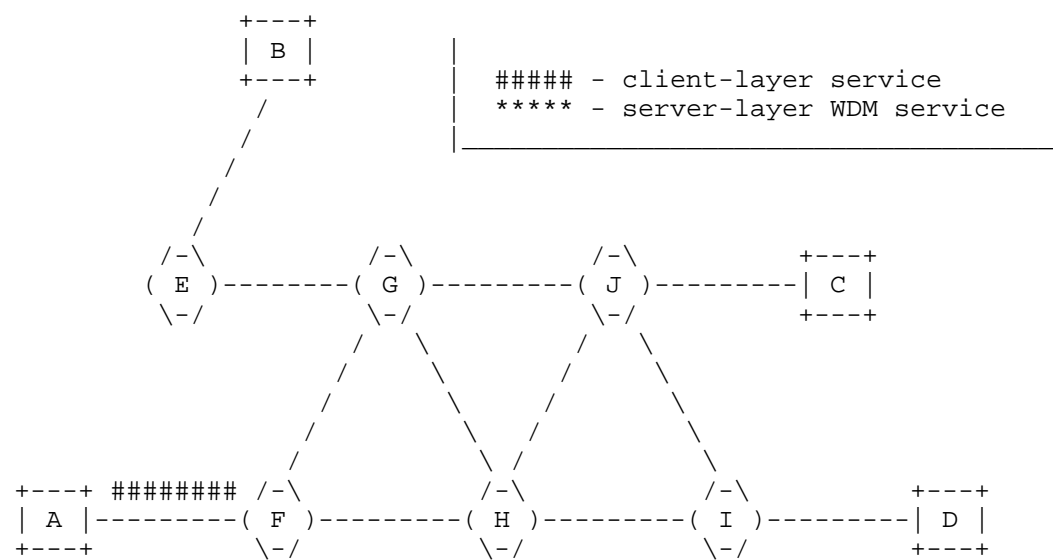


Figure 2a: Hierarchical service activation - Client-layer service setup is initiated.

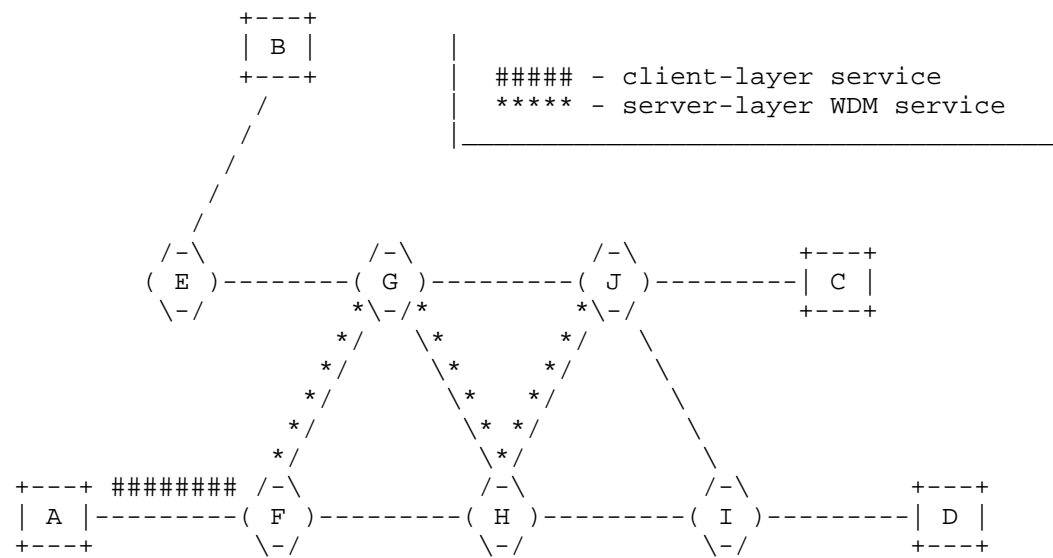


Figure 2b: Hierarchical service activation -
Server-layer WDM service that caters to the
client-layer service is established within the
core.

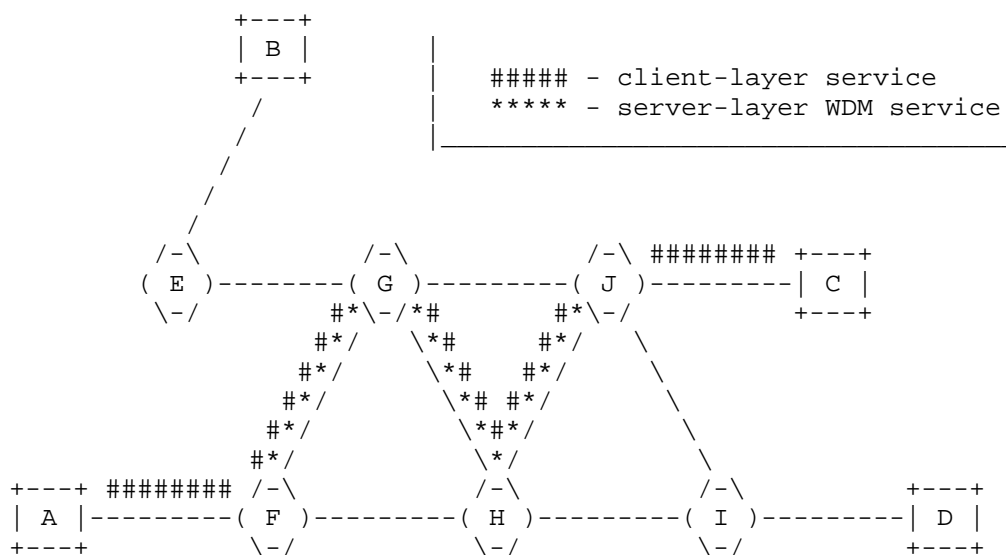


Figure 2c: Hierarchical service activation -
Client-layer service setup is resumed and
the end-to-end connection is established.

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.

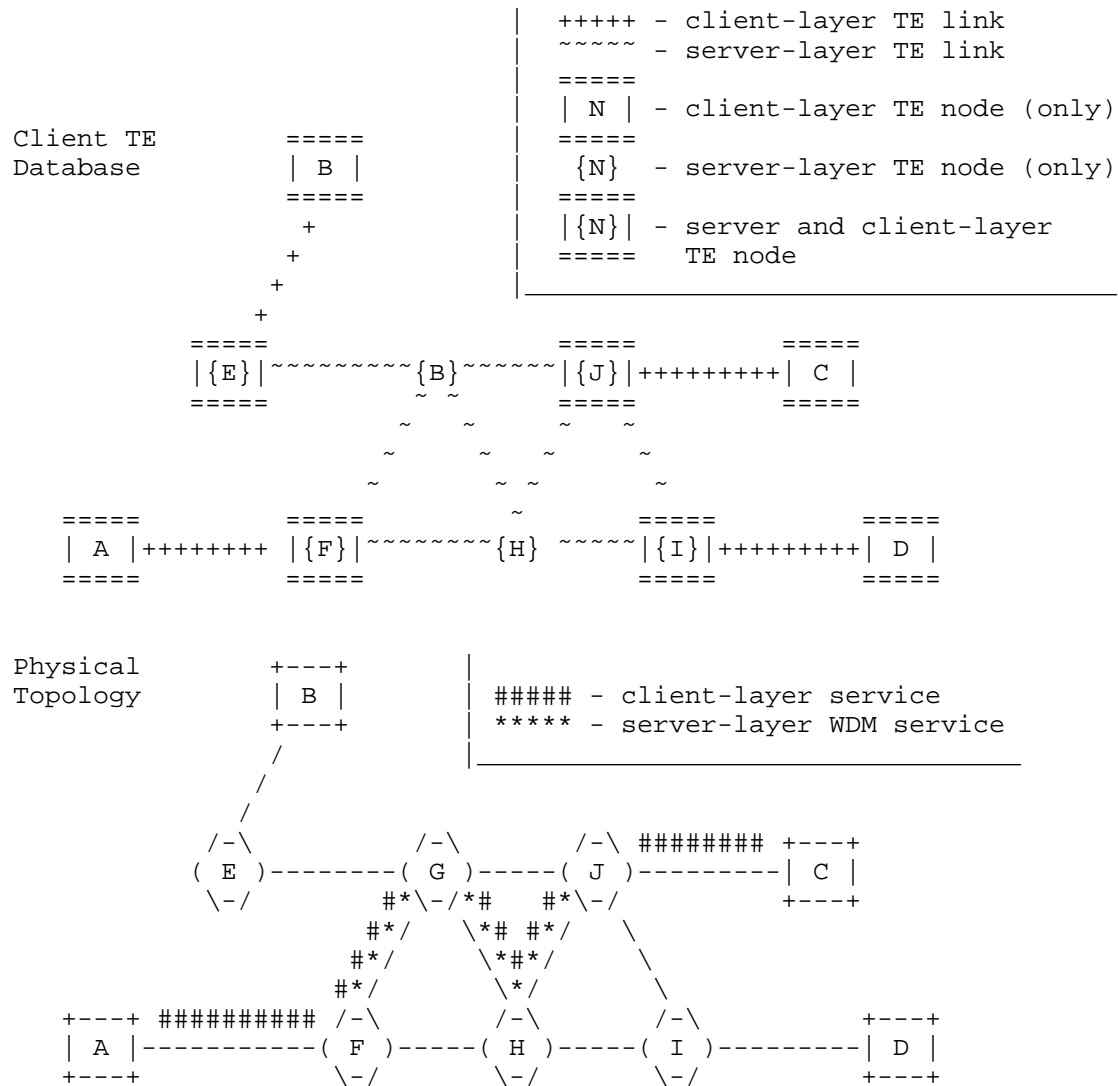


Figure 3: Traffic engineering - ERO with "loose hop"
[Path = {A,F,J,C} (with J loose)].

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.

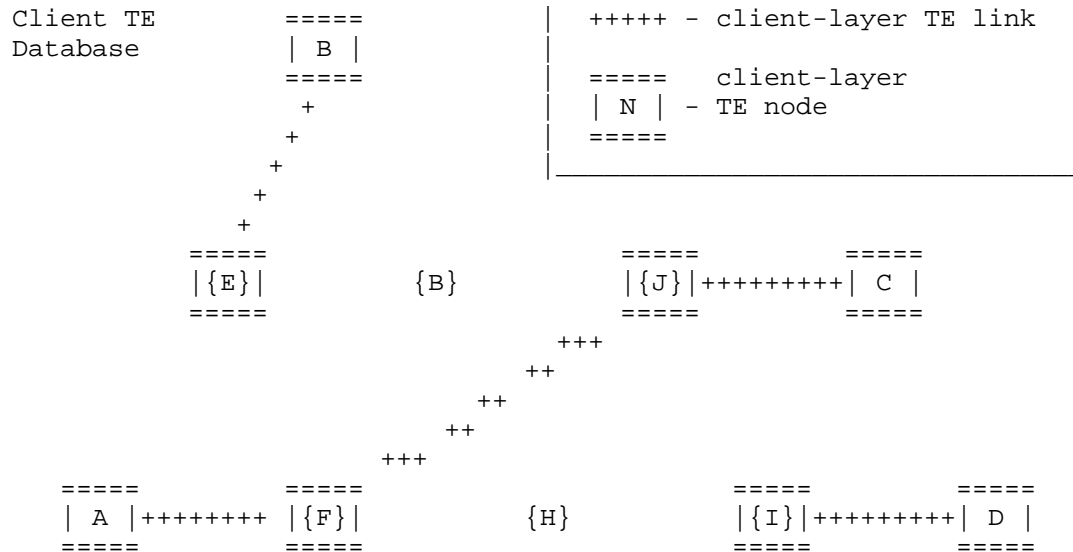
In order to be able to compute an end-to-end path between the two client layer endpoints, the client topology must be sufficiently augmented to indicate where there are paths through the server topology, which can provide connectivity between nodes in the client topology. In other words, in order for 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



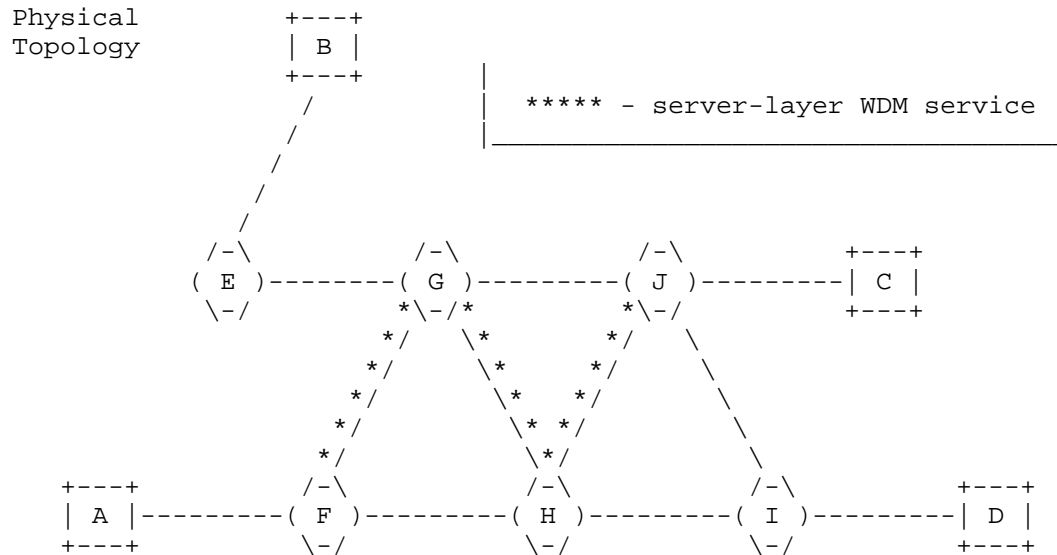


Figure 4: Traffic engineering - end-to-end path computation.[The client layer "TE link" between F and J is produced by creating the underlying server-layer connection; Node A has visibility to end-to-end (A to C) client-layer links and can do CSPF]

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.

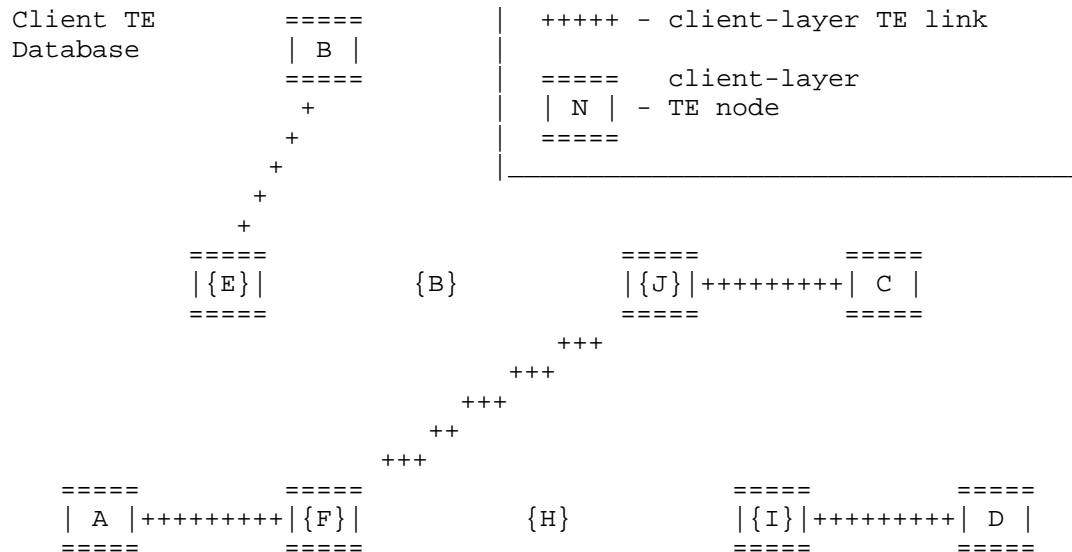
As another scenario, consider a network configuration, where the transponders on nodes E, F, J and I are connected to each other via directionless ROADM technology. In this case it is physically possible to connect any transponder to any other transponder in the server layer network. As there are transport 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.



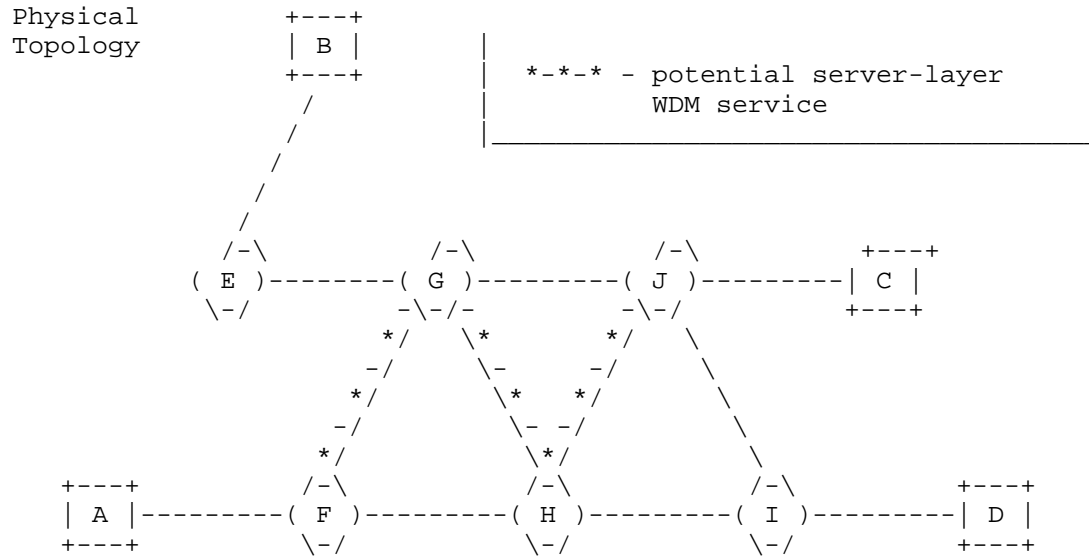


Figure 5: Traffic engineering - end-to-end path computation with "Virtual TE Links". [The "Virtual TE link" between F and J is created in the client layer without actually instantiating the underlying server-layer connection; Node A has visibility to end-to-end client-layer links and can do CSPF]

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.

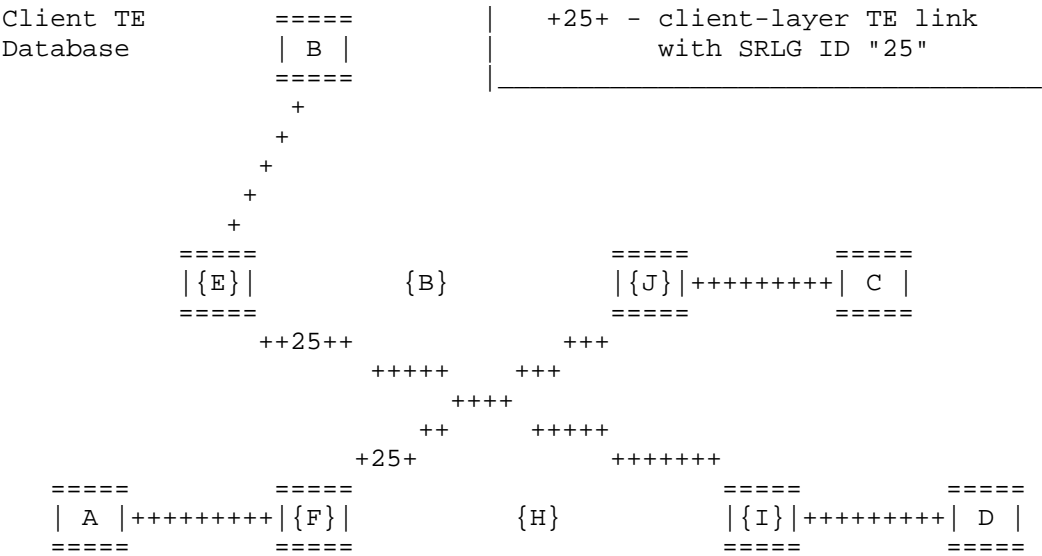
3.2. Macro SRLGs

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.



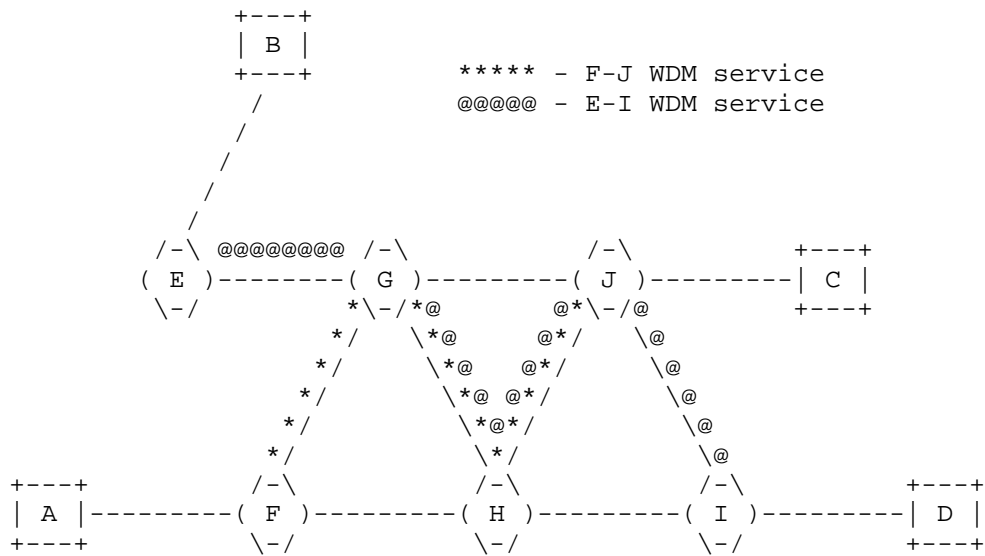


Figure 6: Macro SRLGs - ["TE links" E-I and F-J share fate since the underlying server-layer connections traverse the same path segments [G-H][H-I]. Macro SRLG-ID "25" is assigned to both "TE links"]

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
               +
               +
               +
               +
               +
               =====
               | {E} |      {B}      | {J} | ++++++ | C |
               =====
               ++25/192000++      +++++
               +++++      +++++
               +++++
               ++      ++++++
               +25/192000+      ++++++
               =====
               | A | ++++++ | {F} |      {H}      | {I} | ++++++ | D |
               =====

```

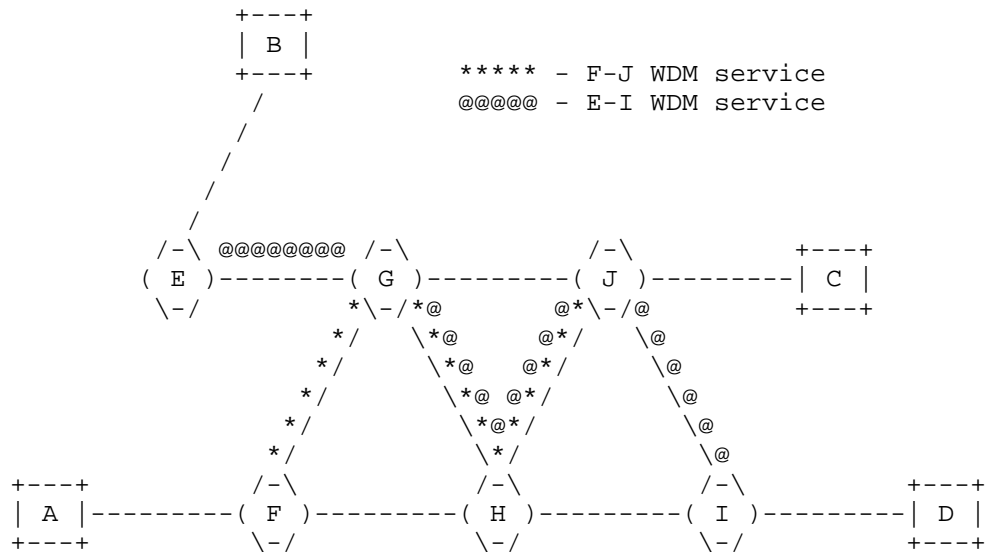


Figure 7: MELGs - ["TE links" E-I and F-J are mutually exclusive (server paths require usage of the same resource: lambda channel 192000). Same MELG ID is assigned to both TE links]

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.

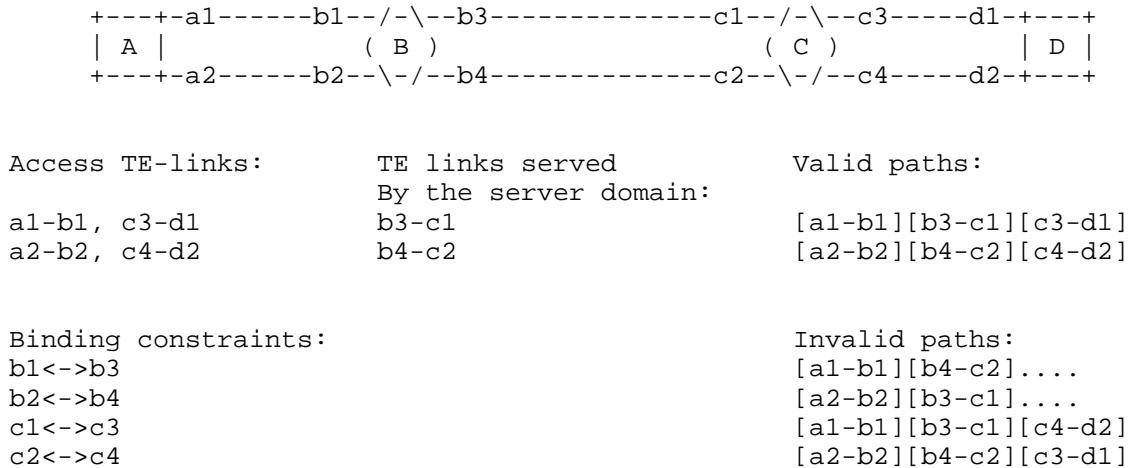


Figure 7: Switching Constraints

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.

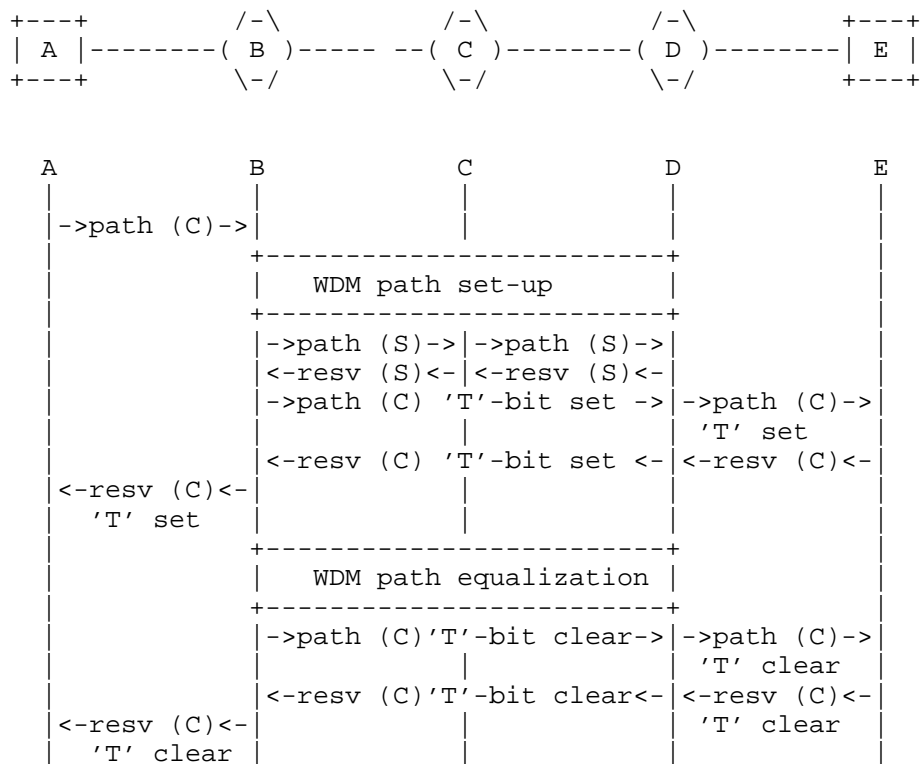


Figure 8: connection set-up

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

Network Working Group
Internet Draft
Intended status: Informational

Sergio Belotti
Don Fedyk
Dimitri Papadimitriou
Alcatel-Lucent

Expires: January 2013

July 9, 2012

Applicability Statement for Layer 1 Virtual Private Network (L1VPN)
Enhanced Mode
draft-belotti-app-statement-llvpn-em-00.txt

Abstract

This document provides an applicability statement on the use of Generalized Multiprotocol Label Switching (GMPLS) protocols and mechanisms to satisfy the requirements of the Layer 1 Virtual Private Network (L1VPN) Enhanced Mode.

L1VPNs provide customer services and connectivity at layer 1 over layer 1 networks. The operation of L1VPNs is divided into the Basic Mode and the Enhanced Mode, where the Enhanced Mode of operation may also include exchange of routing information between the layer 1 network and the customer domain.

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 (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.

1. Introduction.....	3
2. Conventions Used in This Document.....	4
2.1. Terminology.....	4
3. Existing Solutions.....	4
4. General Guidelines.....	4
5. Overlay Extension Service Model.....	6
5.1. Overview of the Service Model.....	6
5.2. Applicability of Existing Solutions.....	6
5.3. Incremental protocol extensions.....	7
6. Virtual Node Service Model.....	7
6.1. Overview of the Service Model.....	7
6.2. Applicability of Existing Solutions.....	7
7. Virtual Link Service Model.....	8
7.1. Overview of the Service Model.....	8
7.2. Applicability of Existing Solutions.....	8
8. Per-VPN Peer Service Model.....	8
8.1. Overview of the Service Model.....	8
8.2. Applicability of Existing Solutions.....	9
9. Manageability Considerations.....	9
10. Security Considerations.....	9
11. References.....	9
11.1. Normative References.....	9
11.2. Informative References.....	10
12. Acknowledgments.....	12
13. Author's Addresses	12

1. Introduction

This document provides an applicability statement on the use of existing Generalized Multiprotocol Label Switching (GMPLS) protocols and mechanisms to the Layer 1 Virtual Private Network (L1VPN) Enhanced Mode.

In particular, this document shows section by section (from Section 5 to 8) the applicability of GMPLS protocols and mechanisms to each sub-model of the Enhanced Mode mentioned in [RFC4847].

Note that discussion in this document is limited to areas where GMPLS protocols and mechanisms are relevant.

As will be described in this document, support of the Overlay Extension service model and the Virtual Node service model are well covered by existing protocol mechanisms already described in other documents, with only minor protocol extensions required. The Virtual Link service model and the Per-VPN Peer service model are not

Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al July 2012
explicitly covered by existing documents, and would require
extension of current GMPLS protocols and mechanisms

Solutions should be scalable and manageable per RFC 4847. Solutions
should not require L1VPN state to be maintained on the P devices as
much as possible.

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. Terminology

The reader is assumed to be familiar with the terminology in
[RFC3031], [RFC3209], [RFC3471], [RFC3473], [RFC4202], [RFC4026] and
[RFC4847], [RFC5251], [RFC5253], [RFC5252] and [RFC4208].

4. Existing Solutions

This section lists existing solution documents that describe how the
L1VPN Enhanced Mode may be constructed using the mechanisms of GMPLS.
This document draws on those solutions and explains their
applicability with respect to the framework described in [RFC4847].
Further solution documents may be listed in a future version of this
document.

- [RFC5251] addresses L1VPN Basic Mode signaling.
- [RFC4208] addresses the application of GMPLS to the Overlay model.
- [RFC5252] describes OSPF based autodiscovery mechanisms.

Note that although [RFC5251] specifies signaling mechanisms for L1VPN
Basic Mode, it is applicable to the L1VPN Enhanced Mode, unless
otherwise specified.

5. General Guidelines

This section provides general guidelines for L1VPN solutions. Note
that applicability to specific sub-models will be separately

Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al July 2012

described in following sections. One important general design guideline is that protocol mechanisms should be re-used where possible. This means that solutions should be incremental, building on existing protocol mechanisms rather than developing wholly new protocols. Further, as service models are extended or developed resulting in the requirement for additional functionalities, deltas should be added to the protocol mechanisms rather than developing new techniques. [RFC4847] describes how the service models can be seen to provide "cascaded" functionality, and this should be leveraged to achieve re-use of protocol extensions so that, for example, it is highly desirable that the same signaling protocols and extensions are used in both the Basic Mode and the Enhanced Mode.

In addition, the following are general guidelines:

- The support of L1VPNs should not necessitate any change to core (P) devices. Therefore, any protocol extensions made to facilitate L1VPNs need to be made in a backward compatible way allowing GMPLS aware P devices to continue to function.
- Customer (C) devices not directly involved in providing L1VPN
- Services should also be protected from protocol extensions made to support L1VPNs. Again, such protocol extensions need to be backwards compatible. Note however, that some L1VPN service models allow for VPN connectivity between C devices rather than between CE devices: in this case, the C devices may need to be aware of protocol extensions.
- Solutions should aim to minimize the protocol extensions on CE devices.
- Solutions should be scalable and manageable per RFC 4847. Solutions should not require L1VPN state to be maintained on the P devices as much as possible.
- Solutions should be secure. Providers should be able to screen and protect information based on their operational policies per RFC 4847.
- Solutions should provide an operational view of the L1VPN for the customer and provider. There should be a common operational and management perspective in regard to other (L2 and L3) VPN services per RFC 4847.

6. Overlay Extension Service Model

6.1. Overview of the Service Model

This service model complements the Basic Mode and may assume all of the requirements, solutions and work items for that model.

In this service model, a CE receives from its attached PEs a list of TE link addresses to which it can request a VPN connection (i.e., membership information).

The CE may also receive some TE information concerning these CE-PE links within the VPN (e.g., switching type).

Further information may be found in [draft-fedyk-ccamp-llvpn-extnd-overlay].

The CE does not receive any of the following from the PE:

- Routing information about the core provider network.
- Information about P device addresses.
- Information about P-P, PE-P or PE-PE TE links.
- Routing information about other customer sites. The CE may have access to routing information about the remainder of the VPN (C-C and CE-C links), but this is exchanged by control plane tunneling on the CE-CE connections and is not passed to the CE in the control plane exchange between PE and CE.

6.2. Applicability of Existing Solutions

The following are required in this service model (in addition to requirements in the L1VPN Basic Mode):

- Interactions between an edge node (CE) and it's adjacent (at the data plane) core-node (PE).
- VPN membership information exchange between a CE and PE.
- CE-PE TE link information exchange between a CE and a PE.

RFC 4208 addresses RSVP-TE procedures between an edge-node and a core-node in the overlay model. RFC5252 enables PE devices using OSPF to dynamically learn about the existence of each other, and attributes of configured CE links and their association with L1VPNs.

Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al July 2012
Furthermore, [RFC5252] allows the exchange of CE-PE TE link information between a CE and a PE.

6.3. Incremental protocol extensions

It can be useful for the ingress node to be able to convey TE metrics (e.g., IGP metric, TE metric, hop counts, latency, etc.) that the path computation algorithm (at the remote node performing route computation or expansion) can optimize for. Similarly, it can be useful for the ingress node to be able to indicate a TE metric bound for the loose segment being expanded by the remote node, (e.g., [DRAFT-TE-OBJ-FUNC-METRIC-BOUND]).

In a similar manner, as described in [DRAFT-TE-METRIC-RECORD], there are 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 for discovering the Shared Risk Link Groups (SRLGs) associated with the route taken by an LSP (e.g., [DRAFT-SRLG-RECORDING]).

It is also possible to improve route diversity for single-homed and dual-homed customer LSPs, which is a common requirement. This may be achieved via signaling extensions that provide shared constraint information for path diversity. Specifically, mechanisms that enable communication to the node computing/expanding the LSP signaled, information to exclude the route taken by a particular LSP or the route taken by all LSPs belonging to a single tunnel (e.g., [DRAFT-EXTENDED-OVERLAY]).

7. Virtual Node Service Model

7.1. Overview of the Service Model

In this service model, there is a private routing exchange between the CE and the PE, or to be more precise between the CE routing protocol instance and the VPN routing protocol instance running on the PE. The provider network is considered as one private node from the customer's perspective. The routing information exchanged between the CE and the PE includes CE-PE TE link information, customer network (i.e., remote CE sites), and may include TE links (Forwarding Adjacencies) connecting CEs (or Cs) across the provider network as well as control plane topology information from the customer network (i.e., CE sites).

7.2. Applicability of Existing Solutions

The following are required in this service model:

- VPN routing
- CE-CE Label Switching Path (LSP) setup, deletion, and modification signaling.

It is possible to use IGP-based auto-discovery (based on [RFC5252]).

Signaling mechanisms are covered by [RFC5251].

8. Virtual Link Service Model

8.1. Overview of the Service Model

In this service model, virtual links are established between PEs. A virtual link is assigned to each VPN and disclosed to the corresponding CEs. The routing information exchanged between the CE and the PE includes CE-PE TE links, customer network (i.e., remote CE sites), virtual links (i.e., PE-PE links) assigned to each VPN, and may include CE-CE (or C-C) Forwarding Adjacencies as well as control plane topology from the customer network (i.e., CE sites).

NOTE - Resource management for a dedicated data plane is a mandatory requirement for the Virtual Link service model. This could be realized by assigning pre-configured FA-LSPs to each VPN routing protocol instance (no protocol extensions needed) in order to instantiate the necessary FAs.

8.2. Applicability of Existing Solutions

Currently, there is no solution document for this type of service model.

9. Per-VPN Peer Service Model

9.1. Overview of the Service Model

In this service model, the provider partitions TE links within the provider network per VPN. The routing information exchanged between the CE and the PE includes CE-PE TE links, customer network (i.e., remote CE sites), as well as partitioned portions of the provider network, and may include CE-CE (or C-C) Forwarding Adjacencies and control plane topology from customer network (i.e., CE sites). Note that PEs may abstract routing information about the provider network and advertise it to CEs.

Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al July 2012
Note scalability must be carefully considered for advertising
provider network routing information to the CE [RFC4726].

9.2. Applicability of Existing Solutions

Currently, there is no solution document for this type of service model.

10. Manageability Considerations

Section 11 of [RFC4847] describes manageability considerations for L1VPNs.

This document defines a following new manageability requirement specific for the L1VPN Enhanced Mode.

MIB modules MUST be available for any protocol extensions for the L1VPN Enhanced Mode.

A future revision of this document may cover more aspects.

11. Security Considerations

Section 12 of [RFC4847] describes security considerations for L1VPNs. This document defines a following new security requirements specific for the L1VPN Enhanced Mode.

In the L1VPN Enhanced Mode, since there is a routing adjacency between a CE and a PE, care must be taken whether the provider network's control plane topology information is leaked to the CE. Due to security concerns, this is not recommended in general, and there must be a mechanism to prevent such operation. A future revision of this document may cover more aspects.

12. IANA Considerations

This document has no actions for IANA.

13. References

13.1. Normative References

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

- Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al. July 2012
- [RFC4208] Swallow, G., et al., "Generalize Multiprotocol Label Switching (GMPLS) User-Network Interface: Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model," RFC 4208, October 2005.
- [RFC4847] Takeda, T., Editor "Framework and Requirements Layer 1 Virtual Private Networks", RFC 4847,
- [RFC5251] Fedyk, D., et al., "Layer 1 VPN Basic Mode", RFC5251, July 2008.
- [RFC5252] Bryskin, I., Berger, L., "OSPF-Based Layer 1 VPN Auto-Discovery", RFC 5252, July 2008.
- [RFC5253] Takeda, T. et al., "Applicability Statement for Layer 1 Virtual Private Network (L1VPN) Basic Mode", RFC 5253, July 2008.

13.2. Informative References

- [Y.1312] Y.1312 - Layer 1 Virtual Private Network Generic requirements and architecture elements, ITU-T Recommendation, September 2003.
- [Y.1313] Y.1313 - Layer 1 Virtual Private Network service and network architectures, ITU-T Recommendation, July 2004.
- [RFC3031] Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol label switching Architecture", RFC 3031, January 2001.
- [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., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., Editor "Generalized Multi-Protocol Label Switching (GMPLS) Signaling - Resource

Internet-Draft Applicability of L1VPN Enhanced Mode Belotti et al. July 2012
ReserVation Protocol-Traffic Engineering (RSVP-TE)
Extensions", RFC 3473, January 2003.

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

[RFC4026] Andersson, L., and Madsen, T., "Provider
Provisioned Virtual Private Network (VPN)
Terminology", RFC 4026, March 2005.

[RFC4726] Farrel, A., et al., "A Framework for Inter-Domain
Multiprotocol Label Switching
Traffic Engineering", RFC 4726, November 2006.

[DRAFT-TE-OBJ-FUNC-METRIC-BOUND]
Ali Z., et al., Resource ReserVation Protocol-
Traffic Engineering (RSVP-TE) extension for signaling
Objective Function and Metric Bound draft-ali-ccamp-
rc-objective-function-metric-bound-01, work in
progress

[DRAFT-SRLG-RECORDING]
F. Zhang, D. Li, O. Gonzalez de Dios, C.
Margaria. C, " RSVP-TE Extensions for Collecting SRLG
Information ", draft-zhang-ccamp-srlg-fa-
configuration-05, work in progress.

[DRAFT-TE-METRIC-RECORD]
ALI Z., et al., "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, work in progress

[DRAFT-EXTENDED-OVERLAY]
Fedyk D., et al., "Layer 1 VPN Enhanced Mode -
Overlay Extension Service Model", draft-fedyk-ccamp-
l1vpn-extnd-overlay-00, work in progress

14. Acknowledgments

The authors would like to thank the editor and authors of draft-ietf-llvpn-applicability-enhanced-mode (Tomonori Takeda, Hamid Ould-Brahim, Adrian Farrel, Deborah Brungard, Dimitri Papadimitriou) and the members of the L1VPN working group for their contribution, in recognition that most of the text in this draft derives from their effort.

The authors would also like to thank Dieter Beller, Lieven Levrau, and Eve Varma for their contributions to this work.

15. Author's Addresses

Sergio Belotti (Alcatel-Lucent)
VIA TRENTO 30
20059 Vimercate
Italy
Phone: +39 039 686 3033
sergio.belotti@alcatel-lucent.com

Don Fedyk (Alcatel-Lucent)
Groton, MA 01450
United States
Phone: +1 978 467 5645
Email: Donald.Fedyk@alcatel-lucent.com

Dimitri Papadimitriou (Alcatel-Lucent)
Copernicuslaan 50
2018 Antwerp
Belgium
Phone: +32 3 2408491
Email: dimitri.papadimitriou@alcatel-lucent.com

Internet Draft
Updates: 3471, 4202, 4203, 5307
Category: Standards Track
Expiration Date: January 4, 2013

Lou Berger (LabN)
Julien Meuric (France Telecom)

July 4, 2012

Revised Definition of The GMPLS Switching Capability and Type Fields

draft-berger-ccamp-swcaps-update-02.txt

Abstract

GMPLS provides control for multiple switching technologies, and hierarchical switching within a technology. GMPLS routing and signaling use common values to indicate switching technology type. These values are carried in routing in the Switching Capability field, and in signaling in the Switching Type field. While the values using in these fields are the primary indicators of the technology and hierarchy level being controlled, the values are not consistently defined and used across the different technologies supported by GMPLS. This document is intended to resolve the inconsistent definition and use of the Switching Capability and Type fields by narrowly scoping the meaning and use of the fields. This document updates any document that uses the GMPLS Switching Capability and Types fields, in particular RFC 3471, RFC 4202, RFC 4203, and RFC 5307.

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/lid-abstracts.html>

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

This Internet-Draft will expire on January 4, 2013

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.

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) provides control for multiple switching technologies. It also supports hierarchical switching within a technology. The original GMPLS Architecture, per [RFC3945], included support for five types of switching capabilities. An additional type was also been defined in [RFC6002]. The switching types defined in these documents include:

1. Packet Switch Capable (PSC)
2. Layer-2 Switch Capable (L2SC)
3. Time-Division Multiplex Capable (TDM)
4. Lambda Switch Capable (LSC)
5. Fiber-Switch Capable (FSC)
6. Data Channel Switching Capable (DCSC)

Support for the original types was defined for routing in [RFC4202], [RFC4203] and [RFC5307], where the types were represented in the Switching Capability (Switching Cap) field. In general, hierarchy within a type is addressed in a type-specific fashion and a single Switching Capability field value is defined per type. The exception to this is PSC which was assigned four values to indicate four levels of hierarchy: PSC-1, PSC-2, PSC-3 and PSC-4. The same values used in routing are defined for signaling in [RFC3471], and are carried in the Switching Type field. Following the IANA registry, we refer to the values used in the routing Switching Capability field and signaling Switching Type field as Switching Types.

In general, a Switching Type does not indicate a specific data plane technology, but rather this needs to be inferred from context. For example L2SC was defined to cover Ethernet and ATM, and TDM was defined to cover both SONET/SDH [RFC4606] and G.709 [RFC4328]. The basic assumption was that different technologies of the same type would never operate within the same control, i.e., signaling and routing, domains.

The past approach in assignment of Switching Types has proven to be

problematic from two perspectives. The first issue is that there are examples of switching technologies where there are different levels of switching that can be performed within the same technology. For example, there are multiple types of Ethernet switching that may occur within a provider network. The second issue is that the Switching Capability field value is used in routing to indicate the format of the Switching Capability-specific information (SCSI) field, and that an implicit mapping of type to SCSI format is impractical for implementations that support multiple switching technologies. These issues led to the introduction of two new types for Ethernet in [RFC6004] and [RFC6060], namely:

7. Ethernet Virtual Private Line (EVPL)
8. 802_1 PBB-TE

An additional value is also envisioned to be assigned in support of G.709v3 by [GMPLS-G709] in order to disambiguate the format of the SCSI field.

While a common representation of hierarchy levels within a switching technology certainly fits the design objectives of GMPLS, the definition of multiple PSC Switching Types has also proven to be of little value. Notably, there are no known uses of PSC-2, PSC-3 and PSC-4.

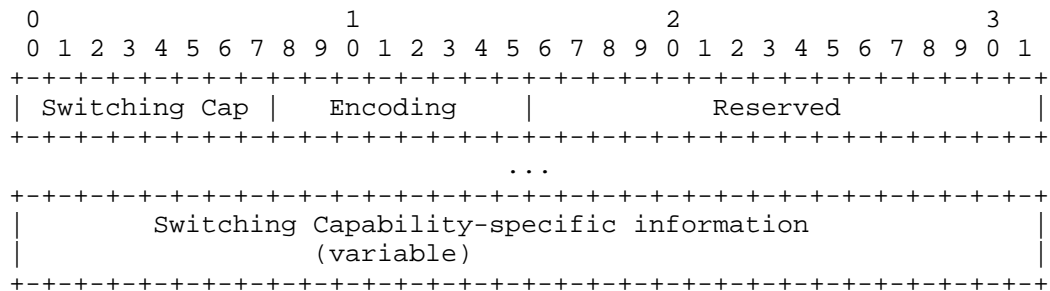
This document proposes to resolve such inconsistent definitions and uses of the Switching Types by reducing the scope of the related fields and narrowing their use. In particular this document proposes deprecating the use of the Switching Types as an identifier of hierarchy levels within a switching technology, and limit its use to identification of a per-switching technology SCSI field format.

This document updates any document that uses the GMPLS Switching Capability and Switching Type fields, in particular RFCs 3471, 4202, 4203, and 5307.

1.1. Current Switching Type Definition

The Switching Type values are carried in both routing and signaling. Values are identified in the IANA GMPLS Signaling Parameters Switching Type registry, which is currently located at <http://www.iana.org/assignments/gmpls-sig-parameters/gmpls-sig-parameters.xml>

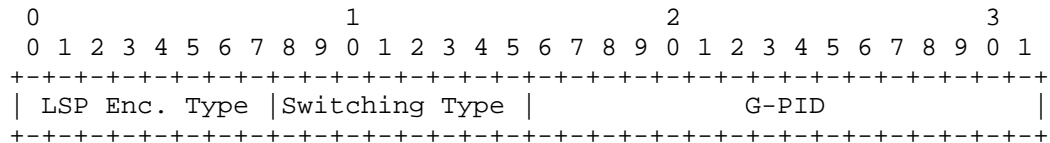
For routing, a common information element is defined to carry switching type values for both OSPF and IS-IS routing protocols in [RFC4202]. Per [RFC4202], switching type values are carried in a Switching Capability (Switching Cap) field in an Interface Switching Capability Descriptor. This information shares a common formatting in both OSPF, as defined by [RFC4203] and in IS-IS, as defined by [RFC5307]:



and

The content of the Switching Capability specific information field depends on the value of the Switching Capability field.

Similarly, the Switching Type field is defined as part of a common format for use by GMPLS signaling protocols in [RFC3471] and is used by [RFC3473]:



Switching Type: 8 bits

Indicates the type of switching that should be performed on a particular link. This field is needed for links that advertise more than one type of switching capability. This field should map to one of the values advertised for the corresponding link in the routing Switching Capability Descriptor ...

1.2. 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].

2. Revised Switching Type Definition

This document modifies the definition of Switching Type. The definitions are slightly different for routing and signaling and are described in the following sections.

2.1. Routing -- Switching Cap Field

For routing, i.e., [RFC4202], [RFC4203] and [RFC5307], the following definition should be used for Switching Cap field:

The Switching Cap field indicates the type of switching being advertised via GMPLS Switching Type values. A different Switching Type value SHOULD be used for each data plane technology even when those technologies share the same type of multiplexing or switching. For example, Time Division Multiplexing (TDM) technologies that have different multiplexing structures, such as SDH [G.707] and OTN [G.709], should use two different Switching Types.

As the format of the Switching Capability specific information field is dependent on the value of this field, a different Switching Type value MUST be used to differentiate between different Switching Capability specific information field formats.

This definition does not modify the format of the Interface Switching Capability Descriptor.

Note that from a practical standpoint, this means that any time a new switching technology might use a different Switching Capability specific information field format, that a new Switching Type SHOULD be used.

2.2. Signaling -- Switching Type Field

For signaling, i.e., [RFC3471] which is used by [RFC3473], the following definition should be used for Switching Type field:

Indicates the type of switching that should be performed on a particular link via GMPLS Switching Type values. This field maps to one of the values advertised for the corresponding link in the routing Switching Capability Descriptor, see [RFC4203] and [RFC5307].

Note that from a practical standpoint, there is no change in the definition of this field.

2.3. Assigned Switching Types

This document deprecates the following Switching Types:

Value	Name
2	Packet-Switch Capable-2 (PSC-2)
3	Packet-Switch Capable-3 (PSC-3)
4	Packet-Switch Capable-4 (PSC-4)

These values SHOULD NOT be treated as reserved values, i.e.,
SHOULD NOT be generated and SHOULD be ignored upon receipt.

3. Compatibility

For existing implementations, the primary impact of this document is deprecating the use of PSC-2, 3 and 4. At the time of publication, there are no known deployments (or even implementations) that make use of these values so there is no compatibility issues for current routing and signaling implementations.

4. Security Considerations

This document impacts the values carried in a single field in signaling and routing. As no new protocol formats or mechanisms are defined, there are no particular security implications raised by this document.

For a general discussion on MPLS and GMPLS related security issues, see the MPLS/GMPLS security framework [RFC5920].

5. IANA Considerations

IANA needs to deprecate and redefine the registry. In particular the Switching Types portion of the Generalized Multi-Protocol Label Switching (GMPLS) Signaling Parameters should be revised to read:

Switching Types

Registration Procedures

Standards Action

Reference

[RFC3471][RFC4328][This.draft]

Value	Name	Reference
0	Unassigned	
1	Packet-Switch Capable-1 (PSC-1)	[RFC3471]
2	Deprecated	[This.draft]
3	Deprecated	[This.draft]
4	Deprecated	[This.draft]
5-29	Unassigned	
30	Ethernet Virtual Private Line (EVPL)	[RFC6004]
31-39	Unassigned	
40	802_1 PBB-TE	[RFC6060]
41-50	Unassigned	

51	Layer-2 Switch Capable (L2SC)	[RFC3471]
52-99	Unassigned	
100	Time-Division-Multiplex Capable (TDM)	[RFC3471]
101-124	Unassigned	
125	Data Channel Switching Capable (DCSC)	[RFC6002]
126-149	Unassigned	
150	Lambda-Switch Capable (LSC)	[RFC3471]
151-199	Unassigned	
200	Fiber-Switch Capable (FSC)	[RFC3471]
201-255	Unassigned	

6. Acknowledgments

We thank John Drake for highlighting the current inconsistent definitions associated with the Switching Capability and Type Fields. Daniele Ceccarelli provided valuable feedback on this document.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "RFC Key Words Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC4202] Kompella, K., Rekhter, Y., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K., Rekhter, Y., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC5307] Kompella, K., Rekhter, Y., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, October 2008.

7.2. Informative References

- [G.707] ITU-T Recommendation G.707/Y.1322 (2007), "Network node interface for the synchronous digital hierarchy (SDH)".
- [G.709] ITU-T Recommendation G.709/Y.1331 (2009), "Interfaces for the Optical Transport Network (OTN)".

- [GMPLS-G709] Zhang, F., Li, D., Li, H., Belotti, S., Ceccarelli, D., "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", work in progress, draft-ietf-ccamp-gmpls-g709-framework.
- [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.
- [RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [RFC4606] Mannie, E., Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 4606, August 2006.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.
- [RFC6002] Berger, L., Fedyk, D., "Generalized MPLS (GMPLS) Data Channel Switching Capable (DCSC) and Channel Set Label Extensions", RFC 6002, October 2010.
- [RFC6004] Berger, L., Fedyk, D., "Generalized MPLS (GMPLS) Support for Metro Ethernet Forum and G.8011 Ethernet Service Switching", RFC 6004, front 2010.
- [RFC6060] Fedyk, D., Shah, H., Bitar, N., Takacs, A., "Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)", RFC 6060, March 2011.

8. Authors' Addresses

Lou Berger
LabN Consulting, L.L.C.
Phone: +1-301-468-9228
Email: lberger@labn.net

Julien Meuric
France Telecom
Research & Development
2, avenue Pierre Marzin
22307 Lannion Cedex - France
Phone: +33 2 96 05 28 28
Email: julien.meuric@orange-ftgroup.com

Generated on: Wed, Jul 04, 2012 12:19:48 PM

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 13, 2013

D. Ceccarelli
D. Caviglia
F. Fondelli
Ericsson
F. Zhang
D. Li
Huawei Technologies
D. Beller
Alcatel-Lucent
July 12, 2012

Link Management Protocol (LMP) Test Messages Extensions for Evolutive
Optical Transport Networks (OTN)
draft-ceccarelli-ccamp-gmpls-g709-lmp-test-04

Abstract

This document specifies Link Management Protocol (LMP) extensions for the support of enhanced Optical Transport Networks (OTN). In particular it updates LMP test messages detailing the ITU-T G.709 OTN technology specific information and extends them in order to cover also recently introduced signal types and containers defined by the ITU-T.

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. Terminology	3
2. Introduction	3
3. Verifying Link Connectivity	3
3.1. Encoding Type	4
3.2. Verify Transport Mechanism	5
3.3. Transmission Rate	6
4. Trace Monitoring	7
4.1. TRACE Object for evolutive OTN	7
4.2. Discovery Response Message for Layer Adjacency Discovery	8
5. LMP Behavior Negotiation update	9
6. Security Considerations	10
7. Acknowledgements	10
8. IANA Considerations	10
9. References	10
9.1. Normative References	10
9.2. Informative References	11
Authors' Addresses	11

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. Introduction

[RFC4204] defines the Link Management Protocol (LMP), which is a protocol of the Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] suite used to manage Traffic Engineering (TE) links. A TE link may be made by multiple physical resources interconnecting Label Switched Routers (LSRs), that are combined together for scalability reasons.

Current definition of LMP consists of two mandatory procedures:

- Control channel management: used to maintain control channels connectivity between adjacent LSRs. Such procedure is based on the exchange of a message called Config message followed by a lightweight keep-alive message exchange
- Link property correlation: used to combine multiple physical links into a single TE link.

and two optional procedures:

- Link verification: used to verify the connectivity of the physical links composing a TE link and to exchange their Interface_Ids
- Fault management: used to suppress alarms and locate failures. This feature may not be needed in G.709 networks because fault management mechanisms are provided by the G.709 architecture.

This document defines G.709 technology specific information needed when running LMP. In particular it is focused on 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.

3. 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.

Evolutive 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.

3.1. 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".

3.2. 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 Test message 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.

3.3. 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)

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.

4.2. 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 message attribute	TraceMonitor/LMP message field
<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.

5. 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

6. Security Considerations

TBD

7. Acknowledgements

The authors would like to thank Attila Takacs, Andras Kern, Sergio Belotti and Pietro Grandi for the kind review of the ID and the valuable comments provided.

8. IANA Considerations

A new C-Type value for the Trace Object Class (21) is TBA by IANA.

9. References

9.1. Normative References

- [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.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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.
- [RFC4204] Lang, J., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [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.

9.2. Informative References

[ITUT-G.709]

ITU-T, "Interface for the Optical Transport Network
(OTN)", G.709 Recommendation (and Amendment 1),
February 2001.

Authors' Addresses

Daniele Ceccarelli
Ericsson
Via E. Melen 77
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via E. Melen 77
Genova - Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Francesco Fondelli
Ericsson
Via Moruzzi 1
Pisa
Italy

Email: francesco.fondelli@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

Dieter Beller
Alcatel-Lucent
Lorenzstrasse 10
Stuttgart 70435
Germany

Email: dieter.beller@alcatel-lucent.com

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 10, 2013

D. Hiremagalur, Ed.
G. Grammel, Ed.
J. Drake, Ed.
Juniper
G. Galimberti, Ed.
Z. Ali, Ed.
Cisco
R. Kunze, Ed.
Deutsche Telekom
July 9, 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-00

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 January 10, 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.

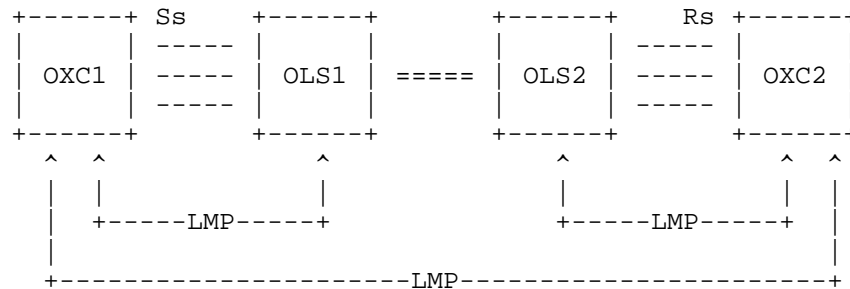
Table of Contents

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-00" and "draft-kunze-g-698-2-management-control-framework-02", for the relevant parameters specified in G.698.2 [ITU.G698.2].

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 Black Link:

- 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-snmplib-00". for more details about these parameters.

The general parameters are

1. Bit-Rate/line coding of optical tributary signals
2. Wavelength - (Hertz) 4 bytes
3. Min Wavelength Range - (Hertz) 4 bytes
4. Max Wavelength Range - (Hertz) 4 bytes
5. Channel Spacing -(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:

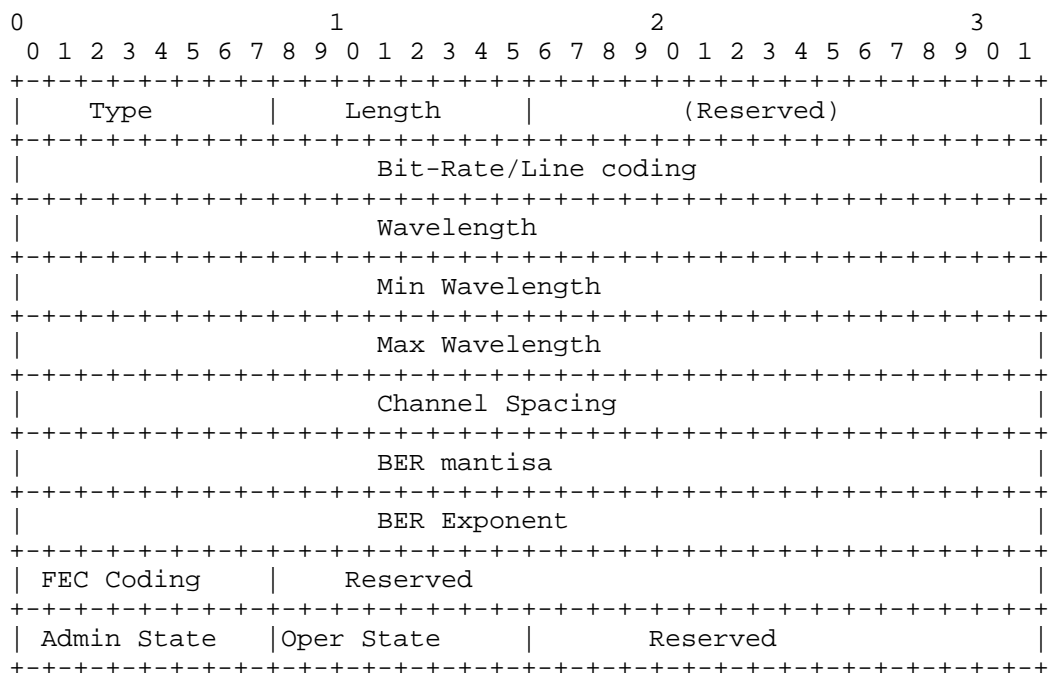


Figure 2: BL_General

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
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:

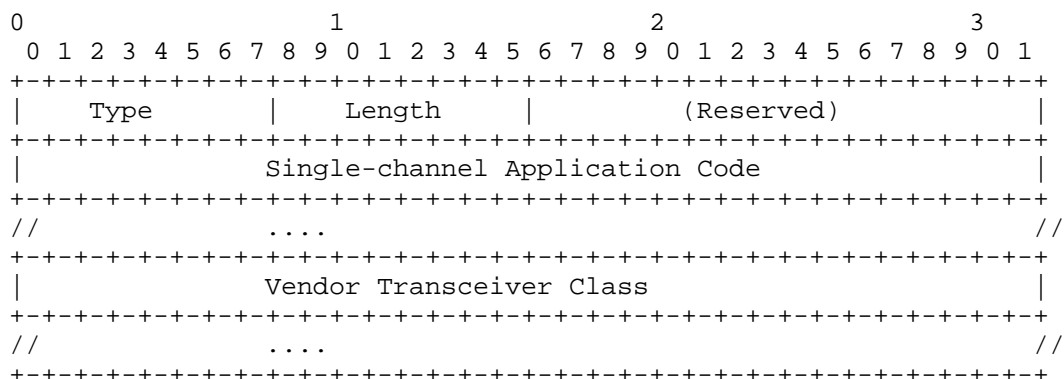


Figure 3: BL_ApplicationCode

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	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	2	3	4	5	6	7	8	9	0	1
Type										Length										(Reserved)																					
										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.

1. Wavelength
The wavelength which has been accepted by the OLS (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:

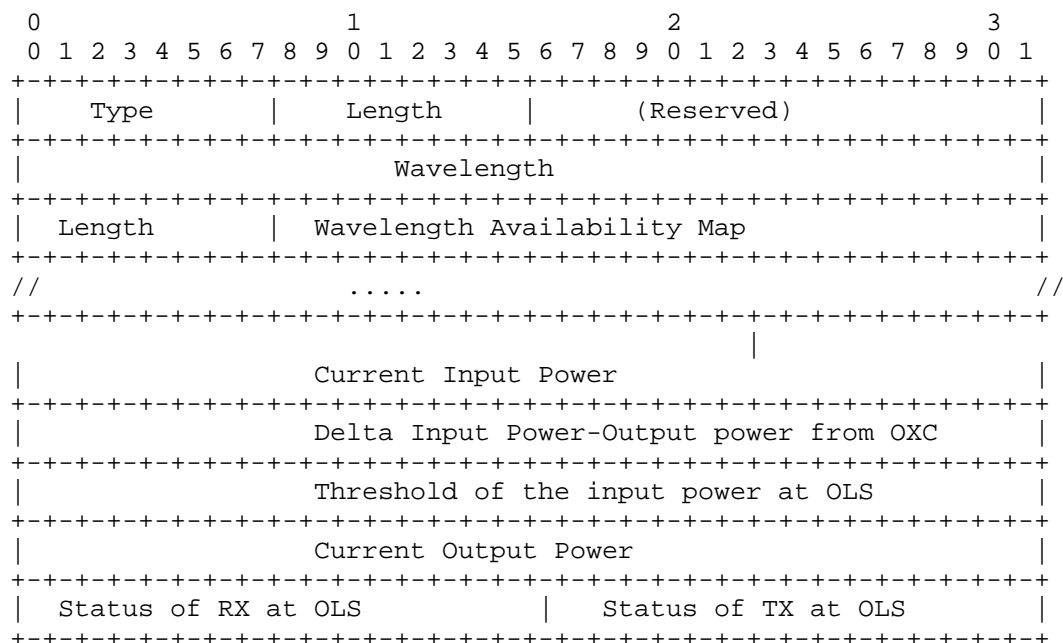


Figure 7: Black Link - OLS_Status

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.694.1, 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

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 29, 2012

J. Dong
M. Chen
Huawei Technologies
Z. Li
China Mobile
May 28, 2012

RSVP-TE Extensions for Lock Instruct and Loopback in MPLS Transport
Profile
draft-dong-ccamp-rsvp-te-mpls-tp-li-lb-03

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 aspects of MPLS as well.

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 November 29, 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.

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	5
5. Security Considerations	6
6. Acknowledgements	6
7. References	6
7.1. Normative References	6
7.2. Informative References	7
Authors' Addresses	7

The requirements of Lock Instruct (LI) and Loopback (LB) are specified in [RFC5860], and the framework of LI and LB is specified in [RFC6371]. [RFC6435] defines in-band Lock Instruct (LI) and Loopback (LB) functions, it leverages the Generic Associated Channel (GACH) and Generic Associated Channel Label (GAL) [RFC5586] and the management plane to perform LI function, and use management plane to perform the LB function. In-band LI and LB are suitable for the scenarios where control plane is not used.

When a control plane is used for establishing MPLS-TP LSPs, it's natural to use and extend the control plane protocol to implement LI and LB functions. Since LI and LB would modify the forwarding plane of an LSP, without the involvement of control plane this may result in inconsistency of the LSP information between control plane and data plane. Besides, with control plane mechanisms, it does not need to rely on the TTL expiration to make the LI/LB commands to reach particular MIP or MEP.

This document specifies extensions to RSVP-TE to implement LI and LB for MPLS-TP LSPs when MPLS-TP control plane is used. The mechanisms defined in this document are complementary to [RFC6435].

The A (Administratively down) bit in ADMIN_STATUS Object [RFC3471] [RFC3473] is used to indicate the lock/unlock of the LSP. One new bit is defined in ADMIN_STATUS Object to indicate the loopback mode.

Format of extended ADMIN_STATUS Object is as below:

										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										Class-Num(196)										C-Type (1)											
-----										-----										-----										-----											
R		Reserved																		B H L I C T A D																					
-----										-----										-----										-----											

Loopback (B): When this bit is set in Notify message sent from the ingress node, it indicates that the target node of this message SHOULD perform loopback function for this LSP. When this bit is set in Notify message sent to the ingress node, it indicates the node originating this message is in "Loopback" mode.

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]
Deletion in progress (D): 1 bit - see [RFC3471]

3. Operations

3.1. Lock Instruct

When a MEP wants to put an LSP in 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. Though the intermediate nodes do not need to take actions during this procedure, they 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 cleared. The intermediate nodes do not need to take action on this message and SHOULD forward it 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.

3.2. Loopback

Notify message is used to support signaling of Loopback request.

When a MEP wants to put particular LSR on the given LSP in loopback mode, it MUST send a Notify message with the Reflect (R) bit, the Loopback (B) Bit and the Administratively down (A) bit in ADMIN_STATUS Object set. The destination address of this Notify message SHOULD be set to the MIP or MEP which is required to loopback the traffic. The ERROR_SPEC object is not relevant in loopback request and MUST carry the Error Code zero ("Confirmation") to indicate that there is no error.

On receipt of this Notify message, the receiver node SHOULD try to put the LSP in loopback mode. If the receiver node puts the LSP into loopback mode successfully, it SHOULD send a Notify message back to the MEP node, with both the Loopback (B) Bit and the Administratively down (A) bit in ADMIN_STATUS Object set, and the ERROR_SPEC object MUST carry the Error Code zero. Otherwise, it SHOULD send a Notify message with the Error Code "OAM Problem" and the new Error Value "Loopback Failure".

When a MEP wants to take the LSP out of the loopback mode, it MUST send a Notify message with the Reflect (R) bit and the Administratively down (A) bit set and the Loopback (B) Bit cleared. The destination address of this Notify message SHOULD be set to the MIP or MEP which is performing the loopback action for this LSP.

On receipt of this Notify message, the receiving node SHOULD try to put the LSP back to normal operation. If the receiving node put the LSP into normal operation successfully, it SHOULD send a Notify message back to the MEP node, with the Administratively down (A) Bit set and the Loopback (B) Bit cleared, and the ERROR_SPEC object MUST carry the Error Code zero. Otherwise, it SHOULD send a Notify message with the Error Code "OAM Problem" and the new Error Value "Exit Loopback Failure".

4. IANA Considerations

One bit ("Loopback" (B)) needs to be allocated in the ADMIN_STATUS Object.

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-07 (work in progress), January 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.
- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", RFC 5586, June 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-08 (work in
progress), April 2012.
- [RFC4974] Papadimitriou, D. and A. Farrel, "Generalized MPLS (GMPLS)
RSVP-TE Signaling Extensions in Support of Calls",
RFC 4974, August 2007.
- [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.

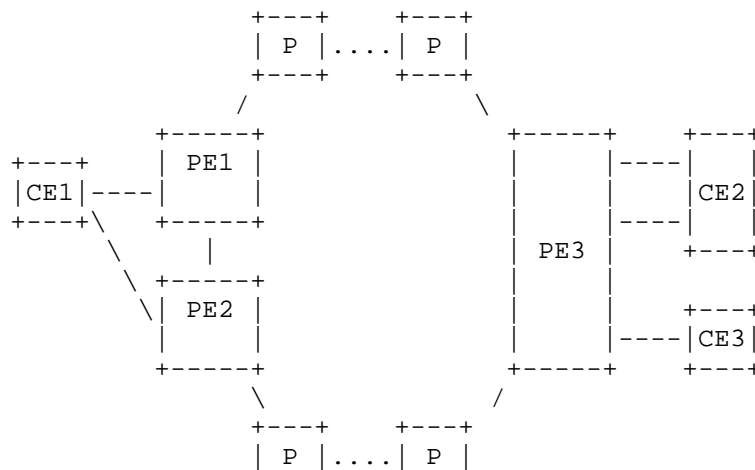


Figure 1: Generalized Layer 1 VPN Reference Model

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                               |

```

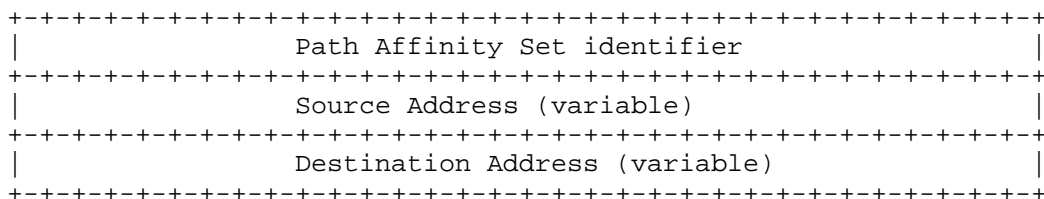


Figure 2 Diverse LSP information

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: December 24, 2012

GMG. G.Galimberti, Ed.
Cisco
RK. R.Kunze, Ed.
Deutsche Telekom
Kam. Lam, Ed.
Alcatel-Lucent
DH. Hiremagalur, Ed.
Juniper
June 22, 2012

A SNMP MIB to manage black-link optical interface parameters of DWDM
applications
draft-galikunze-ccamp-g-698-2-snmp-mib-00

Abstract

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 the Black-Link approach defined in ITU-T Recommendation 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) 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 December 24, 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.

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	8
4.1.3. Optical path from point Ss to Rs	9
4.1.4. Interface at point Rs	11
4.1.5. Alarms and Threshold definition	11
4.1.6. Performance Monitoring (PM) description	12
4.1.7. Generic Parameter description	14
4.2. Use of ifTable	14
4.2.1. Use of ifTable for OPS Layer	16
4.2.2. Use of ifTable for OCh Layer	17
4.2.3. Use of ifStackTable	17
5. Structure of the MIB Module	18
6. Object Definitions	18
7. Relationship to Other MIB Modules	74
7.1. Relationship to the [TEMPLATE TODO] MIB	74
7.2. MIB modules required for IMPORTS	74
8. Definitions	74
9. Security Considerations	75
10. IANA Considerations	75
11. Contributors	77
12. References	77
12.1. Normative References	77
12.2. Informative References	79
Appendix A. Change Log	80
Appendix B. Open Issues	80
Authors' Addresses	80

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 the Black-Link approach 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 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-g698-mgmt-ctrl-framework.

The building of a 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].

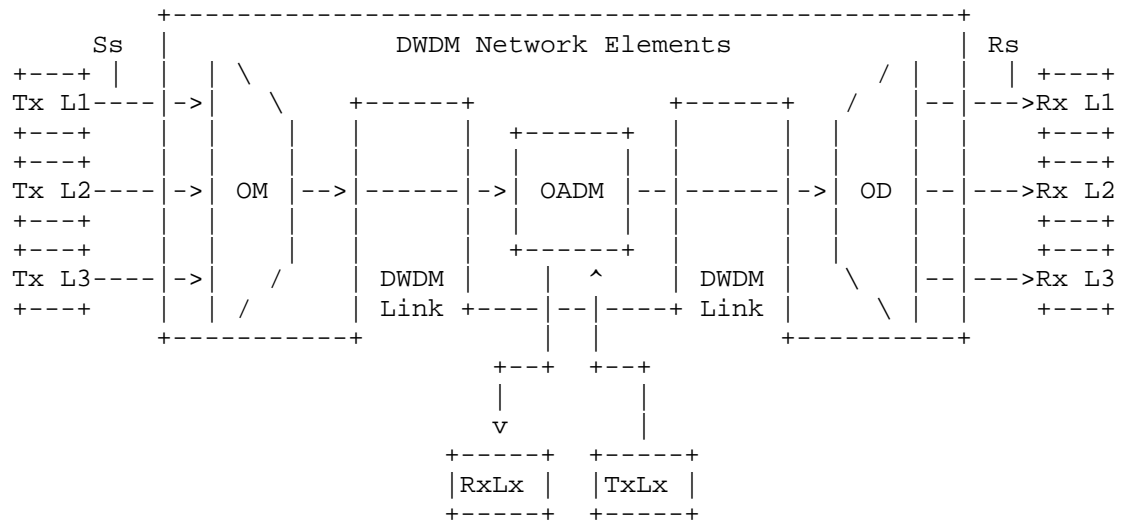
3. Conventions

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 configuratoin 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):

This parameter indicates the wavelength value that Ss and Rs will be set to work (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 5.3 - this parameter can be called Optical Interface Identifier OII as per [draft-martinelli-wson-interface-class] (G, S).

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.

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 polarisation dependent loss:

The polarisation 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.

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).

4.1.4.2. Optional parameters

Current Chromatic Dispersion (CD):

Residual Chromatic Dispersion measured 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).

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.

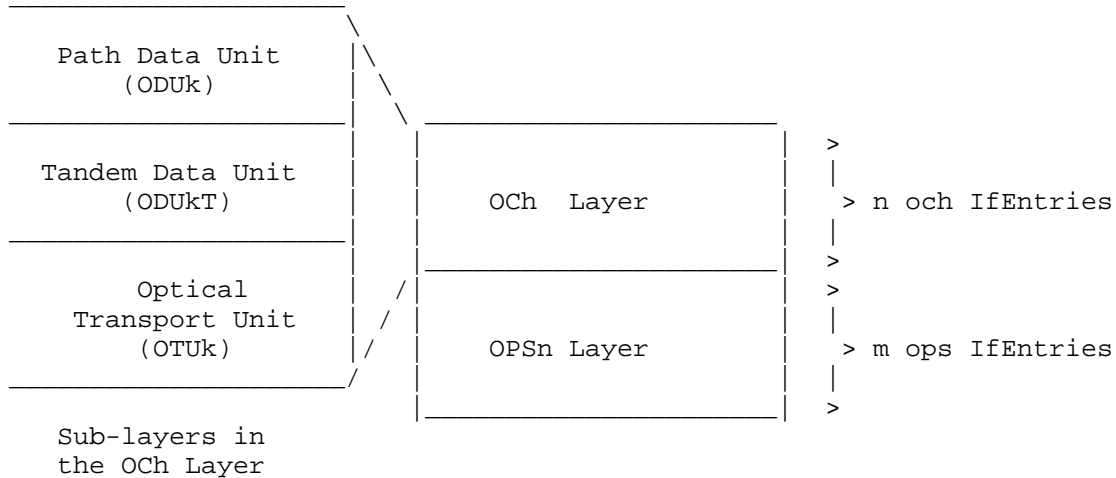


Figure 2: OTN Layers for OPS and OCh

Each opticalChannel IfEntry may be mapped to m opticalPhysicalSection IfEntries, where m is greater than or equal to 1. Conversely, each opticalTransPhysicalSection port entry may be mapped to 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 xxxDirectionality 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
    SnmpAdminString
        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 nominally 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

```

```

        optIfOTNoduTcm15MinThreshUASTCA(10),
        -- OTN OTU Fcs Threshold alarm
        optIfOTNoduTcm15MinThreshFcsTCA(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 {
```

optIfOChMimumumChannelSpacing	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

}

optIfOChMimumumChannelSpacing 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).
          For eg 1528.77 Ghz will be represented as 152877.
        "
    ::= { 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.

"

::= { 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,
        optIfOChSrcSinkMaximumPolarisationDependentLoss 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}

optIfOChSrcSinkMaximumPolarisationDependentLoss       OBJECT-TYPE
SYNTAX         Integer32
UNITS          "0.1 db"
MAX-ACCESS     read-only
STATUS         current
DESCRIPTION
"
    The polarisation 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 extension 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,
        optIfOTNPM15MinESSsThreshold         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}

optIfOTNPMConfigTCMLLevel 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 optIf15MinThreshEsTCA. One notification will be
```

```
        sent per interval per interface if the threshold is exceeded. A
        value of '0' will disable the notification.
    "
    ::= { optIfOTNPMConfigEntry 5 }

optIfOTNPM15MinSESsThreshold 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, optIfOTNPMCurrentTCMLLevel }
    ::= { optIfOTNPMCurrentTable 1 }

OptIfOTNPMCurrentEntry ::=
    SEQUENCE {
        optIfOTNPMCurrentType          OptIfPerformanceDataType,
        optIfOTNPMCurrentLayer          OptIfOTNLayer,
        optIfOTNPMCurrentTCMLLevel      Unsigned32,
        optIfOTNPMCurrentSuspectedFlag  TruthValue,
        optIfOTNPMCurrentInterval       Unsigned32,
        optIfOTNPMCurrentValidIntervals Unsigned32,
        optIfOTNPMCurrentFcS            Unsigned32,
        optIfOTNPMCurrentESs            Unsigned32,
        optIfOTNPMCurrentSESS           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 }

optIfOTNPMIntervalESSs 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,
        optIfOTNPMCurrentDaySEs            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}

optIfOTNPMCurrentDayTCMLLevel 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 current 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
--
optIfOTNPMPrevDayTable OBJECT-TYPE
    SYNTAX      SEQUENCE OF OptIfOTNPMPrevDayEntry
    MAX-ACCESS   not-accessible
    STATUS      current
    DESCRIPTION
        "A Performance monitoring Previous Day Table.
        "
    ::= { optIfOTNPMGroup 6 }

optIfOTNPMPrevDayEntry OBJECT-TYPE
    SYNTAX      OptIfOTNPMPrevDayEntry
    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
        'optIfOTNPMPrevDayLayer' layer.
        "
    INDEX { ifIndex, optIfOTNPMPrevDayType
            ,
            optIfOTNPMPrevDayLayer, optIfOTNPMPrevDayTCMLLevel }
    ::= { optIfOTNPMPrevDayTable 1 }

OptIfOTNPMPrevDayEntry ::=
    SEQUENCE {
        optIfOTNPMPrevDayType          OptIfPerformanceDataType,
        optIfOTNPMPrevDayLayer          OptIfOTNLayer,
        optIfOTNPMPrevDayTCMLLevel      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}

optIfOTNPMPPrevDayTCMLevel 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.

```

```
"
 ::= { optIfOTNPMMPrevDayEntry 3}

optIfOTNPMMPrevDaySuspectedFlag OBJECT-TYPE
    SYNTAX TruthValue
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        If true, the data in this entry may be unreliable.
        "
    ::= { optIfOTNPMMPrevDayEntry 4}

optIfOTNPMMPrevDayFcs OBJECT-TYPE
    SYNTAX Unsigned32
    MAX-ACCESS read-only
    STATUS current
    DESCRIPTION
        "
        Number of pre FEC failures occurred in an observation period.
        "
    ::= { optIfOTNPMMPrevDayEntry 5}

optIfOTNPMMPrevDayESSs 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.
        "
    ::= { optIfOTNPMMPrevDayEntry 6}

optIfOTNPMMPrevDaySESSs 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.
        "
    ::= { optIfOTNPMMPrevDayEntry 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.
        "
    ::= { optIfOTNPMPPrevDayEntry 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,
        optIfOTNPMFECValidIntervals
                                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 threshold
    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, optIfOTNPMIntervalType, optIfOTNPMFECIntervalNumber }
    ::= { optIfOTNPMFECIntervalTable 1 }

OptIfOTNPMFECIntervalEntry ::=
    SEQUENCE {
        optIfOTNPMFECIntervalType
        OptIfPerformanceDataType,

```

```

        optIfOTNPMFECIntervalNumber      Unsigned32,
        optIfOTNPMFECIntervalSuspectedFlag TruthValue,
        optIfOTNPMIntervalFECCorrectedErr Unsigned32,
        optIfOTNPMIntervalFECUncorrectedWords Unsigned32,
        optIfOTNPMIntervalMinFECBERMantissa Unsigned32,
        optIfOTNPMIntervalMinFECBERExponent Unsigned32,
        optIfOTNPMIntervalMaxFECBERMantissa Unsigned32,
        optIfOTNPMIntervalMaxFECBERExponent Unsigned32,
        optIfOTNPMIntervalAvgFECBERMantissa Unsigned32,
        optIfOTNPMIntervalAvgFECBERExponent 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 maximun 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 maximun 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 maximun 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 minimun 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 maximun 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 maximun 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}

optIfOTNPMPrevDayAvgFECBERMantissa 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) ."
  ::= { optIfOTNAlarmEntry 6 }

--
-- 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

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
D, x
New Jersey
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

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: January 2013

G. Bernstein
Grotto Networking
Y. Lee
D. Li
Huawei
W. Imajuku
NTT

July 6, 2012

General Network Element Constraint Encoding for GMPLS Controlled
Networks

draft-ietf-ccamp-general-constraint-encode-08.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 January 6, 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.

Abstract

Generalized Multiprotocol Label Switching can be used to control a wide variety of technologies. In some of these technologies network elements and links may impose additional routing constraints such as asymmetric switch connectivity, non-local label assignment, and label range limitations on links.

This document provides efficient, protocol-agnostic encodings for general information elements representing connectivity and label constraints as well as label availability. It is intended that protocol-specific documents will reference this memo to describe how information is carried for specific uses.

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. Node Switching Asymmetry Constraints.....	3
1.2. Non-Local Label Assignment Constraints.....	4
1.3. Change Log.....	5
2. Encoding.....	5
2.1. Link Set Field.....	5
2.2. Label Set Field.....	7
2.2.1. Inclusive/Exclusive Label Lists	8
2.2.2. Inclusive/Exclusive Label Ranges	9
2.2.3. Bitmap Label Set	9
2.3. Available Labels Sub-TLV.....	10

2.4. Shared Backup Labels Sub-TLV.....	11
2.5. Connectivity Matrix Sub-TLV.....	11
2.6. Port Label Restriction sub-TLV.....	13
2.6.1. SIMPLE_LABEL	14
2.6.2. CHANNEL_COUNT	14
2.6.3. LABEL_RANGE1	15
2.6.4. SIMPLE_LABEL & CHANNEL_COUNT	15
2.6.5. Link Label Exclusivity	16
3. Security Considerations.....	16
4. IANA Considerations.....	16
5. Acknowledgments.....	16
APPENDIX A: Encoding Examples.....	17
A.1. Link Set Field.....	17
A.2. Label Set Field.....	17
A.3. Connectivity Matrix Sub-TLV.....	18
A.4. Connectivity Matrix with Bi-directional Symmetry.....	21
6. References.....	25
6.1. Normative References.....	25
6.2. Informative References.....	25
7. Contributors.....	26
Authors' Addresses.....	27
Intellectual Property Statement.....	28
Disclaimer of Validity.....	28

1. Introduction

Some data plane technologies that wish to make use of a GMPLS control plane contain additional constraints on switching capability and label assignment. In addition, some of these technologies must perform non-local label assignment based on the nature of the technology, e.g., wavelength continuity constraint in WSON [WSON-Frame]. Such constraints can lead to the requirement for link by link label availability in path computation and label assignment.

This document provides efficient encodings of information needed by the routing and label assignment process in technologies such as WSON and are potentially applicable to a wider range of technologies. Such encodings can be used to extend GMPLS signaling and routing protocols. In addition these encodings could be used by other mechanisms to convey this same information to a path computation element (PCE).

1.1. Node Switching Asymmetry Constraints

For some network elements the ability of a signal or packet on a particular ingress port to reach a particular egress port may be

limited. In addition, in some network elements the connectivity between some ingress ports and egress ports may be fixed, e.g., a simple multiplexer. To take into account such constraints during path computation we model this aspect of a network element via a connectivity matrix.

The connectivity matrix (ConnectivityMatrix) represents either the potential connectivity matrix for asymmetric switches or fixed connectivity for an asymmetric device such as a multiplexer. Note that this matrix does not represent any particular internal blocking behavior but indicates which ingress ports and labels (e.g., wavelengths) could possibly be connected to a particular output port. Representing internal state dependent blocking for a node is beyond the scope of this document and due to its highly implementation dependent nature would most likely not be subject to standardization in the future. The connectivity matrix is a conceptual M by N matrix representing the potential switched or fixed connectivity, where M represents the number of ingress ports and N the number of egress ports.

1.2. Non-Local Label Assignment Constraints

If the nature of the equipment involved in a network results in a requirement for non-local label assignment we can have constraints based on limits imposed by the ports themselves and those that are implied by the current label usage. Note that constraints such as these only become important when label assignment has a non-local character. For example in MPLS an LSR may have a limited range of labels available for use on an egress port and a set of labels already in use on that port and hence unavailable for use. This information, however, does not need to be shared unless there is some limitation on the LSR's label swapping ability. For example if a TDM node lacks the ability to perform time-slot interchange or a WSON lacks the ability to perform wavelength conversion then the label assignment process is not local to a single node and it may be advantageous to share the label assignment constraint information for use in path computation.

Port label restrictions (PortLabelRestriction) model the label restrictions that the network element (node) and link may impose on a port. These restrictions tell us 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. Reference [Switch] gives an example where both switch and fixed connectivity matrices are used and both types of constraints occur on the same port.

1.3. Change Log

Changes from 03 version:

- (a) Removed informational BNF from section 1.
- (b) Removed section on "Extension Encoding Usage Recommendations"

Changes from 04,05 versions:

No changes just refreshed document that was expiring.

Changes from 06 version:

Added priority information to available wavelength encodings.

Changes from 07 version:

In port label constraint changed reserved field to Switching Capability and Encoding to allow for self description of labels used and interface capability.

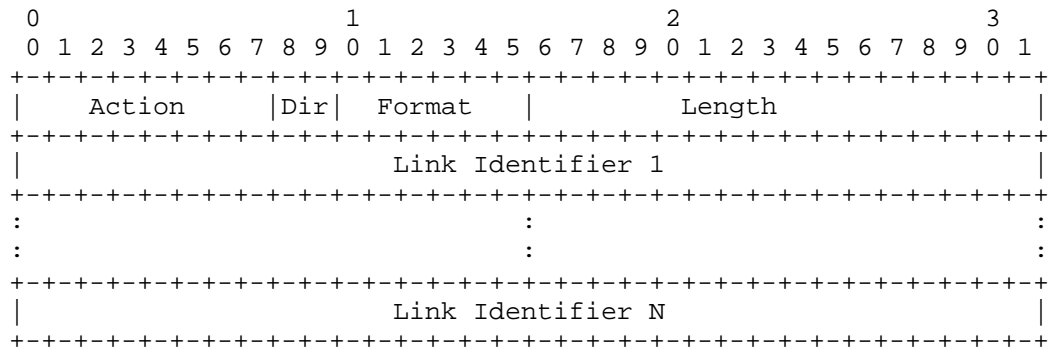
2. Encoding

A type-length-value (TLV) encoding of the general connectivity and label restrictions and availability extensions is given in this section. This encoding is designed to be suitable for use in the GMPLS routing protocols OSPF [RFC4203] and IS-IS [RFC5307] and in the PCE protocol PCEP [PCEP]. Note that the information distributed in [RFC4203] and [RFC5307] is arranged via the nesting of sub-TLVs within TLVs and this document makes use of such constructs. First, however we define two general purpose fields that will be used repeatedly in the subsequent TLVs.

2.1. Link Set Field

We will frequently need to describe properties of groups of links. To do so efficiently we can make use of a link set concept similar to the label set concept of [RFC3471]. This Link Set Field is used

in the <ConnectivityMatrix> sub-TLV, which is defined in Section 2.5. The information carried in a Link Set is defined by:



Action: 8 bits

0 - Inclusive List

Indicates that one or more link identifiers are included in the Link Set. Each identifies a separate link that is part of the set.

1 - Inclusive Range

Indicates that the Link Set defines a range of links. It contains two link identifiers. The first identifier indicates the start of the range (inclusive). The second identifier indicates the end of the range (inclusive). All links with numeric values between the bounds are considered to be part of the set. A value of zero in either position indicates that there is no bound on the corresponding portion of the range. Note that the Action field can be set to 0x02(Inclusive Range) only when unnumbered link identifier is used.

Dir: Directionality of the Link Set (2 bits)

0 -- bidirectional

1 -- ingress

2 -- egress

For example in optical networks we think in terms of unidirectional as well as bidirectional links. For example, label restrictions or

connectivity may be different for an ingress port, than for its "companion" egress port if one exists. Note that "interfaces" such as those discussed in the Interfaces MIB [RFC2863] are assumed to be bidirectional. This also applies to the links advertised in various link state routing protocols.

Format: The format of the link identifier (6 bits)

0 -- Link Local Identifier

Indicates that the links in the Link Set are identified by link local identifiers. All link local identifiers are supplied in the context of the advertising node.

1 -- Local Interface IPv4 Address

2 -- Local Interface IPv6 Address

Indicates that the links in the Link Set are identified by Local Interface IP Address. All Local Interface IP Address are supplied in the context of the advertising node.

Others TBD.

Note that all link identifiers in the same list must be of the same type.

Length: 16 bits

This field indicates the total length in bytes of the Link Set field.

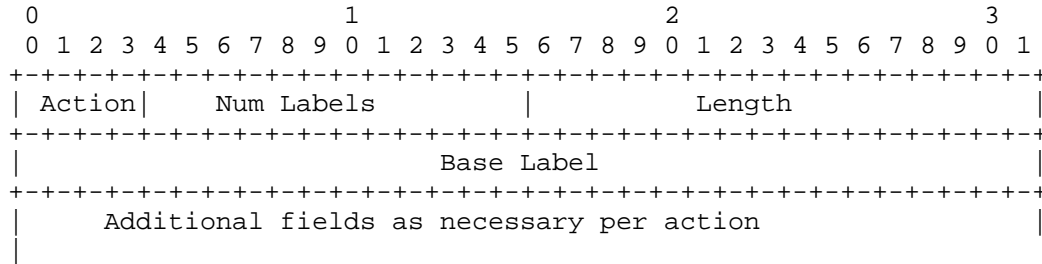
Link Identifier: length is dependent on the link format

The link identifier represents the port which is being described either for connectivity or label restrictions. This can be the link local identifier of [RFC4202], GMPLS routing, [RFC4203] GMPLS OSPF routing, and [RFC5307] IS-IS GMPLS routing. The use of the link local identifier format can result in more compact encodings when the assignments are done in a reasonable fashion.

2.2. Label Set Field

Label Set Field is used within the <AvailableLabels> sub-TLV or the <SharedBackupLabels> sub-TLV, which is defined in Section 2.3. and 2.4. , respectively.

The general format for a label set is given below. This format uses the Action concept from [RFC3471] with an additional Action to define a "bit map" type of label set. The second 32 bit field is a base label used as a starting point in many of the specific formats.



Action:

- 0 - Inclusive List
- 1 - Exclusive List
- 2 - Inclusive Range
- 3 - Exclusive Range
- 4 - Bitmap Set

Num Labels is only meaningful for Action value of 4 (Bitmap Set). It indicates the number of labels represented by the bit map. See more detail in section 3.2.3.

Length is the length in bytes of the entire field.

2.2.1. Inclusive/Exclusive Label Lists

In the case of the inclusive/exclusive lists the wavelength set format is given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|0 or 1 | Num Labels (not used) |                Length                |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                Base Label                |
+-----+-----+-----+-----+-----+-----+-----+-----+
:                                                                 :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                Last Label                |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where:

Num Labels is not used in this particular format since the Length parameter is sufficient to determine the number of labels in the list.

2.2.2. Inclusive/Exclusive Label Ranges

In the case of inclusive/exclusive ranges the label set format is given by:

```

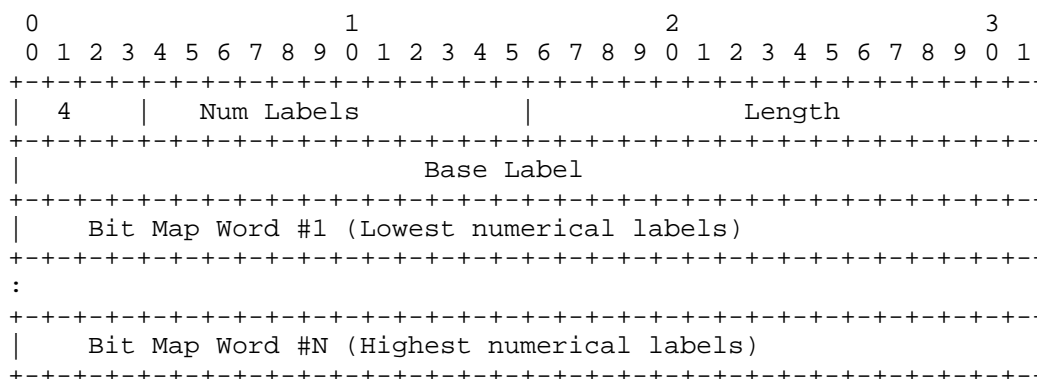
      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|2 or 3 | Num Labels(not used) |                Length                |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                Start Label                |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                End Label                |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Note that the start and end label must in some sense "compatible" in the technology being used.

2.2.3. Bitmap Label Set

In the case of Action = 4, the bitmap the label set format is given by:

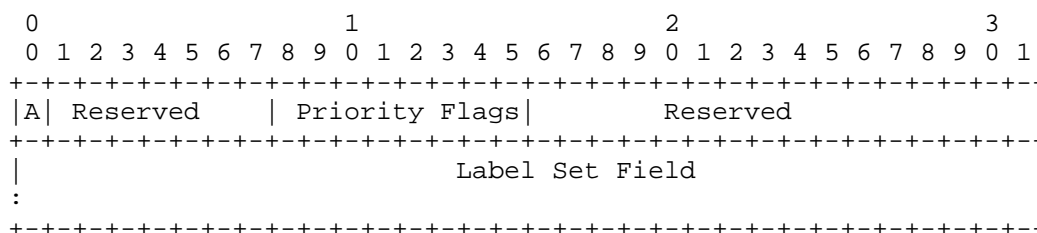


Where Num Labels in this case tells us the number of labels represented by the bit map. Each bit in the bit map represents a particular label with a value of 1/0 indicating whether the label is in the set or not. Bit position zero represents the lowest label and corresponds to the base label, while each succeeding bit position represents the next label logically above the previous.

The size of the bit map is Num Label bits, but the bit map is padded out to a full multiple of 32 bits so that the TLV is a multiple of four bytes. Bits that do not represent labels (i.e., those in positions (Num Labels) and beyond SHOULD be set to zero and MUST be ignored.

2.3. Available Labels Sub-TLV

The Available Labels sub-TLV link consists of an availability flag, priority flags, and a single variable length label set field as follows:



Where

A (Availability bit) = 1 or 0 indicates that the labels listed in the following label set field are available or not available,

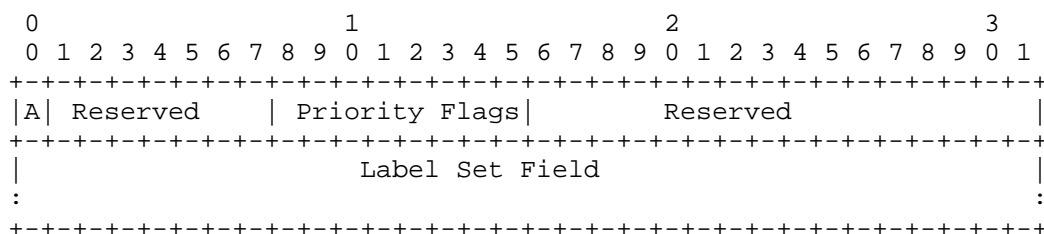
respectively, for use at a given priority level as indicated by the Priority Flags.

Priority Flags: Bit 8 corresponds to priority level 0 and bit 15 corresponds to priority level 7. If a bit is set then the labels in the label set field are available or not available as indicated by the A bit for use at that particular priority level.

Note that Label Set Field is defined in Section 3.2.

2.4. Shared Backup Labels Sub-TLV

The Shared Backup Labels sub-TLV consists of an availability flag, priority flags, and single variable length label set field as follows:



Where

A (Availability bit) = 1 or 0 indicates that the labels listed in the following label set field are available or not available, respectively, for use at a given priority level as indicated by the Priority Flags.

Priority Flags: Bit 8 corresponds to priority level 0 and bit 15 corresponds to priority level 7. If a bit is set then the labels in the label set field are available or not available as indicated by the A bit for use at that particular priority level.

2.5. Connectivity Matrix Sub-TLV

The Connectivity Matrix represents how ingress ports are connected to egress ports for network elements. The switch and fixed connectivity matrices can be compactly represented in terms of a minimal list of ingress and egress port set pairs that have mutual

connectivity. As described in [Switch] such a minimal list representation leads naturally to a graph representation for path computation purposes that involves the fewest additional nodes and links.

A TLV encoding of this list of link set pairs is:

```

      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 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               Link Set A #1               |
:               :                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
|               Link Set B #1               |
:               :                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
|               Additional Link set pairs as needed               |
:               to specify connectivity               :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where

Connectivity is the device type.

0 -- the device is fixed

1 -- the device is switched(e.g., ROADM/OXC)

MatrixID represents the ID of the connectivity matrix and is an 8 bit integer. The value of 0xFF is reserved for use with port wavelength constraints and should not be used to identify a connectivity matrix.

Link Set A #1 and Link Set B #1 together represent a pair of link sets. There are two permitted combinations for the link set field parameter "dir" for Link Set A and B pairs:

- o Link Set A dir=ingress, Link Set B dir=egress

The meaning of the pair of link sets A and B in this case is that any signal that ingresses a link in set A can be potentially switched out of an egress link in set B.

- o Link Set A dir=bidirectional, Link Set B dir=bidirectional

The meaning of the pair of link sets A and B in this case is that any signal that ingresses on the links in set A can potentially egress on a link in set B, and any ingress signal on the links in set B can potentially egress on a link in set A.

See Appendix A for both types of encodings as applied to a ROADM example.

2.6. Port Label Restriction sub-TLV

Port Label Restriction tells us what labels may or may not be used on a link.

The port label restriction of section 1.2. can be encoded as a sub-TLV as follows. More than one of these sub-TLVs may be needed to fully specify a complex port constraint. When more than one of these sub-TLVs are present the resulting restriction is the intersection of the restrictions expressed in each sub-TLV. To indicate that a restriction applies to the port in general and not to a specific connectivity matrix use the reserved value of 0xFF for the MatrixID.

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
MatrixID										RestrictionType										Switching Cap										Encoding									
Additional Restriction Parameters per RestrictionType																																							
:																														:									

Where:

MatrixID: either is the value in the corresponding Connectivity Matrix sub-TLV or takes the value 0xFF to indicate the restriction applies to the port regardless of any Connectivity Matrix.

RestrictionType can take the following values and meanings:

- 0: SIMPLE_LABEL (Simple label selective restriction)
- 1: CHANNEL_COUNT (Channel count restriction)

2: LABEL_RANGE1 (Label range device with a movable center label and width)

3: SIMPLE_LABEL & CHANNEL_COUNT (Combination of SIMPLE_LABEL and CHANNEL_COUNT restriction. The accompanying label set and channel count indicate labels permitted on the port and the maximum number of channels that can be simultaneously used on the port)

4: LINK_LABEL_EXCLUSIVITY (A label may be used at most once amongst a set of specified ports)

Switching Capability is defined in [RFC4203] and Encoding in [RFC3471]. The combination of these fields defines the type of labels used in specifying the port label restrictions as well as the interface type to which these restrictions apply.

2.6.1. SIMPLE_LABEL

In the case of the SIMPLE_LABEL the GeneralPortRestrictions (or MatrixSpecificRestrictions) format is given by:

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
MatrixID										RstType = 0										Reserved																			
										Label Set Field																													

In this case the accompanying label set indicates the labels permitted on the port.

2.6.2. CHANNEL_COUNT

In the case of the CHANNEL_COUNT the format is given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| MatrixID           | RstType = 1   |           MaxNumChannels           |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

In this case the accompanying MaxNumChannels indicates the maximum number of channels (labels) that can be simultaneously used on the port/matrix.

2.6.3. LABEL_RANGE1

In the case of the LABEL_RANGE1 the GeneralPortRestrictions (or MatrixSpecificRestrictions) format is given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| MatrixID           | RstType = 2   |           MaxLabelRange           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Label Set Field                       |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

In this case the accompanying MaxLabelRange indicates the maximum range of the labels. The corresponding label set is used to indicate the overall label range. Specific center label information can be obtained from dynamic label in use information. It is assumed that both center label and range tuning can be done without causing faults to existing signals.

2.6.4. SIMPLE_LABEL & CHANNEL_COUNT

In the case of the SIMPLE_LABEL & CHANNEL_COUNT the format is given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| MatrixID           | RstType = 3   |           MaxNumChannels           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Label Set Field                       |
+-----+-----+-----+-----+-----+-----+-----+-----+

```


APPENDIX A: Encoding Examples

Here we give examples of the general encoding extensions applied to some simple ROADM network elements and links.

A.1. Link Set Field

Suppose that we wish to describe a set of ingress ports that are have link local identifiers number 3 through 42. In the link set field we set the Action = 1 to denote an inclusive range; the Dir = 1 to denote ingress links; and, the Format = 0 to denote link local identifiers. In particular we have:

```

+-----+
| Action=1 | 0 1 | 0 0 0 0 0 0 | Length = 12 |
+-----+
| Link Local Identifier = #3 |
+-----+
| Link Local Identifier = #42 |
+-----+

```

A.2. Label Set Field

Example:

A 40 channel C-Band DWDM system with 100GHz spacing with lowest frequency 192.0THz (1561.4nm) and highest frequency 195.9THz (1530.3nm). These frequencies correspond to $n = -11$, and $n = 28$ respectively. Now suppose the following channels are available:

Frequency (THz)	n Value	bit map position
192.0	-11	0
192.5	-6	5
193.1	0	11
193.9	8	19
194.0	9	20
195.2	21	32
195.8	27	38

With the Grid value set to indicate an ITU-T G.694.1 DWDM grid, C.S. set to indicate 100GHz this lambda bit map set would then be 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
+-----+-----+-----+-----+
|  4      | Num Wavelengths = 40 |   Length = 16 bytes   |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = -11 |
+-----+-----+-----+-----+
|1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0|
+-----+-----+-----+-----+
|1 0 0 0 0 0 1 0|   Not used in 40 Channel system (all zeros)   |
+-----+-----+-----+-----+

```

To encode this same set as an inclusive list we would have:

```

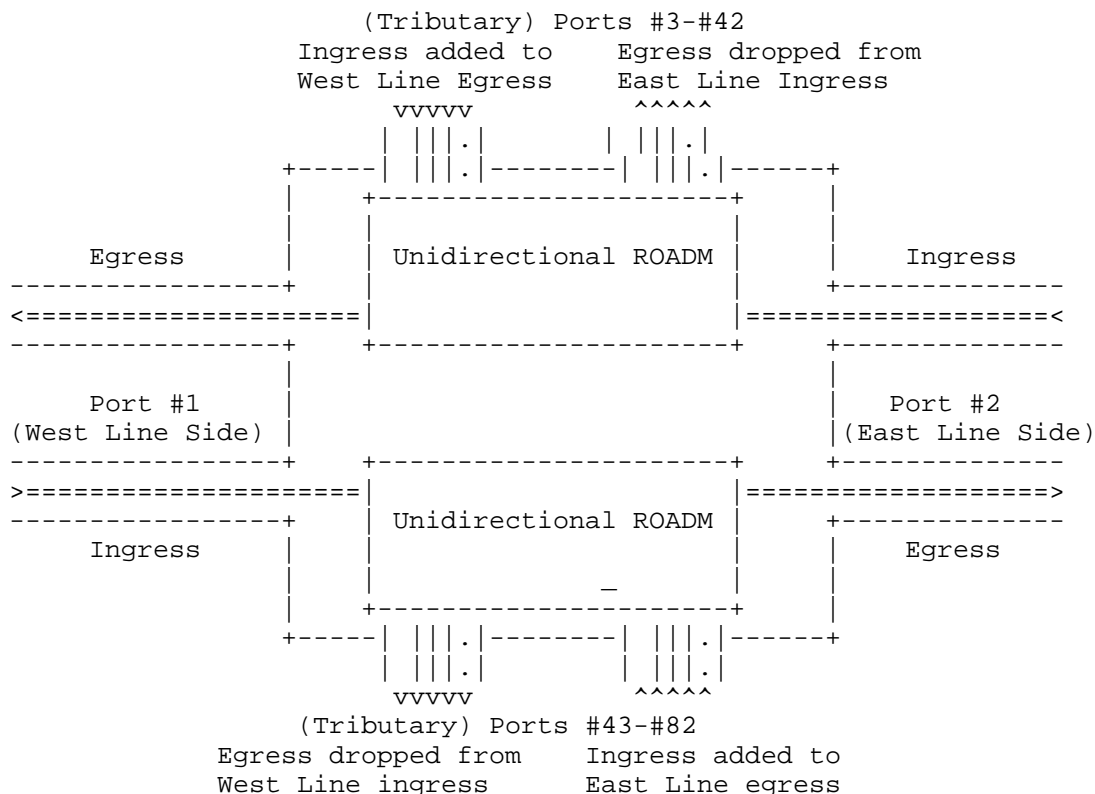
0          1          2          3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
|  0      | Num Wavelengths = 40 |   Length = 20 bytes   |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = -11 |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = -6  |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = -0  |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = 8   |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = 9   |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = 21  |
+-----+-----+-----+-----+
|Grid |  C.S. |   Reserved   | n  for lowest frequency = 27  |
+-----+-----+-----+-----+

```

A.3. Connectivity Matrix Sub-TLV

Example:

Suppose we have a typical 2-degree 40 channel ROADM. In addition to its two line side ports it has 80 add and 80 drop ports. The picture below illustrates how a typical 2-degree ROADM system that works with bi-directional fiber pairs is a highly asymmetrical system composed of two unidirectional ROADM subsystems.



Referring to the figure we see that the ingress direction of ports #3-#42 (add ports) can only connect to the egress on port #1. While the ingress side of port #2 (line side) can only connect to the egress on ports #3-#42 (drop) and to the egress on port #1 (pass through). Similarly, the ingress direction of ports #43-#82 can only connect to the egress on port #2 (line). While the ingress direction of port #1 can only connect to the egress on ports #43-#82 (drop) or port #2 (pass through). We can now represent this potential connectivity matrix as follows. This representation uses only 30 32-bit words.

```

      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
+-----+-----+-----+-----+
| Conn = 1 | MatrixID | Reserved |1
+-----+-----+-----+-----+
      Note: adds to line
+-----+-----+-----+-----+
| Action=1 |0 1|0 0 0 0 0 0| Length = 12 |2
+-----+-----+-----+-----+
| Link Local Identifier = #3 |3
+-----+-----+-----+-----+
| Link Local Identifier = #42 |4
+-----+-----+-----+-----+
| Action=0 |1 0|0 0 0 0 0 0| Length = 8 |5
+-----+-----+-----+-----+
| Link Local Identifier = #1 |6
+-----+-----+-----+-----+
      Note: line to drops
+-----+-----+-----+-----+
| Action=0 |0 1|0 0 0 0 0 0| Length = 8 |7
+-----+-----+-----+-----+
| Link Local Identifier = #2 |8
+-----+-----+-----+-----+
| Action=1 |1 0|0 0 0 0 0 0| Length = 12 |9
+-----+-----+-----+-----+
| Link Local Identifier = #3 |10
+-----+-----+-----+-----+
| Link Local Identifier = #42 |11
+-----+-----+-----+-----+
      Note: line to line
+-----+-----+-----+-----+
| Action=0 |0 1|0 0 0 0 0 0| Length = 8 |12
+-----+-----+-----+-----+
| Link Local Identifier = #2 |13
+-----+-----+-----+-----+
| Action=0 |1 0|0 0 0 0 0 0| Length = 8 |14
+-----+-----+-----+-----+
| Link Local Identifier = #1 |15
+-----+-----+-----+-----+
      Note: adds to line
+-----+-----+-----+-----+
| Action=1 |0 1|0 0 0 0 0 0| Length = 12 |16
+-----+-----+-----+-----+
| Link Local Identifier = #43 |17
+-----+-----+-----+-----+

```

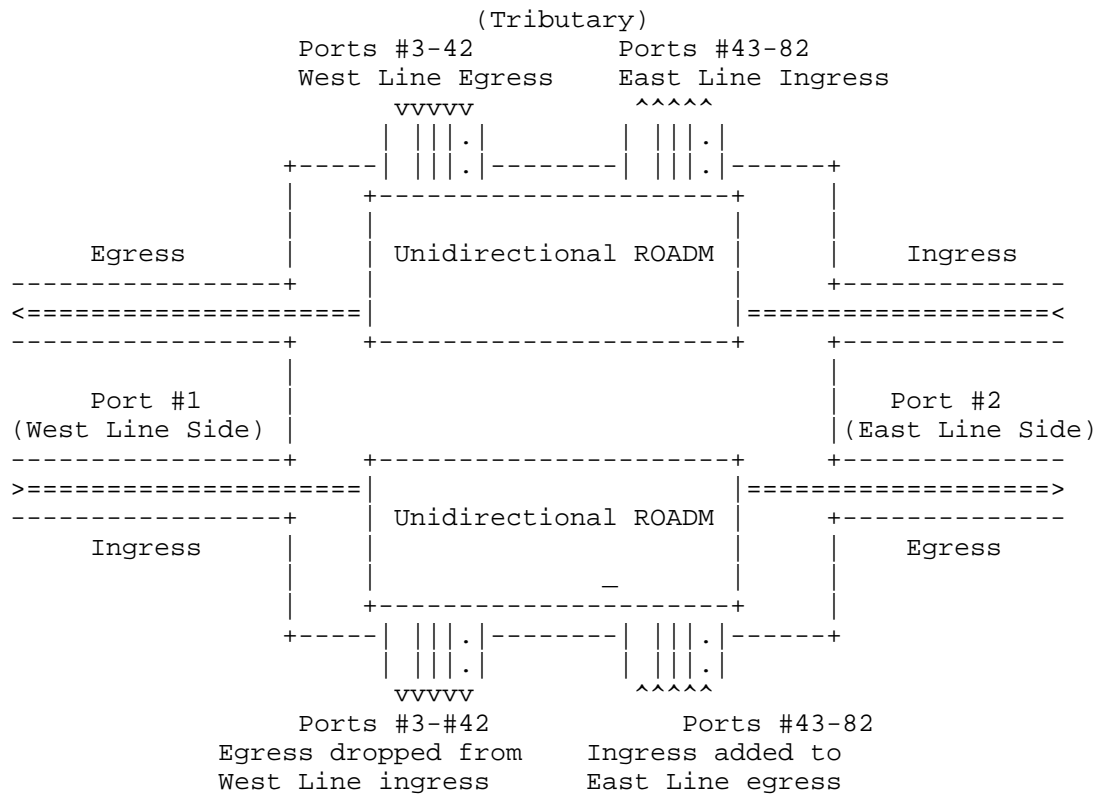
```

|                                     Link Local Identifier = #82                                     |18
+-----+
| Action=0      |1 0|0 0 0 0 0 0|                                     Length = 8                                     |19
+-----+
|                                     Link Local Identifier = #2                                     |20
+-----+
|                                     Note: line to drops                                     |
+-----+
| Action=0      |0 1|0 0 0 0 0 0||                                     Length = 8                                     |21
+-----+
|                                     Link Local Identifier = #1                                     |22
+-----+
| Action=1      |1 0|0 0 0 0 0 0|                                     Length = 12                                    |23
+-----+
|                                     Link Local Identifier = #43                                    |24
+-----+
|                                     Link Local Identifier = #82                                    |25
+-----+
|                                     Note: line to line                                     |
+-----+
| Action=0      |0 1|0 0 0 0 0 0|                                     Length = 8                                     |26
+-----+
|                                     Link Local Identifier = #1                                     |27
+-----+
| Action=0      |1 0|0 0 0 0 0 0|                                     Length = 8                                     |28
+-----+
|                                     Link Local Identifier = #2                                     |29
+-----+

```

A.4. Connectivity Matrix with Bi-directional Symmetry

If one has the ability to renumber the ports of the previous example as shown in the next figure then we can take advantage of the bi-directional symmetry and use bi-directional encoding of the connectivity matrix. Note that we set `dir=bidirectional` in the link set fields.



```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| Conn = 1 | MatrixID | Reserved |1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Add/Drops #3-42 to Line side #1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=1 |0 0|0 0 0 0 0 0| Length = 12 |2
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #3 |3
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #42 |4
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0 |0 0|0 0 0 0 0 0| Length = 8 |5
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #1
|6
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Note: line #2 to add/drops #43-82
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0 |0 0|0 0 0 0 0 0| Length = 8 |7
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #2 |8
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=1 |0 0|0 0 0 0 0 0| Length = 12 |9
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #43
|10
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #82
|11
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Note: line to line
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0 |0 0|0 0 0 0 0 0| Length = 8 |12
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #1
|13
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0 |0 0|0 0 0 0 0 0| Length = 8 |14
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Link Local Identifier = #2
|15
+-----+-----+-----+-----+-----+-----+-----+-----+

```


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.
- [RFC2863] McCloghrie, K. and F. Kastenholtz, "The Interfaces Group MIB", RFC 2863, June 2000.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [G.694.1] ITU-T Recommendation G.694.1, "Spectral grids for WDM applications: DWDM frequency grid", June, 2002.
- [RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005
- [RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

6.2. Informative References

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

[Switch] G. Bernstein, Y. Lee, A. Gavler, J. Martensson, " Modeling WDM Wavelength Switching Systems for Use in GMPLS and Automated Path Computation", Journal of Optical Communications and Networking, vol. 1, June, 2009, pp. 187-195.

[PCEP] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) communication Protocol (PCEP) - Version 1", RFC5440.

7. Contributors

Diego Caviglia
Ericsson
Via A. Negrone 1/A 16153
Genoa Italy

Phone: +39 010 600 3736
Email: diego.caviglia@marconi.com, ericsson.com)

Anders Gavler
Acreo AB
Electrum 236
SE - 164 40 Kista Sweden

Email: Anders.Gavler@acreo.se

Jonas Martensson
Acreo AB
Electrum 236
SE - 164 40 Kista, Sweden

Email: Jonas.Martensson@acreo.se

Itaru Nishioka
NEC Corp.
1753 Simonumabe, Nakahara-ku, Kawasaki, Kanagawa 211-8666
Japan

Phone: +81 44 396 3287
Email: i-nishioka@cb.jp.nec.com

Authors' Addresses

Greg M. Bernstein (ed.)
Grotto Networking
Fremont California, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Young Lee (ed.)
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@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: danli@huawei.com

Wataru Imajuku
NTT Network Innovation Labs
1-1 Hikari-no-oka, Yokosuka, Kanagawa
Japan

Phone: +81-(46) 859-4315
Email: imajuku.wataru@lab.ntt.co.jp

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

Phone: +86-755-28972916
Email: hanjianrui@huawei.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.

Network Working Group
Internet Draft
Category: Informational

Fatai Zhang, Ed.
Dan Li
Huawei
Han Li
CMCC
S. Belotti
Alcatel-Lucent
D. Ceccarelli
Ericsson
June 19, 2012

Expires: December 19, 2012

Framework for GMPLS and PCE Control of
G.709 Optical Transport Networks

draft-ietf-ccamp-gmpls-g709-framework-08.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 December 19, 2012.

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)	18
5.5. Implications for Control Plane Backward Compatibility ...	19
5.6. Implications for Path Computation Elements	20
6. Data Plane Backward Compatibility Considerations	20
7. Security Considerations	21
8. IANA Considerations	21
9. Acknowledgments	21
10. References	22
10.1. Normative References	22
10.2. Informative References	23
11. Authors' Addresses	24
12. Contributors	25
APPENDIX A: ODU connection examples	26

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].

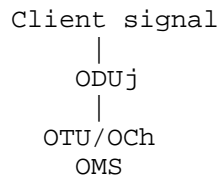


Figure 1 - Basic OTN signal hierarchy

Client signals are mapped into ODUj containers. These ODUj 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) ODUj. 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	$239/238 \times 2\,488\,320$ kbit/s
ODU2	$239/237 \times 9\,953\,280$ kbit/s
ODU3	$239/236 \times 39\,813\,120$ kbit/s
ODU4	$239/227 \times 99\,532\,800$ kbit/s
ODU2e	$239/237 \times 10\,312\,500$ kbit/s
ODUflex for CBR Client signals	$239/238 \times$ client signal bit rate
ODUflex for GFP-F Mapped client signal	Configured bit rate

Table 1 - ODU types and bit rates

NOTE - The nominal ODUk 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.

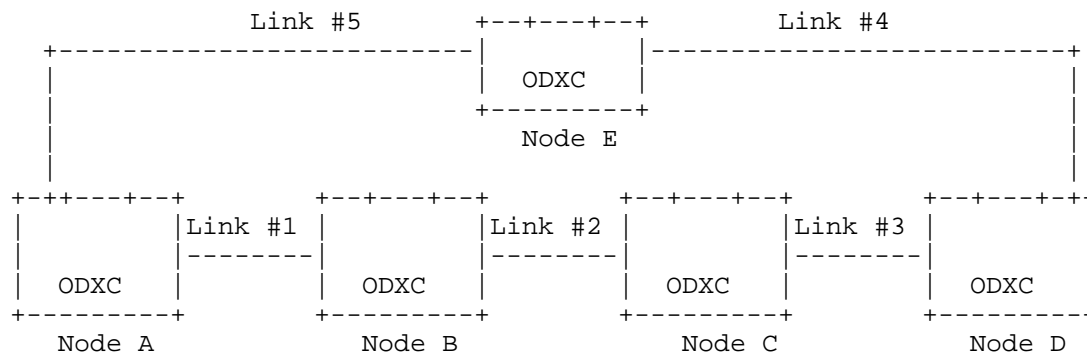


Figure 2 - Example Topology for LO ODU connection management

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.

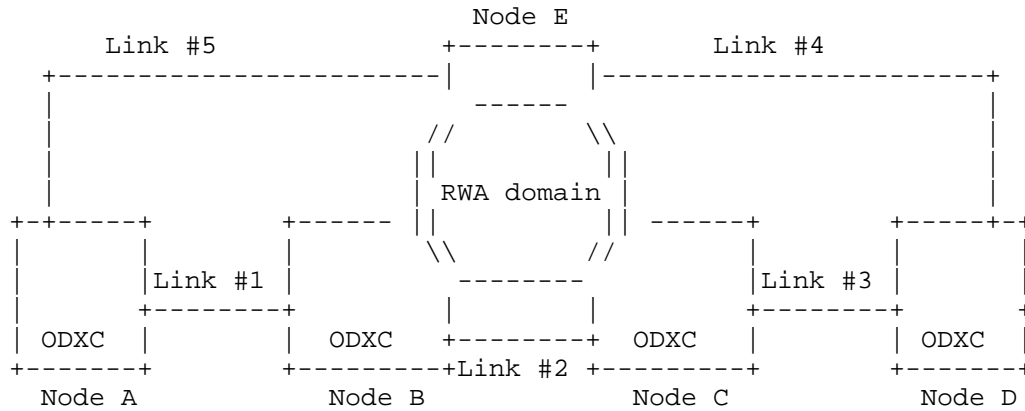


Figure 3 - RWA Hidden Topology for LO ODU connection management

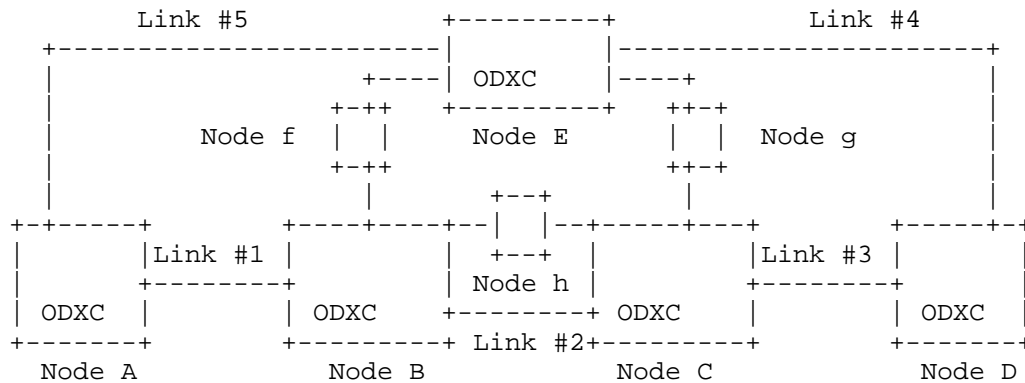


Figure 4 - RWA Visible Topology for LO ODUj connection management

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 ODU_j connection can be nested into another ODU_k (j<k) connection, which forms the LSP hierarchy in ODU layer. The creation of the outer ODU_k connection can be triggered via network planning, or by the signaling of the inner ODU_j connection. For the former case, the outer ODU_k 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 ODU_k 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 (ODU_j) 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 ODU_j 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 ODU_k/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 ODU_k 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 ODU layer multiplexing hierarchy signaling

ODU layer multiplexing hierarchy has been supported by [G709-V3], i.e., a client ODU_j connection can be nested into server layer

ODUk (j<k) connection. Control plane should provide mechanisms to support creation of such ODU hierarchy.

When creating server layer ODU LSP for carrying one specific client LSP, the first and last hop of the server LSP should be capable of selecting the correct link to make sure that both ends of the server LSP can support multiplexing/demultiplexing client signal into / from server LSP.

Therefore, the adaption information (e.g., hierarchical information and TSG) should be carried in the signaling to make the penultimate node of the FA-LSP to select the correct link for carrying the specific client signal.

- 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 ODU_j can be multiplexed into the same OTU_k. For example, both ODU₀ and ODU₁ may be multiplexed into ODU₂. An OTU link may support one or more types of ODU_j 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 ODU_{2e}, 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] (referred to as legacy nodes), while others support [G709-V3] and new OTN control plane characterized in this document (referred to as new nodes). In such case, control plane backward compatibility needs to be taken into consideration (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).

In order to provide backward compatibility, a new Switching Capability type is required for the control of [G709-V3] both in routing and signaling.

From a routing perspective, the advertisement of LSAs carrying new Switching Capability type implies the support of new OTN control plane protocols. A new node must support both legacy routing (i.e., the procedures defined in [RFC4203] with the switching capabilities defined in [RFC4328]) and new routing (i.e., the procedures defined for [G709-V3]), and should use new routing by default. When detecting the presence of a legacy node in the administrative domain (i.e., receiving LSAs carrying legacy Switching Capability type), the new node should advertise its links information by both the new and legacy routing approach, so that the legacy node can obtain the link resource information advertised by the new node.

On the other hand, from a signaling perspective, a new node must support both the legacy signaling procedures defined in [RFC4328] and the new procedures for control of [G709-V3]. Based on the routing information, a new node can determine whether its neighbor node is a legacy one or new one, so that it can determine which signaling procedure (new or legacy signaling procedure) needs to be performed. In case the new node has not enough information to know which signaling procedure its neighbor can support, it can use the new signaling procedure with the new Switching Capability type by default.

Since a legacy node receiving such message will respond with an error message indicating an unsupported Switching Capability type, the new node can perform the signaling again with a procedure [RFC4328] compliant.

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.

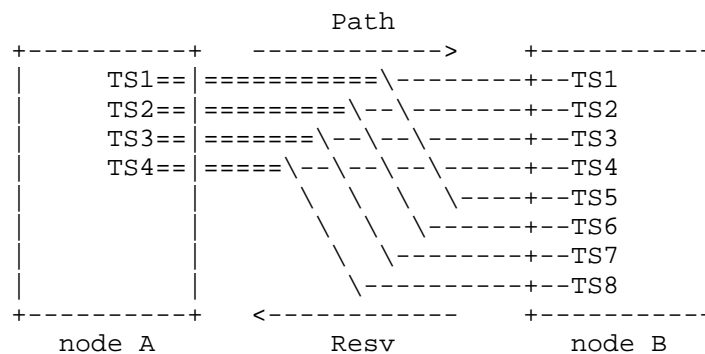


Figure 5 - Interworking between 1.25Gbps TS and 2.5Gbps TS

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 (WSO)", 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", 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 OTU_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.

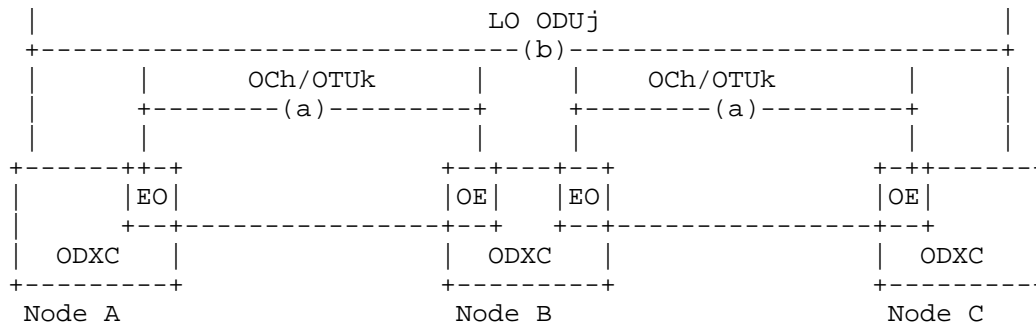


Figure 6 - Connection of LO ODUj (1)

In the case of LO ODUj multiplexing into HO ODUk, Figure 7 gives an example of this kind of LO ODU connection.

In Figure 7, OCh, OTUk, HO ODUk are associated with each other. The LO ODUj 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 ODUk and OTUk are similar to OCh. LO ODUj and OCh/OTUk/HO ODUk have client/server relationships. When a LO ODU connection is setup, it will be using the existing HO ODUk (/OTUk/OCh) connections which have been set up. Those HO ODUk 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.

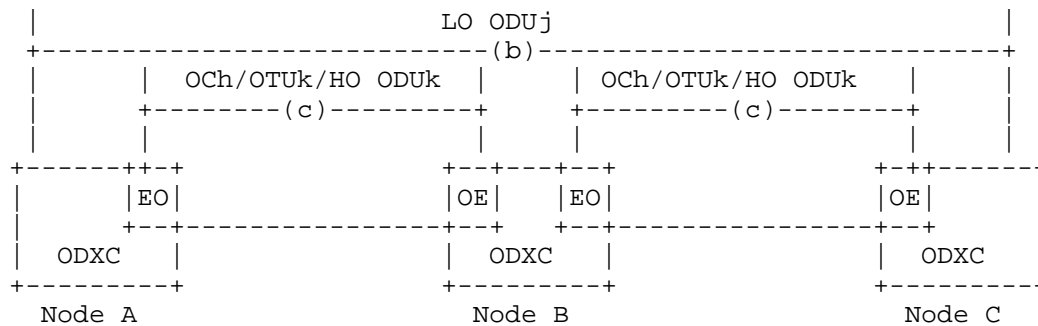


Figure 7 - Connection of LO ODUj (2)

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.

Network work group
Internet Draft
Intended status: Standards Track

Fatai Zhang
Young Lee
Jianrui Han
Huawei
G. Bernstein
Grotto Networking
Yunbin Xu
CATR

Expires: January 6, 2013

July 6, 2012

OSPF-TE Extensions for General Network Element Constraints

draft-ietf-ccamp-gmpls-general-constraints-ospf-te-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 January 6, 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.

Abstract

Generalized Multiprotocol Label Switching can be used to control a wide variety of technologies including packet switching (e.g., MPLS), time-division (e.g., SONET/SDH, OTN), wavelength (lambdas), and spatial switching (e.g., incoming port or fiber to outgoing port or fiber). In some of these technologies network elements and links may impose additional routing constraints such as asymmetric switch connectivity, non-local label assignment, and label range limitations on links. This document describes OSPF routing protocol extensions to support these kinds of constraints under the control of Generalized MPLS (GMPLS).

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. Node Information.....	3
2.1. Connectivity Matrix.....	4
3. Link Information.....	5
3.1. Port Label Restrictions.....	5
4. Routing Procedures.....	6
5. Scalability and Timeliness.....	6
5.1. Different Sub-TLVs into Multiple LSAs.....	7
5.2. Decomposing a Connectivity Matrix into Multiple Matrices..	7
6. Security Considerations.....	7
7. IANA Considerations.....	8
7.1. Node Information.....	8
7.2. Link Information.....	8

8. References.....	8
8.1. Normative References.....	8
8.2. Informative References.....	9
9. Authors' Addresses	9
Acknowledgment.....	11

1. Introduction

Some data plane technologies that wish to make use of a GMPLS control plane contain additional constraints on switching capability and label assignment. In addition, some of these technologies should be capable of performing non-local label assignment based on the nature of the technology, e.g., wavelength continuity constraint in WSON [RFC6163]. Such constraints can lead to the requirement for link by link label availability in path computation and label assignment.

[GEN-Encode] provides efficient encodings of information needed by the routing and label assignment process in technologies such as WSON and are potentially applicable to a wider range of technologies.

This document defines extensions to the OSPF routing protocol based on [GEN-Encode] to enhance the Traffic Engineering (TE) properties of GMPLS TE which are defined in [RFC3630], [RFC4202], and [RFC4203]. The enhancements to the Traffic Engineering (TE) properties of GMPLS TE links can be announced in OSPF TE LSAs. The TE LSA, which is an opaque LSA with area flooding scope [RFC3630], has only one top-level Type/Length/Value (TLV) triplet and has one or more nested sub-TLVs for extensibility. The top-level TLV can take one of three values (1) Router Address [RFC3630], (2) Link [RFC3630], (3) Generic Node Attribute defined in Section 2. In this document, we enhance the sub-TLVs for the Link TLV and define a new top-level TLV (Generic Node Attribute TLV) in support of the general network element constraints under the control of GMPLS.

The detailed encoding of OSPF extensions are not defined in this document. [GEN-Encode] provides encoding detail.

2. Node Information

According to [GEN-Encode], the additional node information representing node switching asymmetry constraints includes Node ID, connectivity matrix. Except for the Node ID which should comply with

Routing Address described in [RFC3630], the other pieces of information are defined in this document.

This document defines a new top TLV named the Generic Node Attribute TLV which carries attributes related to a general network element. This Generic Node Attribute TLV contains one or more sub-TLVs

Per [GEN-Encode], we have identified the following new Sub-TLVs to the Generic Node Attribute TLV. Detail description for each newly defined Sub-TLV is provided in subsequent sections:

Sub-TLV Type	Length	Name
TBD	variable	Connectivity Matrix

In some specific technologies, e.g., WSON networks, Connectivity Matrix sub-TLV may be optional, which depends on the control plane implementations. Usually, for example, in WSON networks, Connectivity Matrix sub-TLV may appear in the LSAs because WSON switches are asymmetric at present. It is assumed that the switches are symmetric switching, if there is no Connectivity Matrix sub-TLV in the LSAs.

2.1. Connectivity Matrix

It is necessary to identify which ingress ports and labels can be switched to some specific labels on a specific egress port, if the switching devices in some technology are highly asymmetric.

The Connectivity Matrix is used to identify these restrictions, which can represent either the potential connectivity matrix for asymmetric switches (e.g. ROADMs and such) or fixed connectivity for an asymmetric device such as a multiplexer as defined in [WSON-Info].

The Connectivity Matrix is a sub-TLV (the type is TBD by IANA) of the Generic Node Attribute TLV. The length is the length of value field in octets. The meaning and format of this sub-TLV are defined in Section 5.3 of [GEN-Encode]. One sub-TLV contains one matrix. The Connectivity Matrix sub-TLV may occur more than once to contain multi-matrices within the Generic Node Attribute TLV. In addition a large connectivity matrix can be decomposed into smaller separate matrices for transmission in multiple LSAs as described in Section 5.

3. Link Information

The most common link sub-TLVs nested to link top-level TLV are already defined in [RFC3630], [RFC4203]. For example, Link ID, Administrative Group, Interface Switching Capability Descriptor (ISCD), Link Protection Type, Shared Risk Link Group Information (SRLG), and Traffic Engineering Metric are among the typical link sub-TLVs.

Per [GEN-Encode], we add the following additional link sub-TLVs to the link-TLV in this document.

Sub-TLV Type	Length	Name
TBD	variable	Port Label Restrictions

Generally all the sub-TLVs above are optional, which depends on the control plane implementations. If it is default no restrictions on labels, Port Label Restrictions sub-TLV may not appear in the LSAs.

3.1. Port Label Restrictions

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.

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. These restrictions represent what wavelength may or may not be used on a link and are relatively static. The detailed information about Port label restrictions is described in [WSON-Info].

The Port Label Restrictions is a sub-TLV (the type is TBD by IANA) of the Link TLV. The length is the length of value field in octets. The meaning and format of this sub-TLV are defined in Section 5.4 of [GEN-Encode]. The Port Label Restrictions sub-TLV may occur more than once to specify a complex port constraint within the link TLV.

4. Routing Procedures

All the sub-TLVs are nested to top-level TLV(s) and contained in Opaque LSAs. The flooding of Opaque LSAs must follow the rules specified in [RFC2328], [RFC5250], [RFC3630], [RFC4203].

Considering the routing scalability issues in some cases, the routing protocol should be capable of supporting the separation of dynamic information from relatively static information to avoid unnecessary updates of static information when dynamic information is changed. A standard-compliant approach is to separate the dynamic information sub-TLVs from the static information sub-TLVs, each nested to top-level TLV ([RFC3630 and RFC5876]), and advertise them in the separate OSPF TE LSAs.

For node information, since the Connectivity Matrix information is static, the LSA containing the Generic Node Attribute TLV can be updated with a lower frequency to avoid unnecessary updates.

For link information, a mechanism MAY be applied such that static information and dynamic information of one TE link are contained in separate Opaque LSAs. For example, the Port Label Restrictions information sub-TLV could be nested to the top level link TLVs and advertised in the separate LSAs.

Note that as with other TE information, an implementation SHOULD take measures to avoid rapid and frequent updates of routing information that could cause the routing network to become swamped. A threshold mechanism MAY be applied such that updates are only flooded when a number of changes have been made to the label availability information (e.g., wavelength availability) within a specific time. Such mechanisms MUST be configurable if they are implemented.

5. Scalability and Timeliness

This document has defined four sub-TLVs for describing generic routing constraints. The examples given in [Gen-Encode] show that very large systems, in terms of label count or ports can be very efficiently encoded. However there has been concern expressed that some possible systems may produce LSAs that exceed the IP Maximum Transmission Unit (MTU) and that methods be given to allow for the splitting of general constraint LSAs into smaller LSA that are under the MTU limit. This section presents a set of techniques that can be used for this purpose.

5.1. Different Sub-TLVs into Multiple LSAs

Two sub-TLVs are defined in this document:

1. Connectivity Matrix (Generic Node Attribute TLV)
2. Port Label Restrictions (Link TLV)

Except for the Connectivity Matrix all these are carried in an Link TLV of which there can be at most one in an LSA [RFC3630]. Of these sub-TLVs the Port Label Restrictions are relatively static, i.e., only would change with hardware changes or significant system reconfiguration.

5.2. Decomposing a Connectivity Matrix into Multiple Matrices

In the highly unlikely event that a Connectivity matrix sub-TLV by itself would result in an LSA exceeding the MTU, a single large matrix can be decomposed into sub-matrices. Per [GEN-Encode] a connectivity matrix just consists of pairs of input and output ports that can reach each other and hence such this decomposition would be straightforward. Each of these sub-matrices would get a unique matrix identifier per [GEN-Encode].

From the point of view of a path computation process, prior to receiving an LSA with a Connectivity Matrix sub-TLV, no connectivity restrictions are assumed, i.e., the standard GMPLS assumption of any port to any port reachability holds. Once a Connectivity Matrix sub-TLV is received then path computation would know that connectivity is restricted and use the information from all Connectivity Matrix sub-TLVs received to understand the complete connectivity potential of the system. Prior to receiving any Connectivity Matrix sub-TLVs path computation may compute a path through the system when in fact no path exists. In between the reception of an additional Connectivity Matrix sub-TLV path computation may not be able to find a path through the system when one actually exists. Both cases are currently encountered and handled with existing GMPLS mechanisms. Due to the reliability mechanisms in OSPF the phenomena of late or missing Connectivity Matrix sub-TLVs would be relatively rare.

6. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC 3630], [RFC 4203].

7. IANA Considerations

[RFC3630] says that the top level Types in a TE LSA and Types for sub-TLVs for each top level Types must be assigned by Expert Review, and must be registered with IANA.

IANA is requested to allocate new Types for the TLV or sub-TLVs as defined in Sections 2 and 3 as follows:

7.1. Node Information

This document introduces a new Top Level Node TLV (Generic Node Attribute TLV) under the OSPF TE LSA defined in [RFC3630].

Value	TLV Type
TBA	Generic Node Attribute

This document also introduces the following sub-TLVs of Generic Node Attribute TLV:

Type	sub-TLV
TBD	Connectivity Matrix

7.2. Link Information

This document introduces the following sub-TLV of TE Link TLV (Value 2):

Type	sub-TLV
TBD	Port Label Restrictions

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.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC5250] L. Berger, I. Bryskin, A. Zinin, R. Coltun "The OSPF Opaque LSA Option", RFC 5250, July 2008.

- [RFC3630] Katz, D., Kompella, K., and Yeung, D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005
- [RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [GEN-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, " General Network Element Constraint Encoding for GMPLS Controlled Networks", work in progress: draft-ietf-ccamp-general-constraint-encode, May 2011.
- [RFC6205] T. Otani, H. Guo, K. Miyazaki, D. Caviglia, " Generalized Labels for Lambda-Switching Capable Label Switching Routers", RFC 6205, January 2011.

8.2. Informative References

- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks (WSN)", RFC 6163, February 2011.
- [WSN-Info] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-info, September 2011.

9. 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

Young Lee
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

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

Phone: +86-755-28977943
Email: hanjianrui@huawei.com

Greg Bernstein
Grotto Networking
Fremont CA, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.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.rtt.com.cn

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

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: danli@huawei.com

Ming Chen
European Research Center
Huawei Technologies
Riesstr. 25, 80992 Munchen, Germany

Phone: 0049-89158834072
Email: minc@huawei.com

Yabin Ye
European Research Center
Huawei Technologies
Riesstr. 25, 80992 Munchen, Germany

Phone: 0049-89158834074
Email: yabin.ye@huawei.com

Acknowledgment

We thank Ming Chen and Yabin Ye from DICONNET Project who provided valuable information for this 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.

Full Copyright Statement

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: October 15, 2012

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
April 13, 2012

Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS)
Control of Evolving G.709 OTN Networks
draft-ietf-ccamp-gmpls-ospf-g709v3-02

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 October 15, 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.

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	17
5.4. Example of single stage muxing	19
5.5. Example of multi stage muxing - Unbundled link	21
5.6. Example of multi stage muxing - Bundled links	22
5.7. Example of component links with non homogeneous hierarchies	24
6. Compatibility	26
7. Security Considerations	26
8. IANA Considerations	26
9. Contributors	27
10. Acknowledgements	28
11. References	28
11.1. Normative References	28
11.2. Informative References	29
Authors' Addresses	30

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].

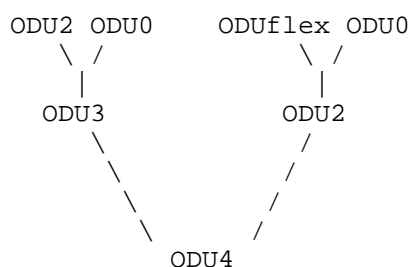
As discussed in [OTN-FWK] and [OTN-INFO], the OSPF-TE must be extended so to be able to advertise the termination and switching capabilities related to each different ODUj and ODUk/OTUk and the advertisement of related multiplexing capabilities. This leads to the need to define a new Switching Capability 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 ODUk TE-links.

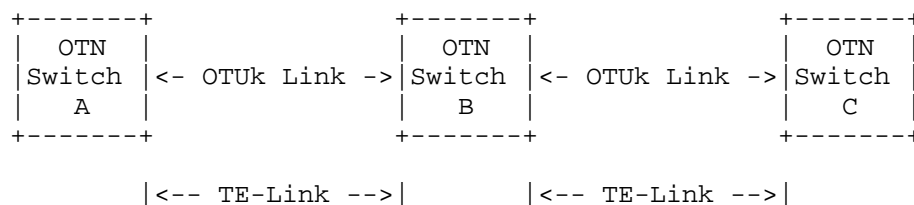


Figure 1: ODUk TE-Links

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 ODUk-TE-Links. As in the one hop case, these types of ODUk-TE-Links also advertise ODUj switching capacity. The advertised capacity could include ODUk switching capacity.

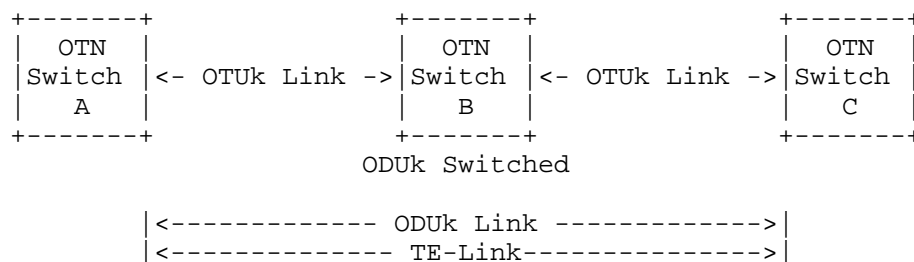


Figure 2: Multiple hop TE-Link

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. $0 \leq \text{Max LSP Bandwidth} \leq \text{ODUk/OTUk}$ 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

The format of the SCSI MUST be as depicted in the following figure:

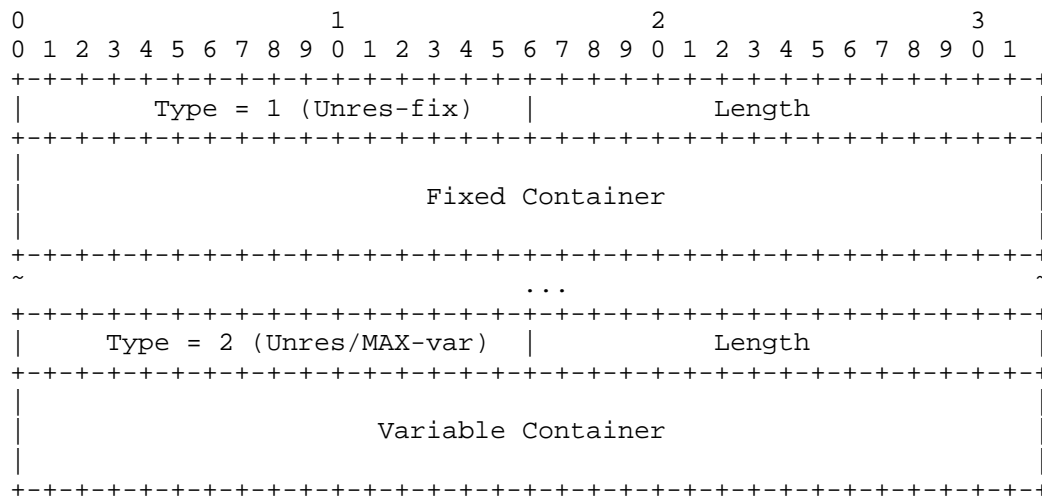


Figure 3: SCSI format

The format of the two different types of Bandwidth TLV are depicted in the following figures:

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								
+-----+																																							

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

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.

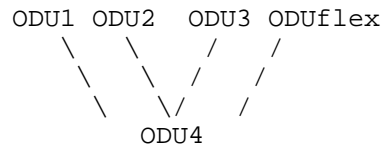
In the case of ODUflex advertisement the Type 2 Bandwidth TLV is used.

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								
+-----+																																							

Figure 11: 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:

```

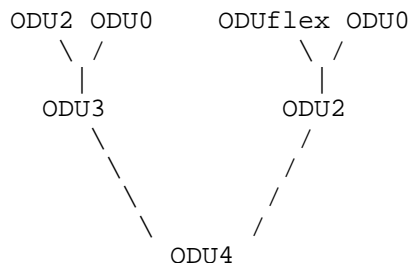
0                                     1                                     2                                     3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (Unres-fix)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|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 12: 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:

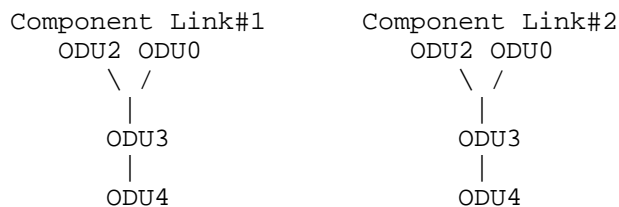
0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Type = 1 (Unres-fix)										Length																													
Sig type=ODU4										#stages= 0										T S TSG										Res 1 0 0 1 0 0 0 0									
Unres ODU4 at Prio 0 =1										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= 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 13: 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:



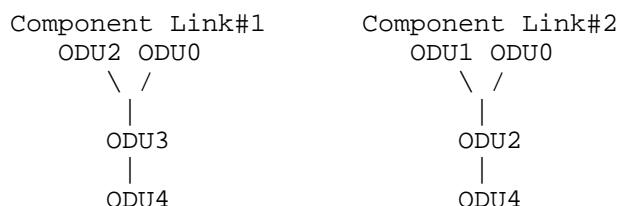
Considering only supported priorities 0 and 3, the advertisement is as follows:

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 14: 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Type = 1 (Unres-fix)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|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)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|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)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|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 15: Example 7 - Multi stage muxing - Non homogeneous hierarchies

6. Compatibility

All implementations of this document MUST support advertisements as defined in this document and [RFC4328]. Implementations SHOULD support the configuration of which advertisement procedures are followed. This approach enables the use of one type of advertisements when only type is needed, and the differentiation of legacy nodes and nodes supporting this document when both are operating in the same network.

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, Altran Italia

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

10. Acknowledgements

The authors would like to thank Fred Gruman 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
July 13, 2012

Expires: January 13, 2013

Generalized Multi-Protocol Label Switching (GMPLS) Signaling
Extensions for the evolving G.709 Optical Transport Networks Control

draft-ietf-ccamp-gmpls-signaling-g709v3-03.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 January 13, 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	6
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 Multiplexing Hierarchy	17
7.1. Extension to LSP_ATTRIBUTES Object	18
7.2. ODU FA-LSP Creation	19
8. Supporting Hitless Adjustment of ODUFlex (GFP)	20
9. Control Plane Backward Compatibility Considerations.....	21
10. Security Considerations.....	22
11. IANA Considerations.....	22

12. References	23
12.1. Normative References.....	23
12.2. Informative References.....	25
13. Contributors	25
14. Authors' Addresses	26
15. Acknowledgment	28

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 needed:

- 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 needs to 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 is 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 =
 ODTUk.ts nominal bit rate * (1 - HO OPUk bit rate tolerance)

Maximum bit rate of ODTUk.ts =
 ODTUk.ts nominal bit rate * (1 + 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.

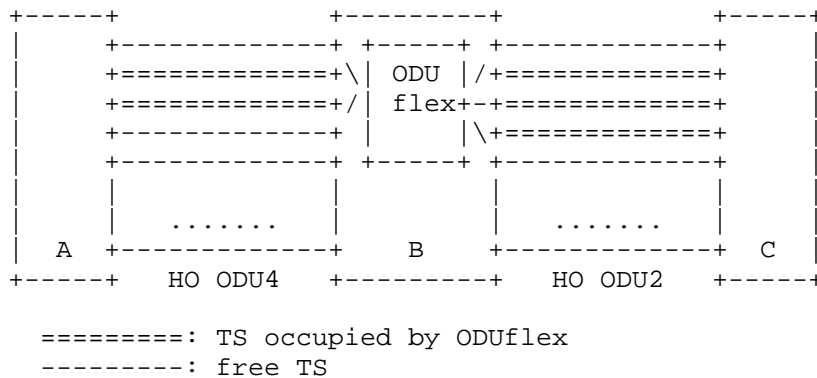


Figure 1 - Example of ODUflex(CBR) Traffic Parameters

- 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, we usually need 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 is not 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 can 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 does 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 is filled into the label.

In case of ODUk to OTUk mapping, the node only needs to fill the ODUj and the ODUk fields with corresponding values in the label. Other bits are reserved and MUST be set to 0.

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 Multiplexing Hierarchy

As described in [OTN-FWK], one ODUj connection can be nested into another ODUk (j < k) connection, which forms the multiplexing hierarchy in the ODU layer. This is useful if there are some intermediate nodes in the network which only support ODUk but not ODUj switching.

For example, in Figure 2, assume that N3 is a legacy node which only supports [G709-V1] and does not support ODU0 switching. If an ODU0 connection between N1 and N5 is required, then we can create an ODU2 connection between N2 and N4 (or ODU1 / ODU3 connection, depending on policies and the capabilities of the two ends of the connection), and nest the ODU0 into the ODU2 connection. In this way, N3 only needs to perform ODU2 switching and does not need to be aware of the ODU0 connection.

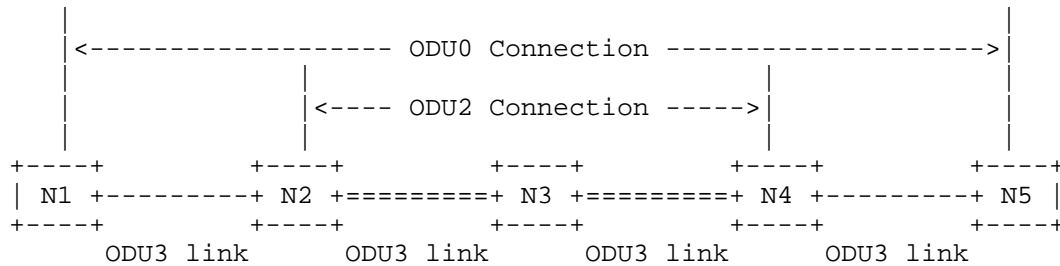


Figure 2 - Example of multiplexing hierarchy

The control plane signaling should support the provisioning of hierarchical multiplexing. Two methods are provided below (taking Figure 2 as example):

- Using the multi-layer network signaling described in [RFC4206], [RFC6107] and [RFC6001] (including related modifications, if needed). That is, when the signaling message for ODU0 connection arrives at N2, a new RSVP session between N2 and N4 is triggered to create the ODU2 connection. This ODU2 connection is treated as a Forwarding Adjacency (FA) after it is created. And then the signaling procedure for the ODU0 connection can be continued using the resource of the ODU2 FA.
- The ODU2 FA-LSP is created in advance based on network planning, which is treated as an FA. Then the ODU0 connection can be created using the resource of the ODU2 FA. In this case, the ODU2 FA-LSP and inner ODU0 connections are created separately.

For both methods, when creating an FA-LSP(e.g., ODU2 FA-LSP), the penultimate hop needs to choose a correct outgoing interface for the ODU2 connection, so that the destination node can support multiplexing and de-multiplexing LO ODU signal(e.g., ODU0). In order to choose a correct outgoing interface for the penultimate hop of the FA-LSP, multiplexing capability (i.e., what client signal type that can be adapted directly to this FA-LSP) should be carried in the signaling to setup this FA-LSP. In addition, when Auto_Negotiation in the data plane is not enabled, TS granularity may also be needed.

7.1. Extension to LSP_ATTRIBUTES Object

In order to indicate the adaptation information for a requested FA-LSP (i.e., the server layer LSP) to carry the client LSP, a new type of Attributes TLV of the LSP_ATTRIBUTES Object (Class-Num = 197, C-Type = 1, defined in [RFC5420]) 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								
+-----+																																							

One or more ODU adaptation TLVs can be carried to indicate the desired adaptation capabilities. Each of an ODU adaptation TLV for each branch of the client signal multiplexing supported by the server LSP MUST be used. Inside each TLV a row for each stage of the hierarchy MUST be included.

A row for the server stage MUST NOT be included as it is already signaled via the Traffic Parameters.

The number of stages is implicitly inferred from the length value.

Signal Type: as defined in [RFC4328] and this document.

For example, in order to create ODU3 FA-LSP passing through a set of ODU4 links to perform ODU1->ODU2->ODU3 hierarchy, the ODU adaptation TLV can be used to indicate the ODU2 into ODU3 multiplexing and ODU1 into ODU2 multiplexing stages.

Type = 2 (ODU adaptation)										Length = 8									
Reserved										Sig. = ODU2					Reserved				
Reserved										Sig. = ODU1					Reserved				

7.2. ODU FA-LSP Creation

When creating an ODU FA-LSP to carry lower ODU, the source node (e.g., node N2 in Figure 2) can include the LSP_ATTRIBUTES object to specify the desired ODU adaptation capabilities.

On receiving the Path message, the penultimate node on the FA-LSP (e.g., node N3 in Figure 2) MUST select an outgoing link which can support the TS granularity (indicated in the G-PID filed in Section 4) and the multiplexing hierarchy (listed in the LSP_ATTRIBUTES object). If no link supporting the specified hierarchy capabilities or TSG, a ParhErr message with Error Code = 38 (LSP Hierarchy Issue) and Error Value = y1(TBA) MUST be sent back to upstream.

Intermediate nodes (except end points and penultimate node) along the FA-LSP don't need to process the ODUflex adaptation TLV, which SHOULD be forwarded to the next node in the Path message without any modification.

8. 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 is 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" will be generated when performing bandwidth adjustment.

If any node along the ODUflex connection doesn't support hitless resizing, a Notify message with Error Code = x2 and Error Value = y1 will be sent to the source node. The source node MAY keep the connection and treat it as a non resizable ODUflex connection, or MAY tear it down, depending on the local policy.

- 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 can 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.

9. Control Plane Backward Compatibility Considerations

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 when the new nodes (supporting [G709-V3] and RSVP-TE extensions defined in this document) and the legacy nodes (supporting [G709-V1] and [RFC4328]) are interworking.

The backward compatibility needs to be considered only when controlling ODU1 or ODU2 or ODU3 connection, because legacy nodes can only support these three ODU signal types. In such case, new nodes can fall back to use signaling message defined in [RFC4328] when detecting legacy node on the path. More detailedly:

- o When receiving Path message using [RFC4328] (i.e., Switching Type = 100), a new node SHOULD follow [RFC4328] to process and reply it.
- o A source node of an ODU LSP can send Path message using new OTN control message (with new Switching Type = 101, TBA by IANA). If

there is legacy node on the LSP, it will fail to process the Generalized Label Request Object because of unknown of the new Switching Type, and reply a PathErr message indicating unknown of this object. The source node MAY re-signal the Path message using [RFC4328], depending on local policies.

- o Alternatively, if a new node has known that its neighbor only supports [RFC4328] in advance (e.g., through manual configuration or auto discovery mechanism), the new node MAY act as an RSVP agent to translate new RSVP-TE message into old one before sending to its neighbor.

No special compatibility consideration needs to be taken if the legacy device has updated its control plane to support this document.

10. 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.

11. 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.

- LSP_ATTRIBUTES Object:

New TLV with Type = 2 (TBA). This TLV is carried in the LSP_ATTRIBUTES Object (Class-Num = 197, C-Type = 1). See Section 7 for the detail definition.

- Error Code = 38 (LSP Hierarchy Issue, referring to [RFC6107]):

A new Error Value is added to the Error Code "LSP Hierarchy Issue":

Error Value	Error case
y1	Last hop of an ODU FA-LSP doesn't support specified adaptation capabilities (Section 7.2).

- Error Code = x2:

New Error Code, indicating errors occurring when controlling a resizable ODUflex connection.

Error Value	Error case
y1	Do not support hitless assignment of ODUflex (GFP) (Section 8).

12. References

12.1. 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.
- [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.
- [RFC5420] A. Farrel, Ed., "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC5420, February 2009.
- [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, June 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, January 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.
- [OTN-LMP] Fatai Zhang, Ed., "Link Management Protocol (LMP) extensions for G.709 Optical Transport Networks", draft-zhang-ccamp-gmpls-g.709-lmp-discovery, Work in Progress, July 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.

12.2. 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.

13. 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

14. 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

15. 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: Standards Track
Expires: September 7, 2012

F. Zhang, Ed.
ZTE
R. Jing
China Telecom
March 06, 2012

RSVP-TE Extensions for Associated Bidirectional LSPs
draft-ietf-ccamp-mppls-tp-rsvpte-ext-associated-lsp-03

Abstract

The MPLS Transport Profile (MPLS-TP) requirements document [RFC5654], describes that MPLS-TP MUST support associated bidirectional point-to-point LSPs.

This document provides a method to bind two unidirectional Label Switched Paths (LSPs) into an associated bidirectional LSP. The association is achieved by defining the new Association Types in the Extended ASSOCIATION object.

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 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.

Table of Contents

1. Introduction	3
2. Conventions used in this document	4
3. Overview	4
3.1. Provisioning Model	4
3.2. Signaling Procedure	4
3.2.1. Single Sided Provisioning Model	5
3.2.2. Double Sided Provisioning Model	5
3.2.3. Asymmetric Bandwidth LSPs	6
3.2.4. Recovery Considerations	6
3.2.5. Teardown of associated bidirectional LSPs	7
4. Association of LSPs	7
4.1. Association Types	7
4.2. REVERSE_LSP Object	8
4.2.1. Format	8
4.2.1.1. Subobjects	9
4.2.2. LSP Control	9
4.2.3. Updated RSVP Message Formats	9
4.2.4. Compatibility	10
5. IANA Considerations	10
5.1. Association Type	10
5.2. REVERSE_LSP Object	10
6. Security Considerations	10
7. Acknowledgement	11
8. References	11
8.1. Normative references	11
8.2. Informative References	12
Authors' Addresses	12

1. Introduction

The MPLS Transport Profile (MPLS-TP) requirements document [RFC5654] describes that MPLS-TP MUST support associated bidirectional point-to-point LSPs. Furthermore, an associated bidirectional LSP is useful for protection switching, for Operations, Administrations and Maintenance (OAM) messages that require a reply path.

The requirements described in [RFC5654] are specifically mentioned in Section 2.1. (General Requirements), and are repeated below:

7. MPLS-TP MUST support associated bidirectional point-to-point LSPs.

11. The end points of an associated bidirectional LSP MUST be aware of the pairing relationship of the forward and reverse LSPs used to support the bidirectional service.

12. Nodes on the LSP of an associated bidirectional LSP where both the forward and backward directions transit the same node in the same (sub)layer as the LSP SHOULD be aware of the pairing relationship of the forward and the backward directions of the LSP.

14. MPLS-TP MUST support bidirectional LSPs with asymmetric bandwidth requirements, i.e., the amount of reserved bandwidth differs between the forward and backward directions.

50. The MPLS-TP control plane MUST support establishing associated bidirectional P2P LSP including configuration of protection functions and any associated maintenance functions.

The above requirements are also repeated in [RFC6373].

The notion of association, as well as the corresponding Resource reSerVation Protocol (RSVP) ASSOCIATION object, is defined in [RFC4872], [RFC4873] and [I-D.ietf-ccamp-assoc-info]. In that context, the object is used to associate recovery LSPs with the LSP they are protecting. This object also has broader applicability as a mechanism to associate RSVP state, and [I-D.ietf-ccamp-assoc-ext] defines the Extended ASSOCIATION object that can be more generally applied.

This document provides a method to bind two reverse unidirectional Label Switched Paths (LSPs) into an associated bidirectional LSP. The association is achieved by defining the new Association Types in the Extended ASSOCIATION object.

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

3.1. Provisioning Model

The associated bidirectional LSP's forward and backward directions are set up, monitored, and protected independently as required by [RFC5654]. Configuration information regarding the LSPs can be sent to one end or both ends of the LSP. Depending on the method chosen, there are two models of signaling associated bidirectional LSP. The first model is the single sided provisioning, the second model is the double sided provisioning.

For the single sided provisioning, the configurations are sent to one end. Firstly, a unidirectional tunnel is configured on this end, then a LSP under this tunnel is initiated with the Extended ASSOCIATION object carried in the Path message to trigger the peer end to set up the corresponding reverse TE tunnel and LSP.

For the double sided provisioning, the two unidirectional TE tunnels are configured independently, then the LSPs under the tunnels are signaled with the Extended ASSOCIATION objects carried in the Path message to indicate each other to associate the two LSPs together to be an associated bidirectional LSP.

A number of scenarios exist for binding LSPs together to be an associated bidirectional LSP. These include: (1) both of them do not exist; (2) both of them exist; (3) one LSP exists, but the other one need to be established. In all scenarios described, the provisioning models discussed above are applicable.

3.2. Signaling Procedure

This section describes the signaling procedures for associating bidirectional LSPs.

Consider the topology described in Figure 1. (An example of associated bidirectional LSP). The LSP1 [via nodes A,D,B] (from A to B) and LSP2 [via nodes B,D,C,A] (from B to A) are being established or have been established, which can form an associated bidirectional LSP between node A and node B.

LSP1 and LSP2 are referenced at the data plane level by the identifiers: A-Node_ID::A-Tunnel_Num::A-LSP_Num::B-Node_ID and B-Node_ID::B-Tunnel_Num::B-LSP_Num::A-Node_ID, respectively [RFC6370].

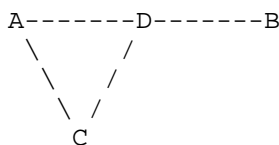


Figure 1: An example of associated bidirectional LSP

3.2.1. Single Sided Provisioning Model

For the single sided provisioning model, LSP1 is triggered by LSP2 or LSP2 is triggered by LSP1. When LSP2 is triggered by LSP1, LSP1 is initialized or refreshed (if LSP1 already exists) at node A with the Extended ASSOCIATION object inserted in the Path message, Association Type is set to "Single Sided Associated Bidirectional LSPs", Association ID set to a value that uniquely identifies the sessions to be associated within the context of the Association Source field (like A-Tunnel_Num), Association Source set to A-Node_ID, Global Association Source set to A-Global_ID. The Extended Association ID field must be included when the Association ID field is insufficient to uniquely identify association. As described in [I-D.ietf-ccamp-assoc-ext], when included, this field must be set to a value that, together with the other fields in the object, uniquely identifies the sessions to be associated. Terminating node B is triggered to set up LSP2 by the received Extended ASSOCIATION object with the Association Type set to the value "Single Sided Associated Bidirectional LSPs", the Association Object inserted in LSP2's Path message is the same as in LSP1's Path message.

When LSP1 is triggered by LSP2, the same rules are applicable. Based on the same values of the Association objects in the two LSPs' Path message, the two LSPs can be bound together to be an associated bidirectional LSP.

3.2.2. Double Sided Provisioning Model

For the double sided provisioning model, the Association Type must be set to "Double Sided Associated bidirectional LSPs" and the other values used in the Extended ASSOCIATION object are outside the scope of this document. For example, they may be communicated via the management plane. No matter how the values are communicated,

identification of the LSPs as being Associated Bidirectional LSPs occurs based on the identical contents in the LSPs' Extended ASSOCIATION objects.

3.2.3. Asymmetric Bandwidth LSPs

A variety of applications, such as internet services and the return paths of OAM messages, exist and which MAY have different bandwidth requirements for each direction. Additional [RFC5654] also specifies an asymmetric bandwidth requirement. This requirement is specifically mentioned in Section 2.1. (General Requirements), and is repeated below:

14. MPLS-TP MUST support bidirectional LSPs with asymmetric bandwidth requirements, i.e., the amount of reserved bandwidth differs between the forward and backward directions.

The approach for supporting asymmetric bandwidth co-routed bidirectional LSPs is defined in [RFC6387]. As to the asymmetric bandwidth associated bidirectional LSPs, the existing SENDER_TSPEC object must be carried in the REVERSE_LSP object as a sub-object in the initialized LSP's Path message to specify the reverse LSP's traffic parameters in case that single sided provisioning model is adopted. Consider the topology described in Figure 1 in the context of asymmetric associated bidirectional LSP, and take LSP2 triggered by LSP1 as an example. Node B is triggered to set up the reverse LSP2 with the corresponding asymmetric bandwidth by the Extended ASSOCIATION object with Association Type "Single Sided Associated Bidirectional LSPs" and the SENDER_TSPEC sub-object in LSP1's Path message, and the SENDER_TSPEC object in the LSP2' Path message is the same as the the SENDER_TSPEC sub-object in LSP1's Path message. When double sided provisioning model is used, the two opposite LSPs with asymmetric bandwidths are concurrently initialized, and this requirement will be satisfied simultaneously.

3.2.4. Recovery Considerations

Consider the topology described in Figure 1, LSP1 and LSP2 form the associated bidirectional LSP. Under the scenario of recovery, a third LSP (LSP3) may be used to protect LSP1. LSP3 can be established before or after the failure occurs, it can share the same TE tunnel with LSP1 or not.

When node A detects that LSP1 is broken, LSP3 will be initialized or refreshed with the Extended ASSOCIATION object inherited from LSP1's Path message. In this way, based on the same Extended ASSOCIATION object, LSP2 and LSP3 will compose the new associated bidirectional LSPs.

3.2.5. Teardown of associated bidirectional LSPs

Associated bidirectional LSPs teardown also follows standard procedures defined in [RFC3209] and [RFC3473] either without or with the administrative status. Note that teardown procedures of the associated bidirectional LSPs are independent of each other, so it is possible that while one LSP1 follows graceful teardown with administrative status, the other LSP2 is torn down without administrative status (using PathTear/ResvTear/PathErr with state removal). However, for the double sided associated bidirectional LSPs, the teardown of LSP1 does not mean that LSP2 must be deleted, which depends on the local policy. While for the single sided associated bidirectional LSPs, the teardown of the initialized LSP should induce the teardown of the trigger-established LSP, but the teardown of the trigger-established LSP (using PathErr with state removal) should not induce the teardown of the initialized LSP (which depends on the local policy).

4. Association of LSPs

4.1. Association Types

The Extended ASSOCIATION object is defined in [I-D.ietf-ccamp-assoc-ext], which enables MPLS-TP required identification. In order to bind two reverse unidirectional LSPs to be an associated bidirectional LSP, the new Association Types are defined in this document:

- o Association Types:

Value	Type
-----	-----
4 (TBD)	Doubled Sided Associated Bidirectional LSPs (D)
5 (TBD)	Single Sided Associated Bidirectional LSPs (A)

See [I-D.ietf-ccamp-assoc-ext] for the definition of other fields and values.

As described in [I-D.ietf-ccamp-assoc-ext], association is always done based on matching Path state or Resv state. Upstream initialized association is represented in Extended ASSOCIATION objects carried in Path message and downstream initialized association is represented in Extended ASSOCIATION objects carried in Resv messages. The new defined Association Types here are only used

in upstream initialized association. Thus they can only appear in Extended ASSOCIATION objects signaled in Path message.

The rules associated with the processing of the Extended ASSOCIATION objects in RSVP message are discussed in [I-D.ietf-ccamp-assoc-ext]. It said that in the absence of Association Type-specific rules for identifying association, the included Extended ASSOCIATION objects MUST be identical. This document adds no specific rules, the association will always operate based on the same Extended ASSOCIATION objects.

4.2. REVERSE_LSP Object

Path Computation Element (PCE)-based approaches, see [RFC4655], may be used for path computation of a GMPLS LSP, and consequently an associated bidirectional LSP, across domains and in a single domain. The ingress Label Switching Router (LSR), maybe serve as a PCE or Path Computation Client (PCC), has more information about the reverse LSP. When the forward LSP is signaled, the reverse LSP's traffic parameters, explicit route, LSP attributes, etc, can be carried in the REVERSE_LSP object of the forward LSP's Path message. The egress LSR can be triggered to establish the reverse LSP according to the control information.

4.2.1. Format

The information of the reverse LSP is specified via the REVERSE_LSP object, which is optional with class numbers in the form 11bbbbbbb has the following format:

Class = TBD (of the form 11bbbbbbb), C_Type = 1 (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
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                                                                    |
  //                               (Subobjects)                               //
  |                                                                    |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

```

This object MUST NOT be used when the Extended ASSOCIATION object do not exist or exist but the Association Type is not "Associated Bidirectional LSPs".

4.2.1.1. Subobjects

The contents of a REVERSE_LSP object are a series of variable-length data items called subobjects, which can be SENDER_TSPCE, EXPLICIT_ROUTE object (ERO), Session Attribute Object, Admin Status Object, LSP_ATTRIBUTES Object, LSP_REQUIRED_ATTRIBUTES Object, PROTECTION Object, ASSOCIATION Object, Extended ASSOCIATION Objects, etc.

4.2.2. LSP Control

The signaling procedure without the REVERSE_LSP object carried in the LSP1's Path message is described in section 3.2.1, which is the default option. A node includes a REVERSE_LSP object and Extended ASSOCIATION object with an "Associated Bidirectional LSPs" Association Type in an outgoing Path message when it wishes to control the reverse LSP, and the receiver node B MUST convert the subobjects of the REVERSE_LSP object into the corresponding objects that carried in LSP2's Path message. The case of a non-supporting egress node is outside of this document. If node A want to tear down the associated bidirectional LSP, a PathTear message will be sent out and Node B is triggered to tear down LSP2.

4.2.3. Updated RSVP Message Formats

This section presents the RSVP message-related formats as modified by this document. Unmodified RSVP message formats are not listed.

The format of a Path message is as follows:

```
<Path Message> ::= <Common Header> [ <INTEGRITY> ]
                    [ [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ] ... ]
                    [ <MESSAGE_ID> ]
                    <SESSION> <RSVP_HOP>
                    <TIME_VALUES>
                    [ <EXPLICIT_ROUTE> ]
                    <LABEL_REQUEST>
                    [ <PROTECTION> ]
                    [ <LABEL_SET> ... ]
                    [ <SESSION_ATTRIBUTE> ]
                    [ <NOTIFY_REQUEST> ... ]
                    [ <ADMIN_STATUS> ]
                    [ <EXTENDED_ASSOCIATION> ... ]
                    [ <REVERSE_LSP> ]
                    [ <POLICY_DATA> ... ]
                    <sender descriptor>
```

The format of the <sender descriptor> is not modified by the present document.

4.2.4. Compatibility

The REVERSE_LSP object is defined with class numbers in the form 11bbbbbb, which ensures compatibility with non-supporting nodes. Per [RFC2205], nodes not supporting this extension will ignore the object but forward it, unexamined and unmodified, in all messages resulting from this message. Especially, this object received in PathTear, or PathErr messages should be forwarded immediately in the same message, but should be saved with the corresponding state and forwarded in any refresh message resulting from that state when received in Path message.

5. IANA Considerations

IANA is requested to administer assignment of new values for namespace defined in this document and summarized in this section.

5.1. Association Type

Within the current document, two new Association Types are defined in the Extended ASSOCIATION object.

Value	Type
-----	-----
4 (TBD)	Double Sided Associated Bidirectional LSPs (D)
5 (TBD)	Single Sided Associated Bidirectional LSPs (A)

5.2. REVERSE_LSP Object

A new class named REVERSE_LSP has been created in the 11bbbbbb rang (TBD) with the following definition:

Class Types or C-types (1, TBD):

There are no other IANA considerations introduced by this document.

6. Security Considerations

This document introduces two new Association Types, and except this, there are no security issues about the Extended ASSOCIATION object are introduced here.

The procedures defined in this document result in an increase in the amount of topology information carried in signaling messages since the presence of the REVERSE_LSP object necessarily means that there is more information about associated bidirectional LSPs. Thus, in the event of the interception of a signaling message, slightly more could be deduced about the state of the network than was previously the case, but this is judged to be a very minor security risk as this information is already available via routing.

Otherwise, 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].

7. Acknowledgement

The authors would like to thank Lou Berger for his great guidance in this work, George Swallow and Jie Dong for the discussion of recovery, Lamberto Sterling for his valuable comments on the section of asymmetric bandwidths, Daniel King for the review of the document, Attila Takacs for the discussion of the provisioning model. At the same time, the authors would also like to acknowledge the contributions of Bo Wu, Xihua Fu, Lizhong Jin for the initial discussions, and Wenjuan He for the prototype implementation.

8. References

8.1. Normative references

- [I-D.ietf-ccamp-assoc-ext]
Berger, L., Faucheur, F., and A. Narayanan, "RSVP Association Object Extensions",
draft-ietf-ccamp-assoc-ext-02 (work in progress),
February 2012.
- [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.
- [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.

- [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.

8.2. Informative References

- [I-D.ietf-ccamp-assoc-info] Berger, L., "Usage of The RSVP Association Object", draft-ietf-ccamp-assoc-info-03 (work in progress), October 2011.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.
- [RFC6370] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport Profile (MPLS-TP) Identifiers", RFC 6370, September 2011.
- [RFC6373] Andersson, L., Berger, L., Fang, L., Bitar, N., and E. Gray, "MPLS Transport Profile (MPLS-TP) Control Plane Framework", RFC 6373, September 2011.
- [RFC6387] Takacs, A., Berger, L., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", RFC 6387, September 2011.

Authors' Addresses

Fei Zhang (editor)
ZTE

Email: zhang.fei3@zte.com.cn

Ruiquan Jing
China Telecom

Email: jingrq@ctbri.com.cn

Fan Yang
ZTE

Email: yang.fan5@zte.com.cn

Weilian Jiang
ZTE

Email: jiang.weilian@zte.com.cn

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.

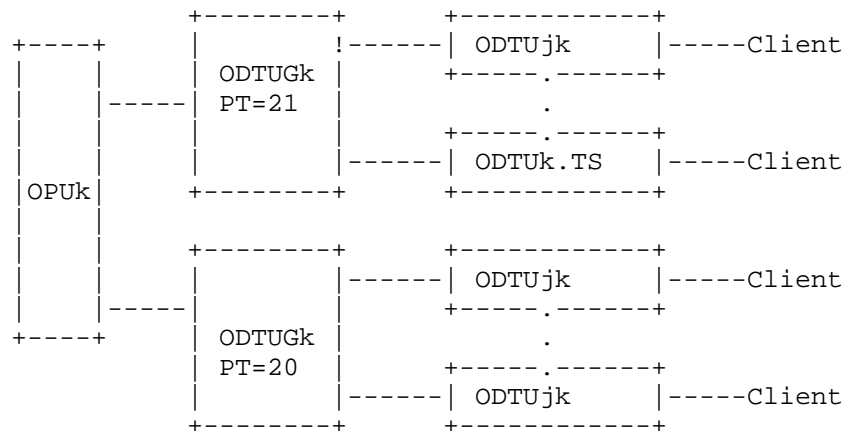


Figure 3: OTN client multiplexing

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.

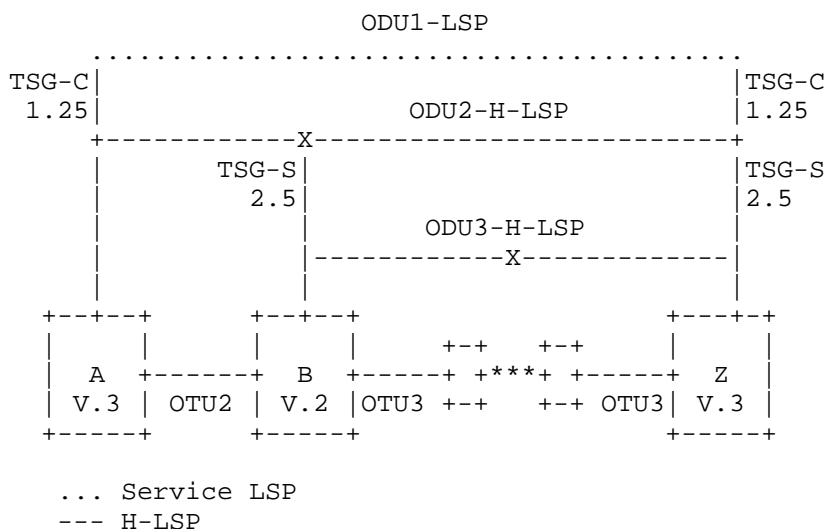


Figure 4: Client-Server TSG example

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.

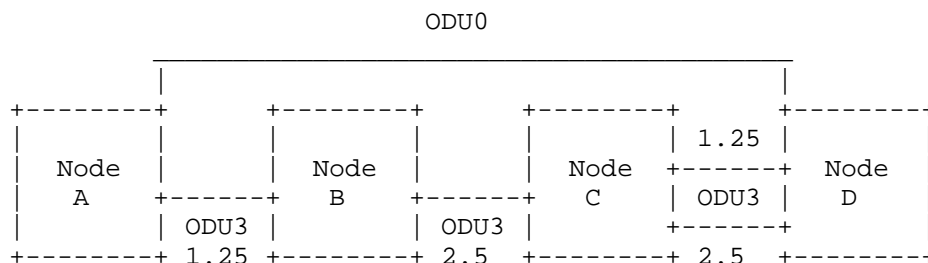


Figure 5: TSG in signaling

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	GMP PT=21
ODU1	AMP PT=20	AMP PT=20	-	AMP PT=21	AMP PT=21	GMP PT=21	GMP PT=21
ODU2	-	AMP PT=20	-	-	AMP PT=21	GMP PT=21	GMP PT=21
ODU2e	-	-	-	-	GMP PT=21	GMP PT=21	GMP PT=21
ODU3	-	-	-	-	-	-	GMP PT=21
ODUf1	-	-	-	GMP PT=21	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.

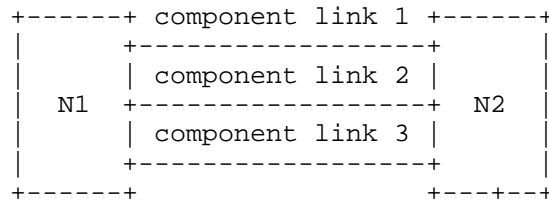


Figure 7: Concurrent path computation

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.

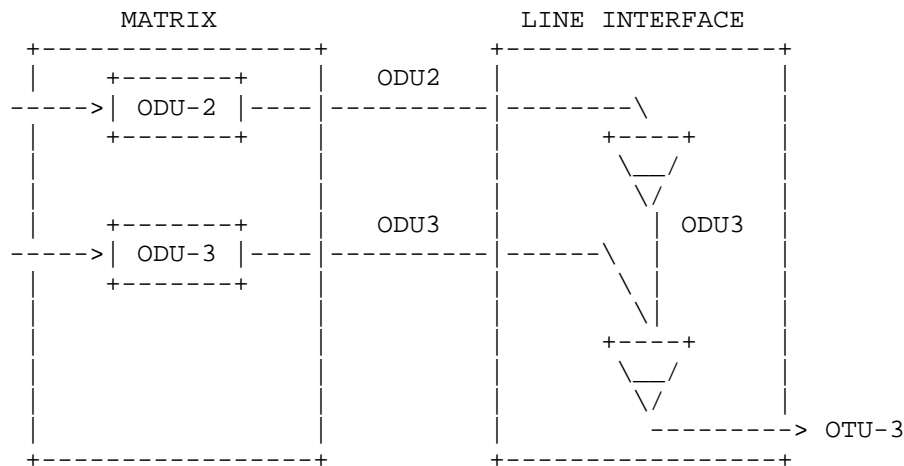


Figure 8: Switching and Terminating capabilities

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.

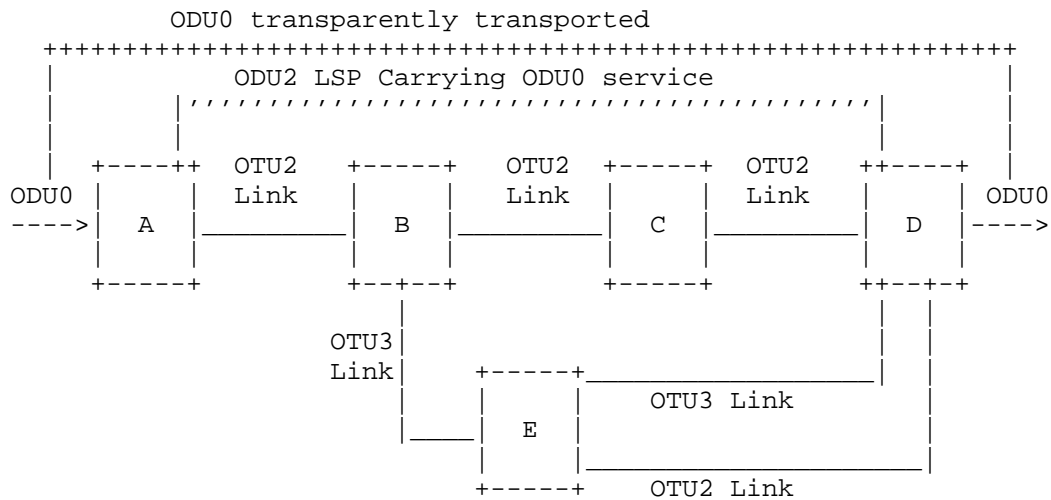


Figure 9: Switching and Terminating capabilities - Example 1

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.

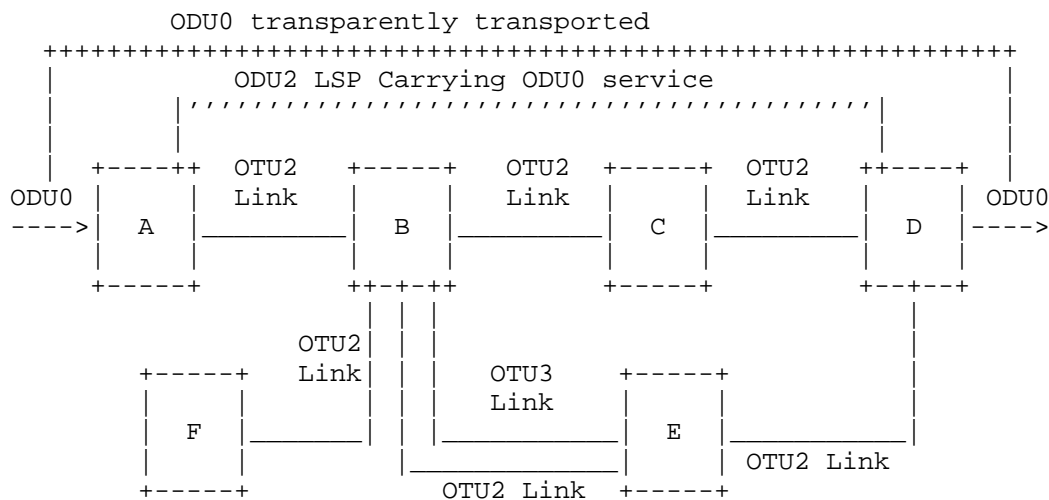


Figure 10: Switching and Terminating capabilities - Example 2

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: Informational
Expires: September 2012

Y. Lee
Huawei
G. Bernstein
Grotto Networking
D. Li
Huawei
W. Imajuku
NTT

March 6, 2012

Routing and Wavelength Assignment Information Model for Wavelength
Switched Optical Networks

draft-ietf-ccamp-rwa-info-14.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 September 6, 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.

Abstract

This document provides a model of information needed by the routing and wavelength assignment (RWA) process in wavelength switched optical networks (WSONs). The purpose of the information described in this model is to facilitate constrained lightpath computation in WSONs. This model takes into account compatibility constraints between WSON signal attributes and network elements but does not include constraints due to optical impairments. Aspects of this information that may be of use to other technologies utilizing a GMPLS control plane are discussed.

Table of Contents

1. Introduction.....	3
1.1. Revision History.....	4
1.1.1. Changes from 01.....	4
1.1.2. Changes from 02.....	4
1.1.3. Changes from 03.....	5
1.1.4. Changes from 04.....	5
1.1.5. Changes from 05.....	5
1.1.6. Changes from 06.....	5
1.1.7. Changes from 07.....	5
1.1.8. Changes from 08.....	5
1.1.9. Changes from 09.....	5
1.1.10. Changes from 10.....	6
1.1.11. Changes from 11.....	6
1.1.12. Changes from 12.....	6
1.1.13. Changes from 13.....	6
2. Terminology.....	6
3. Routing and Wavelength Assignment Information Model.....	7
3.1. Dynamic and Relatively Static Information.....	7

4. Node Information (General).....	8
4.1. Connectivity Matrix.....	8
4.2. Shared Risk Node Group.....	9
5. Node Information (WSON specific).....	9
5.1. Resource Accessibility/Availability.....	11
5.2. Resource Signal Constraints and Processing Capabilities..	14
5.3. Compatibility and Capability Details.....	15
5.3.1. Shared Input or Output Indication.....	15
5.3.2. Modulation Type List.....	15
5.3.3. FEC Type List.....	15
5.3.4. Bit Rate Range List.....	15
5.3.5. Acceptable Client Signal List.....	16
5.3.6. Processing Capability List.....	16
6. Link Information (General).....	16
6.1. Administrative Group.....	17
6.2. Interface Switching Capability Descriptor.....	17
6.3. Link Protection Type (for this link).....	17
6.4. Shared Risk Link Group Information.....	17
6.5. Traffic Engineering Metric.....	17
6.6. Port Label (Wavelength) Restrictions.....	18
6.6.1. Port-Wavelength Exclusivity Example.....	19
7. Dynamic Components of the Information Model.....	21
7.1. Dynamic Link Information (General).....	21
7.2. Dynamic Node Information (WSON Specific).....	21
8. Security Considerations.....	22
9. IANA Considerations.....	22
10. Acknowledgments.....	22
11. References.....	23
11.1. Normative References.....	23
11.2. Informative References.....	24
12. Contributors.....	25
Author's Addresses.....	26
Intellectual Property Statement.....	26
Disclaimer of Validity.....	27

1. Introduction

The purpose of the following information model for WSONs is to facilitate constrained lightpath computation and as such is not a general purpose network management information model. This constraint is frequently referred to as the "wavelength continuity" constraint, and the corresponding constrained lightpath computation is known as the routing and wavelength assignment (RWA) problem. Hence the information model must provide sufficient topology and wavelength restriction and availability information to support this computation. More details on the RWA process and WSON subsystems and their properties can be found in [RFC6163]. The model defined here

includes constraints between WSON signal attributes and network elements, but does not include optical impairments.

In addition to presenting an information model suitable for path computation in WSON, this document also highlights model aspects that may have general applicability to other technologies utilizing a GMPLS control plane. The portion of the information model applicable to other technologies beyond WSON is referred to as "general" to distinguish it from the "WSN-specific" portion that is applicable only to WSON technology.

1.1. Revision History

1.1.1. Changes from 01

Added text on multiple fixed and switched connectivity matrices.

Added text on the relationship between SRNG and SRLG and encoding considerations.

Added clarifying text on the meaning and use of port/wavelength restrictions.

Added clarifying text on wavelength availability information and how to derive wavelengths currently in use.

1.1.2. Changes from 02

Integrated switched and fixed connectivity matrices into a single "connectivity matrix" model. Added numbering of matrices to allow for wavelength (time slot, label) dependence of the connectivity. Discussed general use of this node parameter beyond WSON.

Integrated switched and fixed port wavelength restrictions into a single port wavelength restriction of which there can be more than one and added a reference to the corresponding connectivity matrix if there is one. Also took into account port wavelength restrictions in the case of symmetric switches, developed a uniform model and specified how general label restrictions could be taken into account with this model.

Removed the Shared Risk Node Group parameter from the node info, but left explanation of how the same functionality can be achieved with existing GMPLS SRLG constructs.

Removed Maximum bandwidth per channel parameter from link information.

1.1.3. Changes from 03

Removed signal related text from section 3.2.4 as signal related information is deferred to a new signal compatibility draft.

Removed encoding specific text from Section 3.3.1 of version 03.

1.1.4. Changes from 04

Removed encoding specific text from Section 4.1.

Removed encoding specific text from Section 3.4.

1.1.5. Changes from 05

Renumbered sections for clarity.

Updated abstract and introduction to encompass signal compatibility/generalization.

Generalized Section on wavelength converter pools to include electro optical subsystems in general. This is where signal compatibility modeling was added.

1.1.6. Changes from 06

Simplified information model for WSON specifics, by combining similar fields and introducing simpler aggregate information elements.

1.1.7. Changes from 07

Added shared fiber connectivity to resource pool modeling. This includes information for determining wavelength collision on an internal fiber providing access to resource blocks.

1.1.8. Changes from 08

Added PORT_WAVELENGTH_EXCLUSIVITY in the RestrictionType parameter. Added section 6.6.1 that has an example of the port wavelength exclusivity constraint.

1.1.9. Changes from 09

Section 5: clarified the way that the resource pool is modeled from blocks of identical resources.

Section 5.1: grammar fixes. Removed reference to "academic" modeling pre-print. Clarified RBNF resource pool model details.

Section 5.2: Formatting fixes.

1.1.10. Changes from 10

Enhanced the explanation of shared fiber access to resources and updated Figure 2 to show a more general situation to be modeled.

Removed all 1st person idioms.

1.1.11. Changes from 11

Replace all instances of "ingress" with "input" and all instances of "egress" with "output". Added clarifying text on relationship between resource block model and physical entities such as line cards.

1.1.12. Changes from 12

Section 5.2: Clarified RBNF optional elements for several definitions.

Section 5.3.6: Clarified RBNF optional elements for <ProcessingCapabilities>.

Editorial changes for clarity.

Update the contributor list.

1.1.13. Changes from 13

Section 7.1: Clarified that this information model does not dictate placement of information elements in protocols. In particular, added a caveat that the available label information element may be placed within the ISCD information element in the case of OSPF.

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 input and output line side ports as well as add/drop side ports.

RWA: Routing and Wavelength Assignment.

Wavelength Conversion. 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 Network (WSON): A WDM based optical network in which switching is performed selectively based on the center wavelength of an optical signal.

3. Routing and Wavelength Assignment Information Model

The following WSON RWA information model is grouped into four categories regardless of whether they stem from a switching subsystem or from a line subsystem:

- o Node Information
- o Link Information
- o Dynamic Node Information
- o Dynamic Link Information

Note that this is roughly the categorization used in [G.7715] section 7.

In the following, where applicable, the reduced Backus-Naur form (RBNF) syntax of [RBNF] is used to aid in defining the RWA information model.

3.1. Dynamic and Relatively Static Information

All the RWA information of concern in a WSON network is subject to change over time. Equipment can be upgraded; links may be placed in or out of service and the like. However, from the point of view of RWA computations there is a difference between information that can change with each successive connection establishment in the network

and that information that is relatively static on the time scales of connection establishment. A key example of the former is link wavelength usage since this can change with connection setup/teardown and this information is a key input to the RWA process. Examples of relatively static information are the potential port connectivity of a WDM ROADM, and the channel spacing on a WDM link.

This document separates, where possible, dynamic and static information so that these can be kept separate in possible encodings and hence allowing for separate updates of these two types of information thereby reducing processing and traffic load caused by the timely distribution of the more dynamic RWA WSON information.

4. Node Information (General)

The node information described here contains the relatively static information related to a WSON node. This includes connectivity constraints amongst ports and wavelengths since WSON switches can exhibit asymmetric switching properties. Additional information could include properties of wavelength converters in the node if any are present. In [Switch] it was shown that the wavelength connectivity constraints for a large class of practical WSON devices can be modeled via switched and fixed connectivity matrices along with corresponding switched and fixed port constraints. These connectivity matrices are included with the node information while the switched and fixed port wavelength constraints are included with the link information.

Formally,

```
<Node_Information> ::= <Node_ID> [<ConnectivityMatrix>...]
```

Where the Node_ID would be an appropriate identifier for the node within the WSON RWA context.

Note that multiple connectivity matrices are allowed and hence can fully support the most general cases enumerated in [Switch].

4.1. Connectivity Matrix

The connectivity matrix (ConnectivityMatrix) represents either the potential connectivity matrix for asymmetric switches (e.g. ROADMs and such) or fixed connectivity for an asymmetric device such as a multiplexer. Note that this matrix does not represent any particular internal blocking behavior but indicates which input/output ports and wavelengths could possibly be connected to a particular output port.

Representing internal state dependent blocking for a switch or ROADM is beyond the scope of this document and due to its highly implementation dependent nature would most likely not be subject to standardization in the future. The connectivity matrix is a conceptual M by N matrix representing the potential switched or fixed connectivity, where M represents the number of inputinput ports and N the number of outputoutput ports. This is a "conceptual" matrix since the matrix tends to exhibit structure that allows for very compact representations that are useful for both transmission and path computation [Encode].

Note that the connectivity matrix information element can be useful in any technology context where asymmetric switches are utilized.

ConnectivityMatrix ::= <MatrixID> <ConnType> <Matrix>

Where

<MatrixID> is a unique identifier for the matrix.

<ConnType> can be either 0 or 1 depending upon whether the connectivity is either fixed or potentially switched.

<Matrix> represents the fixed or switched connectivity in that Matrix(i, j) = 0 or 1 depending on whether inputinput port i can connect to outputoutput port j for one or more wavelengths.

4.2. Shared Risk Node Group

SRNG: Shared risk group for nodes. The concept of a shared risk link group was defined in [RFC4202]. This can be used to achieve a desired "amount" of link diversity. It is also desirable to have a similar capability to achieve various degrees of node diversity. This is explained in [G.7715]. Typical risk groupings for nodes can include those nodes in the same building, within the same city, or geographic region.

Since the failure of a node implies the failure of all links associated with that node a sufficiently general shared risk link group (SRLG) encoding, such as that used in GMPLS routing extensions can explicitly incorporate SRNG information.

5. Node Information (WSON specific)

As discussed in [RFC6163] a WSON node may contain electro-optical subsystems such as regenerators, wavelength converters or entire switching subsystems. The model present here can be used in

characterizing the accessibility and availability of limited resources such as regenerators or wavelength converters as well as WSON signal attribute constraints of electro-optical subsystems. As such this information element is fairly specific to WSON technologies.

A WSON node may include regenerators or wavelength converters arranged in a shared pool. As discussed in [RFC6163] this can include OEO based WDM switches as well. There are a number of different approaches used in the design of WDM switches containing regenerator or converter pools. However, from the point of view of path computation the following need to be known:

1. The nodes that support regeneration or wavelength conversion.
2. The accessibility and availability of a wavelength converter to convert from a given inputinput wavelength on a particular inputinput port to a desired outputoutput wavelength on a particular outputoutput port.
3. Limitations on the types of signals that can be converted and the conversions that can be performed.

Since resources tend to be packaged together in blocks of similar devices, e.g., on line cards or other types of modules, the fundamental unit of identifiable resource in this document is the "resource block". A resource block may contain one or more resources. As resources are the smallest identifiable unit of processing resource, one can group together resources into blocks if they have similar characteristics relevant to the optical system being modeled, e.g., processing properties, accessibility, etc.

This leads to the following formal high level model:

```
<Node_Information> ::= <Node_ID> [<ConnectivityMatrix>...]  
[<ResourcePool>]
```

Where

```
<ResourcePool> ::= <ResourceBlockInfo>...  
[<ResourceAccessibility>...] [<ResourceWaveConstraints>...]  
[<RBPoolState>]
```

First the accessibility of resource blocks is addressed then their properties are discussed.

5.1. Resource Accessibility/Availability

A similar technique as used to model ROADMs and optical switches can be used to model regenerator/converter accessibility. This technique was generally discussed in [RFC6163] and consisted of a matrix to indicate possible connectivity along with wavelength constraints for links/ports. Since regenerators or wavelength converters may be considered a scarce resource it is desirable that the model include, if desired, the usage state (availability) of individual regenerators or converters in the pool. Models that incorporate more state to further reveal blocking conditions on input or output to particular converters are for further study and not included here.

The three stage model is shown schematically in Figure 1 and Figure 2. The difference between the two figures is that Figure 1 assumes that each signal that can get to a resource block may do so, while in Figure 2 the access to sets of resource blocks is via a shared fiber which imposes its own wavelength collision constraint. The representation of Figure 1 can have more than one input to each resource block since each input represents a single wavelength signal, while in Figure 2 shows a single multiplexed WDM input/output or output, e.g., a fiber, to/from each set of block.

This model assumes N input ports (fibers), P resource blocks containing one or more identical resources (e.g. wavelength converters), and M output ports (fibers). Since not all input ports can necessarily reach each resource block, the model starts with a resource pool input matrix $RI(i,p) = \{0,1\}$ whether input port i can reach potentially reach resource block p .

Since not all wavelengths can necessarily reach all the resources or the resources may have limited input wavelength range the model has a set of relatively static input port constraints for each resource. In addition, if the access to a set of resource blocks is via a shared fiber (Figure 2) this would impose a dynamic wavelength availability constraint on that shared fiber. The resource block input port constraint is modeled via a static wavelength set mechanism and the case of shared access to a set of blocks is modeled via a dynamic wavelength set mechanism.

Next a state vector $RA(j) = \{0, \dots, k\}$ is used to track the number of resources in resource block j in use. This is the only state kept in the resource pool model. This state is not necessary for modeling "fixed" transponder system or full OEO switches with WDM interfaces, i.e., systems where there is no sharing.

After that, a set of static resource output wavelength constraints and possibly dynamic shared output fiber constraints maybe used. The static constraints indicate what wavelengths a particular resource block can generate or are restricted to generating e.g., a fixed regenerator would be limited to a single lambda. The dynamic constraints would be used in the case where a single shared fiber is used to output the resource block (Figure 2).

Finally, to complete the model, a resource pool output matrix $RE(p,k) = \{0,1\}$ depending on whether the output from resource block p can reach output port k , may be used.

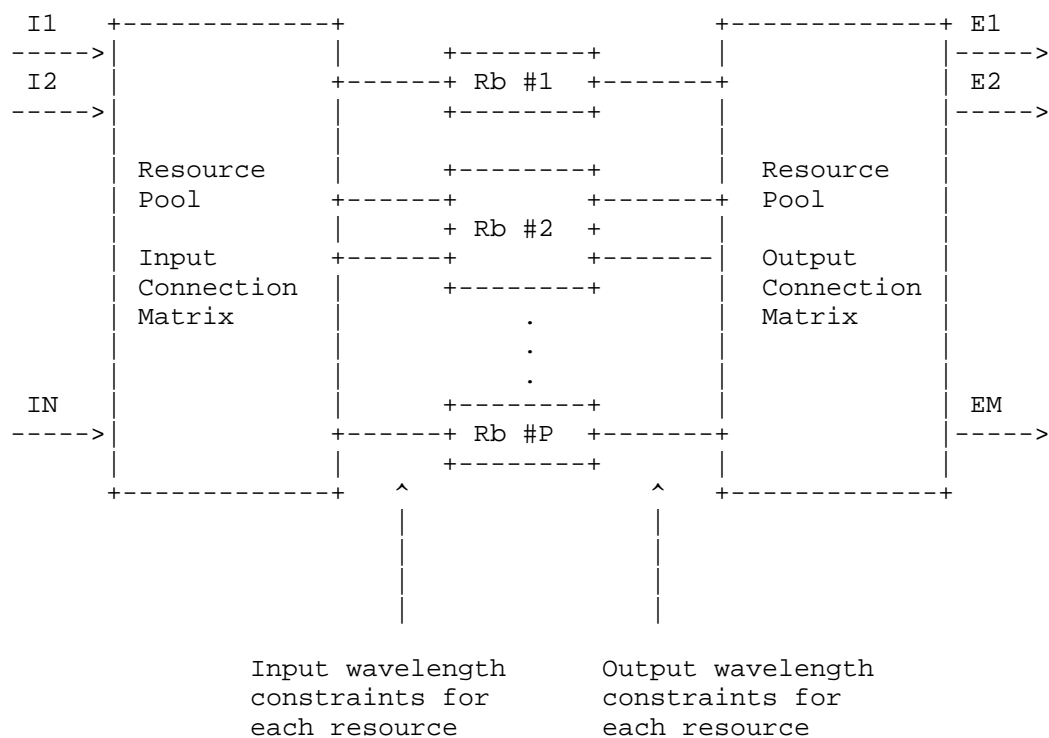


Figure 1 Schematic diagram of resource pool model.

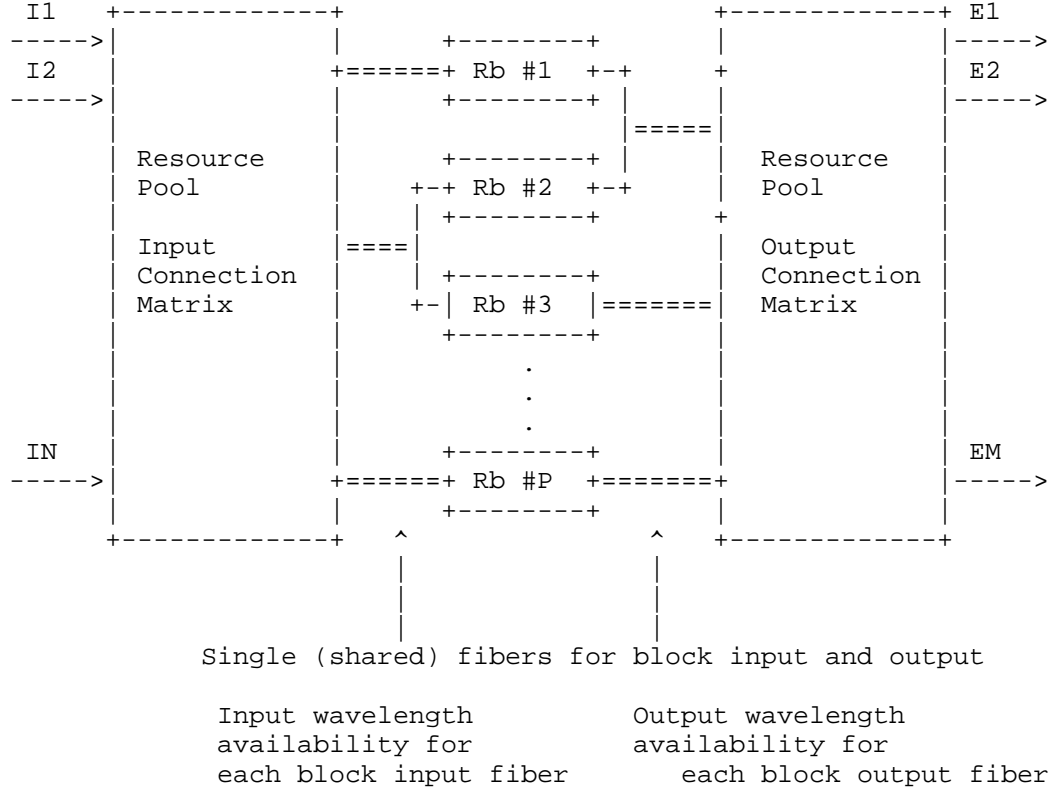


Figure 2 Schematic diagram of resource pool model with shared block accessibility.

Formally the model can be specified as:

```
<ResourceAccessibility ::= <PoolInputMatrix> <PoolOutputMatrix>
```

```
<ResourceWaveConstraints> ::= <InputWaveConstraints>  
<OutputOutputWaveConstraints>
```

```
<RBPoolState>  
::=(<ResourceBlockID><NumResourcesInUse><InAvailableWavelengths><Out  
AvailableWavelengths>)...
```

Note that except for <ResourcePoolState> all the other components of <ResourcePool> are relatively static. Also the <InAvailableWavelengths> and <OutAvailableWavelengths> are only used in the cases of shared input or output access to the particular block. See the resource block information in the next section to see how this is specified.

5.2. Resource Signal Constraints and Processing Capabilities

The wavelength conversion abilities of a resource (e.g. regenerator, wavelength converter) were modeled in the <OutputWaveConstraints> previously discussed. As discussed in [RFC6163] the constraints on an electro-optical resource can be modeled in terms of input constraints, processing capabilities, and output constraints:

```
<ResourceBlockInfo> ::= ([<ResourceSet>] <InputConstraints>
[<ProcessingCapabilities>] <OutputConstraints>)*
```

Where <ResourceSet> is a list of resource block identifiers with the same characteristics. If this set is missing the constraints are applied to the entire network element.

The <InputConstraints> are signal compatibility based constraints and/or shared access constraint indication. The details of these constraints are defined in section 5.3.

```
<InputConstraints> ::= <SharedInput> [<ModulationTypeList>]
[<FECTypeList>] [<BitRateRange>] [<ClientSignalList>]
```

The <ProcessingCapabilities> are important operations that the resource (or network element) can perform on the signal. The details of these capabilities are defined in section 5.3.

```
<ProcessingCapabilities> ::= [<NumResources>]
[<RegenerationCapabilities>] [<FaultPerfMon>] [<VendorSpecific>]
```

The <OutputConstraints> are either restrictions on the properties of the signal leaving the block, options concerning the signal properties when leaving the resource or shared fiber output constraint indication.

```
<OutputConstraints> := <SharedOutput> [<ModulationTypeList>]
[<FECTypeList>]
```

5.3. Compatibility and Capability Details

5.3.1. Shared Input or Output Indication

As discussed in the previous section and shown in Figure 2 the input or output access to a resource block may be via a shared fiber. The <SharedInput> and <SharedOutput> elements are indicators for this condition with respect to the block being described.

5.3.2. Modulation Type List

Modulation type, also known as optical tributary signal class, comes in two distinct flavors: (i) ITU-T standardized types; (ii) vendor specific types. The permitted modulation type list can include any mixture of standardized and vendor specific types.

```
<modulation-list>::=
[<STANDARD_MODULATION>|<VENDOR_MODULATION>]...
```

Where the STANDARD_MODULATION object just represents one of the ITU-T standardized optical tributary signal class and the VENDOR_MODULATION object identifies one vendor specific modulation type.

5.3.3. FEC Type List

Some devices can handle more than one FEC type and hence a list is needed.

```
<fec-list>::= [<FEC>]
```

Where the FEC object represents one of the ITU-T standardized FECs defined in [G.709], [G.707], [G.975.1] or a vendor-specific FEC.

5.3.4. Bit Rate Range List

Some devices can handle more than one particular bit rate range and hence a list is needed.

```
<rate-range-list>::= [<rate-range>]...
```

```
<rate-range>::=<START_RATE><END_RATE>
```

Where the START_RATE object represents the lower end of the range and the END_RATE object represents the higher end of the range.

5.3.5. Acceptable Client Signal List

The list is simply:

```
<client-signal-list> ::= [<GPID>]...
```

Where the Generalized Protocol Identifiers (GPID) object represents one of the IETF standardized GPID values as defined in [RFC3471] and [RFC4328].

5.3.6. Processing Capability List

The ProcessingCapabilities were defined in Section 5.2 as follows:

```
<ProcessingCapabilities> ::= [<NumResources>]  
[<RegenerationCapabilities>] [<FaultPerfMon>] [<VendorSpecific>]
```

The processing capability list sub-TLV is a list of processing functions that the WSON network element (NE) can perform on the signal including:

1. Number of Resources within the block
2. Regeneration capability
3. Fault and performance monitoring
4. Vendor Specific capability

Note that the code points for Fault and performance monitoring and vendor specific capability are subject to further study.

6. Link Information (General)

MPLS-TE routing protocol extensions for OSPF and IS-IS [RFC3630], [RFC5305] along with GMPLS routing protocol extensions for OSPF and IS-IS [RFC4203, RFC5307] provide the bulk of the relatively static link information needed by the RWA process. However, WSON networks bring in additional link related constraints. These stem from WDM line system characterization, laser transmitter tuning restrictions, and switching subsystem port wavelength constraints, e.g., colored ROADM drop ports.

In the following summarize both information from existing GMPLS route protocols and new information that maybe needed by the RWA process.

```
<LinkInfo> ::= <LinkID> [<AdministrativeGroup>]  
[<InterfaceCapDesc>] [<Protection>] [<SRLG>]...  
[<TrafficEngineeringMetric>] [<PortLabelRestriction>]
```

6.1. Administrative Group

AdministrativeGroup: Defined in [RFC3630]. Each set bit corresponds to one administrative group assigned to the interface. A link may belong to multiple groups. This is a configured quantity and can be used to influence routing decisions.

6.2. Interface Switching Capability Descriptor

InterfaceSwCapDesc: Defined in [RFC4202], lets us know the different switching capabilities on this GMPLS interface. In both [RFC4203] and [RFC5307] this information gets combined with the maximum LSP bandwidth that can be used on this link at eight different priority levels.

6.3. Link Protection Type (for this link)

Protection: Defined in [RFC4202] and implemented in [RFC4203, RFC5307]. Used to indicate what protection, if any, is guarding this link.

6.4. Shared Risk Link Group Information

SRLG: Defined in [RFC4202] and implemented in [RFC4203, RFC5307]. This allows for the grouping of links into shared risk groups, i.e., those links that are likely, for some reason, to fail at the same time.

6.5. Traffic Engineering Metric

TrafficEngineeringMetric: Defined in [RFC3630]. This allows for the definition of one additional link metric value for traffic engineering separate from the IP link state routing protocols link metric. Note that multiple "link metric values" could find use in optical networks, however it would be more useful to the RWA process to assign these specific meanings such as link mile metric, or probability of failure metric, etc...

6.6. Port Label (Wavelength) Restrictions

Port label (wavelength) restrictions (PortLabelRestriction) model the label (wavelength) restrictions that the link and various optical devices such as OXCs, ROADMs, and waveband multiplexers may impose on a port. These restrictions tell us what wavelength may or may not be used on a link and are relatively static. This plays an important role in fully characterizing a WSON switching device [Switch]. Port wavelength restrictions are specified relative to the port in general or to a specific connectivity matrix (section 4.1. Reference [Switch] gives an example where both switch and fixed connectivity matrices are used and both types of constraints occur on the same port. Such restrictions could be applied generally to other label types in GMPLS by adding new kinds of restrictions.

```
<PortLabelRestriction> ::= [<GeneralPortRestrictions>...]  
[<MatrixSpecificRestrictions>...]  
  
<GeneralPortRestrictions> ::= <RestrictionType>  
[<RestrictionParameters>]  
  
<MatrixSpecificRestriction> ::= <MatrixID> <RestrictionType>  
[<RestrictionParameters>]  
  
<RestrictionParameters> ::= [<LabelSet>...] [<MaxNumChannels>]  
[<MaxWaveBandWidth>]
```

Where

MatrixID is the ID of the corresponding connectivity matrix (section 4.1.

The RestrictionType parameter is used to specify general port restrictions and matrix specific restrictions. It can take the following values and meanings:

SIMPLE_WAVELENGTH: Simple wavelength set restriction; The wavelength set parameter is required.

CHANNEL_COUNT: The number of channels is restricted to be less than or equal to the Max number of channels parameter (which is required).

PORT_WAVELENGTH_EXCLUSIVITY: A wavelength can be used at most once among a given set of ports. The set of ports is specified as a parameter to this constraint.

WAVEBAND1: Waveband device with a tunable center frequency and passband. This constraint is characterized by the `MaxWaveBandWidth` parameters which indicates the maximum width of the waveband in terms of channels. Note that an additional wavelength set can be used to indicate the overall tuning range. Specific center frequency tuning information can be obtained from dynamic channel in use information. It is assumed that both center frequency and bandwidth (Q) tuning can be done without causing faults in existing signals.

Restriction specific parameters are used with one or more of the previously listed restriction types. The currently defined parameters are:

`LabelSet` is a conceptual set of labels (wavelengths).

`MaxNumChannels` is the maximum number of channels that can be simultaneously used (relative to either a port or a matrix).

`MaxWaveBandWidth` is the maximum width of a tunable waveband switching device.

`PortSet` is a conceptual set of ports.

For example, if the port is a "colored" drop port of a ROADM then there are two restrictions: (a) `CHANNEL_COUNT`, with `MaxNumChannels` = 1, and (b) `SIMPLE_WAVELENGTH`, with the wavelength set consisting of a single member corresponding to the frequency of the permitted wavelength. See [Switch] for a complete waveband example.

This information model for port wavelength (label) restrictions is fairly general in that it can be applied to ports that have label restrictions only or to ports that are part of an asymmetric switch and have label restrictions. In addition, the types of label restrictions that can be supported are extensible.

6.6.1. Port-Wavelength Exclusivity Example

Although there can be many different ROADM or switch architectures that can lead to the constraint where a lambda (label) maybe used at most once on a set of ports Figure 3 shows a ROADM architecture based on components known as a Wavelength Selective Switch (WSS)[OFC08]. This ROADM is composed of splitters, combiners, and WSSes. This ROADM has 11 output ports, which are numbered in the

diagram. Output ports 1-8 are known as drop ports and are intended to support a single wavelength. Drop ports 1-4 output from WSS #2, which is fed from WSS #1 via a single fiber. Due to this internal structure a constraint is placed on the output ports 1-4 that a lambda can be only used once over the group of ports (assuming uni-cast and not multi-cast operation). Similarly the output ports 5-8 have a similar constraint due to the internal structure.

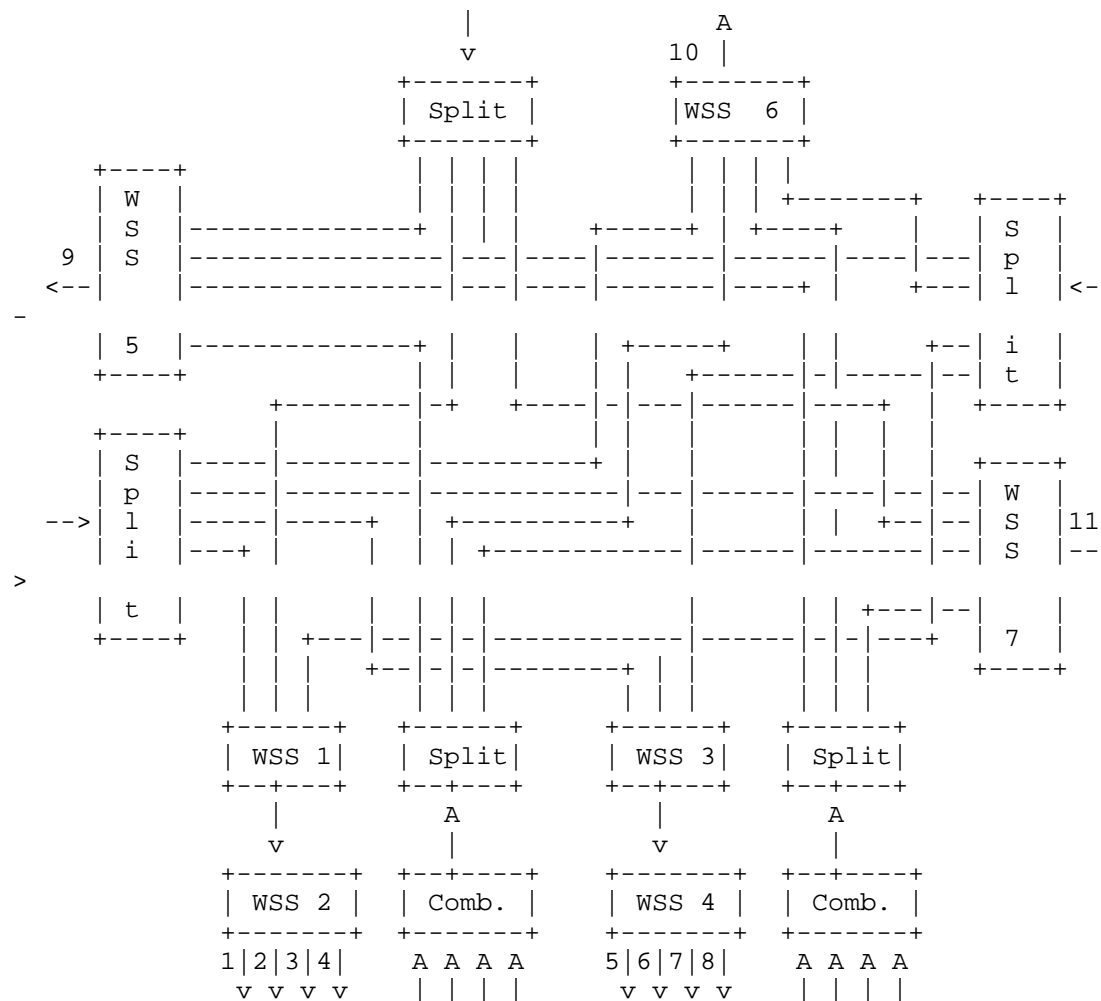


Figure 3 A ROADM composed from splitter, combiners, and WSSs.

7. Dynamic Components of the Information Model

In the previously presented information model there are a limited number of information elements that are dynamic, i.e., subject to change with subsequent establishment and teardown of connections. Depending on the protocol used to convey this overall information model it may be possible to send this dynamic information separate from the relatively larger amount of static information needed to characterize WSON's and their network elements.

7.1. Dynamic Link Information (General)

For WSON links wavelength availability and wavelengths in use for shared backup purposes can be considered dynamic information and hence are grouped with the dynamic information in the following set:

```
<DynamicLinkInfo> ::= <LinkID> <AvailableLabels>  
[<SharedBackupLabels>]
```

AvailableLabels is a set of labels (wavelengths) currently available on the link. Given this information and the port wavelength restrictions one can also determine which wavelengths are currently in use. This parameter could potential be used with other technologies that GMPLS currently covers or may cover in the future.

SharedBackupLabels is a set of labels (wavelengths) currently used for shared backup protection on the link. An example usage of this information in a WSON setting is given in [Shared]. This parameter could potential be used with other technologies that GMPLS currently covers or may cover in the future.

Note that the above does not dictate a particular encoding or placement for available label information. In some routing protocols it may be advantageous or required to place this information within another information element such as the interface switching capability descriptor (ISCD). Consult routing protocol specific extensions for details of placement of information elements.

7.2. Dynamic Node Information (WSON Specific)

Currently the only node information that can be considered dynamic is the resource pool state and can be isolated into a dynamic node information element as follows:

```
<DynamicNodeInfo> ::= <NodeID> [<ResourcePoolState>]
```

8. Security Considerations

This document discussed an information model for RWA computation in WSONs. Such a model is very similar from a security standpoint of the information that can be currently conveyed via GMPLS routing protocols. Such information includes network topology, link state and current utilization, and well as the capabilities of switches and routers within the network. As such this information should be protected from disclosure to unintended recipients. In addition, the intentional modification of this information can significantly affect network operations, particularly due to the large capacity of the optical infrastructure to be controlled.

9. IANA Considerations

This informational document does not make any requests for IANA action.

10. Acknowledgments

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

11. References

11.1. Normative References

- [Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-wson-encode.
- [G.707] ITU-T Recommendation G.707, Network node interface for the synchronous digital hierarchy (SDH), January 2007.
- [G.709] ITU-T Recommendation G.709, Interfaces for the Optical Transport Network(OTN), March 2003.
- [G.975.1] ITU-T Recommendation G.975.1, Forward error correction for high bit-rate DWDM submarine systems, February 2004.
- [RBNF] A. Farrel, "Reduced Backus-Naur Form (RBNF) A Syntax Used in Various Protocol Specifications", RFC 5511, April 2009.
- [RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005
- [RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4328] Papadimitriou, D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

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

11.2. Informative References

- [OFC08] P. Roorda and B. Collings, "Evolution to Colorless and Directionless ROADM Architectures," Optical Fiber communication/National Fiber Optic Engineers Conference, 2008. OFC/NFOEC 2008. Conference on, 2008, pp. 1-3.
- [Shared] G. Bernstein, Y. Lee, "Shared Backup Mesh Protection in PCE-based WSON Networks", iPOP 2008, http://www.grotto-networking.com/wson/iPOP2008_WSON-shared-mesh-poster.pdf .
- [Switch] G. Bernstein, Y. Lee, A. Gavler, J. Martensson, " Modeling WDM Wavelength Switching Systems for Use in GMPLS and Automated Path Computation", Journal of Optical Communications and Networking, vol. 1, June, 2009, pp. 187-195.
- [G.Sup39] ITU-T Series G Supplement 39, Optical system design and engineering considerations, February 2006.
- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", RFC 6163, April 2011.

12. Contributors

Diego Caviglia
Ericsson
Via A. Negrone 1/A 16153
Genoa Italy

Phone: +39 010 600 3736
Email: diego.caviglia@marconi.com, ericsson.com)

Anders Gavler
Acreo AB
Electrum 236
SE - 164 40 Kista Sweden

Email: Anders.Gavler@acreo.se

Jonas Martensson
Acreo AB
Electrum 236
SE - 164 40 Kista, Sweden

Email: Jonas.Martensson@acreo.se

Itaru Nishioka
NEC Corp.
1753 Simonumabe, Nakahara-ku, Kawasaki, Kanagawa 211-8666
Japan

Phone: +81 44 396 3287
Email: i-nishioka@cb.jp.nec.com

Lyndon Ong
Ciena
Email: lyong@ciena.com

Cyril Margaria
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany
Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com

Author's Addresses

Greg M. Bernstein (ed.)
Grotto Networking
Fremont California, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Young Lee (ed.)
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@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: danli@huawei.com

Wataru Imajuku
NTT Network Innovation Labs
1-1 Hikari-no-oka, Yokosuka, Kanagawa
Japan

Phone: +81-(46) 859-4315
Email: imajuku.wataru@lab.ntt.co.jp

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.

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: October 2012

G. Bernstein
Grotto Networking
Y. Lee
D. Li
Huawei
W. Imajuku
NTT

April 24, 2012

Routing and Wavelength Assignment Information Encoding for
Wavelength Switched Optical Networks

draft-ietf-ccamp-rwa-wson-encode-14.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 September 24, 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.

Abstract

A wavelength switched optical network (WSON) requires that certain key information elements are made available to facilitate path computation and the establishment of label switching paths (LSPs). The information model described in "Routing and Wavelength Assignment Information for Wavelength Switched Optical Networks" shows what information is required at specific points in the WSON. Part of the WSON information model contains aspects that may be of general applicability to other technologies, while other parts are fairly specific to WSONs.

This document provides efficient, protocol-agnostic encodings for the WSON specific information elements. It is intended that protocol-specific documents will reference this memo to describe how information is carried for specific uses. Such encodings can be used to extend GMPLS signaling and routing protocols. In addition these encodings could be used by other mechanisms to convey this same information to a path computation element (PCE).

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.....	4
1.1. Revision History.....	4
1.1.1. Changes from 00 draft.....	4
1.1.2. Changes from 01 draft.....	5
1.1.3. Changes from 02 draft.....	5

1.1.4. Changes from 03 draft.....	5
1.1.5. Changes from 04 draft.....	5
1.1.6. Changes from 05 draft.....	5
1.1.7. Changes from 06 draft.....	5
1.1.8. Changes from 07 draft.....	6
1.1.9. Changes from 08 draft.....	6
1.1.10. Changes from 09 draft.....	6
1.1.11. Changes from 10 draft.....	6
1.1.12. Changes from 11 draft.....	6
1.1.13. Changes from 12 draft.....	6
1.1.14. Changes from 13 draft.....	6
2. Terminology.....	7
3. Resources, Blocks, Sets, and the Resource Pool.....	7
3.1. Resource Block Set Field.....	8
4. Resource Pool Accessibility/Availability.....	10
4.1. Resource Pool Accessibility Sub-TLV.....	10
4.2. Resource Block Wavelength Constraints Sub-TLV.....	12
4.3. Resource Pool State Sub-TLV.....	14
4.4. Block Shared Access Wavelength Availability sub-TLV.....	15
5. Resource Properties Encoding.....	16
5.1. Resource Block Information Sub-TLV.....	17
5.2. Modulation Format List Sub-Sub-TLV.....	17
5.2.1. Modulation Format Field.....	18
5.3. FEC Type List Sub-Sub-TLV.....	20
5.3.1. FEC Type Field.....	20
5.4. Input Bit Range List Sub-Sub-TLV.....	22
5.4.1. Bit Range Field.....	22
5.5. Input Client Signal List Sub-Sub-TLV.....	23
5.6. Processing Capability List Sub-Sub-TLV.....	24
5.6.1. Processing Capabilities Field.....	24
6. Security Considerations.....	26
7. IANA Considerations.....	26
8. Acknowledgments.....	26
APPENDIX A: Encoding Examples.....	27
A.1. Wavelength Converter Accessibility Sub-TLV.....	27
A.2. Wavelength Conversion Range Sub-TLV.....	29
A.3. An OEO Switch with DWDM Optics.....	29
9. References.....	32
9.1. Normative References.....	32
9.2. Informative References.....	32
10. Contributors.....	34
Authors' Addresses.....	35
Intellectual Property Statement.....	36
Disclaimer of Validity.....	36

1. Introduction

A Wavelength Switched Optical Network (WSON) is a Wavelength Division Multiplexing (WDM) optical network in which switching is performed selectively based on the center wavelength of an optical signal.

[RFC6163] describes a framework for Generalized Multiprotocol Label Switching (GMPLS) and Path Computation Element (PCE) control of a WSON. Based on this framework, [WSON-Info] describes an information model that specifies what information is needed at various points in a WSON in order to compute paths and establish Label Switched Paths (LSPs).

This document provides efficient encodings of information needed by the routing and wavelength assignment (RWA) process in a WSON. Such encodings can be used to extend GMPLS signaling and routing protocols. In addition these encodings could be used by other mechanisms to convey this same information to a path computation element (PCE). Note that since these encodings are relatively efficient they can provide more accurate analysis of the control plane communications/processing load for WSONs looking to utilize a GMPLS control plane.

Note that encodings of information needed by the routing and label assignment process applicable to general networks beyond WSON are addressed in a separate document [Gen-Encode]. This document makes use of the Label Set Field encoding of [Gen-Encode] and refers to it as a Wavelength Set Field.

1.1. Revision History

1.1.1. Changes from 00 draft

Edits to make consistent with update to [RFC6205], i.e., removal of sign bit.

Clarification of TBD on connection matrix type and possibly numbering.

New sections for wavelength converter pool encoding: Wavelength Converter Set Sub-TLV, Wavelength Converter Accessibility Sub-TLV, Wavelength Conversion Range Sub-TLV, WC Usage State Sub-TLV.

Added optional wavelength converter pool TLVs to the composite node TLV.

1.1.2. Changes from 01 draft

The encoding examples have been moved to an appendix. Classified and corrected information elements as either reusable fields or sub-TLVs. Updated Port Wavelength Restriction sub-TLV. Added available wavelength and shared backup wavelength sub-TLVs. Changed the title and scope of section 6 to recommendations since the higher level TLVs that this encoding will be used in is somewhat protocol specific.

1.1.3. Changes from 02 draft

Removed inconsistent text concerning link local identifiers and the link set field.

Added E bit to the Wavelength Converter Set Field.

Added bidirectional connectivity matrix example. Added simple link set example. Edited examples for consistency.

1.1.4. Changes from 03 draft

Removed encodings for general concepts to [Gen-Encode].

Added in WSON signal compatibility and processing capability information encoding.

1.1.5. Changes from 04 draft

Added encodings to deal with access to resource blocks via shared fiber.

1.1.6. Changes from 05 draft

Revised the encoding for the "shared access" indicators to only use one bit each for input and output.

1.1.7. Changes from 06 draft

Removed section on "WSON Encoding Usage Recommendations"

1.1.1.8. Changes from 07 draft

Section 3: Enhanced text to clarify relationship between pools, blocks and resources. Section 3.1, 3.2: Change title to clarify Pool-Block relationship. Section 3.3: clarify block-resource state.

Section 4: Deleted reference to previously removed RBNF element. Fixed TLV figures and descriptions for consistent sub-sub-TLV nomenclature.

1.1.1.9. Changes from 08 draft

Fixed ordering of fields in second half of sub-TLV example in Appendix A.1.

Clarifying edits in section 3 on pools, blocks, and resources.

1.1.1.10. Changes from 09 draft

Fixed the "Block Shared Access Wavelength Availability sub-TLV" of section 3.4 to use an "RB set field" rather than a single RB ID. Removed all 1st person idioms.

1.1.1.11. Changes from 10 draft

Removed remaining 1st person idioms. Updated IANA section. Update references for newly issued RFCs.

1.1.1.12. Changes from 11 draft

Fixed length fields in section 4 to be 16 bits, correcting errors in TLV and field figures. Added a separate section on resources, blocks, sets and the resource pool. Moved definition of the resource block set field to this new section.

1.1.1.13. Changes from 12 draft

Replaced all instances of "ingress" with "input" and all instances of "egress" with "output".

1.1.1.14. Changes from 13 draft

C bit of Resource Block Set Field is redundant and was removed, i.e., has been returned to "Reserved" block and appendix examples were updated to reflect the change.

Enhanced section 4.2 encoding to allow for optionality of input or output wavelength set fields.

Clarified that wavelength set fields use the Label Set field encoding from [Gen-Encode].

Enhanced section 5.1 encoding to simplify the Modulation and FEC input and output cases.

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 input and output line side ports as well as add/drop side ports.

RWA: Routing and Wavelength Assignment.

Wavelength Conversion. 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 Network (WSO): A WDM based optical network in which switching is performed selectively based on the center wavelength of an optical signal.

3. Resources, Blocks, Sets, and the Resource Pool

The optical system to be encoded may contain a pool of resources of different types and properties for processing optical signals. For the purposes here a "resource" is an individual entity such as a wavelength converter or regenerator within the optical node that acts on an individual wavelength signal.

Since resources tend to be packaged together in blocks of similar devices, e.g., on line cards or other types of modules, the fundamental unit of identifiable resource in this document is the "resource block". A resource block may contain one or more resources. As resource blocks are the smallest identifiable unit of processing resource, one should group together resources into blocks if they have similar characteristics relevant to the optical system being modeled, e.g., processing properties, accessibility, etc.

This document defines the following sub-TLVs pertaining to resources within an optical node:

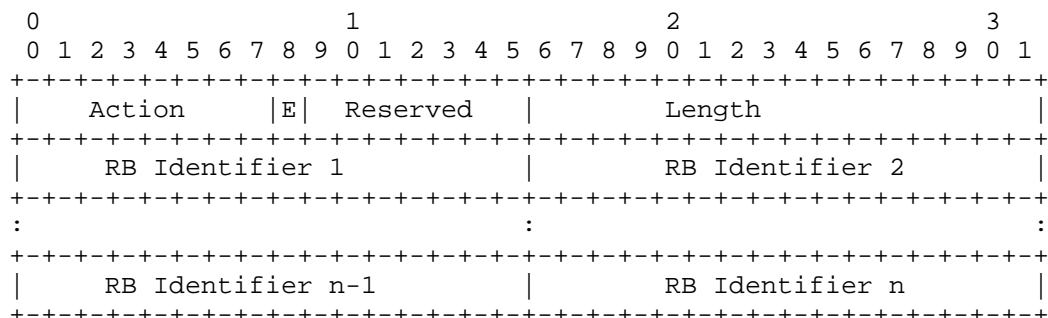
- . Resource Pool Accessibility Sub-TLV
- . Resource Block Wavelength Constraints Sub-TLV
- . Resource Pool State Sub-TLV
- . Block Shared Access Wavelength Availability Sub-TLV
- . Resource Block Information Sub-TLV

Each of these sub-TLVs works with one or more sets of resources rather than just a single resource block. This motivates the following field definition.

3.1. Resource Block Set Field

In a WSON node that includes resource blocks (RB), denoting subsets of these blocks allows one to efficiently describe common properties of the blocks and to describe the structure and characteristics, if non-trivial, of the resource pool. The RB Set field is defined in a similar manner to the label set concept of [RFC3471].

The information carried in a RB set field is defined by:



Action: 8 bits

0 - Inclusive List

Indicates that the TLV contains one or more RB elements that are included in the list.

2 - Inclusive Range(s)

Indicates that the TLV contains one or more ranges of RBs. Each individual range is denoted by two 16 bit RB identifiers in a 32 bit word. The first 16 bits is the RB identifier for the start of the range and the next 16 bits is the RB identifier for the end of the range. Note that the Length field is used to determine the number of ranges.

E (Even bit): Set to 0 denotes an odd number of RB identifiers in the list (last entry zero pad); Set to 1 denotes an even number of RB identifiers in the list (no zero padding). This applies only if Action == 0.

Reserved: 7 bits

This field is reserved. It MUST be set to zero on transmission and MUST be ignored on receipt.

Length: 16 bits

The total length of this field in bytes.

RB Identifier:

The RB identifier represents the ID of the resource block which is a 16 bit integer.

Usage Note: the inclusive range "Action" can result in very compact encoding of resource sets and it can be advantages to number resource blocks in such a way so that status updates (dynamic information) can take advantage of this efficiency.

4. Resource Pool Accessibility/Availability

This section defines the sub-TLVs for dealing with accessibility and availability of resource blocks within a pool of resources. These include the ResourceBlockAccessibility, ResourceWaveConstraints, and RBPoolState sub-TLVs.

4.1. Resource Pool Accessibility Sub-TLV

This sub-TLV describes the structure of the resource pool in relation to the switching device. In particular it indicates the ability of an input port to reach sets of resources and of a sets of resources to reach a particular output port. This is the PoolInputMatrix and PoolOutputMatrix of [WSON-Info].

The resource pool accessibility sub-TLV is defined by:

```

      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 |                               Reserved |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Input Link Set Field A #1 |
:                                                         :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               RB Set Field A #1         |
:                                                         :
+-----+-----+-----+-----+-----+-----+-----+-----+
| Additional Link set and RB set pairs as needed to      |
: specify PoolInputMatrix                                :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Output Link Set Field B #1 |
:                                                         :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               RB Set B Field #1 (for output connectivity) |
:                                                         :
+-----+-----+-----+-----+-----+-----+-----+-----+
| Additional Link Set and RB set pairs as needed to      |
: specify PoolOutputMatrix                                :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where

Connectivity indicates how the input/output ports connect to the resource blocks.

0 -- the device is fixed (e.g., a connected port must go through the resource block)

1 -- the device is switched (e.g., a port can be configured to go through a resource but isn't required)

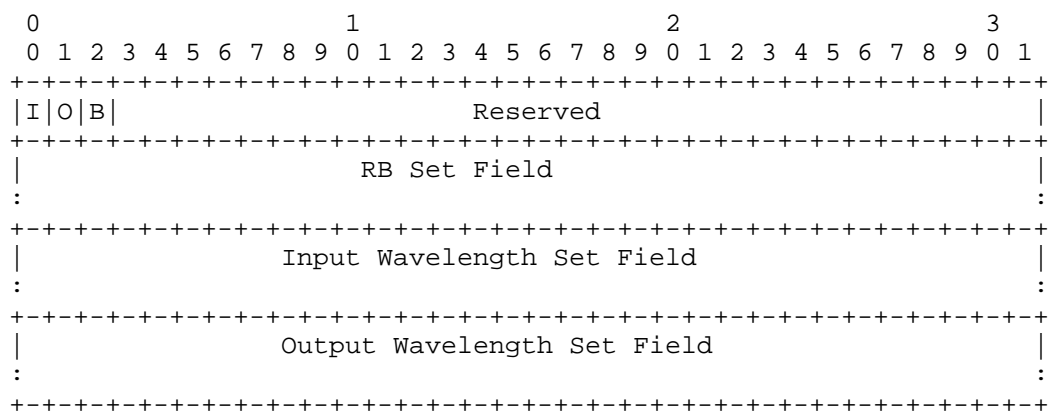
The For the Input and Output Link Set Fields, the Link Set Field encoding defined in [Gen-Encode] is to be used.

Note that the direction parameter within the Link Set Field is used to indicate whether the link set is an input or output link set, and the bidirectional value for this parameter is not permitted in this sub-TLV.

See Appendix A.1 for an illustration of this encoding.

4.2. Resource Block Wavelength Constraints Sub-TLV

Resources, such as wavelength converters, etc., may have a limited input or output wavelength ranges. Additionally, due to the structure of the optical system not all wavelengths can necessarily reach or leave all the resources. These properties are described by using one or more resource wavelength restrictions sub-TLVs as defined below:



I = 1 or 0 indicates the presence or absence of the Input Wavelength Set Field.

O = 1 or 0 indicates the presence or absence of the Output Wavelength Set Field.

B = 1 indicates that a single wavelength set field represents both input and output wavelength constraints.

Currently the only valid combinations of (I,O,B) are (1,0,0), (0,1,0), (1,1,0), (0,0,1).

RB Set Field:

A set of resource blocks (RBs) which have the same wavelength restrictions.

Input Wavelength Set Field:

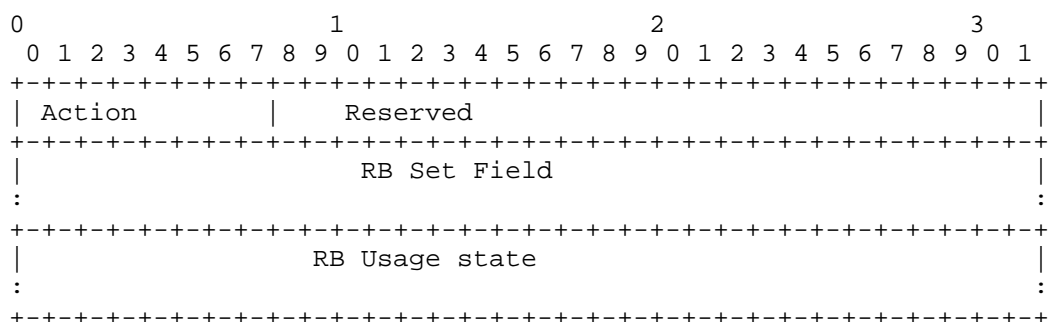
Indicates the wavelength input restrictions of the RBs in the corresponding RB set. This field is encoded via the Label Set field of [Gen-Encode].

Output Wavelength Set Field:

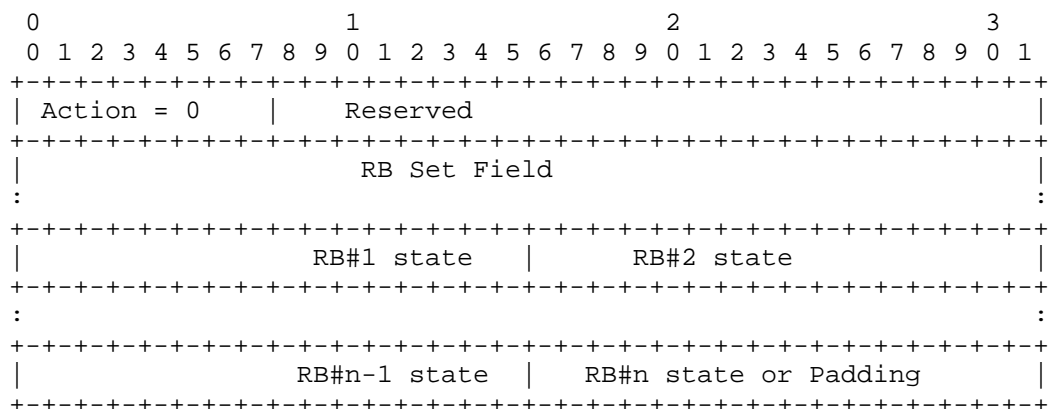
Indicates the wavelength output restrictions of RBs in the corresponding RB set. This field is encoded via the Label Set field of [Gen-Encode].

4.3. Resource Pool State Sub-TLV

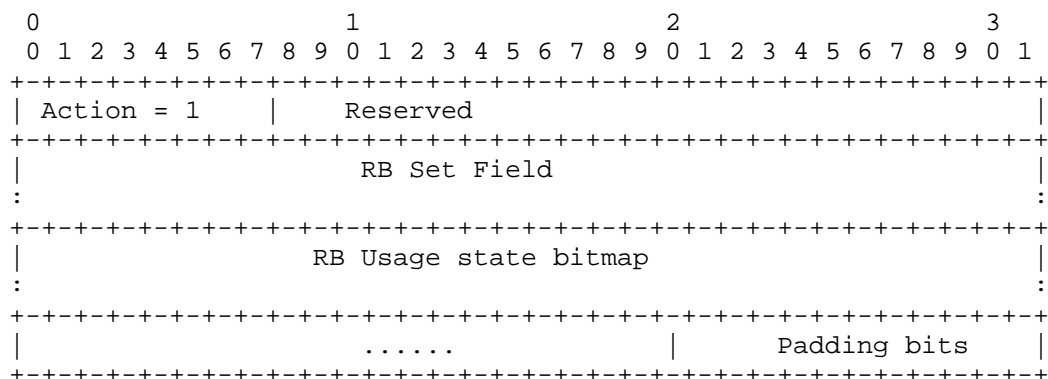
The state of the pool is given by the number of resources available with particular characteristics. A resource block set is used to encode all or a subset of the resources of interest. The usage state of resources within a resource block set is encoded as either a list of 16 bit integer values or a bit map indicating whether a single resource is available or in use. The bit map encoding is appropriate when resource blocks consist of a single resource. This information can be relatively dynamic, i.e., can change when a connection (LSP) is established or torn down.



Where Action = 0 denotes a list of 16 bit integers and Action = 1 denotes a bit map. In both cases the elements of the RB Set field are in a one-to-one correspondence with the values in the usage RB usage state area.



Whether the last 16 bits is a wavelength converter (RB) state or padding is determined by the number of elements in the RB set field.



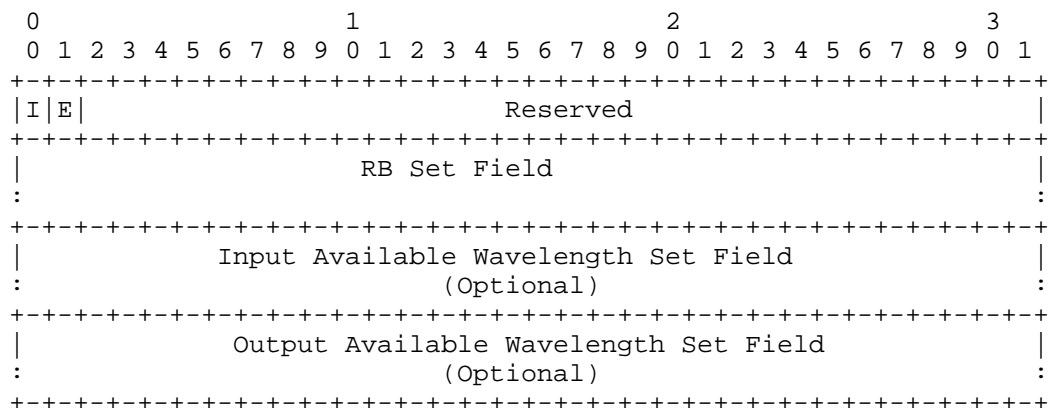
RB Usage state: Variable Length but must be a multiple of 4 bytes.

Each bit indicates the usage status of one RB with 0 indicating the RB is available and 1 indicating the RB is in used. The sequence of the bit map is ordered according to the RB Set field with this sub-TLV.

Padding bits: Variable Length

4.4. Block Shared Access Wavelength Availability sub-TLV

Resources blocks may be accessed via a shared fiber. If this is the case, then wavelength availability on these shared fibers is needed to understand resource availability.



I bit:

Indicates whether the input available wavelength set field is included (1) or not (0).

E bit:

Indicates whether the output available wavelength set field is included (1) or not (0).

RB Set Field:

A Resource Block set in which all the members share the same input or output fiber or both.

Input Available Wavelength Set Field:

Indicates the wavelengths currently available (not being used) on the input fiber to this resource block. This field is encoded via the Label Set field of [Gen-Encode].

Output Available Wavelength Set Field:

Indicates the wavelengths currently available (not being used) on the output fiber from this resource block. This field is encoded via the Label Set field of [Gen-Encode].

5. Resource Properties Encoding

Within a WSON network element (NE) there may be resources with signal compatibility constraints. These resources be regenerators, wavelength converters, etc... Such resources may also constitute the network element as a whole as in the case of an electro optical switch. This section primarily focuses on the signal compatibility and processing properties of such a resource block.

The fundamental properties of a resource block, such as a regenerator or wavelength converter, are:

- (a) Input constraints (shared input, modulation, FEC, bit rate, GPID)
- (b) Processing capabilities (number of resources in a block, regeneration, performance monitoring, vendor specific)

(c) Output Constraints (shared output, modulation, FEC)

5.1. Resource Block Information Sub-TLV

Resource Block descriptor sub-TLVs are used to convey relatively static information about individual resource blocks including the resource block compatibility properties, processing properties, and the number of resources in a block.

This sub-TLV has the following format:

[illegible]

Where I and E, the shared input/output indicator, is set to 1 if the resource blocks identified in the RB set field utilized a shared fiber for input/output access and set to 0 otherwise.

5.2. Modulation Format List Sub-Sub-TLV

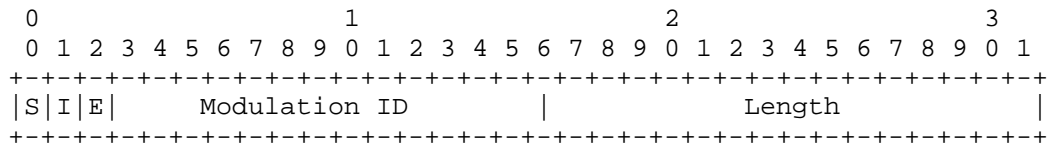
This sub-sub-TLV contains a list of acceptable modulation formats.

```
Type := Modulation Format List
```

Value := A list of Modulation Format Fields

5.2.1. Modulation Format Field

Two different types of modulation format fields are defined: a standard modulation field and a vendor specific modulation field. Both start with the same 32 bit header shown below.



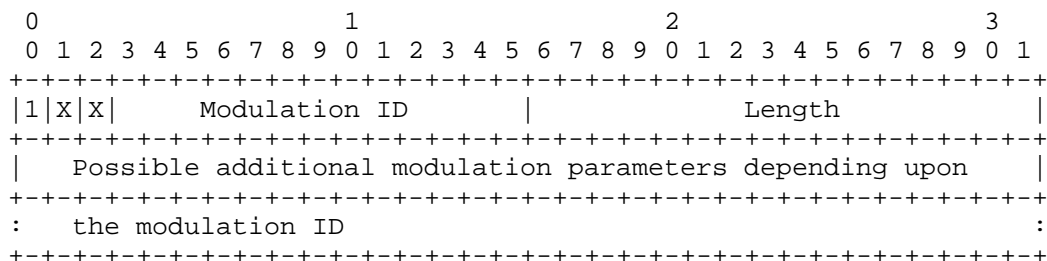
Where S bit set to 1 indicates a standardized modulation format and S bit set to 0 indicates a vendor specific modulation format. The length is the length in bytes of the entire modulation type field.

Where I bit set to 1 indicates it is an input modulation constraint and E bit set to 1 indicates it is an output modulation constraint.

The following I and E bit combination are defined:

I	E	
0	0	Invalid
1	0	Input modulation
0	1	Output modulation
1	1	Input and output modulation

The format for the standardized type for the modulation is given by:



Takes on the following currently defined values:

0	Reserved
1	optical tributary signal class NRZ 1.25G
2	optical tributary signal class NRZ 2.5G
3	optical tributary signal class NRZ 10G
4	optical tributary signal class NRZ 40G
5	optical tributary signal class RZ 40G

Note that future modulation types may require additional parameters in their characterization.

The format for vendor specific modulation field is given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|0|X|X| Vendor Modulation ID |                               Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Enterprise Number                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
: Any vendor specific additional modulation parameters :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Vendor Modulation ID

This is a vendor assigned identifier for the modulation type.

Enterprise Number

A unique identifier of an organization encoded as a 32-bit integer. Enterprise Numbers are assigned by IANA and managed through an IANA registry [RFC2578].

Vendor Specific Additional parameters

There can be potentially additional parameters characterizing the vendor specific modulation.

5.3. FEC Type List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable FEC types.

Type := FEC Type field List

Value := A list of FEC type Fields

5.3.1. FEC Type Field

The FEC type Field may consist of two different formats of fields: a standard FEC field or a vendor specific FEC field. Both start with the same 32 bit header shown below.

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
S I E										FEC ID										Length																			
Possible additional FEC parameters depending upon																																							
: the FEC ID										:																													

Where S bit set to 1 indicates a standardized FEC format and S bit set to 0 indicates a vendor specific FEC format. The length is the length in bytes of the entire FEC type field.

Where I bit set to 1 indicates it is an input FEC constraint and E bit set to 1 indicates it is an output FEC constraint.

The length is the length in bytes of the entire FEC type field.

The format for standard FEC field is given by:

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|X|X|      FEC ID      |      Length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Possible additional FEC parameters depending upon |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:   the FEC ID   :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

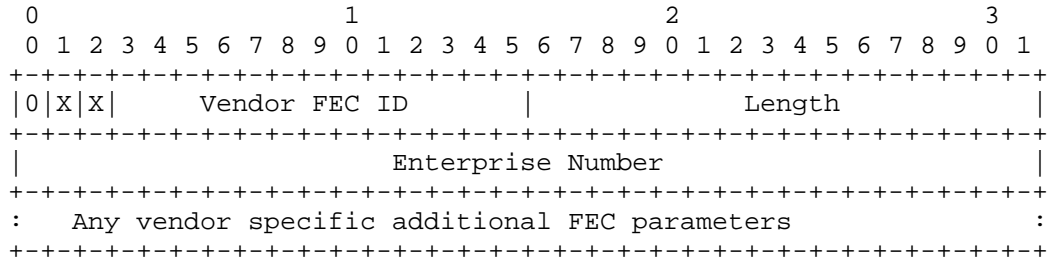
```

Takes on the following currently defined values for the standard FEC ID:

- | | |
|----|---|
| 0 | Reserved |
| 1 | G.709 RS FEC |
| 2 | G.709V compliant Ultra FEC |
| 3 | G.975.1 Concatenated FEC
(RS(255,239)/CSOC(n0/k0=7/6,J=8)) |
| 4 | G.975.1 Concatenated FEC (BCH(3860,3824)/BCH(2040,1930)) |
| 5 | G.975.1 Concatenated FEC (RS(1023,1007)/BCH(2407,1952)) |
| 6 | G.975.1 Concatenated FEC (RS(1901,1855)/Extended Hamming
Product Code (512,502)X(510,500)) |
| 7 | G.975.1 LDPC Code |
| 8 | G.975.1 Concatenated FEC (Two orthogonally concatenated
BCH codes) |
| 9 | G.975.1 RS(2720,2550) |
| 10 | G.975.1 Concatenated FEC (Two interleaved extended BCH
(1020,988) codes) |

Where RS stands for Reed-Solomon and BCH for Bose-Chaudhuri-Hocquengham.

The format for input vendor-specific FEC field is given by:



Vendor FEC ID

This is a vendor assigned identifier for the FEC type.

Enterprise Number

A unique identifier of an organization encoded as a 32-bit integer. Enterprise Numbers are assigned by IANA and managed through an IANA registry [RFC2578].

Vendor Specific Additional FEC parameters

There can be potentially additional parameters characterizing the vendor specific FEC.

5.4. Input Bit Range List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable input bit rate ranges.

Type := Input Bit Range List

Value := A list of Bit Range Fields

5.4.1. Bit Range Field

The bit rate range list sub-sub-TLV makes use of the following bit rate range field:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Starting Bit Rate                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Ending Bit Rate                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The starting and ending bit rates are given as 32 bit IEEE floating point numbers in bits per second. Note that the starting bit rate is less than or equal to the ending bit rate.

The bit rate range list sub-TLV is then given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Bit Range Field #1                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     :                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Bit Range Field #M                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

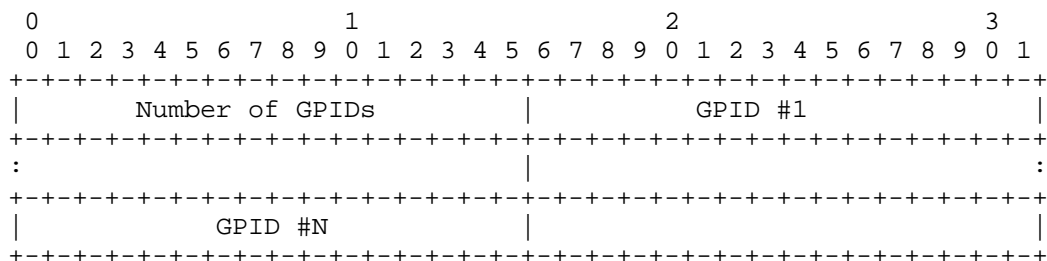
5.5. Input Client Signal List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable input client signal types.

Type := Input Client Signal List

Value := A list of GPIDs

The acceptable client signal list sub-TLV is a list of Generalized Protocol Identifiers (GPIDs). GPIDs are assigned by IANA and many are defined in [RFC3471] and [RFC4328].



Where the number of GPIDs is an integer greater than or equal to one.

5.6. Processing Capability List Sub-Sub-TLV

This sub-sub-TLV contains a list of resource processing capabilities.

Type := Processing Capabilities List

Value := A list of Processing Capabilities Fields

The processing capability list sub-sub-TLV is a list of capabilities that can be achieved through the referred resources::

1. Regeneration capability
2. Fault and performance monitoring
3. Vendor Specific capability

Note that the code points for Fault and performance monitoring and vendor specific capability are subject to further study.

5.6.1. Processing Capabilities Field

The processing capability field is then given by:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Processing Cap ID           |           Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Possible additional capability parameters depending upon |
+-----+-----+-----+-----+-----+-----+-----+-----+
:   the processing ID                               :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

When the processing Cap ID is "regeneration capability", the following additional capability parameters are provided in 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|  T  |  C  |           Reserved           |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where T bit indicates the type of regenerator:

T=0: Reserved

T=1: 1R Regenerator

T=2: 2R Regenerator

T=3: 3R Regenerator

Where C bit indicates the capability of regenerator:

C=0: Reserved

C=1: Fixed Regeneration Point

C=2: Selective Regeneration Point

Note that when the capability of regenerator is indicated to be Selective Regeneration Pools, regeneration pool properties such as input and output restrictions and availability need to be specified. This encoding is to be determined in the later revision.

6. Security Considerations

This document defines protocol-independent encodings for WSON information and does not introduce any security issues.

However, other documents that make use of these encodings within protocol extensions need to consider the issues and risks associated with, inspection, interception, modification, or spoofing of any of this information. It is expected that any such documents will describe the necessary security measures to provide adequate protection.

7. IANA Considerations

This document provides general protocol independent information encodings. There is no IANA allocation request for the TLVs defined in this document. IANA allocation requests will be addressed in protocol specific documents based on the encodings defined here.

8. Acknowledgments

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

APPENDIX A: Encoding Examples

A.1. Wavelength Converter Accessibility Sub-TLV

Example:

Figure 1 shows a wavelength converter pool architecture know as "shared per fiber". In this case the input and output pool matrices are simply:

$$WI = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline 1 & 1 \\ \hline \end{array}, \quad WE = \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & 1 \\ \hline \end{array}$$

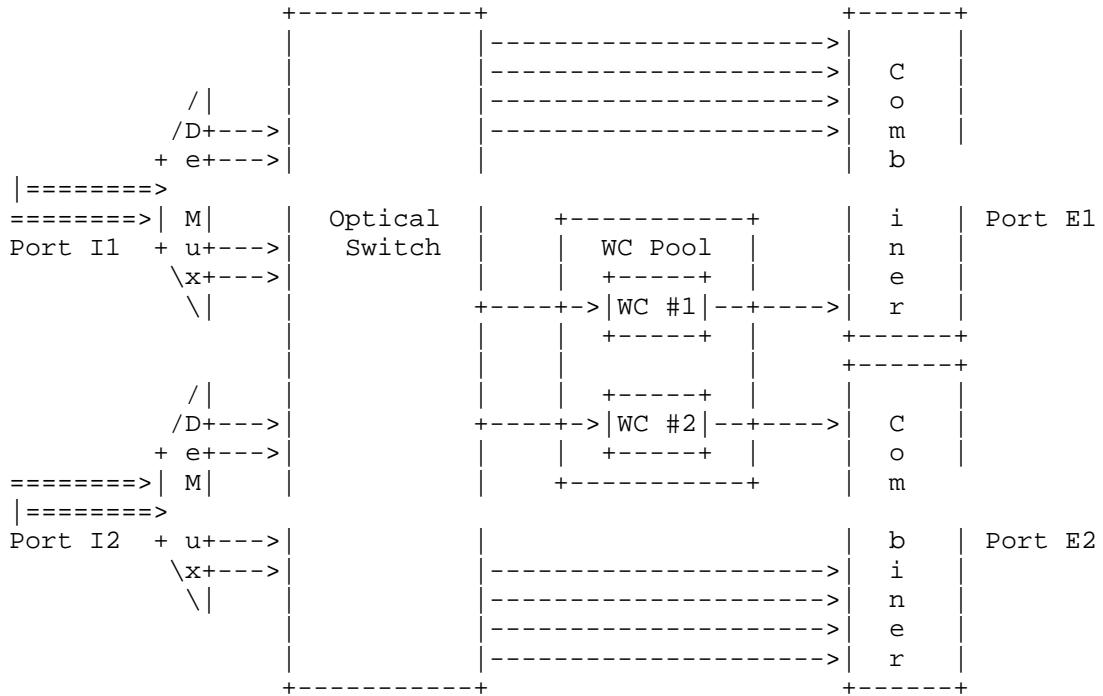


Figure 1 An optical switch featuring a shared per fiber wavelength converter pool architecture.

This wavelength converter pool can be 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
+-----+-----+-----+-----+
| Connectivity=1|                               Reserved          |
|               Note: I1,I2 can connect to either WC1 or WC2      |
+-----+-----+-----+-----+
| Action=0      |0| Reserved          | Length = 12              |
+-----+-----+-----+-----+
|                               Link Local Identifier = #1         |
+-----+-----+-----+-----+
|                               Link Local Identifier = #2         |
+-----+-----+-----+-----+
| Action=0      |1| Reserved          | Length = 8              |
+-----+-----+-----+-----+
|               RB ID = #1          | RB ID = #2                  |
+-----+-----+-----+-----+
|               Note: WC1 can only connect to E1                  |
+-----+-----+-----+-----+
| Action=0      |1| Reserved          | Length = 8              |
+-----+-----+-----+-----+
|                               Link Local Identifier = #1         |
+-----+-----+-----+-----+
| Action=0      |0| Reserved          | Length = 8              |
+-----+-----+-----+-----+
|               RB ID = #1          | zero padding                |
+-----+-----+-----+-----+
|               Note: WC2 can only connect to E2                  |
+-----+-----+-----+-----+
| Action=0      |1| Reserved          | Length = 8              |
+-----+-----+-----+-----+
|                               Link Local Identifier = #2         |
+-----+-----+-----+-----+
| Action=0      |0|                               | Length = 8      |
+-----+-----+-----+-----+
|               RB ID = #2          | zero padding                |
+-----+-----+-----+-----+

```

A.2. Wavelength Conversion Range Sub-TLV

Example:

This example, based on figure 1, shows how to represent the wavelength conversion range of wavelength converters. Suppose the wavelength range of input and output of WC1 and WC2 are {L1, L2, L3, L4}:

```

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
Note: WC Set
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      |1| Reserved      |      Length = 8      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|      WC ID = #1      |      WC ID = #2      |
+-----+-----+-----+-----+-----+-----+-----+
Note: wavelength input range
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2      | Num Wavelengths = 4      |      Length = 8      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Grid |   C.S. |      Reserved      |   n for lowest frequency = 1   |
+-----+-----+-----+-----+-----+-----+-----+-----+
Note: wavelength output range
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2      | Num Wavelengths = 4      |      Length = 8      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|Grid |   C.S. |      Reserved      |   n for lowest frequency = 1   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

A.3. An OEO Switch with DWDM Optics

Figure 2 shows an electronic switch fabric surrounded by DWDM optics. In this example the electronic fabric can handle either G.709 or SDH signals only (2.5 or 10 Gbps). To describe this node, the following information is needed:

```
<Node_Info> ::= <Node_ID>[Other GMPLS sub-
TLVs][<ConnectivityMatrix>...] [<ResourcePool>][<RBPooolState>]
```

In this case there is complete port to port connectivity so the <ConnectivityMatrix> is not required. In addition since there are sufficient ports to handle all wavelength signals the <RBPooolState> element is not needed.

Hence the attention will be focused on the <ResourcePool> sub-TLV:


```

<ResourcePool> ::=
<ResourceBlockInfo>[<ResourceBlockAccessibility>...][<ResourceWaveCo
nstraints>...]

```

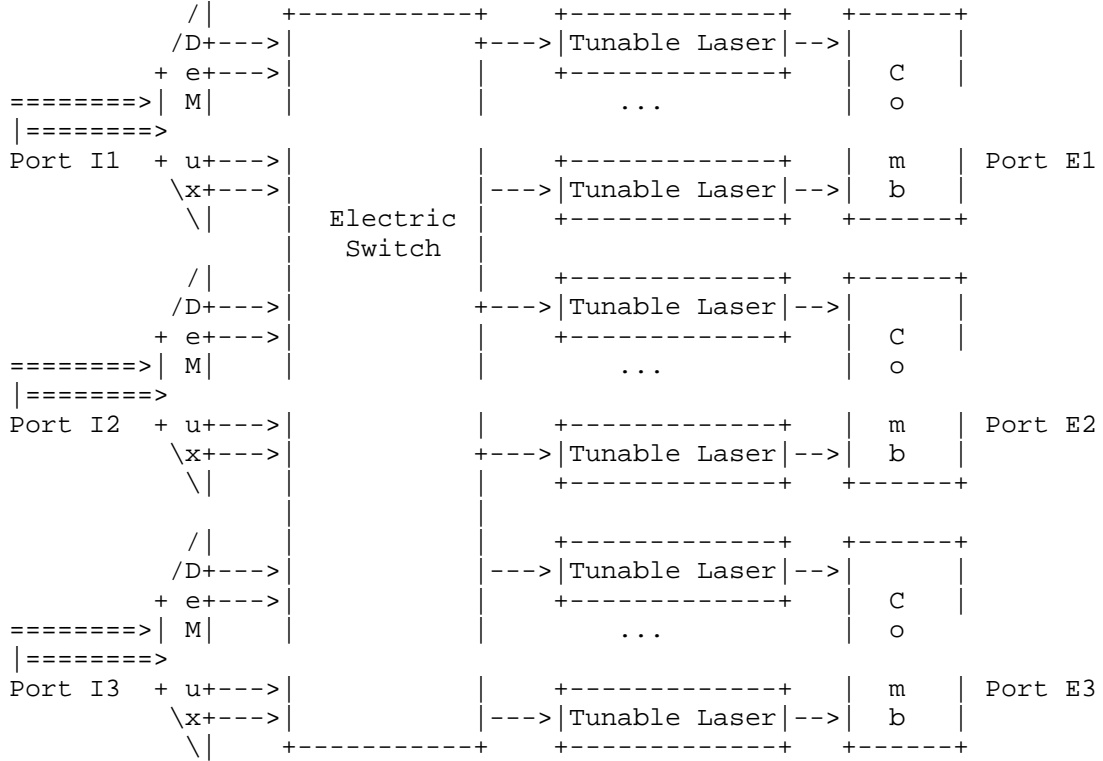


Figure 2 An optical switch built around an electronic switching fabric.

The resource block information will tell us about the processing constraints of the receivers, transmitters and the electronic switch. The resource availability information, although very simple, tells us that all signals must traverse the electronic fabric (fixed connectivity). The resource wavelength constraints are not needed since there are no special wavelength constraints for the resources that would not appear as port/wavelength constraints.

```

<ResourceBlockInfo>:

```

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     RB Set Field                                     |
| : (only one resource block in this example with shared                          |
| | input/output case)                                                           |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0|1|1|                               Reserved                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Modulation Type List Sub-Sub-TLV               |
| :                                     NRZ                                       |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     FEC Type List Sub-Sub-TLV                   |
| :                                     Standard SDH, G.709 FEC                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Input Client Signal Type Sub-TLV             |
| :                                     (GPIDs for SDH and G.709)                   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Input Bit Rate Range List Sub-Sub-TLV         |
| :                                     (2.5Gbps, 10Gbps)                         |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Processing Capabilities List Sub-Sub-TLV     |
| :                                     Fixed (non optional) 3R regeneration         |
| :                                     :                                           |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Since there is fixed connectivity to resource blocks (the electronic switch) the <ResourceBlockAccessibility> is:

```

      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=0|Reserved|                                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Input Link Set Field A #1                   |
| :                                     (All input links connect to resource)       |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     RB Set Field A #1                           |
| :                                     (trivial set only one resource block)       |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Output Link Set Field B #1                  |
| :                                     (All output links connect to resource)     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

9. References

9.1. 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.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC4328] Papadimitriou, D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [G.694.1] ITU-T Recommendation G.694.1, "Spectral grids for WDM applications: DWDM frequency grid", June, 2002.

9.2. Informative References

- [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.
- [Gen-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", work in progress: draft-ietf-ccamp-general-constraint-encode.
- [RFC6205] T. Otani, H. Guo, K. Miyazaki, D. Caviglia, "Generalized Labels for G.694 Lambda-Switching Capable Label Switching Routers", RFC 6205, March 2011.
- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", RFC 6163, April 2011.

[WSO-Info] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-info, March 2009.

10. Contributors

Diego Caviglia
Ericsson
Via A. Negrone 1/A 16153
Genoa Italy

Phone: +39 010 600 3736
Email: diego.caviglia@marconi.com, ericsson.com)

Anders Gavler
Acreo AB
Electrum 236
SE - 164 40 Kista Sweden

Email: Anders.Gavler@acreo.se

Jonas Martensson
Acreo AB
Electrum 236
SE - 164 40 Kista, Sweden

Email: Jonas.Martensson@acreo.se

Itaru Nishioka
NEC Corp.
1753 Simonumabe, Nakahara-ku, Kawasaki, Kanagawa 211-8666
Japan

Phone: +81 44 396 3287
Email: i-nishioka@cb.jp.nec.com

Pierre Peloso
ALU

Email: pierre.peloso@alcatel-lucent.com

Cyril Margaria
NSN

Email: cyril.margaria@nsn.com

Authors' Addresses

Greg M. Bernstein (ed.)
Grotto Networking
Fremont California, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Young Lee (ed.)
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@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: danli@huawei.com

Wataru Imajuku
NTT Network Innovation Labs
1-1 Hikari-no-oka, Yokosuka, Kanagawa
Japan

Phone: +81-(46) 859-4315
Email: imajuku.wataru@lab.ntt.co.jp

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

Phone: +86-755-28972916
Email: hanjianrui@huawei.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.

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: January 2013

Y. Lee
Huawei
G. Bernstein
Grotto Networking

July 16, 2012

GMPLS OSPF Enhancement for Signal and Network Element Compatibility
for Wavelength Switched Optical Networks

draft-ietf-ccamp-wson-signal-compatibility-ospf-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 January 16, 2007.

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 document provides GMPLS OSPF routing enhancements to support signal compatibility constraints associated with WSON network elements. These routing enhancements are required in common optical or hybrid electro-optical networks where not all of the optical signals in the network are compatible with all network elements participating in the network.

This compatibility constraint model is applicable to common optical or hybrid electro optical systems such as OEO switches, regenerators, and wavelength converters since such systems can be limited to processing only certain types of WSON signals.

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. Revision History.....	3
2. The Optical Node Property TLV.....	4
2.1. Sub-TLV Details.....	5
2.1.1. Resource Pool Accessibility.....	6
2.1.2. Resource Block Wavelength Constraints.....	6
2.1.3. Resource Pool State.....	6
2.1.4. Block Shared Access Wavelength Availability.....	6
3. ISCD format extensions.....	6
3.1. Switch Capability Specific Information.....	7
4. WSON Specific Scalability and Timeliness.....	8
5. Security Considerations.....	9
6. IANA Considerations.....	9

7. References.....	11
7.1. Normative References.....	11
7.2. Informative References.....	11
8. Authors and Contributors.....	12
Authors' Addresses.....	12
Intellectual Property Statement.....	12
Disclaimer of Validity.....	13

1. Introduction

The documents [RFC6163, WSON-Info, WSON-Encode] explain how to extend the wavelength switched optical network (WSON) control plane to allow both multiple WSON signal types and common hybrid electro optical systems as well hybrid systems containing optical switching and electro-optical resources. In WSON, not all of the optical signals in the network are compatible with all network elements participating in the network. Therefore, signal compatibility is an important constraint in path computation in a WSON.

This document provides GMPLS OSPF routing enhancements to support signal compatibility constraints associated with general WSON network elements. These routing enhancements are required in common optical or hybrid electro-optical networks where not all of the optical signals in the network are compatible with all network elements participating in the network.

This compatibility constraint model is applicable to common optical or hybrid electro optical systems such as OEO switches, regenerators, and wavelength converters since such systems can be limited to processing only certain types of WSON signals.

1.1. Revision History

From 00 to 01: The details of the encodings for compatibility moved from this document to [WSON-Encode].

From 01 to 02: Editorial changes.

From 02 to 03: Add a new Top Level Node TLV, Optical Node Property TLV to carry WSON specific node information.

From 03 to 04: Add a new sub-TLV, Block Shared Access Wavelength Availability TLV to be consistent with [WSON-Encode] and editorial changes.

From 04 to 05: Add a new section that discusses OSPF scalability and timeliness and editorial changes.

From 05 to 06: Change the title of the draft to "GMPLS OSPF Enhancement" from "OSPF Enhancement" to make sure the changes apply to the GMPLS OSPF rather than the base OSPF. Add specific OSPF procedures on how sub-TLVs are packaged per [RFC3630] and editorial changes.

From 06 to 07: Add clarifying texts on how to sub-divide the Optical Node TLV in case it exceeds the IP MTU fragmentation limit. Delete Section 3.2. to avoid multiple rules so as to avoid confusion.

From 07 to 08: Clean some old texts in Section 3. Align with [WSON-Encode] on the modulation and FEC type.

From 08 to 09: Added ISCD extensions for available labels and shared backup labels.

2. The Optical Node Property TLV

[RFC3630] defines OSPF TE LSA using an opaque LSA. This document adds a new top level TLV for use in the OSPF TE LSA: the Optical Node Property TLV. The Optical Node property TLV describes a single node. It is constructed of a set of sub-TLVs. There are no ordering requirements for the sub-TLVs. Only one Optical Node TLV shall be advertised in each LSA.

The Optical Node Property TLV contains all WSON-specific node properties and signal compatibility constraints. The detailed encodings of these properties are defined in [WSON-Encode].

The following sub-TLVs of the Optical Node Property TLV are defined:

Value	Length	Sub-TLV Type
TBA	variable	Resource Block Information
TBA	variable	Resource Pool Accessibility
TBA	variable	Resource Block Wavelength Constraints
TBA	variable	Resource Pool State
TBA	variable	Block Shared Access Wavelength Availability

The detail encodings of these sub-TLVs are found in [WSON-Encode] as indicated in the table below.

Sub-TLV Type	Section [WSON-Encode]
Resource Block Information	4.1
Resource Pool Accessibility	3.1
Resource Block Wavelength Constraints	3.2
Resource Pool State	3.3
Block Shared Access Wavelength Availability	3.4

All sub-TLVs defined here may occur at most once in any given Optical Node TLV. "At most once" means that if there is sub-TLV related information, it should be always included. These restrictions need not apply to future sub-TLVs. Unrecognized sub-TLVs are ignored.

2.1. Sub-TLV Details

Among the sub-TLVs defined above, the Resource Pool State sub-TLV and Block Shared Access Wavelength Availability are dynamic in nature while the rest are static. As such, they can be separated out from the rest and be advertised with multiple TE LSAs per OSPF router, as described in [RFC3630] and [RFC5250]. Resource Block Information

Resource Block Information sub-TLVs are used to convey relatively static information about individual resource blocks including the resource block properties and the number of resources in a block.

There are five nested sub-TLVs defined in the Resource Block Information sub-TLV.

Value	Length	Sub-TLV Type
TBA	variable	Modulation Format List
TBA	variable	FEC Type List
TBA	variable	Input Bit Range List
TBA	variable	Input Client Signal List
TBA	variable	Processing Capability List

The detail encodings of these sub-TLVs are found in [WSON-Encode] as indicated in the table below.

Name	Section [WSON-Encode]
Modulation Format List	4.2
FEC Type List	4.3
Input Bit Range List	4.4
Input Client Signal List	4.5
Processing Capability List	4.6

2.1.1. Resource Pool Accessibility

This sub-TLV describes the structure of the resource pool in relation to the switching device. In particular it indicates the ability of an ingress port to reach a resource block and of a resource block to reach a particular egress port.

2.1.2. Resource Block Wavelength Constraints

Resources, such as wavelength converters, etc., may have a limited input or output wavelength ranges. Additionally, due to the structure of the optical system not all wavelengths can necessarily reach or leave all the resources. Resource Block Wavelength Constraints sub-TLV describe these properties.

2.1.3. Resource Pool State

This sub-TLV describes the usage state of a resource that can be encoded as either a list of 16 bit integer values or a bit map indicating whether a single resource is available or in use. This information can be relatively dynamic, i.e., can change when a connection is established or torn down.

2.1.4. Block Shared Access Wavelength Availability

Resources blocks may be accessed via a shared fiber. If this is the case then wavelength availability on these shared fibers is needed to understand resource availability.

3. 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 WSON as follows:

Where the Available Label Sub-TLV and Shared Backup Label sub-TLV are defined in [Gen-Encode].

This document has defined five sub-TLVs specific to WSON. The examples given in [WSON-Encode] show that very large systems, in terms of channel count, ports, or resources, can be very efficiently encoded.

If the size of this LSA is greater than the MTU, then these sub-TLVs can be packed into separate LSAs. From the point of view of path computation, the presence of the Resource Block Information sub-TLV indicates that resources exist in the system and may have signal compatibility or other constraints. The other four sub-TLVs indicate constraints on access to, and availability of those resources.

[Page 8]

Note that the connection establishment mechanism (signaling or management) is ultimately responsible for the establishment of the connection, and this implies that such mechanisms must insure signal compatibility.

5. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC3630], [RFC4203].

6. IANA Considerations

This document introduces a new Top Level Node TLV (Optical Node Property TLV) under the OSPF TE LSA defined in [RFC3630].

Value	TLV Type
-------	----------

TBA	Optical Node Property
-----	-----------------------

IANA is to allocate a new TLV Type and its Value for this Top Level Node TLV.

This document also introduces the following sub-TLVs associated with the Optical Node Property TLV as defined in Section 2.1 as follows:

Value	Length	Sub-TLV Type
TBA	variable	Resource Block Information
TBA	variable	Resource Pool Accessibility
TBA	variable	Resource Block Wavelength Constraints
TBA	variable	Resource Pool State
TBA	variable	Block Shared Access Wavelength Availability

IANA is to allocate new sub-TLV Types and their Values for these sub-TLVs defined under the Optical Node Property TLV.

There are five nested sub-TLVs defined in the Resource Block Information sub-TLV as follows:

Value	Length	Sub-TLV Type
-------	--------	--------------

TBA	variable	Modulation Format List
TBA	variable	FEC Type List
TBA	variable	Input Bit Range List
TBA	variable	Input Client Signal List
TBA	variable	Processing Capability List

IANA is to allocate new Sub-TLV Types and their Values for these Sub-TLVs defined under the Resource Block Information Sub-TLV.

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
151 (suggested)	WSON-LSC capable (WSON-LSC)	[This.I-D]

This document defines the following sub-TLVs of the ISCD TLV:

Value	Sub-TLV
1	Available Labels
2	Shared Backup Labels

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.
- [RFC3630] Katz, D., Kompella, K., and Yeung, D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [G.694.1] ITU-T Recommendation G.694.1, "Spectral grids for WDM applications: DWDM frequency grid", June, 2002.
- [RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [WSON-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-wson-encode, work in progress.
- [Gen-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", draft-ietf-ccamp-general-constraint-encode, work in progress.

7.2. Informative References

- [WSON-Info] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info, work in progress.
- [RFC6250] T. Otani, Ed., D. Li, Ed., "Generalized Labels for G.694 Lambda-Switching Capable Label Switching Routers", RFC 6250, March 2011.
- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", RFC 6163, April 2011.

[RFC5250] Berger, L., et al., "The OSPF Opauqe LSA option", RFC 5250, July 2008.

8. Authors and Contributors

Authors' Addresses

Young Lee (ed.)
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

Greg M. Bernstein (ed.)
Grotto Networking
Fremont California, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.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
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

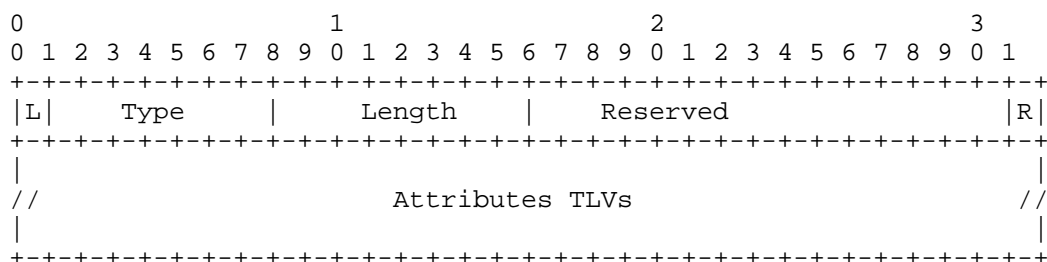
3.1. ERO LSP Attribute Subobject

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.

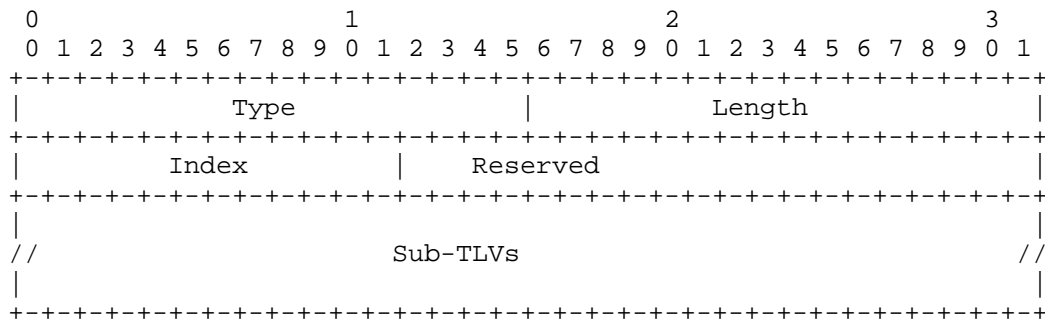
Index A value used to refer to an LSP Attribute Hop Info TLV containing information targeted at the node processing this ERO.

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

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 17, 2013

G. Martinelli, Ed.
G. Galimberti
Cisco
L. Ong
Ciena Corporation
D. Ceccarelli
Ericsson
C. Margaria
Nokia Siemens Networks
July 16, 2012

WSON Optical Interface Class
draft-martinelli-wson-interface-class-03

Abstract

Current work on wavelength switched optical network includes several considerations regarding the interface signal compatibility. In particular ingress and egress optical interfaces will require a check on several optical parameters to assess if the signal generated by the ingress interface can be compatible with the receiving interface. Current solution available encode all parameters in WSON protocol extensions while in this draft will propose an alternative method to keep into account the signal compatibility issue at protocol level.

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 17, 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. Existing WSOON Signal Compatibility protocol extension	3
3. Optical Interface Class	4
3.1. Concept	4
3.2. Procedures	5
3.3. Encoding	6
4. Optical Interface Class Semantic	8
5. Acknowledgements	8
6. IANA Considerations	8
7. Security Considerations	9
8. References	9
8.1. Normative References	9
8.2. Informative References	10
Appendix A. Application Code Mapping	10
A.1. ITU-G.698.1 Application Code Mapping	10
A.2. ITU-G.698.2 Application Code Mapping	12
A.3. ITU-G.959.1 Application Code Mapping	12
Appendix B. Encoding example	14
Authors' Addresses	16

1. Introduction

The current work on Wavelength Switched Optical Network (WSON) define the need of assessing the signal compatibility during the routing and wavelength assignment (RWA) process. In details, the [RFC6163] reports the ingress and egress interfaces and the regeneration points as places where the optical signal compatibility must be assured. On how to evaluate this compatibility, there are several parameters identified according to ITU specification (e.g. [ITU-G.698.1], [ITU-G.698.2] and [ITU-G.959.1] among others). In particular the following set of parameters has been chosen: signal bit rate, modulation format, forward error correction (FEC).

At the current state of art new high bit rates (40G/100G/flexgrid) and new modulation formats (e.g. QPSK flavors) are already deployed in field. Some of them are under standardization at ITU but it is not clear if and when there will be a dominating technology. With the current realistic scenario, DWDM optical networks will have to manage different bit-rates as well as different modulation formats over the same link since different signal characteristics will coexist at the same time.

To a further extent, WSON Control Plane needs consider the case with optical impairments awareness as detailed in [RFC6566]. In such a case, control plane will require some additional interface parameters to assess the optical feasibility path and, as a consequence, further WSON protocols extensions.

Scope of this draft is to propose an Optical Interface Class identifier as a solution for the WSON signal compatibility problem. To some extent the idea is have protocol extensions independent from optical technology evolution by keeping the semantic of optical characteristics separated from protocol scope. The final goal is not an encoding saving but rather a general abstract representation of some physical characteristics.

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. Existing WSON Signal Compatibility protocol extension

Within the current WSON activity the signal compatibility encoding is defined within the [I-D.ietf-ccamp-rwa-wson-encode].

In details, the following set of parameters is considered:

- o Modulation Format. Only NRZ currently defined.
- o FEC, according to G.709 and G.975.
- o Bit Rate.

This list of parameters is defined by ITU and might be subject to change for two reasons: new values for existing parameters, new parameters required by new technology or control plane evolution.

At the current status, the above encoding is going to be used within several WSON specific protocol extensions.

- o OSPF [I-D.ietf-ccamp-wson-signal-compatibility-ospf] since the path computation function need to consider optical interface parameters as a constrain during the RWA process.
- o RSVP [I-D.ietf-ccamp-wson-signaling] since during the signaling phase there is the need to know optical ingress and egress interface properties (and eventually interfaces at regeneration point).
- o In addition in case of PCE control plane solutions, PCEP extension needs similar parameters as envisaged here [I-D.lee-pce-wson-rwa-ext].

In case of any update from ITU standards regarding optical signals and interfaces all the above drafts making use of the same information needs an update.

3. Optical Interface Class

3.1. Concept

The Optical Interface Class is a unique number that identify all information related to optical characteristic's of a physical interface. The class may include other optical parameters related to other interface properties. A class MUST always include signal compatibility information.

The content of each class is out of the scope of this draft and can be defined by other entities (e.g. ITU, optical equipment vendors, etc.).

Since event current implementation of physical interfaces may support

different optical characteristics, a single interface may support multiple interface classes. Which optical interface class is used among all the one available for each interfaces is out of the scope of this draft but likely is an output of the RWA process.

3.2. Procedures

In term of RWA process one operation required is to assess the optical compatibility (LSP endpoint interfaces or regeneration intermediate endpoints). This is done by checking if the two optical endpoint have the same Optical Interface Class value. Note that, if a lightpath implementing an LSP needs regeneration, the complete LSP may have some additional intermediate optical endpoints where regenerations happens.

The procedure of signal compatibility assessment become just a numbers comparison: if two Optical Interface Class are equals the signal compatibility constrain is satisfied. GMPLS protocols don't have to implement any logic or optical knowledge related to signal compatibility.

The procedure is easily generalized in case more than one class is available for each interface since it's sufficient to find two matching values between each optical segment of a WSON LSP.

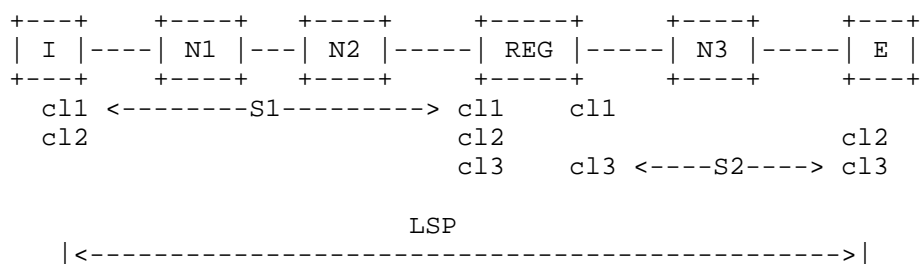


Figure 1

As an example Figure 1 shows a case where the RWA process end up in a path that need a wavelength conversion (I, N1, N2, REG, N3, E). Optical interfaces are reported as cl1, cl2 and cl3. Each optical interface involved in the path at nodes I, REG, E must satisfy the Optical Interface Class constrain. The LSP is then composed by two optical segment: S1 using optical interface class cl1 and S2 using optical interface class cl3.

By using the Optical Interface Class concept every protocol extensions supporting WSON does not need to care about DWDM signal details and does not need to consider technology specific evolution. If a new values are standardized (e.g. new modulation formats become standard) or new parameters are required, wson protocols don't need any extensions. In addition, in case PCE is used within a WSON control plane, this allows all optical parameters to be fully known only by the PCE and does not require the other element any knowledge of them.

3.3. Encoding

The following Optical Interface Class must be be used in proper TLVs for different WSON protocol extensions.

In case an optical interface or a regeneration point will support multiple optical capabilities, a list of Interface Classes can be used. Note that the concept of list is already defined in [I-D.ietf-ccamp-rwa-wson-encode].

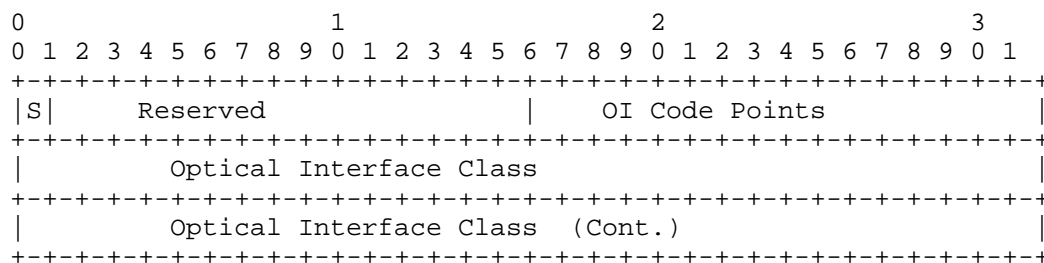


Figure 2: Optical Interface Class

Where the first 32 bist of the encoding shall be used to indentify the semantic of the Optical Interface Class in the following way:

S Standard bit.

S=0, identify not ITU code points

S=1, identify ITU application codes

With S=0, the OI Code Points field can take the following values:

0: reserved

1: Vendor Specific Optical Interface Class.

With S=1, the OI Code Points field can take the following values:

0: reserved

1: [ITU-G.698.1] application code.

2: [ITU-G.698.2] application code.

3: [ITU-G.959.1] application code.

In case of ITU Application Code, there should be a mapping between string defining the application code and the 64 bits number implementing the optical interface class.

As an example, Figure 3 shows how the encoding looks like in case of S bit equals to 1 and Code point equals to 3.

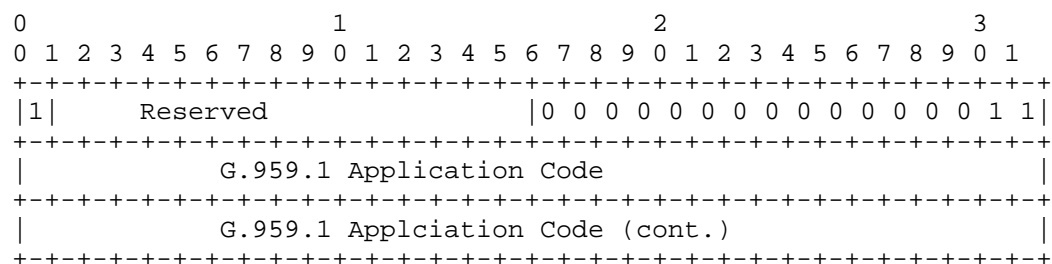


Figure 3: Optical Interface Class example with ITU application code

The mapping of the ITU Application Code over the Optical Interface Class is provided in Appendix A.

It is worthwhile to have a rough estimation if the 64 bits are enough for matching the currently defined ITU application code and with this purpose we provide the current analisys. The application code consist of the following 8 elements: DScW-ytz(v). D is fixed, S has two values (1 bit), c has currently 4 values (for reference [RFC6205] has 4 bit reserved for this), W takes two values (1 bit), t is a placeholder with only one value defined, z has only 3 values defined (2 bits) and v has 3 values (2 bits). In a rough esitimation, the number of information bits is at minimum 10 bit. So 64 bit are enough also for future application code evolutions.

4. Optical Interface Class Semantic

The semantic of the Optical Interface Class must be defined outside the control plane but it must be unique for all control plane for every control plane function or network node. Within this hypothesis, we need to solve the problem on how to make any network element aware of the semantic behind the Optical Interface Class and make sure it can figure out the right value for its interfaces.

An example of semantic is the "Application Code" within [ITU-G.698.1] and [ITU-G.698.2]. The Application Code could be easily represented by the Optical Interface Class index (or number). This number might be used as an index to access a table containing all the values associated with a specific interface using mechanisms like Directory Services. Note that each single interface parameter could be retrieved through a MIB. As an example, [I-D.galikunze-ccamp-g-698-2-snmp-mib] gives another example on the Optical parameter specification includes the OII definition in compliance with [ITU-G.698.2] Chapter 5.3.

Every time a new optical interface is defined or introduced into the market, only a MIB update will be required but there will be no impact on WSON protocols.

Note also that the Control Plane may become aware of the Optical Interface Class semantic by a various of other ways like the network management system or manual provisioning.

As a matter of fact in current WSON technology, standard and proprietary information must co-exist. The introduction of the Optical Interface Class does not change or limit this possibility since the class identifier can be a means to access either public or vendor specific information. In term of protocol encoding however, this solution has the advantage to limit eventually proprietary information in a fixed size field.

5. Acknowledgements

6. IANA Considerations

Optical Interface code points needs to be assigned by IANA?

All drafts are required to have an IANA considerations section (see the update of RFC 2434 [I-D.narten-iana-considerations-rfc2434bis] for a guide). If the draft does not require IANA to do anything, the section contains an explicit statement that this is the case (as

above). If there are no requirements for IANA, the section will be removed during conversion into an RFC by the RFC Editor.

7. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

8. References

8.1. Normative References

[I-D.ietf-ccamp-rwa-wson-encode]

Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-wson-encode-14 (work in progress), April 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-wson-signaling]

Bernstein, G., Xu, S., Lee, Y., Martinelli, G., and H. Harai, "Signaling Extensions for Wavelength Switched Optical Networks", draft-ietf-ccamp-wson-signaling-03 (work in progress), March 2012.

[ITU-G.698.1]

International Telecommunications Union, "Multichannel DWDM applications with single-channel optical interfaces", ITU-T Recommendation G.698.1, November 2009.

[ITU-G.698.2]

International Telecommunications Union, "Amplified multichannel DWDM applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.

[ITU-G.959.1]

International Telecommunications Union, "Optical transport networks physical layer interfaces", ITU-T Recommendation G.659.1, February 2012.

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

8.2. Informative References

- [I-D.galikunze-ccamp-g-698-2-snmp-mib]
Kunze, R. and D. Hiremagalur, "A SNMP MIB to manage black-link optical interface parameters of DWDM applications", draft-galikunze-ccamp-g-698-2-snmp-mib-00 (work in progress), June 2012.
- [I-D.lee-pce-wson-rwa-ext]
Lee, Y. and R. Casellas, "PCEP Extension for WSON Routing and Wavelength Assignment", draft-lee-pce-wson-rwa-ext-04 (work in progress), April 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.
- [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.
- [RFC6205] Otani, T. and D. Li, "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, March 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. Application Code Mapping

A.1. ITU-G.698.1 Application Code Mapping

This recommendation defines (see Section 5.3) the Application Codes: DScW-ytz(v) and B-DScW-ytz(v). Where:

- o B: means Bidirectionals.
- o D: means a DWDM application.
- o S: take values N (narrow spectral excursion), W (wide spectral excursion).
- o c: Channel Spacing (GHz).
- o W: take values S (short-haul), L (long-haul).
- o y: take values 1 (NRZ 2.5G), 2 (indicating NRZ 10G).
- o t: take only D value is defined (link does not contain optical amplifier)
- o z: take values 2 (ITU-T G.652 fibre), 3 (ITU-T G.653 fibre), 5 (indicating ITU-T G.655 fibre).
- o v: take values S (Short wavelength), C (Conventional), L (Long wavelength).

An Optional F can be added indicating a FEC Encoding.

Considering the 64 bits left to encode the application code for each ITU recommendation, the following encoding is proposed:

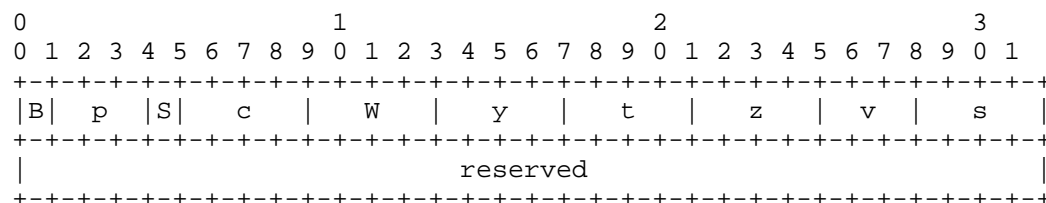


Figure 4: G.689.1 Application Code Encoding

Where (values between parenthesis refer to ITU defined values as reported above):

- B: = 1 bidirectional, 0 otherwise
- p (prefix): = 0 reserved, = 1 (D)
- S: = 0 (N), = 1 (W)

c: Channel Spacing, 4 bits mapped according to same definition in [RFC6205] (note that DWDM spacing apply here)

W: = 0 reserved, = 2 (S), = 3 (L)

y: = 0 reserved, = 1 (1), = 2 (2)

t: = 0 reserved, = 4 (D)

z: = 0 reserved, = 2 (2), = 3 (3), = 5 (5)

v: = 0 reserved, = 1 (S), = 2 (C), = 3 (L)

s (suffix): = 0 reserved, = 1 Fec Encoding

With the following notes: values not mentioned here are not allowed in this application code, the last 32 bits are reserved and shall always be zero.

A.2. ITU-G.698.2 Application Code Mapping

This recommendation defines (see Section 5.3) the Application Codes: DScW-ytz(v) and B-DScW-ytz(v). Since the format is exactly similar to Appendix A.1, this section only reports differences.

W: take values C (link is dispersion compensated), U (link is dispersion uncompensated).

t: take value A (link may contains optical amplifier)

Also here an optical F can be added to indicate FEC encoding.

In term of encoding applications codes follow exactly Figure 4 apart from the following differences:

W: = 0 reserved, = 10 (C), = 11 (U)

t: = 0 reserved, = 1 (A)

A.3. ITU-G.959.1 Application Code Mapping

This recommendation defines (see Section 5.3) the Application Codes: PnWx-ytz and BnWx-ytz. Where:

- o P,B: when presents they indicate Plural or Bidirectional
- o n: maximum number of channels supported by the application code (i.e. an integer number)

- o W: take values I (intra-office), S (short-haul), L (long-haul), V (very long-haul), U (ultra long-haul).
- o x: maximum number of spans allowed within the application code (i.e. an integer number)
- o y: take values 1 (NRZ 2.5G), 2 (NRZ 10G), 9 (NRZ 25G), 3 (NRZ 40G), 7 (RZ 40G).
- o t: take values A (power levels suitable for a booster amplifier in the originating ONE and power levels suitable for a pre-amplifier in the terminating ONE), B (booster amplifier only), C (pre-amplifier only), D (no amplifiers).
- o z: take values 1 (1310 nm sources on ITU-T G.652 fibre), 2 (1550 nm sources on ITU-T G.652 fibre), 3 (1550 nm sources on ITU-T G.653 fibre), 5 (1550 nm sources on ITU-T G.655 fibre).

The following list of suffix can be added to these application codes:

- o F: FEC encoding.
- o D: Adaptative dispersion compensation.
- o E: receiver capable of dispersion compensation.
- o r: reduced target distance.
- o a: power levels appropriate to APD receivers.
- o b: power levels appropriate to PIN receivers.

The following encoding is proposed:

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								
+-----+																																							

Figure 5: G.689.1 Application Code Encoding

Where (values between parenthesis refer to ITU defined values as reported above):

B: = 1 bidirectional, = 0 otherwise.

p (prefix): = 0 reserved, = 2 (P).

n: maximum number of channels (10 bits, up to 1024 channels)

W: = 0 reserved, = 1 (I), = 2 (S), = 3 (L), = 4 (V), = 5 (U)

x: = number of spans (6 bits, up to 64 spans)

y: = 0 reserved, = 1 (1), = 2 (2), = 3 (3), = 7 (7), = 9 (9)

t: = 0 reserved, = 1 (A), = 2 (B), = 3 (C), = 4 (D)

z: = 0 reserved, = 1 (1), = 2 (2), = 3 (3), = 5 (5)

s (suffix): = 0 reserved, = 1 (F), = 2 (D), = 3 (e), = 4 (r), = 5 (a), = 6 (b)

[EDITOR NOTE] if suffixes may be used together, the field has to be redefined as bitfield.

Appendix B. Encoding example

In this section we try to represent how the encoding will change considering the Optical Interface Class. The main result of the Optical interface class will be not encoding saving in term of bytes but a simplified protocol support for new optical technologies.

According to Section 5 of [I-D.ietf-ccamp-rwa-wson-encode] the encoding shall foresee a list of: Input Modulation Type, Input FEC Type, Input Client Signal Types. All the basic objects has a lenght dependent on values carried on. For example if the modulation format is a standard one, the related sub TLV will be 32 bits, if the modulation formart is a proprietary one the length is not known a priori.

Using the concept of interface class the same object will likely become as the one represented in Figure 6.

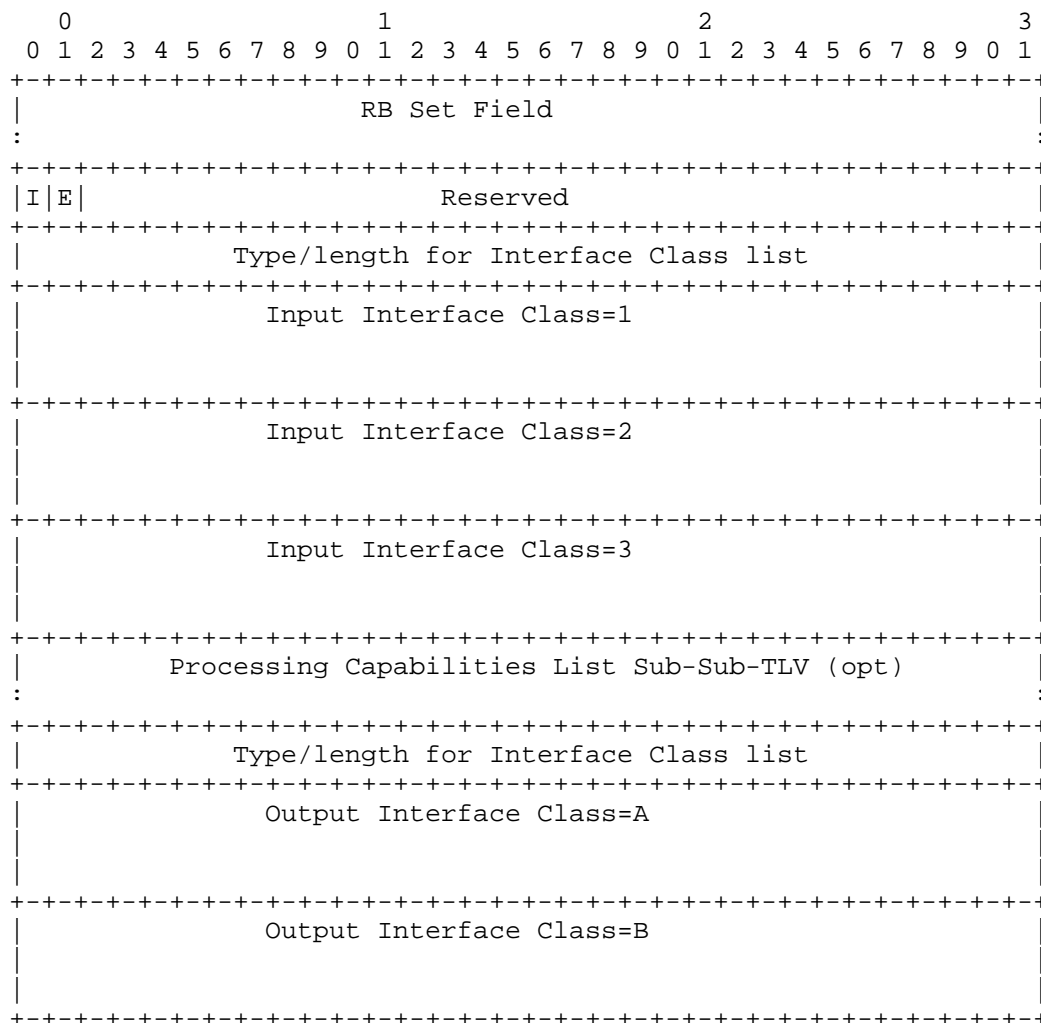


Figure 6

With the following notes:

- o Current draft just defines the Optical interface class encoding as 3 words of 32 bits but, for usage within WSON protocol extensions a proper TLV header shall be defined. In this case we represent a list since the original example in [I-D.ietf-ccamp-rwa-wson-encode] use lists.

- o Current example just represent input and output classes by numbers (1,2,3) and letters (A,B) since example only shows how encoding is simplified.
- o Optical interface classes has a fixed size while basic encoding blocks of [I-D.ietf-ccamp-rwa-wson-encode] have sizes that varies depending on proprietary informations.

As in the example above, the concept of Optical interface class is not to save encoding bytes but to keep the optical semantic outside of GMPLS protocols.

Authors' Addresses

Giovanni Martinelli (editor)
Cisco
via Philips 12
Monza 20900
IT

Phone: +39 039 209 2044
Email: giomarti@cisco.com

Gabriele M Galimberti
Cisco
Via Philips,12
20052 - Monza
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

Lyndon Ong
Ciena Corporation
US

Email: lyong@ciena.com

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

Email: daniele.ceccarelli@ericsson.com

Cyril Margaria
Nokia Siemens Networks
St-Martin str. 76
Mun_{chen}, 81541
Germany

Phone: +49-89-5159-16934
Email: cyril.margaria@nsn.com

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.]

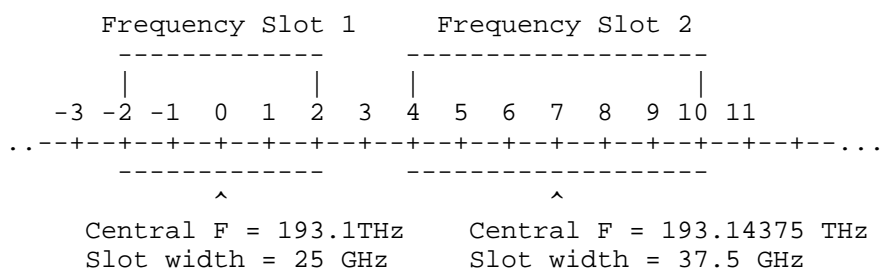
- o Optical Channel Slot (definition in the scope of a fixed grid DWDM network, to be adapted to a flexi-grid). The optical spectrum frequency range (portion of optical spectrum) allocated / occupied by a single optical channel. Each optical channel signal has a defined carrier central frequency and 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

```


- o Slot Width Granularity: 12.5 GHz, as defined in [G.694.1].
- o Slot Width: The slot width determines the "amount" of optical spectrum regardless of its actual "position" in the frequency axis. A slot width is constrained to be $m \times \text{SWG}$ (that is, $m \times 12.5 \text{ GHz}$), where m is an integer greater than or equal to 1.
- o Frequency Slot: The frequency range allocated to a slot within the flexible grid. A frequency slot is defined by its nominal central frequency and its slot width. Assuming a fixed and known central nominal frequency granularity, and assuming a fixed and known slot width granularity, a frequency slot is fully characterized by the values of 'n' and 'm'. Note that an equivalent characterization of a frequency slot is given by the start and end frequencies (i.e., a frequency range) which can, in turn, be defined by their respective values of 'n'. Note that a bidirectional optical transmission section layer network connection may be supported by one optical fiber for both directions (single fiber), or each direction of the connection may be supported by different fibers (pair of fibers). Since a frequency slot is a unidirectional entity (the same nominal central frequency cannot be used in two directions of transmission), the single fiber case is carried out by a pair of unidirectional frequency slots on the same fiber, and the pair of fibers case may have frequency slots that use the same nominal central frequencies.



The symbol '+' represents the allowed nominal central frequencies, the '--' represents the nominal central frequency granularity, and the '^' represents the slot nominal central frequency. The number on the top of the '+' symbol represents the 'n' in the frequency calculation formula. The nominal central frequency is 193.1 THz when n equals zero. Note that over a single frequency slot, one

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?].

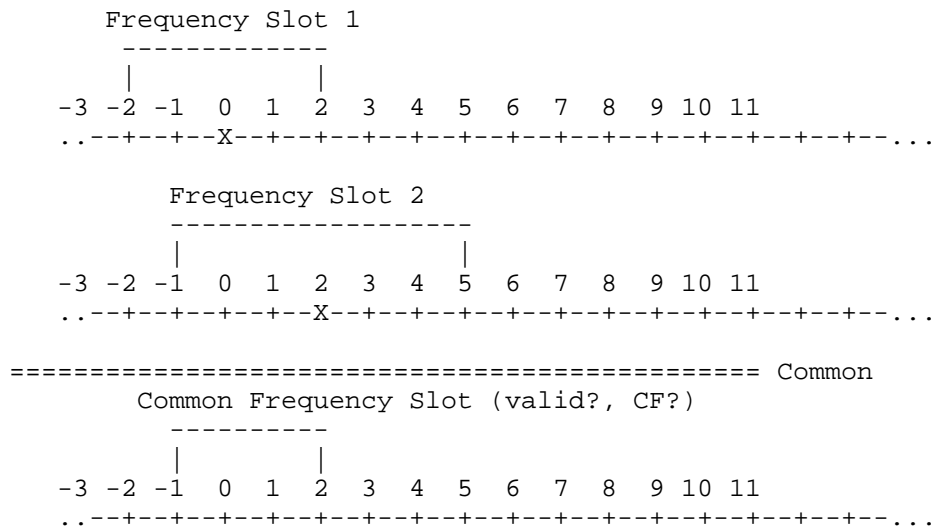


Figure 4. Common Frequency Slot

- 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.

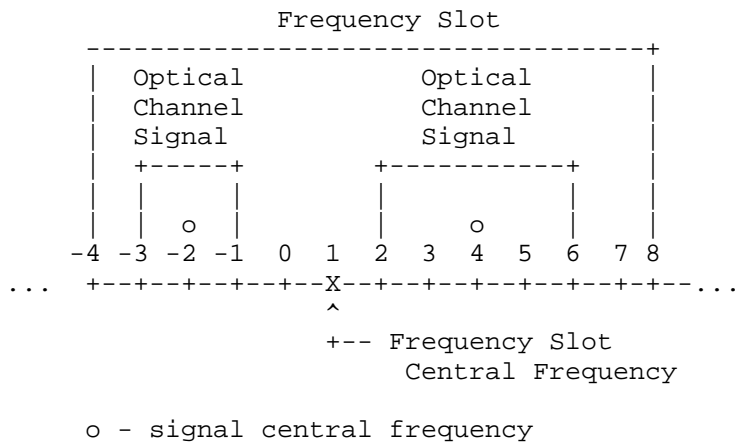


Figure 3. Frequency slot with 2 Optical channel signals

- 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

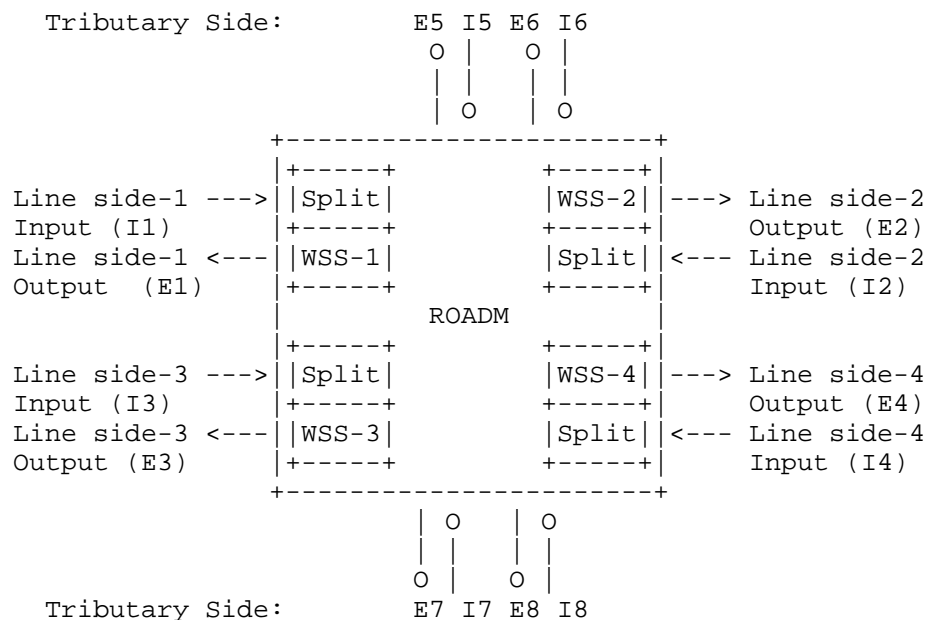


Figure 5. Simplified ROADM model with Line Sides and Tributaries

[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.

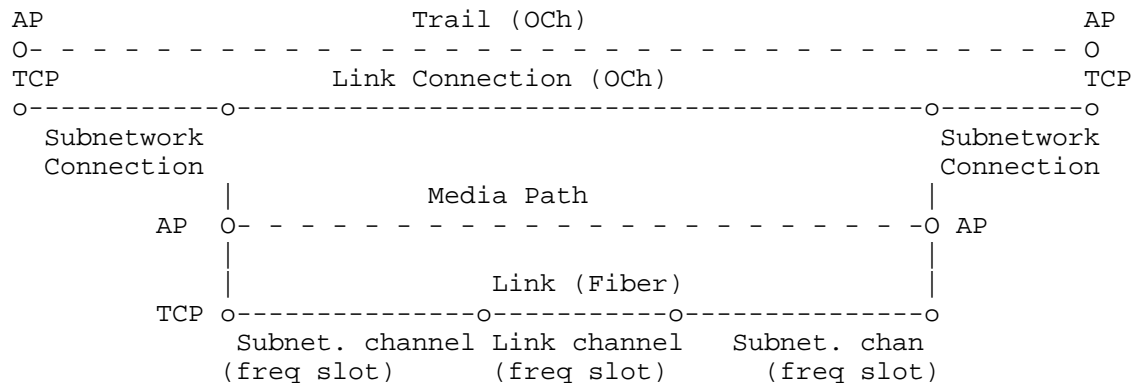


Figure 6. Layered Network Model G.805

[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.]

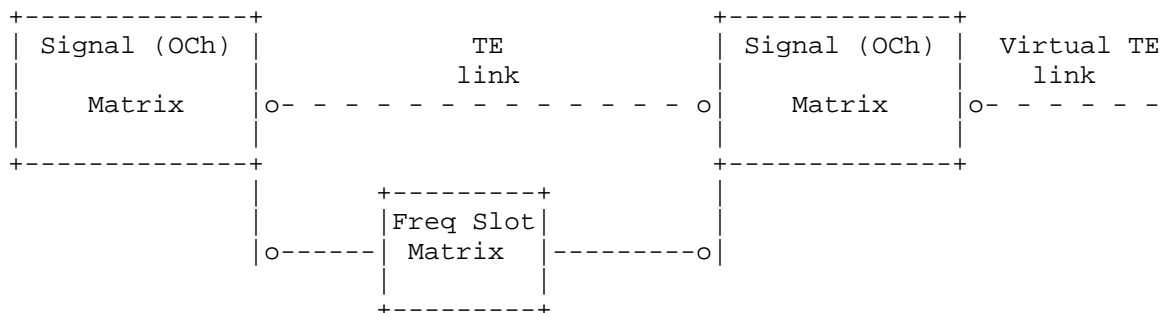


Figure 7. MRN/MLN topology view with TE link / FA

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 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.

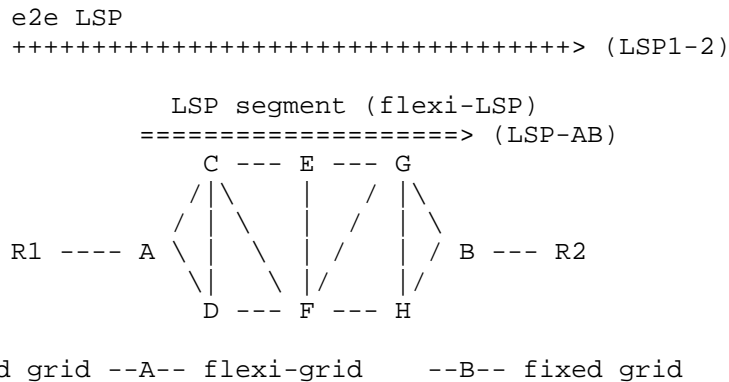


Figure 9. LSP Stitching [RFC5150] and relationship with fixed-flexi

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@ritt.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 (WSOs)", 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

Iftekhar Hussain
Infinera
140 Caspian Ct.
Sunnyvale, 94089
USA

Phone: 408-572-5233
Email: ihussain@infinera.com

Network Working Group
Internet Draft
Intended Status: Standards Track

Fatai Zhang
Xian Zhang
Huawei
D. Ceccarelli
D. Caviglia
Ericsson
Guoying Zhang
CATR
P. Grandi
S. Belotti
Alcatel-Lucent

Expires: January 13 2013

July 12, 2012

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

draft-zhang-ccamp-gmpls-g709-lmp-discovery-06.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 January 13, 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.

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 Requirements	5
4.1. Discovering the Granularity of the TS	5
4.2. Discovering the Supported LO ODU Signal Types	5
5. Extensions: LMP Link Summary Message	6
5.1. Message Extension.....	7
5.1.1. LinkSummary Message.....	7
5.1.2. LinkSummaryAck Message.....	7
5.1.3. LinkSummaryNack Message.....	7
5.2. Object Definitions.....	8
5.3. Procedures	10
6. Security Considerations.....	11
7. IANA Considerations	11
8. Acknowledgments	11
9. References	12
9.1. Normative References.....	12
9.2. Informative References.....	12
10. Authors' Addresses	12

11. Contributors	14
------------------------	----

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.

This document extends the LMP and describes the solution of discovering HO ODU link capability.

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 [ITUT-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)

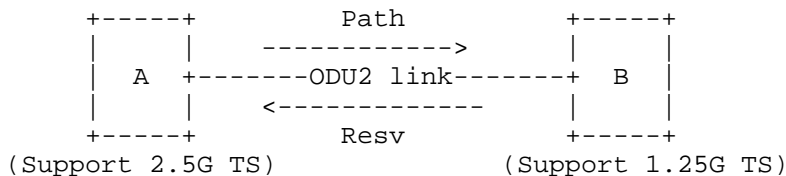
Note that If TS auto-negotiation is supported, a node supporting 1.25Gbps TS can interwork with the other nodes that supporting 2.5Gbps TS by combining specific TSs together in data plane, as descirbied in [OTN-frwk].

4. Link Capability Discovery Requirements

4.1. 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.



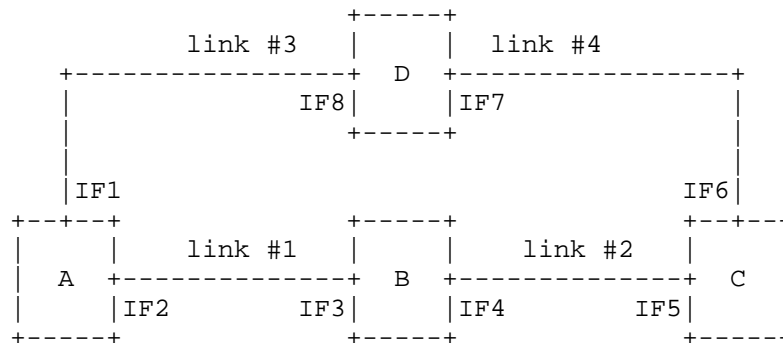
Therefore, for an ODU2 or ODU3 link, in order to reserve TS resources correctly for a LO ODU connection, the control plane of the two ends MUST know which granularity the other end can support before creating the LO ODU connection.

4.2. 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.

5. 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.

5.1. 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].

5.1.1. 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.

5.1.2. LinkSummaryAck Message

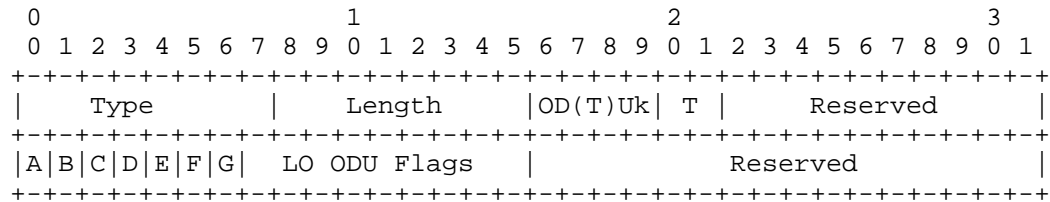
The LinkSummaryAck message is used to tell the remote end that it has the same capability as the remote end after the LinkSummary message is received by the local end.

5.1.3. LinkSummaryNack Message

The LinkSummaryNack message is used to tell the remote end that it has different capability from the remote end after the LinkSummary message is received by the local end. The LinkSummaryNack message also carries the HO ODU Link Capability subobject in the DATA_LINK object to tell the remote end the exact capability of the HO ODU link after negotiation, i.e., the granularity of TS and the types of LO ODU that both side of the HO ODU link can support.

5.2. Object Definitions

A new HO ODU Link Capability subobject type is introduced to the DATA LINK object to carry the HO ODU link capability information. The format of the new subobject is defined as follow:



Type (8 bits):

The value of this subobject type is TBD.

Length (8 bits):

The Length field contains the total length of the subobject in bytes, including the Type and Length fields. As for RFC 4204, the Length MUST be at least 4, and MUST be a multiple of 4. Value of this field is 8.

OD(T)Uk (4 bits):

This field is used to indicate the HO ODU link type (in case of LO ODU_j multiplexing into HO ODU_k, wherein j<k) or the OTU link type (in case of LO ODU_k mapping into OTU_k).

OD(T)Uk field	Signal type of HO ODU _k or OTU _k
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 ODUk (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 ODUj (j<k) flag not corresponding to the signal type HO ODUk/OTUk is set to 1 then the signal type OD(T)Uk has to be intended as HO ODUk and multiplexing of LO ODUj over HO ODUk 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 ODU2e 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.

5.3. Procedures

The Link Summary messages used for capability discovery for HO ODUk or OTUk 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.

6. Security Considerations

TBD.

7. IANA Considerations

TBD.

8. Acknowledgments

TBD.

9. References

9.1. 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] Fatai Zhang 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.
- [ITU-T-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.

9.2. 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.

10. 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

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

11. 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.

