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

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.
Abstract

When using GMPLS control plane, the ingress node may need to request remote node to perform route computation or expansion. In such cases, ingress node needs to convey the required objective function for the path computation algorithm to the remote node, so that the remote node can perform the route computation or expansion. Similarly, there are cases the ingress needs to indicate a TE metric bound for the loose segment which is expanded by the remote node(s). 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.6
   2.2. Metric subobject........................................7
   2.3. Processing Rules for the OF Subobjects................8
   2.4. Processing Rules for the Metric subobject.............9
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), network performance criteria (e.g. latency) are becoming as critical to data path selection as other metrics. In other words, such networks often require finding a path that is computed to minimize end-to-end latency. Even if paths are computed to minimize some other TE metric, it is often required to specify an acceptable latency bound as a constraint. In summary, there is a requirement to be able to find an end-to-end path with different optimization criteria and with performance bound(s) on the TE metric.

When the entire route for an LSP is computed at the ingress node, the above-mentioned 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 TE metrics (e.g., IGP metric, TE metric, hop counts, latency, etc.) to be optimized by the path computation algorithm at the remote 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 the 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 a similar mechanism to
the RSVP-TE protocol to allow an ingress node to indicate in a
Path request the desired objective function at the loose hops.
The objective function is used by the node performing route
expansion to find the "best" candidate paths. The draft also
defines a RSVP-TE protocol mechanism to allow an ingress node to
indicate TE metric bound(s) associated with the route expansion
request.

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|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:

```
+-----------------+-----------------+-----------------+-----------------+
| 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 I-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:

1. 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].

2. 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):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBA (suggest value: 66)</td>
<td>Objective Function (OF) subobject</td>
</tr>
<tr>
<td>TBA (suggest value: 67)</td>
<td>Metric subobject</td>
</tr>
</tbody>
</table>

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


6.2. Informative References


Authors' Addresses

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

George Swallow
Cisco Systems.
swallow@cisco.com
Include Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)
draft-ali-ccamp-rsvp-te-include-route-01.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). 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.
Abstract

There are scenarios in which it is required that two or more LSPs or segments of LSPs 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........................................4
   2.1. Explicit Inclusion Route Subobject (EIRS).....................5
   2.2. EIRS Subobject Processing Rule................................7
   2.3. Processing of EIRS with XRO and EXRS........................9
3. Security Considerations..........................................9
4. IANA Considerations.............................................9
5. Acknowledgments................................................9
6. References...................................................10
   6.1. Normative References......................................10
   6.2. Informative References..................................11
1. Introduction

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, including but not limited to 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 that two Label Switched Paths (LSPs) follow the same route but has no knowledge of how a loose hop of a reference LSP was expanded. There are scenarios in which it is required that two or more LSPs follow same route in the network. E.g., in many deployments it is required that 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: it is sometimes required 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 the 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 ingress node requires certain SRLGs 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 route is computed by the ingress node, the aforementioned requirements can be met by a local decision at the ingress node. However, there are scenarios when a route computation is not performed at the ingress and instead are performed by remote nodes, in which case there is 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 the above-mentioned requirements/scenarios and defines procedures that may be used to signal LSPs such that the entire LSP or segments of LSP 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:

```
+---------------+---------------+---------------+---------------+
|   L           |     Type     |     Length   |     Reserved  |
+---------------+---------------+---------------+---------------+
// one or more EIRS subobjects //
+---------------+---------------+---------------+---------------+
```

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

```
+---------------+---------------+---------------+---------------+
|   L           |     Type     |     Length   |     Reserved  |
+---------------+---------------+---------------+---------------+
|                    | SRLG Id 1 (4 bytes) |
+---------------+---------------------+---------------+
| SRLG Id 1 (continued) |     Reserved        |
+---------------+---------------------+---------------+
|                    | SRLG Id 2 (4 bytes) |
+---------------+---------------------+---------------+
| SRLG Id 2 (continued) |     Reserved        |
+---------------+---------------------+---------------+
```

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):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBA (suggest value: 68)</td>
<td>Explicit Inclusion Route Subobject (EIRS)</td>
</tr>
</tbody>
</table>

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


6.2. Informative References


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

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

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.

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
Internet-Draft    draft-ali-ccamp-te-metric-recording-01.txt

material may not have granted the IETF Trust the right to allow
modifications of such material outside the IETF Standards Process.
Without obtaining an adequate license from the person(s)
controlling the copyright in such materials, this document may not
be modified outside the IETF Standards Process, and derivative
works of it may not be created outside the IETF Standards Process,
except to format it for publication as an RFC or to translate it
into languages other than English.

Abstract

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

Conventions used in this document

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

Table of Contents

Copyright Notice..................................................1
1. Introduction....................................................................3
2. RSVP-TE Requirement..................................................3
  2.1. Cost, Latency and Latency Variation Collection Indication....3
  2.2. Cost, Latency and Latency Variation Collection.................4
  2.3. Cost, Latency and Latency Variation Update....................4
3. RSVP-TE signaling extensions.......................................4
  3.1. Cost Collection Flag............................................4
  3.2. Latency Collection Flag........................................4
  3.3. Latency Variation Collection Flag.............................5
  3.4. Cost subobject................................................5
  3.5. Latency subobject...............................................6
  3.6. Latency Variation subobject....................................6
  3.7. Signaling Procedures..........................................7
4. Security Considerations............................................9
5. IANA Considerations................................................9
  5.1. RSVP Attribute Bit Flags.....................................9
  5.2. New RSVP error sub-code.....................................10
6. Acknowledgments...................................................10
7. References........................................................10
  7.1. Normative References..........................................10

Ali, Swallow, Filsfils         Expires September 2012     [Page 2]
1. Introduction

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

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

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

2. RSVP-TE Requirement

This section outlines RSVP-TE requirements for the support of the automatic discovery of cost, latency and latency variation attributes of an LSP. These requirements are very similar to the requirement of discovering the Shared Risk Link Groups (SRLGs) associated with the route taken by an LSP [DRAFT-SRLG-RECORDING].

2.1. Cost, Latency and Latency Variation Collection Indication

The ingress and egress nodes of the LSP must be capable of indicating whether the cost, latency and latency variation
attributes of the LSP should be collected during the signaling procedure of setting up the LSP.

2.2. Cost, Latency and Latency Variation Collection

The endpoints of the LSP may collect the cost, latency and latency variation information and use it for routing, flooding, and TE link configuration purposes.

2.3. Cost, Latency and Latency Variation Update

When the cost, latency and latency variation property of a TE link along the LSP route changes, e.g., if the administrator changes cost of a TE link, the endpoints of the LSP need to be capable of updating the cost, latency and latency variation information of the path. Similarly, if a path segment of the LSP is rerouted, the endpoints of the LSP need to be capable of updating the cost, latency and latency variation information of the path. In summary, the signaling should be capable of updating the new cost, latency and latency variation information to the endpoints.

3. RSVP-TE signaling extensions

3.1. Cost Collection Flag

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

Cost Collection flag (to be assigned by IANA, recommended bit position 9)

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

The rules of the processing of the Attribute Flags TLV follows [RFC5420].

3.2. Latency Collection Flag

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

Latency Collection flag (to be assigned by IANA, recommended bit position 10)

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

The rules of the processing of the Attribute Flags TLV follows
[RFC5420].

3.3. Latency Variation Collection Flag

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

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

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

The rules of the processing of the Attribute Flags TLV follows
[RFC5420].

3.4. Cost subobject

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Length</th>
<th>Reserved (must be zero)</th>
<th>COST Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

Length: The Length value is set to 8.

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

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

The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and RRO are not changed.

3.5. Latency subobject

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

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

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

Length: The Length value is set to 8.

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

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

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

The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and RRO are not changed.

3.6. Latency Variation subobject

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

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

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

Length: The Length value is set to 8.

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

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

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

The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and RRO are not changed.
Type: The type of the subobject, to be assigned by IANA (recommended value 37).

Length: The Length value is set to 8.

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

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

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

The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and RRO are not changed.

3.7. Signaling Procedures

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

When a node receives a Path message which carries an LSP_REQUIRED_ATTRIBUTES Object and the Cost, Latency or/ and Latency Variation Collection Flag(s) is (are) set, if local policy disallows providing the requested information to the endpoints, the node SHOULD return a Path Error message with error code "Policy Control Failure (2)" and one of the following error subcodes:

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

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

When a node receives a Path message which carries an LSP_REQUIRED_ATTRIBUTES Object and the Cost, Latency or/ and Latency Variation Collection Flag(s) is (are) set, if local policy allows providing the requested information to the endpoints, the node MUST add the requested subobject(s) with the cost, latency or/ and latency variation metric value(s) associated with the local hop to the Path RRO. Then it forwards the Path message to the next node in the downstream direction.

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

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

Typically, cost and latency are additive metrics, but latency variation is not an additive metric. How the ingress and egress nodes computes the end-to-end cost, latency or/ and latency variation metric from information recorded in the RRO is beyond the scope of this document.

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

Based on the local policy, a transit node (e.g. the edge node of a domain) may edit the RRO to remove the route information (e.g. node, interface identifier information) before forwarding it and can summarize the cost, latency or/ and latency variation as a single number for the loose hop that is summarized by the edge node. How a transit node calculates the cost, latency or/ and latency variation metric for the segment summarized by the transit node is beyond the scope of this document.
4. Security Considerations

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

5. IANA Considerations

5.1. RSVP Attribute Bit Flags

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

This document introduces the following three new Attribute Bit Flag:

- Bit number: TBD (recommended bit position 9)
- Defining RFC: this I-D
- Name of bit: Cost Collection Flag

- Bit number: TBD (recommended bit position 9)
- Defining RFC: this I-D
- Name of bit: Latency Collection Flag

- Bit number: TBD (recommended bit position 9)
- Defining RFC: this I-D
- Name of bit: Latency Variation Flag

5.2. ROUTE_RECORD subobject

This document introduces the following three new RRO subobject:

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>--------------</td>
<td>----------------------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
5.2. New RSVP error sub-code

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

<table>
<thead>
<tr>
<th>Sub-code</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Recoding Rejected</td>
<td>To be assigned by IANA.</td>
</tr>
<tr>
<td>Suggested Value:</td>
<td>105.</td>
</tr>
<tr>
<td>Latency Recoding Rejected</td>
<td>To be assigned by IANA.</td>
</tr>
<tr>
<td>Suggested Value:</td>
<td>106.</td>
</tr>
<tr>
<td>Latency Variation Recoding Rejected</td>
<td>To be assigned by IANA.</td>
</tr>
<tr>
<td>Suggested Value:</td>
<td>107.</td>
</tr>
</tbody>
</table>

6. Acknowledgments

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

7. References

7.1. Normative References


7.2. Informative References


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

Rudiger Kunze
Deutsche Telekom AG
Ruediger.Kunze@telekom.de
CCAMP Working Group
Internet Draft
Intended status: Standard Track
Expires: September 11, 2012

Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) LSP Route Diversity using Exclude Routes
draft-ali-ccamp-xro-lsp-subobject-01.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). 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.
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 ...................................4
2. RSVP-TE signaling extensions ...............................5
   2.1. Terminology ...........................................5
   2.2. LSP Subobject ..........................................5
   2.3. Processing rules for the LSP subobject ...............8
   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

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.
1.1. Requirements Notation

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. RSVP-TE signaling extensions

This section describes the signaling extensions required to address the aforementioned requirements. Specifically, this document defines a new LSP subobject to be signaled in the EXCLUDE_ROUTE object (XRO) and/ or Explicit Exclusion Route Subobject (EXRS) defined in [RFC4874]. Inclusion of the LSP subobject in any other RSVP object is not defined.

2.1. Terminology

In this document, the following terminology is adapted:

LSP1/TUNNEL1: LSP1/TUNNEL1 is the LSP/tunnel from which diversity is required.

LSP2/TUNNEL2: The term LSP2/TUNNEL2 is used to refer the LSP being signaled with XRO/ EXRS containing the LSP subobject referencing LSP1/TUNNEL1.

CircuitID: The term CircuitID refers to the LSP Forwarding Equivalence Class (FEC) (LSP ID field of the FEC may be ignored depending on the context the CircuitID term is used).

CircuitIDx: The term CircuitIDx refers to CircuitID of LSPx/TUNNELx.

2.2. LSP Subobject

A new IPv4 Point-to-Point (P2P) LSP subobject is defined by this document as follows (IPv6 P2P LSP subobject will be defined in a later version of the document).

```
+----------------+----------------+----------------+----------------+----------------+----------------+----------------+----------------+
|     L     |     Type     |     Length    |Attribute Flags|Exclusion Flags|
+----------------+----------------+----------------+----------------+----------------+----------------+----------------+
|                 IPv4 tunnel end point address                 |
```
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
A new subobject type is introduced; the subobject type is to be defined 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].

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 to 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 proceeds 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/TUNNELL1) 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/TUNNELL1) referenced in the LSP subobject becomes known (e.g. when LSP1 is signaled) or the TUNNELL1 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/ TUNNELL1) 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |Attribute Flags|Exclusion Flags|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IPv4 tunnel end point address                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Must Be Zero         |     Tunnel ID                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Extended Tunnel ID                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   IPv4 tunnel sender address                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Must Be Zero         |            LSP ID             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The meaning of respective fields in EXRS header is as defined in [RFC4874]. Similarly, the meaning of respective fields in IPv4 P2P LSP subobject is as defined earlier in this document. This is with the exceptions that:
- Processing node exception applies to the node processing the ERO.

- If L bit in the ERO header is not set (ERO.L = 0), the IPv4 P2P LSP subobject is processed against the LSPs for which the processing node is ingress, egress or a transit node.

- Penultimate node exception applies to the penultimate node of the loose hop. This flag is only processed if EXRS.L bit is set, i.e., in the loose ERO hop case.

- Destination node exception applies to the abstract node to which the route is expanded. This flag is only processed if EXRS.L bit is set, i.e., in the loose ERO hop case.

2.4.1. Processing Rules for the EXRS with LSP subobject

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

3. Security Considerations

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

4. IANA Considerations

4.1. New XRO subobject type

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

<table>
<thead>
<tr>
<th>Subobject Type</th>
<th>Description</th>
<th>Subobject</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>IPv4 P2P LSP subobject</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Sub-code</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route to XRO LSP unknown</td>
<td>To be assigned by IANA.</td>
</tr>
<tr>
<td></td>
<td>Suggested Value: 13.</td>
</tr>
<tr>
<td>Failed to respect Exclude Route</td>
<td>To be assigned by IANA.</td>
</tr>
<tr>
<td></td>
<td>Suggested Value: 14.</td>
</tr>
</tbody>
</table>

5. Acknowledgement

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

6. References

6.1. Normative References


6.2. Informative References


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: matt@matthartley.net

Ori Gerstel
Cisco Systems
ogerstel@cisco.com

Gabriele Maria Galimberti
Cisco Systems
ggalimbe@cisco.com

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

Julien Meuric
France Telecom Orange
Email: julien.meuric@orange.com
CCAMP Working Group                                       V.Beeram (Ed)
Internet Draft                                                I.Bryskin
Intended status: Standards Track                               W.Doonan
ADVA Optical Networking                                      J.Drake (Ed)
G.Grammel                                                   Manuel Paul
Juniper Networks                                           Ruediger Kunze
Deutsche Telekom                                          Friedrich Armbruster
Cyril Magaria                                    Oscar Gonzalez de Dios
NSN                                                        Telefonica
Expires: September 12, 2012                              March 12, 2012

GMPLS-UNI BCP                                      draft-beeram-ccamp-gmpls-uni-bcp-01.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September, 2012.

Copyright Notice

Abstract

This document pools together the best current practices that are being used to apply the GMPLS Overlay model at the User-Network Interface (UNI) reference point (as defined in [G.8080])

Table of Contents

1. Introduction ................................................... 3
2. Multi-Layered Approach ........................................ 3
3. Traffic Engineering ............................................ 4
   3.1. Augmenting the Client-Layer Topology .................... 6
       3.1.1. Virtual TE Links .................................... 7
   3.2. Macro SRLGs ................................................. 8
   3.3. MELGs ...................................................... 9
   3.4. Switching Constraints .................................... 10
4. Connection Setup ................................................ 10
5. Path computation aspects ....................................... 11
6. L1VPNs .......................................................... 12
7. Use cases ........................................................ 13
   7.1. Service optimization and restoration in Multi-Layer Networks ............................................. 13
    7.2. IP/MPLS Offloading with UNI automation .................. 14
    7.3. Use of PCE and VNTM in Multilayer Network Operation ................................................. 15
8. Security Considerations .......................................... 15
9. IANA Considerations ............................................. 15
10. References ...................................................... 15
    10.1. Normative References .................................... 15
    10.2. Informative References .................................. 16
11. Acknowledgments ................................................. 16
1. Introduction

Generalized Multiprotocol Label Switching (GMPLS) provides tools to create end-to-end services in various transport technologies. These tools can be used to support service management in different types of deployment models. [RFC 4208] discusses how GMPLS can be applied to the overlay model. There are a good number of implementations that have built on the basic concepts discussed in [RFC 4208] and have successfully demonstrated interoperability. This document is an attempt to pool together the best current practices that are being used to apply the GMPLS Overlay model at the User-Network Interface (UNI) reference point (as defined in [G.8080]).

[RFC 4208] recommends the use of hierarchical service activation when GMPLS is used for the core network and section 7.3.3 of [RFC4847], "Virtual Link Service Model" augments this by introducing a representation of server-layer network resources into a client-layer network topology. This memo explains how this augmentation enhances client-layer networking in an overlay model. The concepts discussed in this document are based primarily on experiences drawn from interoperating GMPLS-enabled IP routers with Optical Transport elements, but any GMPLS supported technology may be used in the client and server-layer networks.

2. Multi-Layered Approach

When an end-to-end service crosses a boundary between two regions of dissimilar transport technology, it is necessary to execute distinct forms of service activation within each region.

Fig 1: Sample Hybrid Topology
(See PDF version)

For example, in the hybrid network illustrated in Fig 1, provisioning a transport service between two GMPLS-enabled IP routers on either side of the optical WDM transport topology 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 along the determined path, per the set of figures shown in Fig 2.

Fig 2: Hierarchical Service Action
(see PDF version)

3. Traffic Engineering

The previous section outlines the basic method for activating end-to-end services across a multi-layer network. As a necessary part of that process an initial path selection process was performed, whereby an appropriate path between the desired endpoints was 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 will provide automated mechanisms for computing the desired path or paths between network endpoints.

In particular, operators do not expect under normal circumstances to be required to explicitly specify the end-to-end path; rather, operators 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 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) selected by ingress element of an end-to-end service. In order to be able to compute and qualify paths, the PCE must SHOULD be provided with information regarding the traffic engineering capabilities of the layer network to which it is associated, 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 elements in the client-layer is quite distinct from that required by the elements in the server-layer.
network. Hence, the server-layer traffic engineering links are of no importance for the client-layer network, and it is actually desirable to block their advertisements into the client TE domain by the server-layer border nodes.

For example, in the sample hybrid network (Fig 1) there are multiple optical transport elements supporting the connection between the GMPLS-enabled IP routers, and hence the physical topology between them includes several nodes and links. However, the optical elements between the IP routers are not able to switch traffic within the client-layer network of routers (e.g. IP/MPLS), as the optical elements are lambda switches, not IP/MPLS switches. Hence while the intervening optical elements may physically exist along the path, they are not a part of the topology available to the IP/MPLS routers 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.

Fig 3: Traffic Engineering - ERO with "loose hop"
(See PDF version)

In this example, the TE topology associated with the client-layer network is indicated by the links and nodes colored yellow, whereas the TE topology associated with the server-layer network is indicated by the links and nodes colored green. The nodes at the edge of the server-layer network are visible in both the topologies. The yellow topology is capable of switching traffic within the client-layer, whereas the green 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 yellow topology to which the B and D routers is connected does not include a fully-yellow path 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; and [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 yellow topology MUST be sufficiently augmented to indicate where there are paths through the green topology which can provide connectivity between nodes in the yellow 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 periodically computed by the server-layer network and 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.

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 are in the same layer network. For example, IP routers that are connected via GMPLS UNI to a WDM network may be capable of terminating optical trails that are lambda switched by the network. Because the network domain normally would not want to leak its actual topology information into the client domain, clients would not be able to compute end-to-end paths across the network domain despite that client and network links belong to the same (WDM) layer network. The method described in the following sections of this document solves the problem of partitioned client topology for this case as well.

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 at node F can be connected to the transponder in node J via the optical path F-G-H-J. A lambda LSP can be provisioned in the server-layer along this path, and then advertised as a TE link into the client-layer. With the availability of this link, the path computation function at node A is able to compute an end-to-end path from A to C.

Fig 4: Traffic Engineering - End to End Path Computation
(See PDF version)
In this case, in order for the TE link to be made available in the client-layer network topology, the network resources corresponding to the underlying server-layer LSP MUST be fully provisioned beforehand.

As another scenario, consider a network configuration where the transponders at nodes E, F, J and I are connected to each other via directionless ROADM components. 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 element containing an adaptation function to the client-layer network, the operator in this case would not wish to reserve any network resources in the server-layer network until a client LSP is signaled. The next section proposes a method to address this common operational requirement.

3.1.1. Virtual TE Links

A "Virtual TE Link" as defined in section 7.3.3 of [RFC4847] is a TE link that is advertised into the client-layer network, with the available but not necessarily reserved/commuted 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 sharing of server-layer network resources with other Virtual TE links.

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 (yellow 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 connection, 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 resources with the TE-link in question start advertising "zero" available bandwidth. Likewise, the TE network image reverts back to the original form as soon as the last client-layer connection, going through the TE link in question, is released, i.e. Virtual TE Link becomes "virtual" again.

3.2. Macro SRLGs

The Virtual TE links that are advertised into the client-layer network topology cannot be assumed to be totally 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 (real 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 to use SRLGs as a part of Virtual TE Link advertisements.

A traditional SRLG 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. However, there is a scalability issue with physical SRLGs in multi-layer environments. For example, if a WDM layer LSP serves an IP layer link, every WDM link and node traversed by the LSP MUST be considered as a separate SRLG. The number of SRLGs to be advertised to client (e.g. IP) layer per TE link would be directly proportional to the number of hops traversed by the underlying server-layer LSP.

The notion of Macro SRLGs addresses this scaling problem. Macro SRLGs have the same protocol format as their physical counterparts and can be assigned automatically for each Virtual TE Link that is advertised into the client-layer network as a result of the existence of an underlying server-layer LSP (instantiated or
otherwise). A Macro SRLG represents a set of shared path segments that are traversed by two or more of the underlying server-layer LSPs. Each shared path segment can be viewed as a sequence of shared resources where each individual resource has a physical SRLG associated with it (example depicted in Fig 6). The actual procedure for deriving these Macro SRLGs is beyond the scope of this document.

Fig 6: Macro SRLGs
(See PDF version)

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 could be mapped onto 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 could be mutually exclusive.

One example of mutually exclusive Virtual TE Links is when the paths for the network domain LSPs supporting the Virtual TE Links not only intersect, but also require usage of the same resource (e.g. lambda channel) on the intersection. 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 end-to-end paths) it is important to know about such mutually exclusive relationship between Virtual TE Links. This memo 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 computer will be able to decide on whether two or more Virtual TE Links are mutually exclusive or not by finding the 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 bit of information is important for a path computer: committing new Virtual TE links (vs. re-using committed ones) has a consequence of committing more network resources and disabling other Virtual TE links that have common MELGs with newly committed Virtual TE Link.

Exact format of the MELGs sub-TLV is described in [MELG]

[TBD: MELG Figure/Example]

3.4. Switching Constraints

Certain types of network configurations necessitate the specification of connectivity constraints in the Virtual TE Link advertisements. If the switching constraints associated with the binding of Virtual and access TE links terminated on a given network border node do not get advertised into the client domain, there is a risk of an invalid path being computed (Fig 7). This document recommends the use of the extensions specified in [GEN_CNSTR] to address this issue.

Fig 7: Switching Constraints
(See PDF version)

4. Connection Setup

Experience with control plane operations in multi-layer networks indicates there are benefits to coordinating certain signaling operations, in the following manner. Consider the scenario where the network domain 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-layer service would need to be somehow coordinated with that of the server-layer service. To avoid this operationally awkward issue, a phased connection setup process as depicted in Fig 8 is proposed.
As long as the LSP segment across the server-layer network is not completely "UP" (e.g., Fully Power Equalized), the nodes at the edge of the server-layer network through which the LSP passes would signal the client-layer PATH/RESV messages with the T (Testing) bit set in the ADMIN_STATUS. The T bit would be cleared in these messages only after the LSP segment across the server-layer network is deemed fully operable.

5. Path computation aspects

It is assumed that a client domain path computation function makes use of advertised client domain 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, (Corresponding to RFC4655 Composite PCE node) or remote (i.e. network/External PCEs). Path computations could be triggered by client nodes or NMS. Generally speaking, the responsibility of the client domain path computation function is to compute one or two paths for each source-destination pair of the TE-LSPs. Path computation SHOULD be subject to one or more path optimization criterions (such as shortest path, minimal latency, etc.) and path computation constraints (e.g. link unreserved bandwidth, link colors, layer-specific constraints, explicit exclusions, etc.)

As the augmented topology does hide server layer links and nodes, it is RECOMMENDED to support SRLG diverse path computation.

Furthermore the path computation SHOULD consider the connectivity and switching constraint in addition to all usual TE path computation constraints (e.g. unreserved bandwidth, link colors, layer-specific constraint) when available.

When using PCE architecture and PCEP protocol those aspects are covered by RFC5440, RFC5521 and RFC5541.

As described in section 3.3. Virtual TE link may not only share risk but may also depend on the same also server layer resources, thus creating mutual exclusivity between Virtual TE Links. Therefore, network topologies containing Virtual TE links have an increased probability of LSP setup failures. In such topologies concurrent path computation that takes in consideration MELG will reduce
signaling failures (Not considering MELGs may result, for example, in two LSPs routed on two Virtual TE-Links sharing the same server layer resource). PCEP supports concurrent path computation per RFC5440, expressing MELG constraint is out of scope of this document (defined in [MELG])

Core domain path computation and Inter-PCE path computation is out of scope for this document.

6. L1VPNs

[RFC4208] implies that multiple independent sets of clients, located in the same or different layer networks, could be connected to the same network domain, providing the connectivity between the clients within each set, while blocking the connectivity between the clients from different sets (i.e. allows for the L1VPNs application).

This document suggests:

- New sub-TLV - VPN IDs sub-TLV - to be added to the top level TE Link TLV. Exact format of the VPN IDs sub-TLV is described in [GMPLS UNI RTG]
- Configuring on the network end of each access TE link zero, one or more network unique VPN IDs and adding the configured information as VPN IDs sub-TLV to the TE link advertisement;
- Configuring zero, one or more network unique VPN IDs for each Virtual TE Link and adding the configured information as VPN IDs sub-TLV to the TE link advertisement;
- Making the network responsible for proper filtering of the TE Link advertisements, so that the information pertinent to VPN X is leaked only to the clients that are members of the said VPN X

This approach would achieve the following:

- Automatic VPN member auto-discovery;
- Providing to the clients VPN specific view of the network;
- Partitioning network resources between VPNs;
- Ensuring successful path computations (and therefore connectivity) only between members of the same VPN

[RFC4208] implies that access TE Links could be named from a single or a separate (per-VPN) name space. This draft takes the former approach, that is, regardless of the associated VPNs, all access and Virtual TE Links MUST be named from the same (specifically, network) name space. Apart from simplicity, one reason for such choice is the following consideration: a GMPLS LSP established between a pair of
clients is likely to be advertised as a TE Link into the client’s layer TE domain. For example, a GMPLS LSP established between a pair of IP routers is likely to be advertised as a TE Link into IP/MPLS layer TE domain. This means that neither access nor Virtual TE Links belong to the "real" client layer network. Hence assigning addresses for access and Virtual TE links from the network name space would not cause address collisions/re-configurations in the client layer.

[TBD: L1VPN Figure/Example]

7. Use cases

7.1. Service optimization and restoration in Multi-Layer Networks

Multi-layer networks, as described in this document, are a reality today and they are operated by different groups following different operational procedures.

This requires an independent optimization of the client and server layer networks, and this could lead to the situation where the re-routing of a client layer LSP fails because some of the resources on the selected alternate path share fate with some of the resources on the LSP’s failed path. This would happen due to lack of knowledge of the server layer network when the client layer path computation function selects the alternative path.

The high percentage of IP traffic in operator networks today makes it necessary that client and server layer share sufficient information to enable an optimized transport for IP/MPLS services and address existing inefficiencies. One important point from the carrier perspective is that the usage of server-layer SRLG information by the client layer path computation is essential to address these issues.

In a typical multi-layer network, in which the IP/MPLS network is the client network and the WDM/OTN network is the server network, it is the client layer network that is responsible for the protection of the IP/MPLS traffic using mechanisms such as FRR and/or LFA. Regardless of the mechanism that is used, SRLG information from the server layer network helps to optimize the client layer network with respect to reduced link utilization and reliable and efficient protection of the client traffic.
Today server layer network SRLGs are used mainly to calculate diverse alternative paths for the IP/MPLS client layer network. Therefore the following procedure MUST be periodically performed:

- Build traffic matrix for the server layer network
  (based on IP links)
- Solve traffic engineering problems in the server layer network
- Calculate new SRLGs for the client layer network
- Simulate failure scenarios

GMPLS UNI reduces the OPEX costs of doing these procedures manually by providing:

- the advertisement of server layer network SRLG information into client layer network via common routing protocol
- the client layer network path computation function uses this SRLG information in selecting maximally diverse paths.

7.2. IP/MPLS Offloading with UNI 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 network segment in the IP network by placing IP traffic into GMPLS LSPs in the server layer network in order to reduce the load on intermediate IP routers. The increase in traffic is typically caused by an elevated number of high traffic flows/services traversing an IP network segment, which requires core routers to forward large IP traffic volumes.

The decision process driving IP offloading is complex and is constrained by a set of rules that 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 must establish GMPLS LSPs in the server layer (e.g. optical) network and automatically assign them identifiers, either numbered or unnumbered, in the client layer network. This information can be automatically exchanged using the procedures from [RFC 4203]. However, such procedures are not always implemented in commercial equipment. Consequently, this information may need to be configured manually as part of the initial set-up/installation of these LSPs.
Later, when the GMPLS LSP tunnel needs to be established, the hierarchical TE Link addresses MUST be included in the UNI path request.

7.3. Use of PCE and VNTM in Multilayer 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 GMPLS LSPs, in the server layer network. I.e., in can be used to provide TE links to the client layer network in real time.

Several cooperation modes between PCE, VNTM and the NMS have been proposed in [RFC 5623]. For instance, the operator can request a new MPLS path via the NMS, which consults a PCE with information of the multi-layer network. The PCE, in case that there are enough resources in the MPLS layer, returns a path made of real TE links. On the other hand, if there is a lack of resources at the 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 augmented topology (SRLGs, MELGs, TE metrics of the Virtual TE-Links), which can be provided via GMPLS-UNI.

8. Security Considerations

TBD

9. IANA Considerations

This document has no actions for IANA.

10. References

10.1. Normative References

10.2. Informative References


11. Acknowledgments

TBD

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
Revised Definition of The GMPLS Switching Capability and Type Fields

draft-berger-ccamp-swcaps-update-00.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/1id-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 August 20, 2012
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 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 were 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 issues 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 also defines, for routing, a generic method for identifying a hierarchy levels within a switching technology.

An alternate approach, which is not advocated by this document, is to ensure that Switching Types are assigned for all hierarchy levels within a switching technology as part of any new work, e.g., as part of [GMPLS-G709].

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]:

```
+---------------+-------------------+-------------------+
<table>
<thead>
<tr>
<th>Switching Cap</th>
<th>Encoding</th>
<th>Reserved</th>
</tr>
</thead>
</table>
+---------------+-------------------+-------------------+
```

... 

```
+---------------+-------------------+-------------------+
| Switching Cap |        reserved   | Switching Cap     |
|---------------|-------------------| specific info     |
|               |                   | variable          |
+---------------+-------------------+-------------------+
```

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]:

```
+---------------+-------------------+-------------------+
<table>
<thead>
<tr>
<th>LSP Enc. Type</th>
<th>Switching Type</th>
<th>G-PID</th>
</tr>
</thead>
</table>
+---------------+-------------------+-------------------+
```

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 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 should use two different Switching Types. Additionally, a different Switching Type MUST be used to indicate 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:

<table>
<thead>
<tr>
<th>Value</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Packet-Switch Capable-2 (PSC-2)</td>
</tr>
<tr>
<td>3</td>
<td>Packet-Switch Capable-3 (PSC-3)</td>
</tr>
<tr>
<td>4</td>
<td>Packet-Switch Capable-4 (PSC-4)</td>
</tr>
</tbody>
</table>

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

3. Intra-Layer Hierarchy

Authors note: This section is for discussion and may be dropped. Particularly, need to revisit MLN/IACD/XRO implications to ensure there are no gaiting issues.

Multiple switching technologies support forms of hierarchical switching within a particular data plane technology. As discussed above, GMPLS routing originally envisioned support for such cases for packet networks using PSC-2, 3, 4. In other cases, GMPLS defined support using technology specific mechanisms, for example Signal Type was defined for SONET/SDH, see [RFC4606]. Given that one of the objectives of GMPLS is to generalize control plane protocols, it is reasonable to define a method for supporting hierarchical switching within a particular data plane technology that is not specific to any particular technology. This section defines such a mechanism for routing. No additional mechanism is defined for signaling.

In order to support hierarchical switching within a particular data plane technology in routing, this section defines an Intra-Layer Hierarchy, or ILH, field. This field allows for representation of up to 15 layers of hierarchical switching. The ILH field is carried in a portion of the previously defined reserved field of the Interface Switching Capability Descriptor and has the following format:

```
 0                   1                   2                   3
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap | Encoding |       Reserved       |  ILH  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

For compatibility reasons, an ILH value of 0 indicates that the ILH field is not being used. The mapping of ILH values to specific levels of hierarchy within a data plane technology is specific to each switching technology and is therefore outside the scope of this document.
4. Compatibility

This document has two impacts on existing implementations. Both routing and signaling impacts must be considered.

For existing implementations, the primary impact is deprecating the use of PSC-2, 3 and 4. At the time of publication of this document, 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.

A secondary impact is the use of the previously reserved field of the routing Interface Switching Capability Descriptor. For existing routing implementations, this field should be set to all zeros when generating a Descriptor, and should be ignored on receipt. Furthermore, existing nodes are expected to propagate reserved fields without any modification. Therefore the use of this reserved field is not considered to result in any compatibility issues in routing. As this field is not used in signaling, there are no signaling compatibility issues.

5. 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].

6. IANA Considerations

IANA needs to deprecate and redefine the registry.

7. Acknowledgments

We thank John Drake for highlighting the current inconsistent definitions associated with the Switching Capability and Type Fields.

8. References

8.1. Normative References


8.2. Informative References


9. 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
Abstract

The continuous increase of flexibility and bit rate in optical networks has higher and higher impacts on inter-channel effects (e.g. Cross-phase modulations). This effect leads to the introduction of Guard Bands between adjacent light paths in order to reduce the inter-channel detrimental effects.

This document provides requirements for the development of protocol extensions to support Generalized Multi-Protocol Label Switching (GMPLS) and Path Computation Element (PCE) management of Guard Bands.

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 6, 2012.

Copyright Notice
Table of Contents

1.  Introduction ...................................................... 3
   1.1.  Terminology ................................................ 3
2.  Definitions ........................................................ 3
3.  Scenarios .......................................................... 4
4.  Guard Band definition ............................................. 4
5.  Requirements ....................................................... 5
   5.1.  PCE Requirements ............................................. 5
   5.2.  PCEP Requirements ........................................... 7
   5.3.  GMPLS Requirements .......................................... 7
      5.3.1.  OSPF-TE Requirements .................................. 7
      5.3.2.  RSVP-TE Requirements ................................... 7
6.  Security Considerations ........................................... 8
7.  IANA Considerations ............................................... 8
8.  Contributors ........................................................ 8
9.  Acknowledgements .................................................. 8
10. References ......................................................... 8
    10.1.  Normative References ...................................... 8
    10.2.  Informative References .................................... 9
Authors’ Addresses ..................................................... 9
1. Introduction

Given the advancement of optical transmission technology, optical channels may use thinner granularity of the spectrum, which are configurable depending on the modulation format and bit-rate [G.FLEXIGRID]. Thus, thanks to this flexibility, the capacity of optical networks is strongly increasing. However, the spacing between channels may be limited by the inter-channel effects (e.g., cross-phase modulation - XPM) which can lead to a bit error rate increase. Typically, as the case of XPM or cross-talk, the larger the spectral distance among interfering signals, the less detrimental the effect. Thus, a guard band (i.e., a spectral distance such that detrimental effects are mitigated) may be considered to counteract inter-channel detrimental effects [sambo-jlt].

Guard Bands (GB) may be required in either fixed- [RFC6163] or flexible-grid networks [G.694.1v1] [G.FLEXIGRID]. As an example, in fixed-grid networks, high-speed signals (100Gbit/s and beyond) may be deployed together with low-speed signals (10Gb/s). In such a scenario, high-speed signals utilizing phase-modulated formats (e.g., dual polarization quadrature phase shift keying - DP-QPSK - 100Gb/s) suffer from XPM induced by low-speed signals, exploiting intensity modulation (e.g. on off keying - OOK - 10Gb/s). Thus, GB may be used to avoid problems of XPM between low- and high-speed signals. Similarly, in flex-grid networks, high-speed signals may exploit quadrature amplitude modulation (QAM), which experience both intensity and phase modulation. Also in such a scenario, XPM may be very detrimental.

The value of a guard band may depend on physical properties of the traversed links and on the bit rate and modulation format of the interfering signals. Given two interfering signals, inter-channel effects among the two signals are counteracted if they are separated by GB. This document describes the requirements of PCE and GMPLS control to account for guard bands.

1.1. Terminology

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

2. Definitions

Demeaning LSP: an LSP which induces a detrimental effect on another LPS
Degraded LSP: an LSP which may be degraded by inter-channel effects induced by a demeaning GB: guard band

Working LSP: an active LSP

RSA: routing and spectrum assignment

IV: impairment validated (e.g., a route is impairment validated if its bit error rate is acceptable for any channel)

3. Scenarios

Fixed-grid network is here assumed as a particular case of flex-grid network, thus hereafter only the case of flex-grid networks will be treated. Similarly, RWA is assumed as a particular case of RSA and only RSA will be treated. The following PCE scenarios are considered [draft-flexible]:

- IV and RSA PCE: From a GB point of view there is no difference between IV+RSA and IV&RSA, so a general IV+RSA case will be considered. In this case the PCE provides the ingress node with an impairment-validated route and a set of frequency slots.

- IV PCE: PCE provides ingress node with an impairment-validated route. Then, slot assignment is distributed and performed by the egress node which may rely on collecting link status through the signaling protocol (RSVP-TE).

- IV Candidate path PCE: PCE provides ingress node with a set of candidate routes (i.e., a set of impairment-validated routes). Then, a route is selected by the ingress node. Slot assignment is distributed and performed by the egress node through the signaling protocol (RSVP-TE).

4. Guard Band definition

GB is defined as the minimum frequency range which separates two contiguous signals, S1 at bit rate B1 and modulation format M1 and S2 at a bit rate B2 and modulation format M2, such that detrimental effects are negligible.
Assuming fixed-grid networks, a number of channels (e.g. of a grid spacing of 50 GHz), instead of a number of slots would be considered for GB.

The computation of GB may require the knowledge of:

a. bit rate B and modulation format M of the interfering signals

b. the power P values of the signals at each span. In order to limit the stored and exchanged information, an unique value of P (worst-case scenario) may be considered for the demeaning LSP: i.e., the maximum value P experienced by an LSP of type (B2,M2). Similarly, an unique value P (worst-case scenario) may be considered for the degraded LSP: the minimum value P experienced by an LSP of type (B1,M1).

c. Fiber parameters: e.g. fiber attenuation, dispersion parameter, and fiber nonlinear Kerr coefficient

Bit rate and modulation format should be mandatory information for GB computation (e.g., PCE may select the value of GB from a stored set of GB values, each one associated to a bit rate and modulation format pair), thus treated in the rest of the document.

5. Requirements

5.1. PCE Requirements

- IV+RSA PCE: given an LSP request, by exploiting a TED, PCE may account for GB in the IV+RWA (or IV&RSA) process, if needed. In the case of:

  + Stateful PCE: PCE has a TED (in simple terms as disseminated by OSPF-TE) plus an LSP-DB which are the active LSPs state. (e.g., the route and the slot used by a working LSP). In order
to identify the required GB, the TED plus the LSP-DB exploited by the PCE should be extended to store the following information:

++ Bit rate $B$ of any working LSP in the network
++ Modulation format $M$ of any working LSP in the network
++ Allocated central frequency and slot width for any active LSP in the network.

+ Stateless PCE: PCE exploits a TED which includes per-link information regarding the usage of the optical spectrum resource (e.g., Available Frequency Ranges). If the PCE obtains the TED via e.g. OSPF-TE this may also add additional requirements to OSPF-TE as detailed later on. In order to identify the required GB, the TED exploited by the PCE should be extended to store the set of required information. An example of such pieces of information could be:

++ Used frequency slots
++ Bit rate $B$ associated to any frequency slot in use
++ Modulation format $M$ associated to any frequency slot in use

- IV PCE: given an LSP request, PCE provides the ingress node with an impairment validated route. Then, wavelength or the slot assignment is distributed, e.g. performed through a signaling protocol (RSVP-TE). In this case, PCE should inform the ingress node about the requirements of GB to separate the given LSP from other LSPs of specific bit rate $B$ and modulation format $M$. Thus, PCEP and RSVP-TE may require extensions to account for GB.

- IV Candidate path PCE: given an LSP request, PCE provides the ingress node with a set of impairment validated routes. A route is selected by the ingress node. Then, wavelength or the slot assignment is distributed, e.g. performed through a signaling protocol (RSVP-TE). In this case, PCE should inform, for each candidate route, the ingress node about the requirements of GB to separate the given LSP from other LSPs of specific bit rate $B$ and modulation format $M$. Thus, PCEP and RSVP-TE may require extensions to account for GB.
5.2. PCEP Requirements

- IV&RSA PCE: in this case, no extensions for GB are required by PCEP because PCEP client (e.g., the ingress node) does not require to know GB information.

- IV PCE: in this case, an extension may be needed in the PCEP Path Computation Reply message to inform the ingress node about required GBs along the route. Then, GB information should be considered in the routing and slot assignment.

- IV candidate path PCE: in this case, an extension may be needed in the PCEP Path Computation Reply message to inform the ingress node, for any candidate route, about required GBs along the candidate routes. Then, GB information should be considered in the routing and slot assignment.

5.3. GMPLS Requirements

5.3.1. OSPF-TE Requirements

- Stateful PCE: the LSP-DB is not filled through OSPF-TE, thus no OSPF-TE extension is required.

- Stateless PCE: the TED may be filled through OSPF-TE, thus OSPF-TE extensions may be required to carry used frequency slot information, such as the associated bit-rate B and modulation format M.

5.3.2. RSVP-TE Requirements

If the PCE only provides the ingress node with a route (IV PCE and IV candidate path PCE), the slot assignment is performed at the egress node. To this aim, RSVP-TE Path message gathers frequency range slot availability information along the route.

- IV&RSA PCE: no extensions for GB are required by RSVP-TE

- IV PCE: extensions to RSVP-TE may be required to enable distributed RSA process which accounts for GB. In particular, extensions to RSVP-TE may be required to identify the frequency spectrum along the route that should be not selected because of GB.

- IV candidate path PCE: extensions to RSVP-TE may be required to enable distributed RSA process which accounts for GB. In particular, extensions to RSVP-TE may be required to identify frequency spectrum along the route that should be not selected.
because of GB.

6. Security Considerations
   TBD

7. IANA Considerations
   TBD

8. Contributors
   Fabio Cavaliere, Ericsson
   Email: fabio.cavalier@ericsson.com
   
   Paola Iovanna, Ericsson
   Email: paola.iovanna@ericsson.com
   
   Piero Castoldi, Scuola Superiore S.Anna
   Email: castoldi@sssup.it

9. Acknowledgements

10. References

10.1. Normative References


10.2. Informative References

[G.694.1v1]

[draft-flexible]

[sambo-jlt]

Authors’ Addresses

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

Email: daniele.ceccarelli@ericsson.com

Giulio Bottari
Ericsson
Via G.Moruzzi, 1
Pisa
Italy

Email: giulio.bottari@ericsson.com
OSPFTE extension to support GMPLS for Flex Grid
draft-dhillon-ccamp-super-channel-ospfte-ext-02.txt

Abstract


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/1id-abstracts.txt

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

This Internet-Draft will expire on April 23, 2012.
1. Introduction

To enable scaling of existing transport systems to ultra high data rates of 1 Tbps and beyond, next generation systems providing super-channel switching capability are currently being developed. To allow efficient allocation of optical spectral bandwidth for such high bit rate systems, International Telecommunication Union Telecommunication Standardization Sector (ITU-T) is extending the
G.694.1 grid standard (termed ‘Fixed-Grid’) to include flexible grid (termed ‘Flex-Grid’) support.

This document defines OSPF-TE extensions in support of flex-grid networks.

Figure 1 shows a network consisting of Network Elements (NEs) with super channel switching capability. User can create super-channel connections using GMPLS through these NEs. To create these super-channel connections, system needs to model TELINKs for routing which are capable of super channel switching and hence there is a need to extend the switching capability descriptor in TELINK for super channel switched networks.

<table>
<thead>
<tr>
<th>SC Switch</th>
<th>---Link ---&gt;</th>
<th>SC Switch</th>
<th>&lt;- Link----&gt;</th>
<th>SC Switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>B</td>
<td></td>
<td>C</td>
</tr>
</tbody>
</table>

Figure 1: TE-Links

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 RFC 2119 [RFC2119].

3. Interface Switching Capability Descriptor

The Interface Switching Capability Descriptor describes switching capability of an interface [RFC 4203]. This document defines a new Switching Capability value for Flex Grid [FLEX-GRID] as follows:

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>102 (TBA by IANA)</td>
<td>Super-Channel-Switch-Capable(SCSC)</td>
</tr>
</tbody>
</table>

Switching Capability and Encoding values MUST be used as follows:
Switching Capability = SCSC
Encoding Type = Lambda [as defined in RFC3471]

The Interface Switching Capability Descriptor is a sub-TLV (of type 15) of the Link TLV. The length is the length of value field in Octets. The format of the value field is as shown below:

```
  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              |
+----------------------------------------------+
| Switching Capability-specific information              |
| (variable)                                      |
+----------------------------------------------+
```

Figure 2: ISCD & SCSI

Max LSP Bandwidth will be based on Max Slot Width field in BW-sub-TLV (Ref to section 3.1 for details on BW sub-TLV) and the modulation format used.
3.1. Switch Capability Specific Information

The technology specific part of the ISCD can include a variable number of sub-TLVs. We propose to encode Slice Information in Bandwidth sub-TLVs under SCSI field. The format of BW sub-TLVs is as shown below.

[Editor’s note: To provide options similar to Label set field defined in [9], we have included 2 variants to advertise slice level information. These are bit-format and list/range format].

3.2. BW sub TLV: Bit Map format

The figure below shows format of Type=1 sub-TLV for encoding slice information in bit-map format. This sub-TLV must be repeated for each priority that is supported on the Te-link.

```
+---------------------------------------------------------------+
<p>| 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       |              Length               |
|---------------------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>Slice Spacing</th>
<th>Pri</th>
<th>Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-Start</td>
<td></td>
<td>Num of Slices</td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>---------------------</td>
</tr>
<tr>
<td>Min Slot Width</td>
<td></td>
<td>Max Slot Width</td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>---------------------</td>
</tr>
<tr>
<td>Bit-Map showing Available Slices</td>
<td>(up to 48 bytes)</td>
<td></td>
</tr>
</tbody>
</table>
+---------------------------------------------------------------+
```

Figure 3: Type=1 BW sub TLV in Bit-Map format

3.2.1. Meaning of sub TLV fields

- **Slice Spacing**: 8-bit field (S.S) which can take one of the values as shown in table below.
  - For e.g., the 12.5GHz spacing is specified by setting this field to value 4.
Table 1: Slice Spacing Values

<table>
<thead>
<tr>
<th>S.S. (GHz)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>12.5</td>
<td>4</td>
</tr>
<tr>
<td>Future use</td>
<td>5 - 15</td>
</tr>
</tbody>
</table>

- **Priority**: 3-bit field
  - 3-bit field to identify one of the 8 priorities for which Slice information (BW) is advertised.
- **N-Start**: 16-bit field
  - Is a two’s complement integer to specify start of the grid
  - Use center freq formula to determine start of spectrum
- **Number of slices**: 16-bit field
  - Total number of slices advertised for the link. This includes (available plus consumed).
- **Minimum Slot Width**: 16-bit field
  - This is a positive integer value
  - This field is similar to Min LSP BW field. The value in this field is used to determine the smallest frequency slot width that the advertising node can allocate for an LSP. This is defined by the following equation:
    \[
    \text{Smallest Frequency slot width} = \text{Slice Spacing} \times \text{integer value in ‘Minimum Slot Width’ field}
    \]
- **Maximum Slot Width**: 16-bit field
  - This is a positive integer value
  - This field is used to determine the Maximum contiguous frequency slot width that the advertising node can allocate for an LSP. This is defined by the following equation:
    \[
    \text{Largest Contiguous Frequency slot width} = \text{Slice Spacing} \times \text{integer value in ‘Maximum Slot Width’ field}
    \]
- **Available slices encoded as bit-map**
  - Each bit represents availability of one slice of width identified by S.S field
  - Zero: Available ; One: occupied
  - Padding MUST be used to align with 32 bit boundary.
3.3. BW sub TLV: List and Rage format

The figure below shows format of Type=2 sub-TLV for encoding slice information in list/range format. This sub-TLV must be repeated for each priority that is supported on the Te-Link.

```
+---------------------------------------------------------------+
<p>|             Type=2            |              Length           |
|---------------------------------------------------------------|
| Slice Spacing | Pri | Res     |         Num of Entries        |
|---------------------------------------------------------------|
|             Min Slot Width        |       Max Slot Width          |
|---------------------------------------------------------------|
|             N-Start-1         |          N-end-1              |
|---------------------------------------------------------------|
|             N-Start-2         |          N-end-2              |
|---------------------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>More Entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-Start-n</td>
</tr>
</tbody>
</table>
+---------------------------------------------------------------+
```

Figure 4: Type=2 BW sub TLV in List/Range format

3.3.1. Meaning of sub TLV fields

- The meaning of above fields is same as in Type=1 BW-sub-TLV. For details refer to section 3.2.1.
  - Slice Spacing,
  - Priority,
  - Maximum Slot Width &
  - Minimum Slot Width
- Number of Entries: 16-bit field
  - Is a positive integer value.
  - Total number of N-start & N-End rows advertised for the link.
- N-Start-x: 16-bit field
  - Is a two's complement integer value (+ve, -ve or zero) to specify start of the grid.
o Use center freq formula to determine start of spectrum

o N-end-x: 16-bit field
  o Is a two’s complement integer value (+ve, -ve or zero) to specify end of the list/range.
  o Use center freq formula to determine end of spectrum

3.4. BW advertisement procedure

This section describes bandwidth advertisement for Te-Links capable SCSC.

o Optical nodes capable of Super Channel Switching advertise slices of certain width available based on the frequency spectrum supported by the node (e.g. C band, extended C-band). For example, node(s) supporting extended C-band will advertise 384 slices.

o The BW advertisement involves an ISCD containing
  o Slice information in bit-map format (Type=1 BW-sub-TLV) where each bit corresponds to a single slice of width as identified by S.S field. OR
  o Slice information in list/range format (Type=2 BW-sub-TLV) where each 32-bit entry represents an individual slice or list or range.

o The slice position/numbering in Type=1 sub-TLV is identified based on N-start field. The N-start field is derived based on ITU center frequency formula.

o The advertising node MUST also set Number of Slices field.

o Minimum & Maximum slot width fields are included to allow for any restrictions on the link for carrying super channel LSPs.

o The BW advertisement is priority based and up to 8 priority levels are allowed.

o The node capable of supporting one or more priorities MUST set the priority field and include BW-sub TLV for each of the priority supported.

4. Examples

4.1. Example: BW advertisement without any service present

Figure 5 shows an example of BW sub-TLV for a te-link which has no service established over it yet. Attributes of BW sub-TLV in the te-link are:
### Example: How to use advertised Bandwidth

Assume user wants to setup Super Channel LSP over a single Flex-Grid link with BW requirement = 200GHz and transponder fully tunable.

- The path computing node performs the following:
  - Determine the number of slices required for the LSP (200/S.S = 16)
  - Look for contiguous spectrum availability on each link from BW adv (both dir)
  - Look for 16 contiguous bits in the BW advertisement TLV
  - If available select the link for LSP creation.
  - Signal for LSP creation. Once LSP is created, update BW available via new advertisement using the same Bandwidth sub-TLV.
5. Security Considerations

<Add any security considerations>

6. IANA Considerations

IANA needs to assign a new Grid field value to represent ITU-T Flex-Grid.

7. References

7.1. Normative References


7.2. Informative References


8. Acknowledgments

The authors would like to thank Khuzema Pithewan, Ashok Kunjidhapatham & Mohit Misra for their valuable comments.
Authors’ Addresses

Abinder Dhillon
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: adhillon@infinera.com

Iftekhar Hussain
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: ihussain@infinera.com

Rajan Rao
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: rrao@infinera.com

Marco Sosa
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: msosa@infinera.com

Contributor’s Addresses

Biao Lu
Email: blu@infinera.com

Subhendu Chattopadhyay
Email: schattopadhyay@infinera.com

Harpreet Uppal
Email: harpreet.uppal@infinera.com
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
RSVP-TE Extensions for Lock Instruct and Loopback in MPLS Transport Profile
draft-dong-ccamp-rsvp-te-mpls-tp-li-lb-02

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 September 10, 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
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
1. Introduction

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].

2. Extensions to RSVP-TE

In this document, Path and Resv message are used to implement LI function, and Notify message is used for LB functions. Two new flags (Lock bit and Loopback bit) are defined in ADMIN_STATUS Object [RFC3471] [RFC3473] that can be carried in Path/Resv and Notify message.

Format of extended ADMIN_STATUS Object is as below:

```
+--------+--------+--------+--------+--------+--------+--------+        |
|        |        |        |        |        |        |        |        |
|Length  | Class-Num(196)| C-Type (1) |
+--------+--------+        +--------+        +--------+        +--------+
|        |        |        |        |        |        |        |        |
|R|      |        |        |        |        |        |        |        |
|Reserved|        |        |        |        |        |        |        |
```

Lock (K): When this bit is set in Path message, it indicates that local actions related to the "Lock" mode should be taken. When this bit is set in Resv or Notify message, it indicates that the LSP is put in "Lock" mode.
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 Lock (K) bit and the Reflect (R) bit in ADMIN_STATUS Object set. 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 take the LSP out of service. If the receiving MEP locks the LSP successfully, it SHOULD send a Resv message with the Lock (K) 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 Lock (K) bit cleared. Though the intermediate nodes do not need to take actions during this procedure, they would be aware of whether the LSP is put in Lock mode or not.

When an LSP is put in lock mode, the subsequent Path and Resv messages SHOULD keep the Lock (K) 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 Lock (K) 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 Lock (K) 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 Lock (K) 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 Lock (K) 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 Lock (K) 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 Lock (K) 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 Lock (K) 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

Two bits ("Lock" (K) and "Loopback" (B)) need 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 for his 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.


[RFC6371] Busi, I. and D. Allan, "Operations, Administration, and
7.2. Informative References

[I-D.ietf-ccamp-rsvp-te-mpls-tp-oam-ext]


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
A SNMP MIB to manage black-link optical interface parameters of DWDM applications
draft-galimbe-kunze-g-698-2-snmp-mib-02

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."
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 ifTable ............................... 17
5. Structure of the MIB Module ............................ 18
6. Object Definitions ..................................... 18
7. Relationship to Other MIB Modules ........................ 73
  7.1. Relationship to the [TEMPLATE TODO] MIB ............ 73
  7.2. MIB modules required for IMPORTS .................... 73
8. Definitions ............................................ 73
9. Security Considerations ................................ 73
10. IANA Considerations .................................. 74
11. Contributors ......................................... 75
12. References ........................................... 76
  12.1. Normative References .............................. 76
  12.2. Informative References ............................ 77
Appendix A. Change Log .................................. 78
Appendix B. Open Issues .................................. 78
Authors’ Addresses ....................................... 78

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 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, 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 blak 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 Level 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 Dispersion (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 range (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 gives 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 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 (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 Seconds (es):
   It is a one-second period in which one or more bits are in error or during which Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected (G)

Severely Errored Seconds (ses):
   It is a one-second period which has a bit-error ratio = 1x10^(-3) or during which Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected (G)

Unavailable Seconds (uas):
   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 (bberr):
   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 Words Error:
   The number of Words 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)
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.

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.

<table>
<thead>
<tr>
<th>ifTable Object</th>
<th>Use for OTN OPS Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ifIndex</td>
<td>The interface index.</td>
</tr>
<tr>
<td>ifDescr</td>
<td>Optical Transport Network (OTN) Optical Physical Section (OPS)</td>
</tr>
<tr>
<td>ifType</td>
<td>opticalPhysicalSection (xxx)</td>
</tr>
</tbody>
</table>

<<<Editor Note: Need new IANA registration value for xxx. >>>

<table>
<thead>
<tr>
<th>ifSpeed</th>
<th>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.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ifPhysAddress</td>
<td>An octet string with zero length. (There is no specific address associated with the interface.)</td>
</tr>
<tr>
<td>ifAdminStatus</td>
<td>The desired administrative state of the interface. Supports read-only access.</td>
</tr>
<tr>
<td>ifOperStatus</td>
<td>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.</td>
</tr>
<tr>
<td>ifLastChange</td>
<td>The value of sysUpTime at the last change in</td>
</tr>
</tbody>
</table>
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 ifTable

Use of ifStackTable

EDITOR NOTE: more to be provided (similar to RFC 3591 Section 2.5)
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

<table>
<thead>
<tr>
<th>HigherLayer</th>
<th>LowerLayer</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>j</td>
</tr>
<tr>
<td>0</td>
<td>k</td>
</tr>
<tr>
<td>j</td>
<td>i</td>
</tr>
<tr>
<td>k</td>
<td>i</td>
</tr>
<tr>
<td>i</td>
<td>0</td>
</tr>
</tbody>
</table>

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-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;
This is the MIB module for the optical parameters associated with the black link end points.

OptIfChannelSpacing ::= TEXTUAL-CONVENTION
STATUS current
DESCRIPTION
"Channel spacing
1 - 6.25GHz
2 - 12.5GHz
3 - 25GHz
4 - 50GHz
5 - 100 Ghz"
SYNTAX INTEGER {
  spacing6-25Ghz(1),
  spacing12-5Ghz(2),
  spacing25Ghz(3),
  spacing50Ghz(4),
  spacing100Ghz(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 {
  rate2-5G(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)
}

OptIfOTNType ::= 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 layers, we include it 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 {
    OTUKLayer(1),
    ODUkLayer(2),
    TCMSubLayer(3)
}

-- Alarm for the OCh and OTUk sublayer
--
OptIfOTNOChAlarms ::= TEXTUAL-CONVENTION
STATUS current
DESCRIPTION "This is the possible alarms from the OCh and OTUk layer."
SYNTAX INTEGER {
    optIfOtnLosAlarm(1), -- OTN Loss of signal alarm
    optIfOtnLofAlarm(2), -- OTN Loss of frame alarm
    optIfOtnLomAlarm(3), -- OTN Loss of multi frame alarm
    optIfOtnOtuSsfAlarm(4), -- OTN SSF alarm
    optIfOtnOtuBdiAlarm(5), -- OTN OTU BDI alarm
    optIfOtnOtuTtimAlarm(6), -- OTN OTU Trail termination mismatch alarm
    optIfOtnOtuIaeAlarm(7), -- OTN OTU IAE alarm
    optIfOtnOtuDegAlarm(8), -- OTN OTU signal degrade alarm
    optIfOptIfOtnOtuFecExcessiveErrsAlarm(9), -- OTN OTU Fec Excessive
Errors alarm

optIf15MinThreshBBETCA(10), -- OTN OTU BBE Threshold alarm
optIf15MinThreshESTCA(11), -- OTN OTU ES Threshold alarm
optIf15MinThreshSESTCA(12), -- OTN OTU SES Threshold alarm
optIf15MinThreshUASTCA(13), -- OTN OTU UAS Threshold alarm
optIf15MinThreshFcsTCA(14), -- OTN OTU Fcs Threshold alarm
optIf15MinThreshFECUnCorrectedWordsTCA(15), -- OTN FEC uncorrected words
-- TCA
optIf15MinThreshPreFECBERTCA(16) -- OTN Pre FEC BER TCA

OptIfOTNODukTcmAlarms ::= TEXTUAL-CONVENTION
STATUS current
DESCRIPTION "This is the alarms from the ODUk and TCM layer."
SYNTAX INTEGER {
    optIfOTNODukTcmOciAlarm(1), -- OTN ODU/TCM OCI alarm
    optIfOTNODukTcmLckAlarm(2), -- OTN ODU/TCM LCK alarm
    optIfOTNODukTcmBdiAlarm(3), -- OTN ODU/TCM BDI alarm
    optIfOTNODukTcmTtimAlarm(4), -- OTN ODU/TCM TTIM alarm
    optIfOTNODukTcmDegAlarm(5), -- OTN ODU/TCM Signal degrade alarm,
    optIfOTNODukTcmSSfAlarm(6), -- OTN ODU/TCM SSF alarm,
    optIfOTNODukTcm15MinThreshBBETCA(7), -- OTN OTU BBE Threshold alarm
    optIfOTNODukTcm15MinThreshESTCA(8), -- OTN OTU ES Threshold alarm
    optIfOTNODukTcm15MinThreshSESTCA(9), -- OTN OTU SES Threshold alarm
    optIfOTNODukTcm15MinThreshUASTCA(10), -- OTN OTU UAS Threshold alarm
    optIfOTNODukTcm15MinThreshPcsTCA(11), -- OTN OTU Fcs Threshold alarm
}

-- Addition to the RFC 3591 objects
optIfOPSmConfigTable OBJECT IDENTIFIER ::= { optIfObjects 10 }
opIfOTNPMoObjects OBJECT-TYPE ::= { optIfObjects 11 }
opIfOTNAlarm OBJECT IDENTIFIER ::= { optIfObjects 12 }
opIfOTNNotifications OBJECT IDENTIFIER ::= { optIfObjects 13 }

optIfOChConfigTable OBJECT IDENTIFIER ::= { optIfOCh 1 }
    // Extended the optIfOChConfigTable

optIfOChSinkCurrentTable OBJECT IDENTIFIER ::= { optIfOCh 1 }
    // Extended this table to add OSNR/CD/Q

optIfOChSrcConfigTable OBJECT IDENTIFIER ::= { optIfOCh 10 }
opIfOChSrcSinkConfigTable OBJECT IDENTIFIER ::= { optIfOCh 11 }
opIfOChSinkConfigTable OBJECT IDENTIFIER ::= { optIfOCh 12 }
opIfOTNPMSconfigTable OBJECT IDENTIFIER ::= { optIfOTNPMoObjects 1 }
opIfOTNPMSCurrentTable OBJECT IDENTIFIER ::= { optIfOTNPMoObjects 2 }
opIfOTNPMSIntervalTable OBJECT IDENTIFIER ::= { optIfOTNPMoObjects 3 }
opIfOTNPMSCurrentDayTable OBJECT IDENTIFIER ::= { optIfOTNPMoObjects 4 }
opIfOTNPMPrevDayTable OBJECT IDENTIFIER ::= { optIfOTNPMoObjects 5 }

optIfOTNPMFEConfigTable OBJECT IDENTIFIER ::= { optIfOTNPObjects 6 }
optIfOTNPMFECCurrentTable OBJECT IDENTIFIER ::= { optIfOTNPObjects 7 }
optIfOTNPMFEIntervalTable OBJECT IDENTIFIER ::= { optIfOTNPObjects 8 }
optIfOTNPMFECCurrentDayTable OBJECT IDENTIFIER ::= { optIfOTNPObjects 9 }
optIfOTNPFECPrevDayTable OBJECT IDENTIFIER ::= { optIfOTNPObjects 10}

-- 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,
  optIfOPSwavelengthsUsed Unsigned32
}

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
::= { 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
   The categories are A, B, C, D and E."
::= { optIfOPSmConfigEntry  3 }

optIfOPSmWavelengthsUsed  OBJECT-TYPE
SYNTAX    Unsigned32
MAX-ACCESS read-write
STATUS  current
DESCRIPTION
   "Number of wavelengths used currently."
::= { optIfOPSmConfigEntry  4 }

-- OCh config table
-- modified the OCh Table group
-- General parameters for the Black Link Ss-Rs will be added to
-- the OchConfigTable

optIfOChConfigTable OBJECT-TYPE
SYNTAX    SEQUENCE OF OptIfOChConfigEntry
MAX-ACCESS not-accessible
STATUS  current
DESCRIPTION
   "A table of Och General config parameters"
::= { optIfOCh 1 }

optIfOChConfigEntry OBJECT-TYPE
SYNTAX    OptIfOChConfigEntry
MAX-ACCESS not-accessible
STATUS  current
DESCRIPTION
   " A conceptual row that contains OCh configuration information
   of an interface. "
INDEX  { ifIndex  }
::= { optIfOChConfigTable 1 }
OptIfOChConfigEntry ::= 
  SEQUENCE {
    optIfOChMinimumChannelSpacing          OptIfChannelSpacing,
    optIfOChBitRateLineCoding             OptIfBitRateLineCoding,
    optIfOChFEC                            Integer32,
    optIfOChSinkMaximumBERMantisa          Integer32,
    optIfOChSinkMaximumBERExponent         Integer32,
    optIfOChMinWavelength                  Integer32,
    optIfOChMaxWavelength                  Integer32,
    optIfOChVendorTransceiverClass         OCTET STRING,
    optIfOChOpticalInterfaceApplicationCode OCTET STRING,
    optIfOChLaserAdminState               Integer,
    optIfOChLaserOperationalState          TruthValue,
    optIfOChAdminState                     Integer,
    optIfOChOperationalState               Integer
  }

optIfOChMinimumChannelSpacing  OBJECT-TYPE
  SYNTAX      OptIfChannelSpacing
  UNITS       "Gigahertz"
  MAX-ACCESS  read-only
  STATUS      current
  DESCRIPTION 
    "A minimum nominal difference in frequency (GHz) between two adjacent channels."
  ::= { optIfOChConfigEntry 3 }

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)
    "
  ::= { optIfOChConfigEntry 4 }

optIfOChFEC  OBJECT-TYPE
  SYNTAX  Integer32
  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 1.1 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
11 - 40G FEC (for new applications under study)
12 - 40G EFEC (for new applications under study)
13 - 100G FEC (for new applications under study)
14 - 100G EFEC (for new applications under study)
99 - Vendor Specific

::= { optIfOChConfigEntry 5 }

optIfOChSinkMaximumBERMantisa OBJECT-TYPE
SYNTAX Integer32
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 ."

::= { optIfOChConfigEntry 6 }

optIfOChSinkMaximumBERExponent OBJECT-TYPE
SYNTAX Integer32
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 ."

::= { optIfOChConfigEntry 7 }

optIfOChMinWavelength OBJECT-TYPE
SYNTAX Integer32
UNITS "hertz"
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"This parameter indicate minimum wavelength spectrum in a
definite wavelength Band (L, C and S)

::= { optIfOChConfigEntry  8}

optIfOChMaxWavelength  OBJECT-TYPE
SYNTAX  Integer32
UNITS "hertz"
MAX-ACCESS read-only
STATUS  current
DESCRIPTION
"    This parameter indicate maximum wavelength spectrum in a
    definite wavelength Band (L, C and S)"

::= { optIfOChConfigEntry  9}

optIfOChWavelength  OBJECT-TYPE
SYNTAX  Integer32
UNITS "hertz"
MAX-ACCESS read-write
STATUS  current
DESCRIPTION
"    This parameter indicates the wavelength value."

::= { optIfOChConfigEntry  10}

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)"

::= { optIfOChConfigEntry  11}

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"
::= { optIfOChConfigEntry 12}

doChLaserAdminState 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"
::= { optIfOChConfigEntry 13}

doChLaserOperationalState OBJECT-TYPE
SYNTAX INTEGER {
    off(0),
    on(1)
}
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" The Operational Status of Laser : 0 - Off 1 - On"
::= { optIfOChConfigEntry 14}

doChAdminState 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.

::= { optIfOChConfigEntry 15}

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"

::= { optIfOChConfigEntry 16}

-- 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."
::= { optIfOCh 10 }

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
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
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
SYNTAX  Integer32
UNITS   "0.01 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 19150"
::= { optIfOChSrcConfigEntry  3}

optIfOChMaximumCentralFrequency
SYNTAX  Integer32
UNITS "0.01 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 19150

::= { optIfOChSrcConfigEntry 4}

optIfOChMaximumSpectralExcursion OBJECT-TYPE
SYNTAX Integer32
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 parameters for the optical path from Src to Sink
(Ss to Rs)."
::= { optIfOCh 11 }

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,
  optIfOChSrcSinkInterFerometricCrosstalk              Integer32,
  optIfOChSrcSinkOpticalPathOSNRPenalty                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 }

optIfOChSrcSinkMinimumSsOpticalReturnLoss
SYNTAX Integer32
UNITS ".1 db"
MAX-ACCESS read-only
STATUS current
DESCRIPTION "These parameters define minimum optical return loss (in dB) of the cable plant at the source reference point (Src/Ss), including any connectors.

::= { optIfOChSrcSinkConfigEntry 3 }

optIfOChSrcSinkMaximumDiscreteReflectanceSrcToSink
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
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
SYNTAX Integer32
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}

optIfOChSrcSinkInterferometricCrosstalk 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}

optIfOChSrcSinkOpticalPathOSNRPenalty OBJECT-TYPE
SYNTAX  Integer32
UNITS "0.1 db"
MAX-ACCESS read-only
STATUS current
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.
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}

optIfOChSinkMinimumOSNRTolerance 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}

-- The OptIfOChSinkCurrentEntry table has been enhanced to add the following optional parameters
-- For current status
-- OptIfOChSinkCurrentEntry

optIfOChSinkCurrentTable OBJECT-TYPE
SYNTAX  SEQUENCE OF OptIfOChSinkCurrentEntry
MAX-ACCESS not-accessible
STATUS  current
DESCRIPTION
  "A table of OCh sink performance monitoring information for the current 15-minute interval."
 ::= { optIfOCh  2}
optIfOChSinkCurrentEntry OBJECT-TYPE
SYNTAX OptIfOChSinkCurrentEntry
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."
INDEX { ifIndex }
::= { optIfOChSinkCurrentTable 1 }

OptIfOChSinkCurrentEntry ::= 
  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."
::= { optIfOChSinkCurrentEntry 7}

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 ."
::= { optIfOChSinkCurrentEntry 8}

optIfOChSinkCurrentQ OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"'Q' factor estimated at Rx Transceiver port."
::= { optIfOChSinkCurrentEntry 9}

-- 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 'optIfOTNPMConfigSublayer' layer."
::= { optIfOTNPMObjects 1 }

optIfOTNPMConfigEntry OBJECT-TYPE
SYNTAX      optIfOTNPMConfigEntry
MAX-ACCESS  not-accessible
STATUS  current
DESCRIPTION
" A conceptual entry in the performance monitoring configuration for the type 'optIfOTNPMConfigSublayer' layer."
" INDEX { ifIndex, optIfOTNPMConfigType, optIfOTNPMConfigSublayer, optIfOTNPMConfigTCMLevel }::= { optIfOTNPMConfigTable 1 }

OptIfOTNPMConfigEntry ::= SEQUENCE {
optIfOTNPMConfigType                   OptIfOTNType,
optIfOTNPMConfigLayer                  OptIfOTNLayer,
optIfOTNPMConfigTCMLevel               Unsigned32,
optIfOTNPMESRInterval                  Integer32,
optIfOTNPMSESRIInterval                Integer32,
optIfOTNPMOTNValidIntervals            Integer32,
optIfOTNPMOTNThresh15MinFcs            Integer32,
optIfOTNPMOTNThresh15MinESs            Integer32,
optIfOTNPMOTNThresh15MinSESs           Integer32,
optIfOTNPMOTNThresh15MinUASs           Integer32,
optIfOTNPMOTNThresh15MinBBEs           Integer32,
}

optIfOTNPMConfigType       OBJECT-TYPE
SYNTAX  OptIfOTNType
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}

optIfOTNPMConfigSublayer 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 sublayer PM is not related to the black link PM
management, but since this is a common PM model for the ODU/TCM layer,
we may include it here.
"
::= { optIfOTNPMConfigEntry  2}

optIfOTNPMConfigTCMLevel OBJECT-TYPE
SYNTAX  Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" This parameter indicates the TCM level (1-6)
if the PM is of the type TCM. This will be 0 for OTUK/ODUK.
"
::= { optIfOTNPMConfigEntry  3}

optIOTNPMESRInterval OBJECT-TYPE
SYNTAX  Integer32
UNITS "seconds"
MAX-ACCESS read-write
STATUS current
DESCRIPTION
" This parameter indicates the measurement interval
for error seconds ratio.
"
::= {optIfOTNPMConfigEntry  4}

optIOTNPMSESRIInterval OBJECT-TYPE
SYNTAX  Integer32
UNITS "seconds"
MAX-ACCESS read-write
STATUS current
DESCRIPTION

This parameter indicates the measurement interval for severely error seconds ratio.

::= {optIfOTNPMConfigEntry 5}

optIfOTNPM15MinThreshFcs OBJECT-TYPE
SYNTAX Integer32
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 6}

optIfOTNPM15MinThreshES OBJECT-TYPE
SYNTAX Integer32
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. A value of '0' will disable the notification.

::= {optIfOTNPMConfigEntry 7}

optIfOTNPM15MinThreshSES OBJECT-TYPE
SYNTAX Integer32
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. A value of '0' will disable the notification.

::= {optIfOTNPMConfigEntry 8}

optIfOTNPM15MinThreshUAS OBJECT-TYPE
SYNTAX  Integer32
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. A value of '0' will disable the
notification."
::= {optIfOTNPMConfigEntry  9}

optIfOTNPM15MinThreshBBE  OBJECT-TYPE
SYNTAX  Integer32
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. A value of '0' will disable the
notification."
::= {optIfOTNPMConfigEntry  10}

--
-- 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."
::= {optIfOTNPMObjects 2}

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 'optIfOTNPMCurrentSublayer' layer."
INDEX  { ifIndex, optIfOTNPMCurrentType,}
optIfOTNPMCurrentSublayer, optIfOTNPMCurrentTCMLevel  
::= { optIfOTNPMCurrentTable 1 }

OptIfOTNPMCurrentEntry ::=  
   SEQUENCE  
      optIfOTNPMCurrentType        OptIfOTNType,  
      optIfOTNPMCurrentLayer       OptIfOTNLayer,  
      optIfOTNPMCurrentTCMLevel    Unsigned32,  
      optIfOTNPMCurrentSuspectedFlag TruthValue,  
      optIfOTNPMCurrentFcs         Integer32,  
      optIfOTNPMCurrentESs         Integer32,  
      optIfOTNPMCurrentSESs        Integer32,  
      optIfOTNPMCurrentUASs        Integer32,  
      optIfOTNPMCurrentBBEs        Integer32,  
      optIfOTNPMCurrentESR         Integer32,  
      optIfOTNPMCurrentSESR        Integer32,  
      optIfOTNPMCurrentBBER        Integer32,  
   }

optIfOTNPMCurrentType OBJECT-TYPE  
SYNTAX  OptIfOTNType  
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}

optIfOTNPMCurrentSublayer 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 sublayer 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}

optIfOTNPMCurrentFcs OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"Number of Failures occurred in an observation period.
"
::= { optIfOTNPMCurrentEntry 5}

optIfOTNPMCurrentESs OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"This is the number of seconds in which one or more bits are in
error or during which Loss of Signal (LOS) or Alarm Indication
Signal (AIS) is detected.
"
::= { optIfOTNPMCurrentEntry 6}

optIfOTNPMCurrentSESs OBJECT-TYPE
SYNTAX   Integer32
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 bit-error ratio =
1x10^{-3} or during which Loss of Signal (LOS) or Alarm
Indication Signal (AIS) is detected.
"
::= { optIfOTNPMCurrentEntry 7}

optIfOTNPMCurrentUASs OBJECT-TYPE
SYNTAX  Integer32
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 8}

optIfOTNPMCurrentBBEs OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS read-only
STATUS  current
DESCRIPTION
"An errored block not occurring as part of an SES.
"
::= { optIfOTNPMCurrentEntry 9}

optIfOTNPMCurrentESR OBJECT-TYPE
SYNTAX  Integer32
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 10}

optIfOTNPMCurrentSESR OBJECT-TYPE
SYNTAX  Integer32
UNIT  "\,0.001"
MAX-ACCESS read-only
The ratio of SES in available time to total seconds in available time during a fixed measurement interval.

::= { optIfOTNPMCurrentEntry  11}
optIfOTNPMIntervalType OBJECT-TYPE
SYNTAX    OptIfOTNType
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}

optIfOTNPMIntervalSublayer 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 sublayer 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    OptIfIntervalNumber
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.

::= { optIfOTNPIMIntervalEntry 4}

optIfOTNPIMIntervalSuspectedFlag OBJECT-TYPE
SYNTAX TruthValue
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" If true, the data in this entry may be unreliable.
"
::= { optIfOTNPIMIntervalEntry 5}

optIfOTNPIMIntervalFcs OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" Number of Failures occurred in an observation period.
"
::= { optIfOTNPIMIntervalEntry 6}

optIfOTNPIMIntervalESs OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" It is a one-second period in which one or more bits are in error or during which Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected.
"
::= { optIfOTNPIMIntervalEntry 7}

optIfOTNPIMIntervalSESs OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" The number of seconds which have a severe error.
It is a one-second period which has a bit-error ratio = 1x10^{-3} or during which Loss of Signal (LOS) or Alarm
Indication Signal (AIS) is detected.

\[ \ ::= \{\ \text{optIfOTNPMIntervalEntry} \ 8\} \]

\text{optIfOTNPMIntervalUASs} \quad \text{OBJECT-TYPE}
SYNTAX Integer32
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."

\[ \ ::= \{\ \text{optIfOTNPMIntervalEntry} \ 9\} \]

\text{optIfOTNPMIntervalBBEs} \quad \text{OBJECT-TYPE}
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"An errored block not occurring as part of an SES."

\[ \ ::= \{\ \text{optIfOTNPMIntervalEntry} \ 10\} \]

\text{optIfOTNPMIntervalESR} \quad \text{OBJECT-TYPE}
SYNTAX Integer32
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."

\[ \ ::= \{\ \text{optIfOTNPMIntervalEntry} \ 11\} \]

\text{optIfOTNPMIntervalSESr} \quad \text{OBJECT-TYPE}
SYNTAX Integer32
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  Integer32
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
--

optIfOTNPMMCurrentDayTable OBJECT-TYPE
SYNTAX  SEQUENCE OF OptIfOTNPMMCurrentDayEntry
MAX-ACCESS  not-accessible
STATUS  current
DESCRIPTION
" A Performance monitoring Current Day Table."
::= {optIfOTNPMMObjects 4}

optIfOTNPMMCurrentDayEntry OBJECT-TYPE
SYNTAX      OptIfOTNPMMCurrentDayEntry
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 'optIfOTNPMMCurrentDaySublayer' layer."
INDEX { ifIndex, optIfOTNPMMCurrentDayType, optIfOTNPMMCurrentDaySublayer, optIfOTNPMMCurrentDayTCMLevel  }
::= { optIfOTNPMMCurrentDayTable 1 }

OptIfOTNPMMCurrentDayEntry  ::==
SEQUENCE { optIfOTNPMMCurrentDayType OptIfOTNType, optIfOTNPMMCurrentDayLayer OptIfOTNLayer, optIfOTNPMMCurrentDayTCMLevel Unsigned32, optIfOTNPMMNECurrentDaySuspectedFlag TruthValue, optIfOTNPMMNECurrentDayFcs Integer32, optIfOTNPMMNECurrentDayESs Integer32, optIfOTNPMMNECurrentDaySESs Integer32, optIfOTNPMMNECurrentDayUASs Integer32,}
optIfOTNPMNECurrentDayBBEs           Integer32,
optIfOTNPMNECurrentDayESR            Integer32,
optIfOTNPMNECurrentDaySESR           Integer32,
optIfOTNPMNECurrentDayBBER           Integer32,
}

optIfOTNPMCurrentDayType OBJECT-TYPE
SYNTAX   OptIfOTNType
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}

optIfOTNPMCurrentDaySublayer 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 sublayer PM is not related to the black link PM management, but since this is a common PM model for the ODU/TCM layer, we may include it here."
 ::= { optIfOTNPMCurrentDayEntry 2}

optIfOTNPMCurrentDayTCMLevel OBJECT-TYPE
SYNTAX   Unsigned32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"This parameter indicates the TCM level (1-6)
if the PM is of the type TCM. This will be 0 for OTUK/ODUK."
 ::= { optIfOTNPMCurrentDayEntry 3}

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

optIfOTNPMCurrentDayFcs  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
  " Number of Failures occurred in an observation period."
 ::= { optIfOTNPMCurrentDayEntry  5}

optIfOTNPMCurrentDayESs  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
  " The number of seconds which have an error.
   It is a one-second period in which one or more bits are in error
   or during which Loss of Signal (LOS) or Alarm Indication Signal
   (AIS) is detected."
 ::= { optIfOTNPMCurrentDayEntry  6}

optIfOTNPMCurrentDaySESs  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
  " The number of seconds which have a severe error.
   It is a one-second period which has a bit-error ratio =
   1x10Eminus3 or during which Loss of Signal (LOS) or Alarm
   Indication Signal (AIS) is detected."
 ::= { optIfOTNPMCurrentDayEntry  7}

optIfOTNPMCurrentDayUASs  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
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  Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
" An errored block not occurring as part of an SES."

::= { optIfOTNPMCurrentDayEntry  9}

optIfOTNPMCurrentDayESR OBJECT-TYPE
SYNTAX  Integer32
UNIT ".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  Integer32
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  Integer32
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
--
opIfOTNMPPrevDayTable OBJECT-TYPE
SYNTAX  SEQUENCE OF OptIfOTNMPPrevDayEntry
MAX-ACCESS not-accessible
STATUS  current
DESCRIPTION
  "A Performance monitoring Previous Day Table."

 ::= {optIfOTNPMObjects 5}
opIfOTNMPPrevDayEntry OBJECT-TYPE
SYNTAX      OptIfOTNMPPrevDayEntry
MAX-ACCESS  not-accessible
STATUS      current
DESCRIPTION
  "A conceptual entry in the Near end or Far End performance
  monitoring previous day table for the type
  'optIfOTNMPPrevDaySublayer' layer."

 INDEX  { ifIndex, optIfOTNMPPrevDayType ,
                 optIfOTNMPPrevDaySublayer, optIfOTNMPPrevDayTCMLevel }

 ::= { optIfOTNMPPrevDayTable 1 }

OptIfOTNMPPrevDayEntry ::==
SEQUENCE {
  optIfOTNMPPrevDayType                        OptIfOTNType,
  optIfOTNMPPrevDayLayer                       OptIfOTNLayer,
  optIfOTNMPPrevDayTCMLevel                    Unsigned32,
  optIfOTNMPNEPrevDaySuspectedFlag             TruthValue,
  optIfOTNMPNEPrevDayFcs                       Integer32,
  optIfOTNMPNEPrevDayESs                       Integer32,
  optIfOTNMPNEPrevDayUESs                      Integer32,
  optIfOTNMPNEPrevDayBBEs                      Integer32,
  optIfOTNMPNEPrevDayESR                       Integer32,
  optIfOTNMPNEPrevDaySESR                      Integer32,
  optIfOTNMPNEPrevDayBBER                      Integer32,
}
optIfOTNPMPrevDayType OBJECT-TYPE
SYNTAX OptIfOTNType
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"
::= { optIfOTNPMPrevDayEntry 1}

optIfOTNPMPrevDaySublayer 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 sublayer 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."
::= { optIfOTNPMPrevDayEntry 2}

optIfOTNPMPrevDayTCMLevel 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."
::= { optIfOTNPMPrevDayEntry 3}

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

optIfOTNPMPrevDayFcs    OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
   " Number of failures occurred in an observation period."
   ::= { optIfOTNPMPrevDayEntry  5}

optIfOTNPMPrevDayESs    OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
   " The number of seconds which have an error. It is a one-second period in which one or more bits are in error or during which Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected."
   ::= { optIfOTNPMPrevDayEntry  6}

optIfOTNPMPrevDaySESs   OBJECT-TYPE
SYNTAX  Integer32
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 which has a bit-error ratio = 1x10^-3 or during which Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected."
   ::= { optIfOTNPMPrevDayEntry  7}

optIfOTNPMPrevDayUAs    OBJECT-TYPE
SYNTAX  Integer32
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.
"
 ::= { optIfOTNPMPrevDayEntry 8}

optIfOTNPMPrevDayBBEs OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS read-only
STATUS  current
DESCRIPTION
" An errored block not occurring as part of an SES."
 ::= { optIfOTNPMPrevDayEntry 9}

optIfOTNPMPrevDayESR  OBJECT-TYPE
SYNTAX  Integer32
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."
 ::= { optIfOTNPMPrevDayEntry 10}

optIfOTNPMPrevDaySESR  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS read-only
STATUS  current
DESCRIPTION
" The ratio of SES in available time to total seconds in available
  time during a fixed measurement interval."
 ::= { optIfOTNPMPrevDayEntry 11}

optIfOTNPMPrevDayBBER  OBJECT-TYPE
SYNTAX  Integer32
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."
::= { optIfOTNPMObjects 6 }

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 OptIfOTNType,
  optIfOTNPMFECValidInterval Integer32,
  optIfOTNPMOTNThresh15MinFECUnCorrectedWords Integer32,
  optIfOTNPMOTNThreshPreFECBERMantissa Integer32,
  optIfOTNPMOTNThreshPreFECBERExponent Integer32,
}

optIfOTNPMFECConfigType OBJECT-TYPE
SYNTAX  OptIfOTNType
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 }

optIfOTNPMFECValidInterval OBJECT-TYPE
SYNTAX Integer32
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}

optIfOTNPM15MinThreshFECUnCorrectedWords  OBJECT-TYPE
SYNTAX  Integer32
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. A value of '0' will disable the notification.

::= {optIfOTNPMFECConfigEntry 3}

optIfOTNPM15MinThreshPreFECBERMantissa  OBJECT-TYPE
SYNTAX  Integer32
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}

optIfOTNPM15MinThreshPreFECBERExponent  OBJECT-TYPE
SYNTAX  Integer32
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}
Internet-Draft  draft-galimbe-kunze-g-698-2-snmp-mib-02      March 2012

--
-- FEC PM Table
--

optIfOTNPMFECCurrentTable OBJECT-TYPE
   SYNTAX  SEQUENCE OF OptIfOTNPMFECCurrentEntry
   MAX-ACCESS not-accessible
   STATUS  current
   DESCRIPTION
      "A Performance monitoring FEC Current Table."
   ::= {optIfOTNPObjects 7}

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}
   ::= {optIfOTNPMECCurrentTable 1}

OptIfOTNPMFECCurrentEntry ::= SEQUENCE {
   optIfOTNPMFECCurrentType            OptIfOTNType,
   optIfOTNPMFECCurrentSuspectedFlag   TruthValue,
   optIfOTNPMCurrentFECCorrectedErr    Integer32,
   optIfOTNPMCurrentFECUncorrectedWords Integer32,
   optIfOTNPMCurrentFECBERMantissa     Integer32,
   optIfOTNPMCurrentFECBERExponent     Integer32,
}

optIfOTNPMFECCurrentType OBJECT-TYPE
   SYNTAX OptIfOTNType
   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"
   ::= {optIfOTNPMECCurrentEntry 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  Integer32
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  Integer32
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  Integer32
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  Integer32
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."
::= {optIfOTNPMObjects 8}

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 OptIfOTNType,
    optIfOTNPMFECIntervalNumber OptIfIntervalNumber,
    optIfOTNPMFECIntervalSuspectedFlag TruthValue,
    optIfOTNPMIntervalFECCorrectedErr Integer32,
    optIfOTNPMIntervalFECUncorrectedErr Integer32,
    optIfOTNPMIntervalMinFECBERMantissa Integer32,
    optIfOTNPMIntervalMinFECBERExponent Integer32,
    optIfOTNPMIntervalMaxFECBERMantissa Integer32,
    optIfOTNPMIntervalMaxFECBERExponent Integer32,
    optIfOTNPMIntervalAvgFECBERMantissa Integer32,
    optIfOTNPMIntervalAvgFECBERExponent Integer32,
}

optIfOTNPMFECIntervalType OBJECT-TYPE
SYNTAX OptIfOTNType
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  OptIfIntervalNumber
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  Integer32
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  Integer32
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 Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"The minimum bit error rate at receiving side before the FEC function counted over one second. This is the minimum Pre FEC BER in the current 24hour period."
::= { optIfOTNPMFECIntervalEntry 6}

optIfOTNPMIntervalMinFECBERExponent OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"The minimum bit error rate at receiving side before the FEC function counted over one second. This is the minimum Pre FEC BER in the current 24hour period."
::= { optIfOTNPMFECIntervalEntry 7}

optIfOTNPMIntervalMaxFECBERMantissa OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"The maximum bit error rate at receiving side before the FEC function counted over one second. This is the maximum Pre FEC BER in the current 24hour period."
::= { optIfOTNPMFECIntervalEntry 8}

optIfOTNPMCurrentMaxFECBERExponent OBJECT-TYPE
SYNTAX Integer32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"The maximum bit error rate at receiving side before the FEC
function counted over one second .. exponent. This is the maximum Pre 
FEC BER in the current 24hour period.  
"  ::= { optIfOTNPMFECIntervalEntry  9}

optIfOTNPMIntervalAvgFECBERMantissa OBJECT-TYPE 
SYNTAX  Integer32 
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  Integer32 
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}
INDEX { ifIndex, optIfOTNPMFECCurrentDayType }
 ::= { optIfOTNPMFECCurrentDayTable 1 }

OptIfOTNPMFECCurrentDayEntry ::= 
 SEQUENCE {
   optIfOTNPMFECCurrentDayType OptIfOTNType,
   optIfOTNPMFECCurrentDaySuspectedFlag TruthValue,
   optIfOTNPMCurrentDayFECCorrectedErr Integer32,
   optIfOTNPMCurrentDayFECUncorrectedWords Integer32,
   optIfOTNPMCurrentDayMinFECBERMantissa Integer32,
   optIfOTNPMCurrentDayMinFECBERExponent Integer32,
   optIfOTNPMCurrentDayMaxFECBERMantissa Integer32,
   optIfOTNPMCurrentDayMaxFECBERExponent Integer32,
   optIfOTNPMCurrentDayAvgFECBERMantissa Integer32,
   optIfOTNPMCurrentDayAvgFECBERExponent Integer32,

   optIfOTNPMFECCurrentDayType        OBJECT-TYPE
   SYNTAX  OptIfOTNType
   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  Integer32
   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  Integer32
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  Integer32
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  Integer32
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  Integer32
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.
optIfOTNPMFECCurrentDayMaxFECBERExponent  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION

The maximum bit error rate at receiving side before the FEC function counted over one second. This is the maximum PreFEC BER in the current 24-hour period.

::= { optIfOTNPMFECCurrentDayEntry  7}

optIfOTNPMCurrentDayAvgFECBERMantissa  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION

The average bit error rate at receiving side before the FEC function counted over one second. This is the average PreFEC BER in the current 24-hour period.

::= { optIfOTNPMFECCurrentDayEntry  8}

optIfOTNPMCurrentdayAvgFECBERExponent  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION

The average bit error rate at receiving side before the FEC function counted over one second. This is the average PreFEC BER in the current 24-hour period.

::= { optIfOTNPMFECCurrentDayEntry  9}

-- 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."
::= {optIfOTNPMObjects 10}

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 OptIfOTNType,
  optIfOTNPMFECPrevDaySuspectedFlag TruthValue,
  optIfOTNPMPrevDayFECCorrectedErr Integer32,
  optIfOTNPMPrevDayFECUncorrectedWords Integer32,
  optIfOTNPMPrevDayMinFECBERMantissa Integer32,
  optIfOTNPMPrevDayMaxFECBERMantissa Integer32,
  optIfOTNPMPrevDayAvgFECBERMantissa Integer32,
  optIfOTNPMPrevDayMaxFECBERExponent Integer32,
  optIfOTNPMPrevDayAvgFECBERExponent Integer32,
}

optIfOTNPMFECPrevDayType OBJECT-TYPE
SYNTAX OptIfOTNType
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."
optIfOTNPMPrevDayFECCorrectedErr OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"The number of bits corrected by the FEC are counted in the previous day."
::= { optIfOTNPMFECPrevDayEntry 2}

optIfOTNPMPrevDayFECUncorrectedWords OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"The number of un-corrected words by the FEC are counted over the previous Day."
::= { optIfOTNPMFECPrevDayEntry 3}

optIfOTNPMPrevDayMinFECBERMantissa OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"The maximum bit error rate at receiving side before the FEC function counted over one second .. mantissa. This is the maximum Pre FEC BER in the previous 24hour period."
::= { optIfOTNPMFECPrevDayEntry 4}

optIfOTNPMPrevDayMinFECBERExponent OBJECT-TYPE
SYNTAX   Integer32
MAX-ACCESS read-only
STATUS   current
DESCRIPTION
"The minimum bit error rate at receiving side before the FEC function counted over one second .. exponent. This is the maximum Pre FEC BER in the previous 24hour period"
::= { optIfOTNPMFECPrevDayEntry 5}
optIfOTNPMPrevDayMaxFECBERMantissa  OBJECT-TYPE
SYNTAX  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"The maxium 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  Integer32
MAX-ACCESS  read-only
STATUS  current
DESCRIPTION
"The maxium 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  Integer32
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}

optIfOTNPMPrevdayAvgFECBERExponent  OBJECT-TYPE
SYNTAX  Integer32
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,
   optIfOTNAlarmSublayer                 OptIfOTNSublayer,
   optIfOTNAlarmTCMLevel                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 }

optIfOTNAlarmSublayer OBJECT-TYPE
SYNTAX      OptIfOTNSublayer
MAX-ACCESS  read-only
STATUS      current
DESCRIPTION
   "This specifies which sublayer this alarm is for."
::= { optIfOTNAlarmEntry 2 }

optIfOTNAlarmTCMLevel OBJECT-TYPE
SYNTAX Unsigned32
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"TCM level 1-6 of the alarm. It will be 0 if alarm sublayer 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 sublayer
‘optIfOTNAlarmSublayer’."
::= { 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 { optIfOTNAlarmSublayer,
          optIfOTNAlarmTCMLevel,
          optIfOTNAlarmType,
          optIfOTNAlarmDate }
STATUS current
DESCRIPTION
"Notification of a recently set OTN alarm of Sublayer
and Type."
::= { optIfOTNNotifications 1 }
optIfOTNAlarmClear NOTIFICATION-TYPE
    OBJECTS { optIfOTNAlarmSublayer,
              optIfOTNAlarmTCMLevel,
              optIfOTNAlarmType,
              optIfOTNAlarmDate }
    STATUS current
    DESCRIPTION "Notification of a recently clear OTN alarm of Sublayer
                  and Type."
    ::= { optIfOTNNotifications 2 }

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:

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
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:

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>OBJECT IDENTIFIER value</th>
</tr>
</thead>
<tbody>
<tr>
<td>sampleMIB</td>
<td>{ mib-2 XXX }</td>
</tr>
</tbody>
</table>

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


[ITU.G874]  International Telecommunications Union, "Management


12.2. 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.

[RFC4054] Strand, J. and A. Chiu, "Impairments and Other Constraints

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: +49xxxxxxxxxxx
Email: RKunze@telekom.de
Hing-Kam Lam (editor)
Alcatel-Lucent
600-700 Mountain Avenue, Murray Hill
New Jersey, 07974
USA
Phone: +19085820672
Email: kam.lam@alcatel-lucent.com

Dharini Hiremagalur (editor)
Juniper
1194 N Mathilda Avenue
Sunnyvale - 94089 California
USA
Phone: +1408
Email: dharinikh@juniper.net
Abstract

To enable scaling of existing transport systems to ultra high data rates of 1 Tbps and beyond, next generation systems providing super-channel switching capability are currently being developed. To allow efficient allocation of optical spectral bandwidth for such high bit rate systems, International Telecommunication Union Telecommunication Standardization Sector (ITU-T) is extending the G.694.1 grid standard (termed "Fixed-Grid") to include flexible grid (termed "Flex-Grid") support (draft revised ITU-T G.694.1, revision 1.4, Oct 2011). This necessitates definition of new label format for the Flex-Grid. This document defines a super-channel label as a Super-Channel Identifier and an associated list of 12.5 GHz slices representing the optical spectrum of the super-channel. The label information can be encoded using a fixed length or variable length format. This label format can be used in GMPLS signaling and routing protocol to establish super-channel based optical label switched paths (LSPs).

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/1id-abstracts.txt

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

This Internet-Draft will expire on September 10, 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. Terminology....................................................6
3. Motivation for Super-Channel Label.............................6
   3.1. Flex-Grid Slice Numbering................................6
   3.2. Super-Channel Label.......................................7
   3.2.1. Super-Channel Label Encoding Format..................8
   3.2.2. LSP Encoding Type, Switching Type, and Generalized-PID (G-PID) in Generalized Label Request..................11
4. Security Considerations.......................................11
5. IANA Considerations...........................................11
6. References....................................................11
1. Introduction

Future transport systems are expected to support service upgrades to data rates of 1 Tbps and beyond. To scale networks beyond 100Gbps, multi-carrier super-channels coupled with advanced multi-level modulation formats and flexible channel spectrum bandwidth allocation schemes have become pivotal for future spectral efficient transport network architectures [1,2].

A super-channel represents an ultra high aggregate capacity channel containing multiple carriers which are co-routed through the network as a single entity from the source transceiver to the sink transceiver [3,7]. By multiplexing multiple carriers, modulating each carrier with multi-level advanced modulation formats (such as PM-QPSK, PM-8QAM, PM-16QAM), allocating an appropriate-sized flexible channel spectral bandwidth slot, and using a coherent receiver for detecting closely packed sub-carriers, a super-channel can support ultra high data rates in a spectrally efficient manner while maintaining required system reach. Figure 1 contrasts channel spectrum bandwidth allocation schemes for various bit rate optical paths on fixed-grid and flex-grid. ITU-T fixed-grid permits allocation of channel spectrum bandwidth in "single" fixed-sized slots (e.g., 50GHz, 100GHz etc) independent of the channel bit rate. In contrast, a flex-grid can allocate "arbitrary" size channel spectral bandwidth as an integer multiple of 12.5 GHz fine granularity slices. This means, a flex-grid can support multiple data rates channels (optical paths) in a spectrally efficient manner as it allocates appropriate-sized spectrum bandwidth slots, as opposed to fixed-sized slots. As in the examples in the figure, the optical spectrum slices assigned will be to a given super-channel in a contiguous manner. However, for flexibility in finding available optical spectrum on fragmented fibers and to reduce signaling message overhead, the two schemes proposed in this document also allow for identification of a split-spectrum super-channel with optical spectral slices that are non-contiguous, spread across multiple slots. Note that the channel capacity available on a given number of optical spectral slices depends on (among other factors) how many contiguous optical slots are used. The definition of the channel capacity available for a split-spectrum super-channel split across multiple slots of different widths is outside the scope of this document.
ITU-T G.694.1
Center frequency (f) = 193.1 THz

\[ n = -3 \quad n = -2 \quad n = -1 \quad n = 0 \quad n = +1 \quad n = +2 \]

\[ \ldots \quad \ldots \]

\[ \begin{array}{ccccccc}
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
50 \text{ GHz} & 50 \text{ GHz} \\
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
\end{array} \]

\( n = -2 \quad n = +1 \)

(10 Gbps channel) (40 Gbps channel)
(a fixed 50 GHz chunk) (a fixed 50 GHz chunk)

\[ \begin{array}{ccccccc}
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 0 & 1 \\
\hline
\end{array} \]

\[ \begin{array}{ccccccc}
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
\begin{array}{cc}
1 \text{ Tbps super-channel} & 100 \text{ Gbps Channel} \\
16 \text{ slices of 12.5 GHz} & 4 \text{ slices} \\
\end{array} \\
\hline
\end{array} \]

\( \begin{array}{ccccccc}
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
\begin{array}{cc}
200 \text{ GHz} & 50 \text{ GHz} \\
\end{array} \\
\hline
\end{array} \]

\( \begin{array}{ccccccc}
\wedge & \wedge & \wedge & \wedge & \wedge & \wedge & \wedge \\
\hline
\begin{array}{cc}
1 \text{ Tbps super-channel} & 100 \text{ Gbps Channel} \\
16 \text{ slices of 12.5 GHz} & 4 \text{ slices} \\
\end{array} \\
\hline
\end{array} \]

(b)

Figure 1 ITU-T (a) 50 GHz fixed-grid (G.694.1) (b) 12.5 GHz granular flex-grid
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 RFC 2119 [RFC2119].

3. Motivation for Super-Channel Label

[RFC3471] defines new forms of MPLS "label" for the optical domain that are collectively referred to as a "generalized label". [RFC6205] defines a standard wavelength label based on ITU-T fixed-grids ([G.694.1] and [G.694.2]) for use by Lambda-Switch-Capable (LSC) LSRs.

A new label format for super-channels assignment on flex-grid is needed because the existing label formats (such as the waveband switching label defined in RFC3471 and the wavelength label defined in RFC6205) either lack necessary fields to carry required flex-grid related information (e.g., channel spacing) or do not allow signaling of arbitrary flexible-size optical spectral bandwidth in an efficient manner (e.g., in terms of integer multiple of fine granularity 12.5GHz slices). For example,

- Waveband switching label format (defined in section 3.3.1 of RFC3471) lacks fields to carry necessary information to support flex-grid.
- Wavelength label allows signaling of single fixed-size optical spectrum bandwidth slot only.
- Wavelength label does not allow signaling of arbitrary flexible-size optical spectrum bandwidth needed for super-channels assignment on flex-grid.

3.1. Flex-Grid Slice Numbering

Given a slice spacing value (e.g., 0.0125 THz) and a slice number "n", the slice left edge frequency can be calculated as follows:

Slice Left Edge Frequency (THz) = 193.1 THz + n*slice spacing (THz).

Where "n" is a two’s-complement integer (i.e., positive, negative, or 0) and "slice spacing" is 0.0125 THz conforming to ITU-T Flex-
Grid. (Editor’s Note: in the future, if necessary the slice numbering scheme will be updated in accordance with the Flex-Grid.)

Figure 2 shows an example using the slice number scheme described earlier.

3.2. Super-Channel Label

In order to setup an optical path manually or dynamically, we need a way to identify and reserve resources (i.e., signal optical spectral bandwidth for the super-channels) along the optical path. For this purpose, this document defines a super-channel label to cover the cases of split-spectrum super-channels as well, such that the label consists of a Super-Channel Identifier and an associated list of contiguous or non-contiguous set of 12.5 GHz slices representing arbitrary size optical spectrum of the super-channels (Note: in the future, slice granularity could be 6.25 GHz.)

\[(n=0 \text{ is } 193.1 \text{ THz})\]

\[
\begin{array}{ccccccc}
  n=-2 & n=-1 & n=0 & n=+1 & n=+2 \\
  ^ & ^ & ^ & ^ & ^ \\
  \ldots & \ldots & \ldots & \ldots & \ldots \\
  8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
  \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
\end{array}
\]

A super-channel with
Spectral BW = 150 GHz
(12 slices of 12.5 GHz)

\[
\begin{array}{c}
  n_{\text{start}} = -7 \\
  n_{\text{end}} = +4 \\
(\text{see label encoding format for details})
\end{array}
\]

Figure 2 flex-grid example of the proposed slice numbering scheme.
3.2.1. Super-Channel Label Encoding Format

This section describes two options (option A and B) for encoding the super-channel label by making extensions to the waveband switching label [RFC3471] and wavelength label [RFC6205] formats.

- Option A: Encode super-channel label as a list of start and end slice numbers corresponding to N slots, each consisting of contiguous slices with each slot denoted by its starting and ending slice number (e.g., "n_start_1" and "n_end_1" represent contiguous slices in slot#1, "n_start_2" and "n_end_2" in slot#2, ..., "n_start_N" and "n_end_N" in slot#N).

```
+----------------+---------+
|   Grid         |  Value  |
+----------------+---------+
| Reserved       |    0    |
| Super-Channel Id (16-bit) | Grid | S.S. | Reserved (9-bit) |
| Reserved (16-bit) | Number of Entries(16-bit) |
| n_start_1(contiguous slot #1) | n_end_1(contiguous slot #1) |
| n_start_2(contiguous slot #2) | n_end_2(contiguous slot #2) |
| n_start_N (contiguous slot #N) | n_end_N (contiguous slot #N) |
```

Super-Channel Id: 16 bits

This field represents a logical identifier for a super-channel or split-spectrum super-channel. To disambiguate waveband switching and super-channel label applications, we propose to rename the Waveband Identifier (32-bit) as a Super-Channel Identifier (16-bit).

Grid: 3 bits

This field indicates the Grid type. The value for Grid should be set to xx (to be assigned by IANA) for the ITU-T flex-grid.
S.S. (slice spacing): 4 bits

This field should be set to a value of 4 to indicate 12.5 GHz in both labels.

<table>
<thead>
<tr>
<th>S.S. (GHz)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>12.5</td>
<td>4</td>
</tr>
<tr>
<td>Future use</td>
<td>5 - 15</td>
</tr>
</tbody>
</table>

Number of Entries: 16-bit

This field represents the number of 32-bit entries in the super-channel label (i.e., number of slots with contiguous slices). For example, in the case of a super-channel with contiguous optical spectrum, this field should have a value of 1 (indicating one slot of contiguous slices).

\( n_{\text{start}_i} \) (\( i=1,2,\ldots N \)): 16 bits

\( n_{\text{end}_i} \) (\( i=1,2,\ldots N \)): 16 bits

A super-channel with contiguous spectrum or a split-spectrum super-channel with non-contiguous optical spectrum can be represented by \( N \) slots of slices where two adjacent slots can be contiguous or non-contiguous, however each slot contains contiguous slices. Each slot...
is denoted by \( n_{\text{start}}_i \) (which indicates the lowest or starting 12.5 GHz slice number of the slot) and \( n_{\text{end}}_i \) (which indicates the highest or ending 12.5 GHz slice number of the slot). "\( n_{\text{start}}_i \)" and "\( n_{\text{end}}_i \)" are two’s-complement integers that can take either a positive, negative, or zero value.

- **Option B:** Encode super-channel label as a first slice number of the grid (denoted as "\( n_{\text{start}} \text{ of Grid} \)") plus the entire list of slices in the grid as a Bitmap

```
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 |
| Super-Channel Id (16-bit)  Grid  S.S.  Reserved (9-bit) |
| +-----------------+-----------------+-----------------+-----------------
| \( n_{\text{start}} \text{ of Grid} \) (16-bit)  Num of Slices in Grid (16-bit) |
| +-----------------+-----------------+-----------------+-----------------
| Bitmap Word #1(first set of 32 slices from the left most edge) |
| +-----------------+-----------------+-----------------+-----------------
| Bitmap Word #2 (next set of 32 contiguous slice numbers) |
| +-----------------+-----------------+-----------------+-----------------
| ... |
| +-----------------+-----------------+-----------------+-----------------
| Bitmap Word #N(last set of 32 contiguous slice numbers) |
| +-----------------+-----------------+-----------------+-----------------+
```

Where:

Super-Channel Id, Grid, and S.S fields are same as described earlier in option A.

**n_start of Grid:** 16-bit

This field indicates the first slice number in Grid for the band being referenced (i.e., the start of the left most edge of the Grid).

**Num of Slices in Grid:** 16-bit

This field represents the total number of slices in the band. The value in this field determines the number of 32-bitmap words required for the grid.

**Bitmap (Word):** 32-bit
Each bit in the 32-bitmap word represents a particular slice with a value of 1 or 0 to indicate whether for that slice reservation is required (1) or not (0). Bit position zero in the first word represents the first slice in the band (Grid) and corresponds to the value indicated in the "n_start of Grid" field.

Both options allow efficient encoding of a super-channel label with contiguous and non-contiguous slices. Option B yields a fixed length format while option A a variable length format. Option B is relatively simpler, more flexible, however, might be less compact than option A for encoding a single super-channel with contiguous optical spectrum. In contrast, option A provides a very compact representation for super-channels with contiguous optical spectrum, however, might be less flexible in encoding split-spectrum super-channels with arbitrary non-contiguous set of slices.

3.2.2. LSP Encoding Type, Switching Type, and Generalized-PID (G-PID) in Generalized Label Request

For requesting a super-channel label in a Generalized Label Request defined in section 3.1.1 of RFC3471, this document proposes to use LSP Encoding Type = Lambda (as defined in RFC4328), Switching Type = Super-Channel-Switch-Capable(SCSC) (as defined in [6]), and a new G-PID type = OTUadapt and a new G-PID value (similar to as defined in section 3.1.3 of RFC4328) to be assigned by IANA.

4. Security Considerations

<Add any security considerations>

5. IANA Considerations

IANA needs to assign a new Grid field value to represent ITU-T Flex-Grid and a new G-PID value.

6. References

6.1. Normative References


6.2. Informative References


7. Acknowledgments

<Add any acknowledgements>
Appendix A. Super-Channel Label Format Example

Suppose node A and Node Z are super-channel switching capable and node A receives a request for establishing a 1 Tbps optical LSP from itself to node Z. Assume the super-channel requires a "contiguous" spectral bandwidth of 200 GHz with left-edge frequency of 191.475 THz for the left-most 12.5 GHz slice and left-edge frequency of 191.6625 THz for the right-most slice. This means $n_{\text{start}} = (191.475 - 193.1)/0.0125 = -130$ and $n_{\text{end}} = (191.6625 - 193.1)/0.0125 = -115$ (i.e. we need 16 slices of 12.5 GHz starting from slice number -130 and ending at slice number -115).

Node A signals the LSP via a Path message including a super-channel label format encoding option A defined in section 3.3:

```
0                   1                   2                   3
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Super-Channel Id (16-bit)    |Grid | S.S.  | Reserved (9-bit)|
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Reserved (16-bit)            | Number of Entries(16-bit) |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |n_start_1 (contiguous slot #1)|  n_end_1(contiguous slot #1) |
 +-------------------------------+-------------------------------+
```

Where:

- Super-Channel Id = 1 : super-channel number 1
- Number of Entries: 1
- Grid = xx : ITU-T Flex-Grid
- S.S. = 4 : 12.5 GHz Slice Spacing
- $n_{\text{start}}_1 = -130$ : left-most 12.5 GHz slice number for slot 1
- $n_{\text{end}}_1 = -115$ : Right-most 12.5 GHz slice number for slot 1
Authors’ Addresses

Iftekhar Hussain
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: ihussain@infinera.com

Abinder Dhillon
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: adhillon@infinera.com

Zhong Pan
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: zpan@infinera.com

Marco Sosa
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: msosa@infinera.com

Bert Basch
Verizon Communications
60 Sylvan Rd., Waltham, MA 02451
Email: bert.e.basch@verizon.com

Steve Liu
Verizon Communications
60 Sylvan Rd., Waltham, MA 02451
Email: steve.liu@verizon.com
Andrew G. Malis  
Verizon Communications  
60 Sylvan Rd., Waltham, MA 02451  
Email: andrew.g.malis@verizon.com

Contributor’s Addresses

Rajan Rao  
Infinera  
140 Caspian Ct., Sunnyvale, CA 94089  
Email: rrao@infinera.com

Biao Lu  
Infinera  
140 Caspian Ct., Sunnyvale, CA 94089  
Email: blu@infinera.com

Subhendu Chattopadhyay  
Infinera  
140 Caspian Ct., Sunnyvale, CA 94089  
Email: schattopadhyay@infinera.com

Harpreet Uppal  
Infinera  
140 Caspian Ct., Sunnyvale, CA 94089  
Email: harpreet.uppal@infinera.com
Super-Channel Optical Parameters GMPLS Routing Extensions
draft-hussain-ccamp-super-channel-param-ospfte-00.txt

Abstract

This document builds on [6][7] and defines GMPLS routing extensions to allow added CSPF constraints for efficient super-channel spectrum assignment on flexible grid networks.

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/drafts.txt
1. Introduction

Future transport systems are expected to support service upgrades to data rates of 1 Tbps and beyond. To scale networks beyond 100Gbps, multi-carrier super-channels coupled with advanced multi-level modulation formats and flexible channel spectrum bandwidth
allocation schemes have become pivotal for future spectral efficient transport network architectures [1,2].

The coexistence of super-channels using different modulation formats on the same optical fiber network infrastructure may have a detrimental effect on the Optical Signal to Noise Ratio (OSNR) of adjacent super-channels due to interference such as cross-phase modulation. Therefore, it may be highly desirable to be able to evaluate the mutual impact of the existing and new super-channels on each other’s quality of transmission (e.g., bit error rate) before establishing new super-channels.

The document [9] defines GMPLS signaling extensions to convey super-channel optical parameters. This document defines GMPLS routing extensions to advertise the above mentioned super-channel parameters via OSPF-TE link LSA using new Super-Channel sub-TLV. This sub-TLV is carried under the Switching Capability-specific information (SCSI) field of the Interface Switching Capability Descriptor (ISCD) with the Super-Channel-Switch-Capable (SCSC) value defined in [6]. This information allows each source node across the network to apply added CSPF constraints and assign new super-channels spectrum efficiently by considering not only the availability of the required number of slices but also the optical signal compatibility of the existing and the new super-channels along the desired path.

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 RFC 2119 [RFC2119].

3. GMPLS Routing Extensions for Super-Channel Optical Parameters

This document defines OSPF-TE extensions for advertising following information using the Super-Channel sub-TLV depicted in Figure 1. For each super-channel this sub-TLV advertises following information:

- Super-Channel In-Use Slices sub-TLV
- Super-Channel Carriers sub-TLV
The Super-Channel sub-TLV is advertised in the OSPF-TE link LSA under the SCSI field of the ISCD using Super-Channel-Switch-Capable (SCSC) value defined in [6].

3.1. Super-Channel In-Use Slices sub-TLV

This sub-TLV contains the in-use slices information of a super-channel. For further information about various fields in this sub-TLV refer to [6][7].
Figure 2: Super-Channel In-Use Slices sub-TLV Format.

[Editor’s Note: encoding of in-use slices in bitmap format is left for a possible future revision]

3.2. Super-Channel Carriers sub-TLV

The format of the Super-Channel Carriers sub-TLV is defined in [9]. In summary, this sub-TLV contains following information.

- Number of Carriers in the Super-Channel
- Carrier sub-TLV
  - Carrier Center Frequency sub-sub-TLV
  - Carrier Modulation sub-sub-TLV
  - Carrier FEC sub-sub-TLV

4. Procedure for OSPF-TE Advertisement

This section describes procedure for advertising the aforementioned information in the OSPF-TE link LSAs.

- The optical parameters of the super-channel are signaled when new super-channels are established (see [9]).
Over time change in the status of in-use slices occurs when new super-channels are setup (or when established super-channels are released).

Each node along the path traversed by the super-channels advertises the current status of the in-use slices for each super-channel in the OSPF-TE link LSA using sub-TLVs described earlier.

Through OSPF-TE LSAs flooding other nodes in the routing domain learn about the current status of in-use slices on each TE link.

5. Possible Applications

The presence of this information across the network topology enables source nodes in the network to apply added CSPF constraints for example to:

- Group super-channels with different modulation formats in different bands (slice ranges)
- Group super-channels with same bit-rate in a band while separating with guard band from super-channels with different bit-rate.

Allows efficient network utilization (e.g., reduces new requests blocking probability) by avoiding excessive worst-case OSNR penalty while preserving desired quality of transmission of the existing super-channels.

6. TLV Encoding Examples

To be added later.

7. Security Considerations

<Add any security considerations>

8. IANA Considerations

IANA needs to assign a new Grid field value to represent ITU-T Flex-Grid.
9. References

9.1. Normative References


9.2. Informative References


10. Acknowledgments

<Add any acknowledgements>
Authors’ Addresses

Iftekhar Hussain
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: ihussain@infinera.com

Abinder Dhillon
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: adhillon@infinera.com

Marco Sosa
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: msosa@infinera.com
Contributor’s Addresses

Rajan Rao
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: rrao@infinera.com

Biao Lu
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: blu@infinera.com

Subhendu Chattopadhyay
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: schattopadhyay@infinera.com

Harpreet Uppal
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: harpreet.uppal@infinera.com

Vinayak Dangui
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: vdangui@infinera.com

Michael VanLeeuwen
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: MVanleeuwen@infinera.com

Zhong Pan
Infinera
140 Caspian Ct., Sunnyvale, CA 94089

Email: zpan@infinera.com
Abstract

This document builds on [6][7] and defines GMPLS signaling extensions to carry super-channel optical parameters for efficient spectrum assignment on flexible grid networks.

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/1id-abstracts.txt

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

Future transport systems are expected to support service upgrades to data rates of 1 Tbps and beyond. To scale networks beyond 100Gbps, multi-carrier super-channels coupled with advanced multi-level modulation formats and flexible channel spectrum bandwidth allocation schemes have become pivotal for future spectral efficient transport network architectures [1,2].
The coexistence of super-channels using different modulation formats on the same optical fiber network infrastructure may have a detrimental effect on the Optical Signal to Noise Ratio (OSNR) of adjacent super-channels due to interference such as cross-phase modulation. Therefore, it may be highly desirable to be able to evaluate the mutual impact of the existing and new super-channels on each other’s quality of transmission (e.g., bit error rate) before establishing new super-channels.

This document defines GMPLS signaling extensions to convey super-channel optical parameters such as number of carriers, each carrier’s center frequency, modulation, and FEC type in the RSVP message. This allows nodes along the super-channel path to learn the aforementioned super-channel optical characteristics and in turn advertise this information to other nodes in the network using GMPLS routing extensions defined in [9].

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 RFC 2119 [RFC2119].

3. GMPLS Signaling Extensions for Super-Channel Optical Parameters

This document defines extensions for signaling super-channel optical parameters including:

- Number of Carriers
- Carrier Center Frequency (THz)
- Carrier Modulation
- Carrier Baudrate (Gbit/s)
- Carrier FEC Type

This document defines two options for encoding this information.

[Editor’s note: to allow full flexibility we have included two encoding options]
3.1. Option 1: Encode Super-Channel Optical Parameters in the RSVP FLOWSPEC or TSPEC Object

Figure 1: Super-Channel Carriers TLV Format

Figure 2: Carrier sub-TLV Format.

Figure 3: Carrier Center Frequency sub-sub-TLV.
<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 4: Carrier Modulation sub-sub-TLV.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>I</td>
<td>TLV Type</td>
<td>Length</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Figure 5: Carrier FEC sub-sub-TLV.

Where:

- When the S bit in a TLV is set to 1 it indicates that the TLV contains standardized fields (e.g., Modulation, FEC Type) and when the S bit is set to 0 in a TLV it indicates a vendor specific TLV (see [8])
- Modulation ID, FEC ID, and I fields are similar to as defined in [8]
- The Length field in the super-channel Carriers TLV specifies the length in octets of the complete set of TLVs including the set of sub-TLVs that follow.
3.2. Option 2: Encode the Aforementioned Information along with the Super-Channel Label

For example use Super-Channel Label defined in [7] to also encode Super-Channel Carriers TLV, the Carrier sub-TLVs, and the associated set of sub-sub-TLVs defined in the previous section.

4. Procedure for Signaling Super-Channel Optical Parameters

- The optical parameters of the super-channel are signaled in the RSVP message using encoding option 1 (or option 2).
- During a new super-channel establishment, each node along the new super-channel setup path allocates the required number of slices and also learns the associated set of signaled super-channel optical parameters.

5. TLV Encoding Examples

To be added later.

6. Security Considerations

<Add any security considerations>

7. IANA Considerations

IANA needs to assign a new Grid field value to represent ITU-T Flex-Grid.

8. References

8.1. Normative References


8.2. Informative References


9. Acknowledgments

<Add any acknowledgements>
Authors’ Addresses

Iftekhar Hussain
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: ihussain@infinera.com

Vinayak Dangui
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: vdangui@infinera.com

Michael VanLeeuwen
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: MVanleeuwen@infinera.com

Marco Sosa
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: msosa@infinera.com
Contributor’s Addresses

Abinder Dhillon
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: adhillon@infinera.com

Rajan Rao
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: rrao@infinera.com

Biao Lu
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: blu@infinera.com

Subhendu Chattopadhyay
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: schattopadhyay@infinera.com

Harpreet Uppal
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: harpreet.uppal@infinera.com

Zhong Pan
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: zpan@infinera.com
General Network Element Constraint Encoding for GMPLS Controlled Networks

draft-ietf-ccamp-general-constraint-encode-07.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/1id-abstracts.txt

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

This Internet-Draft will expire on September 6, 2012.

Copyright Notice

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

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.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Action |    Num Labels         |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Base Label                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Additional fields as necessary per action                 |
```

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:
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 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:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|A| Reserved    | Priority Flags|        Reserved               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Label Set Field                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
+------------------------------------------------------------------+
| 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:

- 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.

- 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.

```
+------------------+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MatrixID | RestrictionType |   Reserved/Parameter  |
|---------------------------------------------------------------|
| 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)

2.6.1. SIMPLE_LABEL

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

```
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 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 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:
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 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                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

In this case the accompanying label set and MaxNumChannels indicate labels permitted on the port and the maximum number of labels that can be simultaneously used on the port.

2.6.5. Link Label Exclusivity

In the case of the SIMPLE_LABEL & CHANNEL_COUNT the format is given by:
In this case the accompanying port set indicate that a label may be used at most once among the ports in the link set field.

3. 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.

4. IANA Considerations

TBD. Once our approach is finalized we may need identifiers for the various TLVs and sub-TLVs.

5. Acknowledgments

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

<table>
<thead>
<tr>
<th>Frequency (THz)</th>
<th>n Value</th>
<th>bit map position</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.0</td>
<td>-11</td>
<td>0</td>
</tr>
<tr>
<td>192.5</td>
<td>-6</td>
<td>5</td>
</tr>
<tr>
<td>193.1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>193.9</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>194.0</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>195.2</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>195.8</td>
<td>27</td>
<td>38</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Grid</th>
<th>C.S.</th>
<th>Reserved</th>
<th>n for lowest frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>Not used in 40 Channel system (all zeros)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
### 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.
(Tributary)

Ports #3-42          Ports #43-82
West Line Egress    East Line Ingress
vvvvv                ^^^^^

+-----| |||.|--------| |||.|------+
|    +----------------------+     |
|    |                      |     |
|    -------------------+    |                      |     +--------------
|                                 |
Port #1                   Port #2
(West Line Side)            (East Line Side)

<===================================

+-----| |||.|--------| |||.|------+
| |||.|        | |||.|
vvvv          ^^^^^

Egress dropped from West Line ingress
Ingress added to East Line egress
<table>
<thead>
<tr>
<th>Conn = 1</th>
<th>MatrixID</th>
<th>Reserved</th>
<th>Add/Drops #3-42 to Line side #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action=1</td>
<td>0 0 0 0 0 0 0</td>
<td>Length = 12</td>
<td></td>
</tr>
<tr>
<td>Link Local Identifier = #3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link Local Identifier = #42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action=0</td>
<td>0 0 0 0 0 0 0</td>
<td>Length = 8</td>
<td></td>
</tr>
<tr>
<td>Link Local Identifier = #1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: line #2 to add/drops #43-82

| Action=0 | 0 0 0 0 0 0 0 | Length = 8 |                               |
| Link Local Identifier = #2 |                     |          |                               |
| Action=1 | 0 0 0 0 0 0 0 | Length = 12 |                               |
| Link Local Identifier = #43 |                     |          |                               |
| Link Local Identifier = #82 |                     |          |                               |

Note: line to line

| Action=0 | 0 0 0 0 0 0 0 | Length = 8 |                               |
| Link Local Identifier = #1 |                     |          |                               |
| Action=0 | 0 0 0 0 0 0 0 | Length = 8 |                               |
| Link Local Identifier = #2 |                     |          |                               |
6. References

6.1. Normative References


6.2. Informative References


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

draft-ietf-ccamp-gmpls-g709-framework-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 September 8, 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 ..................... 7
       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 ........................ 16
   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 .................................................. 21
    10.1. Normative References .................................. 21
    10.2. Informative References ............................... 23
11. Authors’ Addresses .......................................... 24
12. Contributors ................................................ 25
APPENDIX A: ODU connection examples ............................. 25

1. Introduction

OTN has become a mainstream layer 1 technology for the transport network. Operators want to introduce control plane capabilities based on Generalized Multi-Protocol Label Switching (GMPLS) to OTN networks, to realize the benefits associated with a high-function control plane (e.g., improved network resiliency, resource usage efficiency, etc.).

GMPLS extends MPLS to encompass time division multiplexing (TDM) networks (e.g., SONET/SDH, PDH, and G.709 sub-lambda), lambda switching optical networks, and spatial switching (e.g., incoming port or fiber to outgoing port or fiber). The GMPLS architecture is provided in [RFC3945], signaling function and Resource ReserVation
Protocol-Traffic Engineering (RSVP-TE) extensions are described in [RFC3471] and [RFC3473], routing and OSPF extensions are described in [RFC4202] and [RFC4203], and the Link Management Protocol (LMP) is described in [RFC4204].

The GMPLS protocol suite including provision [RFC4328] provides the mechanisms for basic GMPLS control of OTN networks based on the 2001 revision of the G.709 specification [G709-V1]. Later revisions of the G.709 specification, including [G709-V3], have included some new features; for example, various multiplexing structures, two types of TSs (i.e., 1.25Gbps and 2.5Gbps), and extension of the Optical Data Unit (ODU) ODUj definition to include the ODUflex function.

This document reviews relevant aspects of OTN technology evolution that affect the GMPLS control plane protocols and examines why and how to update the mechanisms described in [RFC4328]. This document additionally provides a framework for the GMPLS control of OTN networks and includes a discussion of the implication for the use of the Path Computation Element (PCE) [RFC4655].

For the purposes of the control plane the OTN can be considered as being comprised of ODU and wavelength (OCh) layers. This document focuses on the control of the ODU layer, with control of the wavelength layer considered out of the scope. Please refer to [RFC6163] for further information about the wavelength layer.

2. Terminology

OTN: Optical Transport Network

ODU: Optical Channel Data Unit

OTU: Optical channel transport unit

OMS: Optical multiplex section

MSI: Multiplex Structure Identifier

TPN: Tributary Port Number

LO ODU: Lower Order ODU. The LO ODUj (j can be 0, 1, 2, 2e, 3, 4, flex.) represents the container transporting a client of the OTN that is either directly mapped into an OTUk (k = j) or multiplexed into a server HO ODUk (k > j) container.
HO ODU: Higher Order ODU. The HO ODUk (k can be 1, 2, 2e, 3, 4.) represents the entity transporting a multiplex of LO ODUj tributary signals in its OPUk area.

ODUflex: Flexible ODU. A flexible ODUk can have any bit rate and a bit rate tolerance up to +/-100 ppm.

3. G.709 Optical Transport Network (OTN)

This section provides an informative overview of those aspects of the OTN impacting control plane protocols. This overview is based on the ITU-T Recommendations that contain the normative definition of the OTN. Technical details regarding OTN architecture and interfaces are provided in the relevant ITU-T Recommendations.

Specifically, [G872-2001] and [G872Am2] describe the functional architecture of optical transport networks providing optical signal transmission, multiplexing, routing, supervision, performance assessment, and network survivability. [G709-V1] defines the interfaces of the optical transport network to be used within and between subnetworks of the optical network. With the evolution and deployment of OTN technology many new features have been specified in ITU-T recommendations, including for example, new ODU0, ODU2e, ODU4 and ODUflex containers as described in [G709-V3].

3.1. OTN Layer Network

The simplified signal hierarchy of OTN is shown in Figure 1, which illustrates the layers that are of interest to the control plane. Other layers below OCh (e.g. Optical Transmission Section - OTS) are not included in this Figure. The full signal hierarchy is provided in [G709-V3].

```
Client signal
   | ODUj
   | OTU/OCh
   | OMS
```

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.

<table>
<thead>
<tr>
<th>ODU Type</th>
<th>ODU nominal bit rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU0</td>
<td>1 244 160 kbits/s</td>
</tr>
<tr>
<td>ODU1</td>
<td>239/238 x 2 488 320 kbit/s</td>
</tr>
<tr>
<td>ODU2</td>
<td>239/237 x 9 953 280 kbit/s</td>
</tr>
<tr>
<td>ODU3</td>
<td>239/236 x 39 813 120 kbit/s</td>
</tr>
<tr>
<td>ODU4</td>
<td>239/227 x 99 532 800 kbit/s</td>
</tr>
<tr>
<td>ODU2e</td>
<td>239/237 x 10 312 500 kbit/s</td>
</tr>
<tr>
<td>ODUflex for CBR</td>
<td>239/238 x client signal bit rate</td>
</tr>
<tr>
<td>Mapped client</td>
<td>Configured bit rate</td>
</tr>
</tbody>
</table>

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).
<table>
<thead>
<tr>
<th>ODU Type</th>
<th>ODU bit-rate tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU0</td>
<td>+/- 20 ppm</td>
</tr>
<tr>
<td>ODU1</td>
<td>+/- 20 ppm</td>
</tr>
<tr>
<td>ODU2</td>
<td>+/- 20 ppm</td>
</tr>
<tr>
<td>ODU3</td>
<td>+/- 20 ppm</td>
</tr>
<tr>
<td>ODU4</td>
<td>+/- 20 ppm</td>
</tr>
<tr>
<td>ODU2e</td>
<td>+/- 100 ppm</td>
</tr>
<tr>
<td>ODUflex for CBR</td>
<td></td>
</tr>
<tr>
<td>Client signals</td>
<td>+/- 100 ppm</td>
</tr>
<tr>
<td>ODUflex for GFP-F</td>
<td></td>
</tr>
<tr>
<td>Mapped client signal</td>
<td>+/- 100 ppm</td>
</tr>
</tbody>
</table>

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 bit rate should be set to an integer multiplier of the High Order (HO) Optical Channel Physical Unit (OPU) OPUk TS rate, the tolerance should be +/-100ppm, and the bit rate should be determined by the node that performs the mapping.

[Editors’ Note: As outcome of ITU SG15/q11 expert meeting held in Vimercate in September 2010 it was decided that a resizable ODUflex(GFP) occupies the same number of TS on every link of the path (independently of the High Order (HO) OPUk TS rate). Please see WD07 and the meeting report of this meeting for more information.

The authors will update the above text related to Packet client mapping as soon as new version of G.709 will be updated accordingly with expert meeting decision reported here.]
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 ODUj's 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:

<table>
<thead>
<tr>
<th>OTUk</th>
<th>2.5Gb/s</th>
<th>1.25Gb/s</th>
<th>Nominal Bit rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTU1</td>
<td>1</td>
<td>2</td>
<td>2.5Gb/s</td>
</tr>
<tr>
<td>OTU2</td>
<td>4</td>
<td>8</td>
<td>10Gb/s</td>
</tr>
<tr>
<td>OTU3</td>
<td>16</td>
<td>32</td>
<td>40Gb/s</td>
</tr>
<tr>
<td>OTU4</td>
<td>--</td>
<td>80</td>
<td>100Gb/s</td>
</tr>
</tbody>
</table>

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
  - 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
  - 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
  - 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 <= 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 ODUj (including ODUflex) into the OTUk
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 ODUj into a HO ODUk (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 OPUk 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 ODUj 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 ODUj
transported by the HO ODUk. The TPN is the same for all the TSs
assigned to the transport of the same ODUj 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 ODUk 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 ODUj connections independent of the value of j. Note that when the bit rate of ODUj is less than the server bit rate, ODUj 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\(_j\)s 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 TSSs 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](image-url)
If an ODUj 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 ODUj to OTUk 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](image-url)
Figure 4 - RWA Visible Topology for LO ODUj connection management

In Figure 4, the cloud of previous figure is substitute by the real topology. The nodes f, g, h are nodes with OCH switching capability.

In the examples (i.e., Figure 3 and Figure 4), we have considered the case in which LO-ODUj connections are supported by OCh connection, and the case in which the supporting underlying connection can be also made by a combination of HO-ODU/OCh connections.

In this case, the ODU routing/path selection process will request an HO-ODU/OCh connection between node C and node E from the RWA domain. The connection will appear at ODU level as a Forwarding Adjacency, which will be used to create the ODU connection.

5. GMPLS/PCE Implications

The purpose of this section is to provide a set of requirements to be evaluated for extensions of the current GMPLS protocol suite and the PCE applications and protocols to encompass OTN enhancements and connection management.

5.1. Implications for LSP Hierarchy with GMPLS TE

The path computation for ODU connection request is based on the topology of ODU layer, including OCh layer visibility.

The OTN path computation can be divided into two layers. One layer is OCh/OTUK, the other is ODUj. [RFC4206] and [RFC6107] define the mechanisms to accomplish creating the hierarchy of LSPs. The LSP
management of multiple layers in OTN can follow the procedures defined in [RFC4206], [RFC6107] and related MLN drafts.

As discussed in section 4, the route path computation for OCh is in the scope of WSON [RFC6163]. Therefore, this document only considers ODU layer for ODU connection request.

LSP hierarchy can also be applied within the ODU layers. One of the typical scenarios for ODU layer hierarchy is to maintain compatibility with introducing new [G709-V3] services (e.g., ODU0, ODUflex) into a legacy network configuration (containing [G709-V1] or [G709-V2] OTN equipment). In this scenario, it may be needed to consider introducing hierarchical multiplexing capability in specific network transition scenarios. One method for enabling multiplexing hierarchy is by introducing dedicated boards in a few specific places in the network and tunneling these new services through [G709-V1] or [G709-V2] containers (ODU1, ODU2, ODU3), thus postponing the need to upgrade every network element to [G709-V3] capabilities.

In such case, one ODUj connection can be nested into another ODUk (j<k) connection, which forms the LSP hierarchy in ODU layer. The creation of the outer ODUk connection can be triggered via network planning, or by the signaling of the inner ODUj connection. For the former case, the outer ODUk connection can be created in advance based on network planning. For the latter case, the multi-layer network signaling described in [RFC4206], [RFC6107] and [RFC6001] (including related modifications, if needed) are relevant to create the ODU connections with multiplexing hierarchy. In both cases, the outer ODUk connection is advertised as a Forwarding Adjacency (FA).

5.2. Implications for GMPLS Signaling

The signaling function and Resource reSerVation Protocol-Traffic Engineering (RSVP-TE) extensions are described in [RFC3471] and [RFC3473]. For OTN-specific control, [RFC4328] defines signaling extensions to support G.709 Optical Transport Networks Control as defined in [G709-V1].

As described in Section 3, [G709-V3] introduced some new features that include the ODU0, ODU2e, ODU4 and ODUflex containers. The mechanisms defined in [RFC4328] do not support such new OTN features, and protocol extensions will be necessary to allow them to be controlled by a GMPLS control plane.

[RFC4328] defines the LSP Encoding Type, the Switching Type and the Generalized Protocol Identifier (Generalized-PID) constituting the common part of the Generalized Label Request. The G.709 Traffic
Parameters are also defined in [RFC4328]. The following signaling aspects should be considered additionally since [RFC4328] was published:

- Support for specifying the new signal types and the related traffic information

The traffic parameters should be extended in signaling message to support the new optical Channel Data Unit (ODUj) including:

- ODU0
- ODU2e
- ODU4
- ODUflex

For ODUflex, since it has a variable bandwidth/bit rate BR and a bit rate tolerance T, the (node local) mapping process must be aware of the bit rate and tolerance of the ODUj being multiplexed in order to select the correct number of TS and the fixed/variable stuffing bytes. Therefore, bit rate and bit rate tolerance should also be carried in the Traffic Parameter in the signaling of connection setup request.

For other ODU signal types, the bit rates and tolerances of them are fixed and can be deduced from the signal types.

- Support for LSP setup using different Tributary Slot Granularity (TSG)

The signaling protocol should be able to identify the type of TS (i.e., the 2.5 Gbps TS granularity and the new 1.25 Gbps TS granularity) to be used for establishing an H-LSP which will be used to carry service LSP(s) requiring specific TS type.

- Support for LSP setup of new ODUk/ODUflex containers with related mapping and multiplexing capabilities

New label should be defined to carry the exact TS allocation information related to the extended mapping and multiplexing hierarchy (For example, ODU0 into ODU2 multiplexing (with 1.25Gbps TS granularity)), in order to setting up the ODU connection.

- Support for Tributary Port Number allocation and negotiation

Tributary Port Number needs to be configured as part of the MSI information (See more information in Section 3.1.2.1). A new
extension object has to be defined to carry TPN information if control plane is used to configure MSI information.

- Support for ODU Virtual Concatenation (VCAT) and Link Capacity Adjustment Scheme (LCAS)

GMPLS signaling should support the creation of Virtual Concatenation of ODUk signal with k=1, 2, 3. The signaling should also support the control of dynamic capacity changing of a VCAT container using LCAS ([G.7042]). [RFC6344] has a clear description of VCAT and LCAS control in SONET/SDH and OTN networks.

- Support for ODU layer multiplexing hierarchy signaling

ODU layer multiplexing hierarchy has been supported by [G709-V3], i.e., a client ODUj 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 should select a suitable route for an ODUj connection request. In order to perform the path computation, it must evaluate the available bandwidth on each candidate link. The routing protocol should be extended to convey some information to represent ODU TE topology.

GMPLS Routing [RFC4202] defines Interface Switching Capability Descriptor of TDM which can be used for ODU. However, some issues discussed below, should also be considered.

Interface Switching Capability Descriptors present a new constraint for LSP path computation. [RFC4203] defines the switching capability and related Maximum LSP Bandwidth and the Switching Capability specific information. When the Switching Capability field is TDM the Switching Capability Specific Information field includes Minimum LSP Bandwidth, an indication whether the interface supports Standard or Arbitrary SONET/SDH, and padding. Hence a new Switching Capability value needs to be defined for [G709-V3] ODU switching in order to allow the definition of a new Switching Capability Specific Information field definition. The following requirements should be considered:

- Support for carrying the link multiplexing capability

As discussed in section 3.1.2, many different types of ODUj can be multiplexed into the same OTUk. For example, both ODU0 and ODU1 may be multiplexed into ODU2. An OTU link may support one or more types of ODUj signals. The routing protocol should be capable of carrying this multiplexing capability.

- Support any ODU and ODUflex

The bit rate (i.e., bandwidth) of TS is dependent on the TS granularity and the signal type of the link. For example, the bandwidth of a 1.25G TS in an OTU2 is about 1.249409620 Gbps, while the bandwidth of a 1.25G TS in an OTU3 is about 1.254703729 Gbps.

One LO ODU may need different number of TSs when multiplexed into different HO ODUs. For example, for ODU2e, 9 TSs are needed when multiplexed into an ODU3, while only 8 TSs are needed when multiplexed into an ODU4. For ODUflex, the total number of TSs to
be reserved in a HO ODU equals the maximum of \([\text{bandwidth of ODUflex} / \text{bandwidth of TS of the HO ODU}]\).

Therefore, the routing protocol must be capable of carrying the necessary and sufficient link bandwidth information for performing accurate route computation for any of the fixed rate ODU\(_s\) 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 must 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 must 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

Assume [RFC4328] has been deployed to control the OTN networks supporting [G709-V1], control plane backward compatibility needs to be taken into consideration. Scenarios for backward compatibility are described as follows:

- Legacy OTN devices supporting [G709-V1] may run control plane protocol defined in [RFC4328];
- Legacy OTN devices supporting [G709-V1] may also support new OTN control plane characterized in this document after control plane updating;
- New OTN devices supporting [G709-V3] always support new OTN control plane characterized in this document;
- New OTN devices SHOULD support falling back to [RFC4328] for interworking scenarios.

Based on these scenarios, control plane backward compatibility SHOULD be taken into account when interworking between the new control plane characterized in this document and the legacy control plane defined in [RFC4328].
A new Switching Capability type is required for control of [G709-V3] in the routing and signaling to enable the backward procedure.

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.

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<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]). [GMPLS-SEC] 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 Vissers for his review and useful comments.

10. References

10.1. Normative References


10.2. Informative References


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
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 must first be created. 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\text{j} client that is mapped into an OTU\text{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\text{j} client that is mapped into an ODU\text{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\text{j} represents the container transporting a client of the OTN that is either directly mapped into an OTU\text{k} (k = j) or multiplexed into a server HO ODU\text{k} (k > j) container. Consequently, the HO ODU\text{k} represents the entity transporting a multiplex of LO ODU\text{j} tributary signals in its OPU\text{k} area.

In the case of LO ODU\text{j} mapped into an OTU\text{k} (k = j) directly, Figure 6 give an example of this kind of LO ODU connection.

In Figure 6, The LO ODU\text{j} is switched at the intermediate ODXC node. OCh and OTUk are associated with each other. From the viewpoint of connection management, the management of OTUk is similar with OCh. LO ODU\text{j} and OCh/OTUk have client/server relationships.

For example, one LO ODU1 connection can be setup between Node A and Node C. This LO ODU1 connection is to be supported by OCh/OTU1 connections, which are to be set up between Node A and Node B and between Node B and Node C. LO ODU1 can be mapped into OTU1 at Node A, demapped from it in Node B, switched at Node B, and then mapped into the next OTU1 and demapped from this OTU1 at Node C.

Figure 6 - Connection of LO ODU\text{j} (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.

![Diagram](image)

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
Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolving G.709 OTN Networks
draft-ietf-ccamp-gmpls-ospf-g709v3-01

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 July 21, 2012.
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 ............................................... 26
10. Acknowledgements ........................................... 28
11. References ............................................... 28
  11.1. Normative References .................................. 28
  11.2. Informative References ................................ 29
Authors’ Addresses ............................................. 29
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 [WSON-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:

```
  ODU2  ODU0  ODUflex  ODU0
     \   \     \     /       /
      |   |     |   |     |       |
      ODU3 ODU2
     \   \   /
      \   /  /
       \ /  /
        ODU4
```

The ODU4 is the root of the muxing tree, ODU3 and ODU2 are containers directly multiplexed into the server and then ODU2, ODU0 are the leaves of the ODU3 branch, while ODUflex and ODU0 are the leaves of the ODU2 one. This means that on this traffic card it is possible to have the following multiplexing capabilities:

- ODU2->ODU3->ODU4
- ODU0->ODU3->ODU4
- ODUflex->ODU2->ODU4
- ODU0->ODU2->ODU4

3. TE-Link Representation

G.709 ODUk/OTUk Links are represented as TE-Links in GMPLS Traffic Engineering Topology for supporting ODUj layer switching. These TE-Links can be modeled in multiple ways. Some of the prominent
representations are captured below.

OTUk physical Link(s) can be modeled as a TE-Link(s). The TE-Link is termed as OTUk-TE-Link. The OTUk-TE-Link advertises ODUj switching capacity. The advertised capacity could include ODUk switching capacity. Figure-1 below provides an illustration of one hop OTUk TE-links.

![Figure 1: ODUk TE-Links](image.png)

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](image.png)

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:

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:

```
+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+
| 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 |
+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+
| Switching Cap | Encoding | Reserved | Switching Cap | Encoding | Reserved | Switching Cap | Encoding | Reserved | Switching Cap | Encoding | Reserved | Switching Cap | Encoding | Reserved |
+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+--------------------------------+
```

Maximum LSP Bandwidth

The MAX LSP bandwidth field is used according to [RFC4203]: i.e. 0 <= Max LSP Bandwidth <= ODUk/OTUk and intermediate values are those on...
the branch of OTN switching hierarchy supported by the interface.
E.g. in the OTU4 link it could be possible to have ODU4 as MAX LSP
Bandwidth for some priorities, ODU3 for others, ODU2 for some others
etc. The bandwidth unit MUST be in bytes per second and the encoding
MUST be in IEEE floating point format. The discrete values for
various ODU is shown in the table below.

<table>
<thead>
<tr>
<th>ODU Type</th>
<th>ODU nominal bit rate</th>
<th>Value in Byte/Sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU0</td>
<td>1 244 160 kbits/s</td>
<td>0x4D1450C0</td>
</tr>
<tr>
<td>ODU1</td>
<td>239/238 x 2 488 320 kbit/s</td>
<td>0x4D94F048</td>
</tr>
<tr>
<td>ODU2</td>
<td>239/237 x 9 953 280 kbit/s</td>
<td>0x4E959129</td>
</tr>
<tr>
<td>ODU3</td>
<td>239/236 x 39 813 120 kbit/s</td>
<td>0x4F963367</td>
</tr>
<tr>
<td>ODU4</td>
<td>239/227 x 99 532 800 kbit/s</td>
<td>0x504331E3</td>
</tr>
<tr>
<td>ODU2e</td>
<td>239/237 x 10 312 500 kbit/s</td>
<td>0x4E9AF70A</td>
</tr>
</tbody>
</table>

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:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type = 1 (Unres-fix)   |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                         Fixed Container                       |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
˜                               ...                             ˜
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type = 2 (Unres/MAX-var)  |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                      Variable Container                       |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: SCSI format

The format of the two different types of Bandwidth TLV are depicted in the following figures:
Figure 4: Bandwidth TLV – Type 1 –

The values of the fields shown in figure 4 are explained after figure 6.
Figure 5: Bandwidth TLV - Type 2 -

- Signal Type: Indicates the ODU type being advertised
<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ODU1</td>
</tr>
<tr>
<td>2</td>
<td>ODU2</td>
</tr>
<tr>
<td>3</td>
<td>ODU3</td>
</tr>
<tr>
<td>4</td>
<td>ODU4</td>
</tr>
<tr>
<td>10</td>
<td>ODU0</td>
</tr>
<tr>
<td>11</td>
<td>ODU2e</td>
</tr>
<tr>
<td>20</td>
<td>ODUflex CBR</td>
</tr>
<tr>
<td>21</td>
<td>ODUflex GFP-F resizable</td>
</tr>
<tr>
<td>22</td>
<td>ODUflex GFP-F non resizable</td>
</tr>
<tr>
<td>230-256</td>
<td>Experimental</td>
</tr>
</tbody>
</table>

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 fallen 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 priority (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:

\[
\text{Max LSP BW} = (# \text{ available TS}) \times \text{(ODTuk.ts nominal bit rate)} \times (1 - \text{HO OPUk bit rate tolerance})
\]

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 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 Cap | 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...
The next supported ODU[j] 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 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 at 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 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)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
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:
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:

```
Figure 9: Example 2 - TSG, T and S utilization
```
SCSI of ISCD 1 - TSG=2
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 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 LSPs 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type = 2 (Unres/MAX-var)   |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S. type=ODUflex|  #stages= 1   |T|S| TSG | Res |   Priority    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Stage#1=ODU3 |                   Padding                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 0              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 1              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 2              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 3              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 4              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 5              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 6              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Unreserved Bandwidth at priority 7              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 0              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 1              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 2              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 3              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 4              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 5              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 6              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAX LSP  Bandwidth at priority 7              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 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.

```
ODU1  ODU2  ODU3  ODUflex
    \   \   /    /    /
     \  /  /    /
      \ /  /
       ODU4
```

and the related SCSIs as follows:
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:

```
ODU2 ODU0   ODUflex ODU0
  \ /          \ /          \ /          \ /          \ /          \ /          \ /          \ /          \ /
     |          |            |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |       
```
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:

```
   +---------------------------------------------+
   | Type = 1 (Unres-fix)                       |
   +---------------------------------------------+
   | Sig type = ODU4 | #stages = 0 | T | S | TSG | Res | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
   | Unres ODU4 at Prio 0 = 2                   |
   +---------------------------------------------+
   | Type = 1 (Unres-fix)                       |
   +---------------------------------------------+
   | Sig type = ODU3 | #stages = 1 | T | S | TSG | Res | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
   | Stage#1 = ODU4 | Padding    |
   | Unres ODU3 at Prio 0 = 4                   |
   | Unres ODU3 at Prio 3 = 4                   |
   +---------------------------------------------+
   | Type = 1 (Unres-fix)                       |
   +---------------------------------------------+
   | Sig type = ODU2 | #stages = 2 | T | S | TSG | Res | 1 | 0 | 0 | 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)                       |
   +---------------------------------------------+
   | Sig type = ODU0 | #stages = 2 | T | S | TSG | Res | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
   | Stage#1 = ODU3 | Stage#2 = ODU4 | Padding |
   | Unres ODU0 at Prio 0 = 128                 |
   | Unres ODU0 at Prio 3 = 128                 |
```

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:

```
Component Link#1         Component Link#2
ODU2 ODU0                ODU1 ODU0
\ /                     \ /
|                        |
ODU3                    ODU2
|                        |
ODU4                    ODU4
```

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

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type = 1 (Unres-fix)</td>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig type=ODU4</td>
<td>#stages= 0</td>
<td>T</td>
<td>S</td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unres ODU4 at Prio 0 =1</td>
<td>Unres ODU4 at Prio 3 =1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type = 1 (Unres-fix)</td>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig type=ODU3</td>
<td>#stages= 1</td>
<td>T</td>
<td>S</td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage#1=ODU4</td>
<td>Padding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unres ODU3 at Prio 0 =2</td>
<td>Unres ODU3 at Prio 3 =2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type = 1 (Unres-fix)</td>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig type=ODU2</td>
<td>#stages= 2</td>
<td>T</td>
<td>S</td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage#1=ODU3</td>
<td>Stage#2=ODU4</td>
<td>Padding</td>
<td></td>
</tr>
</tbody>
</table>
```
### SCSI of ISCD 2 - Component Link#2

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Type = 1 (Unres-fix)</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODU4</td>
<td>#stages= 0</td>
</tr>
<tr>
<td>Unres ODU4 at Prio 0 =1</td>
<td>Unres ODU4 at Prio 3 =1</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Type = 1 (Unres-fix)</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODU2</td>
<td>#stages= 1</td>
</tr>
<tr>
<td>Stage#1=ODU4</td>
<td>Padding</td>
</tr>
<tr>
<td>Unres ODU2 at Prio 0 =10</td>
<td>Unres ODU2 at Prio 3 =10</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Type = 1 (Unres-fix)</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODU1</td>
<td>#stages= 2</td>
</tr>
<tr>
<td>Stage#1=ODU2</td>
<td>Stage#2=ODU4</td>
</tr>
<tr>
<td>Unres ODU1 at Prio 0 =40</td>
<td>Unres ODU1 at Prio 3 =40</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Type = 1 (Unres-fix)</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODU0</td>
<td>#stages= 2</td>
</tr>
<tr>
<td>Stage#1=ODU2</td>
<td>Stage#2=ODU4</td>
</tr>
<tr>
<td>Unres ODU0 at Prio 0 =80</td>
<td>Unres ODU0 at Prio 3 =80</td>
</tr>
</tbody>
</table>
Figure 15: Example 7 - Multi stage muxing - Non homogeneous hierarchies

6. Compatibility

In order to achieve backward compatibility with implementations based on [RFC4328] both the [RFC4328] based ISCD and the ISCD defined in this document MUST be advertised.

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

TBD

9. Contributors

Xiaobing Zi, Huawei Technologies
Email: zixiaobing@huawei.com

Francesco Fondelli, Ericsson
Email: francesco.fondelli@ericsson.com

Marco Corsi, Altran Italia
EMail: marco.corsi@altran.it

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 Kunjidhaphatham
akunjidhaphatham@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
10. Acknowledgements

The authors would like to thank Fred Gruman for the precious comments and suggestions.

11. References

11.1. Normative References


11.2. Informative References


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: kpitthewan@infinera.com

John E Drake  
Juniper  

Email: jdrake@juniper.net
Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for the evolving G.709 Optical Transport Networks Control

draft-ietf-ccamp-gmpls-signaling-g709v3-02.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/1id-abstracts.txt.

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

This Internet-Draft will expire on September 9, 2012.

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 ..... 5
   5.1. Usage of ODUflex(CBR) Traffic Parameters ................. 7
   5.2. Usage of ODUflex(GFP) Traffic Parameters ................. 9
6. Generalized Label ............................................. 9
   6.1. New definition of ODU Generalized Label ................. 10
   6.2. Examples ................................................ 12
   6.3. Label Distribution Procedure ............................ 14
       6.3.1. Notification on Label Error ....................... 15
   6.4. Supporting Virtual Concatenation and Multiplication .... 15
7. Supporting Multiplexing Hierarchy ............................ 16
   7.1. ADAPTATION Object .................................... 17
   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
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 superceded 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 (TS) granularity (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].

5. Extensions for Traffic Parameters for the Evolving G.709

The traffic parameters for G.709 are defined as follows:

```
+-----------------+-----------------+-----------------+-----------------+
|     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:

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not significant</td>
</tr>
<tr>
<td>1</td>
<td>ODU1 (i.e., 2.5 Gbps)</td>
</tr>
<tr>
<td>2</td>
<td>ODU2 (i.e., 10 Gbps)</td>
</tr>
<tr>
<td>3</td>
<td>ODU3 (i.e., 40 Gbps)</td>
</tr>
</tbody>
</table>
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:

\[ 2.5\text{Gbps} \times (1 \pm 100\text{ppm}) \]

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:

\[
N = \text{Ceiling of } \frac{\text{ODUflex(CBR) nominal bit rate} \times (1 + \text{ODUflex(CBR) bit rate tolerance})}{\text{ODTUk.ts nominal bit rate} \times (1 - \text{HO OPUk bit rate tolerance})}
\]

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]:

\[
\text{ODUflex(CBR) nominal bit rate} = \text{CBR client bit rate} \times \frac{239}{238} / T
\]

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>
<thead>
<tr>
<th>ODUk.ts</th>
<th>Minimum</th>
<th>Nominal</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU2.ts</td>
<td>1.249 384 632</td>
<td>1.249 409 620</td>
<td>1.249 434 608</td>
</tr>
<tr>
<td>ODU3.ts</td>
<td>1.254 678 635</td>
<td>1.254 703 729</td>
<td>1.254 728 823</td>
</tr>
<tr>
<td>ODU4.ts</td>
<td>1.301 683 217</td>
<td>1.301 709 251</td>
<td>1.301 735 285</td>
</tr>
</tbody>
</table>

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.

```
      +-----+             +---------+             +-----+
      |     |-------------| +-----| +-------------|     |
      |     |-------------|     |     |-------------|     |
      |     |-------------| +-----| +-------------|     |
      |     |             |         |             |     |
      |     |   .......   |         |   .......   |     |
      |  A  |-------------|    B    |-------------|  C  |
      +-----+   HO ODU4   +---------+   HO ODU2   +-----+

=========: TS occupied by ODUflex
----------: free TS
```

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 2.5Gbps * (1 + 100ppm), and the minimum bit rate of the tributary slot of ODU4
equals 1.301 683 217Gbps, so the total number of tributary slots 
N1 to be reserved on this link is:

\[ N1 = \text{ceiling} \left( \frac{2.5 \text{Gbps} \times (1 + 100\text{ppm})}{1.301 683 217\text{Gbps}} \right) = 2 \]

- On the HO ODU2 link between node B and C:

The maximum bit rate of the ODUflex equals 2.5Gbps \times (1 + 100ppm), 
and the minimum bit rate of the tributary slot of ODU2 equals 
1.249 384 632Gbps, so the total number of tributary slots N2 to 
be reserved on this link is:

\[ N2 = \text{ceiling} \left( \frac{2.5 \text{Gbps} \times (1 + 100\text{ppm})}{1.249 384 632\text{Gbps}} \right) = 3 \]

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>
<thead>
<tr>
<th>ODU type</th>
<th>Nominal bit-rate</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODUflex(GFP) of n TS, 1&lt;=n&lt;=8</td>
<td>n * ODU2.ts</td>
<td>+/-100 ppm</td>
</tr>
<tr>
<td>ODUflex(GFP) of n TS, 9&lt;=n&lt;=32</td>
<td>n * ODU3.ts</td>
<td>+/-100 ppm</td>
</tr>
<tr>
<td>ODUflex(GFP) of n TS, 33&lt;=n&lt;=80</td>
<td>n * ODU4.ts</td>
<td>+/-100 ppm</td>
</tr>
</tbody>
</table>

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         TPN           |   Reserved    |        Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
˜             Bit Map         .........                         ˜
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The ODU Generalized Label is used to indicate how the LO ODUj signal is multiplexed into the HO ODUk link. Note that the LO OUDj signal type is indicated by traffic parameters, while the type of HO ODUk 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 ODUj 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 ODUj 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 ODUj mapped into OTUk (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 ODUj is multiplexed into HO ODUk
occupying one or more TSs, a new TPN value is configured at the two ends of the HO ODUk link and is put into the related MSI byte(s) in the OPUk 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 ODUj 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)**

<table>
<thead>
<tr>
<th>HO ODUk</th>
<th>LO ODUj</th>
<th>TPN</th>
<th>TPN Assignment Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU2</td>
<td>ODU1</td>
<td>1^4</td>
<td>Fixed, = TS# occupied by ODU1</td>
</tr>
<tr>
<td></td>
<td>ODU1</td>
<td>1^16</td>
<td>Fixed, = TS# occupied by ODU1</td>
</tr>
<tr>
<td>ODU3</td>
<td>ODU2</td>
<td>1^4</td>
<td>Flexible, != other existing LO ODU2s’ TPNs</td>
</tr>
</tbody>
</table>

**Table 4 - TPN Assignment Rules (1.25Gbps TS granularity)**

<table>
<thead>
<tr>
<th>HO ODUk</th>
<th>LO ODUj</th>
<th>TPN</th>
<th>TPN Assignment Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU1</td>
<td>ODU0</td>
<td>1^2</td>
<td>Fixed, = TS# occupied by ODU0</td>
</tr>
<tr>
<td>ODU2</td>
<td>ODU1</td>
<td>1^4</td>
<td>Flexible, != other existing LO ODU1s’ TPNs</td>
</tr>
<tr>
<td>ODU0 &amp; ODUflex</td>
<td>1^8</td>
<td>Flexible, != other existing LO ODU0s and ODUflexes’ TPNs</td>
<td></td>
</tr>
<tr>
<td>ODU1</td>
<td>1^16</td>
<td>Flexible, != other existing LO ODU1s’ TPNs</td>
<td></td>
</tr>
<tr>
<td>ODU2</td>
<td>1^4</td>
<td>Flexible, != other existing LO ODU2s’ TPNs</td>
<td></td>
</tr>
<tr>
<td>ODU0 &amp; ODU2e &amp; ODUflex</td>
<td>1^32</td>
<td>Flexible, != other existing LO ODU0s and ODU2es and ODUflexes’ TPNs</td>
<td></td>
</tr>
<tr>
<td>ODU4</td>
<td>Any ODU</td>
<td>1^80</td>
<td>Flexible, != ANY other existing LO ODUs’ TPNs</td>
</tr>
</tbody>
</table>

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 filed 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.

(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 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       TPN = 2         |   Reserved    |     Length = 8        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 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 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 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 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.

```
|<------------------- ODU0 Connection -------------------> |
|<----- ODU2 Connection ----->

+----+         +----+         +----+         +----+         +----+
| N1 +---------+ N2 +---------+ N3 +---------+ N4 +---------+ N5 |
|+++++        +++++        +++++        +++++        +++++ |
| ODU3 link  ODU3 link  ODU3 link  ODU3 link  ODU3 link |
```

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. ADAPTATION Object

In order to create ODU FA-LSP (i.e., the server layer LSP) for carrying the client LSP, a new object called ADAPTATION Object is introduced, with two TLVs defined in this document:

- Type 1 = Server TSG signaling
- Type 2 = Hierarchy signaling

(1) Type=1 - Server TSG TLV

<table>
<thead>
<tr>
<th>Type = 1 (TSG)</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSG</td>
<td>Reserved</td>
</tr>
</tbody>
</table>

TSG: Tributary Slot Granularity (8bit): Used for signaling the server layer TSG:

- 0 - Reserved
- 1 - either 1.25Gbps or 2.5Gbps
- 2 - 2.5Gbps
- 3 - 1.25Gbps
- 4˜255 - Reserved

Where value 1 is used where the fallback procedure at the source end of FA is enabled and the default value of 1.25Gbps can be fallen back to 2.5Gbps. This means that either 1.25 Gbps or 2.5 Gbps can be used as the server TSG at the sink end of FA.

Values 2 and 3 are used to signal a 2.5Gbps or 1.25Gbps interfaces respectively and there is no chance to modify it.
Other values are reserved for future extension.

(2) Type=2 - Hierarchy TLV

+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| Type=2 (Hierarchy) |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| Length     |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| LSP Enc. Type | Switching Type | Signal Type | Mapping |
| ...          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| LSP Enc. Type | Switching Type | Signal Type | Mapping |
| ...          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |

A Hierarchy 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.

The meaning of the fields is defined as follow:

LSP Encoding Type and Switching Type: These fields can assume any value inherited from the Generalized Label Request Object in GMPLS signaling, defined in [RFC3471] and following related RFCs and drafts.

Signal Type: In the case of non OTN signal types, this field MUST be set to 0, while in the case of OTN signal types if MUST be filled accordingly to [RFC4328] and this document.

Mapping: This field indicates the mapping function used in each client-server relationship of the hierarchy. The values of this field are listed below:

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
</tr>
<tr>
<td>1</td>
<td>AMP</td>
</tr>
<tr>
<td>2</td>
<td>BMP</td>
</tr>
<tr>
<td>3</td>
<td>GMP</td>
</tr>
</tbody>
</table>

Zhang                      Expires September 2012
For example, in order to create ODU3 FA-LSP passing through a set of ODU4 links to perform ODU1->ODU2->ODU3 hierarchy, the Hierarchy TLV can be used to indicate the ODU2 into ODU3 multiplexing and ODU1 into ODU2 multiplexing stages.

```
+---------------------------------------------+
|    Type = 2 (Hierarchy)       |         Length = 8            |
+---------------------------------------------+
| Enc.=12(ODUk) | Switching=101 |  Sig. = ODU2  | Mapping = AMP |
+---------------------------------------------+
| Enc.=12(ODUk) | Switching=101 |  Sig. = ODU1  | Mapping = AMP |
+---------------------------------------------+
```

7.2. ODU FA-LSP Creation

When creating an ODU FA-LSP, the source node (e.g., node N2 in Figure 2) can include the ADAPTATION object to specify the desired hierarchy 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 has the ability to carry the requested ODU FA-LSP which can support the TS granularity and the multiplexing hierarchy listed in the ADAPTATION object at the remote end of the link (Note that such remote capability information can be obtained through LMP, routing protocol or configuration). Then the penultimate node uses the IF_ID RSVP_HOP Object to indicate the selected link for carrying the FA-LSP, as described in [RFC3473]. If no link supporting the specified hierarchy capabilities, a PathErr message with Error Code = 38 (LSP Hierarchy Issue) and Error Value = y1(new value) MUST be sent back to upstream.

Other intermediate nodes along the FA-LSP don’t need to process the ADAPTATION object, just forwarding it 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:

- When receiving Path message using [RFC4328] (i.e., Switching Type = 100), a new node SHOULD follow [RFC4328] to process and reply it.

- 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.

- 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]:

  - G.709 SENDER_TSPEC Object: Class = 12, C-Type = 5 [RFC4328]
  - 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.

- **ADAPTATION Object:**

  New object with Class-Num = xx, C-Type = xx. 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":

<table>
<thead>
<tr>
<th>Error Value</th>
<th>Error case</th>
</tr>
</thead>
<tbody>
<tr>
<td>y1</td>
<td>Last hop of an ODU FA-LSP doesn’t support specified adaption capabilities (Section 7.2).</td>
</tr>
</tbody>
</table>

- **Error Code = x2:**
New Error Code, indicating errors occurring when controlling a resizable ODUflex connection.

<table>
<thead>
<tr>
<th>Error Value</th>
<th>Error case</th>
</tr>
</thead>
<tbody>
<tr>
<td>y1</td>
<td>Do not support hitless assignment of ODUflex (GFP) (Section 8).</td>
</tr>
</tbody>
</table>

12. References

12.1. Normative References


12.2. Informative References


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: llyong@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
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.
RSVP-TE Extensions for Associated Bidirectional LSPs
draft-ietf-ccamp-mpls-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
1. Introduction

The MPLS TransportProfile (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.iertf-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:

- **Association Types**:

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>4  (TBD)</td>
<td>Doubled Sided Associated Bidirectional LSPs  (D)</td>
</tr>
<tr>
<td>5  (TBD)</td>
<td>Single Sided Associated Bidirectional LSPs   (A)</td>
</tr>
</tbody>
</table>

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 11bbbbbb has the following format:

Class = TBD (of the form 11bbbbbb), C_Type = 1 (TBD)

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.

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Double Sided Associated Bidirectional LSPs (D)</td>
</tr>
<tr>
<td>5</td>
<td>Single Sided Associated Bidirectional LSPs (A)</td>
</tr>
</tbody>
</table>

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]


[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching


8.2. Informative References

[I-D.ietf-ccamp-assoc-info]


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
Information model for G.709 Optical Transport Networks (OTN)
draft-ietf-ccamp-otn-g709-info-model-03

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 July 21, 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

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>1.1</td>
<td>Terminology</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>OSPF-TE requirements overview</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>RSVP-TE requirements overview</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>G.709 Digital Layer Info Model for Routing and Signaling</td>
<td>5</td>
</tr>
<tr>
<td>4.1</td>
<td>Tributary Slot Granularity</td>
<td>8</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Data Plane Considerations</td>
<td>8</td>
</tr>
<tr>
<td>4.1.1.1</td>
<td>Payload Type and TSG relationship</td>
<td>8</td>
</tr>
<tr>
<td>4.1.1.2</td>
<td>Fall-back procedure</td>
<td>10</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Control Plane considerations</td>
<td>10</td>
</tr>
<tr>
<td>4.2</td>
<td>Tributary Port Number</td>
<td>14</td>
</tr>
<tr>
<td>4.3</td>
<td>Signal type</td>
<td>14</td>
</tr>
<tr>
<td>4.4</td>
<td>Bit rate and tolerance</td>
<td>15</td>
</tr>
<tr>
<td>4.5</td>
<td>Unreserved Resources</td>
<td>16</td>
</tr>
<tr>
<td>4.6</td>
<td>Maximum LSP Bandwidth</td>
<td>16</td>
</tr>
<tr>
<td>4.7</td>
<td>Distinction between terminating and switching capability</td>
<td>16</td>
</tr>
<tr>
<td>4.8</td>
<td>Priority Support</td>
<td>19</td>
</tr>
<tr>
<td>4.9</td>
<td>Multi-stage multiplexing</td>
<td>19</td>
</tr>
<tr>
<td>4.10</td>
<td>Generalized Label</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Security Considerations</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>IANA Considerations</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>Contributors</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>Acknowledgements</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>References</td>
<td>21</td>
</tr>
<tr>
<td>9.1</td>
<td>Normative References</td>
<td>21</td>
</tr>
<tr>
<td>9.2</td>
<td>Informative References</td>
<td>22</td>
</tr>
<tr>
<td>Authors’ Addresses</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>
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.
<table>
<thead>
<tr>
<th>ODU client</th>
<th>OTU server</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU 0</td>
<td>-</td>
</tr>
<tr>
<td>ODU 1</td>
<td>OTU 1</td>
</tr>
<tr>
<td>ODU 2</td>
<td>OTU 2</td>
</tr>
<tr>
<td>ODU 2e</td>
<td>-</td>
</tr>
<tr>
<td>ODU 3</td>
<td>OTU 3</td>
</tr>
<tr>
<td>ODU 4</td>
<td>OTU 4</td>
</tr>
<tr>
<td>ODU flex</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 1: OTN mapping capability
<table>
<thead>
<tr>
<th>ODU client</th>
<th>ODU server</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25 Gbps client</td>
<td>ODU 0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5 Gbps client</td>
<td>ODU 1</td>
</tr>
<tr>
<td>ODU 0</td>
<td></td>
</tr>
<tr>
<td>10 Gbps client</td>
<td>ODU 2</td>
</tr>
<tr>
<td>ODU0,ODU1,ODUflex</td>
<td></td>
</tr>
<tr>
<td>10,3125 Gbps client</td>
<td>ODU 2e</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>40 Gbps client</td>
<td>ODU 3</td>
</tr>
<tr>
<td>ODU0,ODU1,ODU2,ODU2e,ODUflex</td>
<td></td>
</tr>
<tr>
<td>100 Gbps client</td>
<td>ODU 4</td>
</tr>
<tr>
<td>ODU0,ODU1,ODU2,ODU2e,ODU3,ODUflex</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>CBR clients from greater than</td>
<td></td>
</tr>
<tr>
<td>2.5 Gbit/s to 100 Gbit/s: or</td>
<td></td>
</tr>
<tr>
<td>GFP-F mapped packet clients from</td>
<td>ODUflex</td>
</tr>
<tr>
<td>1.25 Gbit/s to 100 Gbit/s.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: OTN multiplexing capability

How an ODUk connection service is transported within an operator network is governed by operator policy. For example, the ODUk connection service might be transported over an ODUk path over an OTUk 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 ODUk 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:

- Case A) one or more wavelength sub-networks connected by optical links or
- Case B) one or more ODU links (having sub-lambda and/or lambda bandwidth granularity)
- 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 types 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 consists 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.
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 ODU0 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.

```
+--------+      +--------+      +--------+      +--------+
|        |      |        |      |        | 1.25 |        |
|  Node  |      |  Node  |      |  Node  +------+  Node  |
|   A    +------+   B    +------+   C    | ODU3 |   D    |
|        | ODU3 |        | ODU3 |        +------+        |
+--------+ 1.25 +--------+ 2.5 +--------+ 2.5 +--------+
```

**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.

<table>
<thead>
<tr>
<th>ODUj</th>
<th>2.5G TS</th>
<th>OPU2</th>
<th>OPU3</th>
<th>1.25G TS</th>
<th>OPU1</th>
<th>OPU2</th>
<th>OPU3</th>
<th>OPU4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODU0</td>
<td>-</td>
<td>AMP</td>
<td>PT=20</td>
<td>GMP</td>
<td>PT=21</td>
<td>GMP</td>
<td>PT=21</td>
<td>GMP</td>
</tr>
<tr>
<td>ODU1</td>
<td>AMP</td>
<td>AMP</td>
<td>PT=20</td>
<td>-</td>
<td>AMP</td>
<td>AMP</td>
<td>GMP</td>
<td>AMP</td>
</tr>
<tr>
<td>ODU2</td>
<td>AMP</td>
<td>AMP</td>
<td>PT=20</td>
<td>AMP</td>
<td>GMP</td>
<td>AMP</td>
<td>GMP</td>
<td>GMP</td>
</tr>
<tr>
<td>ODU2e</td>
<td>AMP</td>
<td>AMP</td>
<td>-</td>
<td>AMP</td>
<td>GMP</td>
<td>AMP</td>
<td>GMP</td>
<td>GMP</td>
</tr>
<tr>
<td>ODU3</td>
<td>-</td>
<td>AMP</td>
<td>AMP</td>
<td>AMP</td>
<td>GMP</td>
<td>GMP</td>
<td>GMP</td>
<td>GMP</td>
</tr>
<tr>
<td>ODUfl</td>
<td>-</td>
<td>AMP</td>
<td>AMP</td>
<td>GMP</td>
<td>GMP</td>
<td>GMP</td>
<td>GMP</td>
<td>GMP</td>
</tr>
</tbody>
</table>

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 an 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 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.
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*4*8=992 bits) to be allocated while an ODUflex into an ODU4 may need up to eighty labels (80*4*8=2560 bits).

A new flexible and scalable ODUk label format needs to be defined.

5. Security Considerations

TBD

6. IANA Considerations

TBD

7. Contributors

Jonathan Sadler, Tellabs

EMail: jonathan.sadler@tellabs.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


9.2. Informative References


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
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/1id-abstracts.txt

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

This Internet-Draft will expire on September 6, 2012.

Copyright Notice

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

- Node Information
- Link Information
- Dynamic Node Information
- 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,

\[
\text{<Node\_Information>} ::= \text{<Node\_ID>} [\text{<ConnectivityMatrix>}...]
\]

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 input input ports and N the number of output output 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 input input port i can connect to output output 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 input input wavelength on a particular input input port to a desired output output wavelength on a particular output output 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, 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,...,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 may be 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.

![Diagram of resource pool model](image)

**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>
<ResourceOutputWaveConstraints>

<RBPoolState> ::= (<ResourceBlockID><NumResourcesInUse><InAvailableWavelengths><OutAvailableWavelengths>)...
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:

\(<\text{client-signal-list}>::=[<\text{GPID}>]...\)

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:

\(<\text{ProcessingCapabilities}>::=[<\text{NumResources}>] [<\text{RegenerationCapabilities}>] [<\text{FaultPerfMon}>] [<\text{VendorSpecific}>]\)

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.

\[
\text{LinkInfo} :: = \text{LinkID} \[\text{AdministrativeGroup}\] \\
[\text{InterfaceCapDesc}] [\text{Protection}] [\text{SRLG}]... \\
[\text{TrafficEngineeringMetric}] [\text{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, ROADM, 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 unicast 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


11.2. Informative References


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
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.
Signaling Extensions for Wavelength Switched Optical Networks
draft-ietf-ccamp-wson-signaling-03.txt

Abstract

This memo provides extensions to Generalized Multi-Protocol Label
Switching (GMPLS) signaling for control of wavelength switched
optical networks (WSON). Such extensions are necessary in WSONs
under a number of conditions including: (a) when optional
processing, such as regeneration, must be configured to occur at
specific nodes along a path, (b) where equipment must be configured
to accept an optical signal with specific attributes, or (c) where
equipment must be configured to output an optical signal with
specific attributes. In addition this memo provides mechanisms to
support distributed wavelength assignment with bidirectional LSPs,
and choice in distributed wavelength assignment algorithms. These
extensions build on previous work for the control of lambda and
G.709 based networks.

Status of this Memo

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

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

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

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

Expires September 7, 2012
1. Introduction

This memo provides extensions to Generalized Multi-Protocol Label Switching (GMPLS) signaling for control of wavelength switched optical networks (WSON). Fundamental extensions are given to permit simultaneous bi-directional wavelength assignment while more advanced extensions are given to support the networks described in [RFC6163] which feature connections requiring configuration of input, output, and general signal processing capabilities at a node along a LSP.

These extensions build on previous work for the control of lambda and G.709 based networks.

2. Terminology

CWDM: Coarse Wavelength Division Multiplexing.

DWDM: Dense Wavelength Division Multiplexing.

FOADM: Fixed Optical Add/Drop Multiplexer.

ROADM: Reconfigurable Optical Add/Drop Multiplexer. A reduced port count wavelength selective switching element featuring ingress and egress line side ports as well as add/drop side ports.

RWA: Routing and Wavelength Assignment.

Wavelength Conversion/Converters: The process of converting an information bearing optical signal centered at a given wavelength to one with "equivalent" content centered at a different wavelength. Wavelength conversion can be implemented via an optical-electronic-optical (OEO) process or via a strictly optical process.

WDM: Wavelength Division Multiplexing.
Wavelength Switched Optical Networks (WSON): WDM based optical networks in which switching is performed selectively based on the center wavelength of an optical signal.

AWG: Arrayed Waveguide Grating.

OXC: Optical Cross Connect.

Optical Transmitter: A device that has both a laser tuned on certain wavelength and electronic components, which converts electronic signals into optical signals.

Optical Responder: A device that has both optical and electronic components. It detects optical signals and converts optical signals into electronic signals.

Optical Transponder: A device that has both an optical transmitter and an optical responder.

Optical End Node: The end of a wavelength (optical lambdas) lightpath in the data plane. It may be equipped with some optical/electronic devices such as wavelength multiplexers/demultiplexer (e.g. AWG), optical transponder, etc., which are employed to transmit/terminate the optical signals for data transmission.

3. Requirements for WSON Signaling

The following requirements for GMPLS based WSON signaling are in addition to the functionality already provided by existing GMPLS signaling mechanisms.

3.1. WSON Signal Characterization

WSON signaling MUST convey sufficient information characterizing the signal to allow systems along the path to determine compatibility and perform any required local configuration. Examples of such systems include intermediate nodes (ROADMs, OXCs, Wavelength converters, Regenerators, OEO Switches, etc...), links (WDM systems) and end systems (detectors, demodulators, etc...). The details of any local configuration processes are out of the scope of this document.

From [RFC6163] we have the following list of WSON signal characteristic information:
List 1. WSON Signal Characteristics

1. Optical tributary signal class (modulation format).
2. FEC: whether forward error correction is used in the digital stream and what type of error correcting code is used.
3. Center frequency (wavelength)
4. Bit rate
5. G-PID: General Protocol Identifier for the information format.

The first three items on this list can change as a WSON signal traverses a network with regenerators, OEO switches, or wavelength converters. An ability to control wavelength conversion already exists in GMPLS signaling along with the ability to share client signal type information (G-PID). In addition, bit rate is a standard GMPLS signaling traffic parameter. It is referred to as Bandwidth Encoding in [RFC3471]. This leaves two new parameters: modulation format and FEC type, needed to fully characterize the optical signal.

3.2. Per LSP Network Element Processing Configuration

In addition to configuring a network element (NE) along an LSP to input or output a signal with specific attributes, we may need to signal the NE to perform specific processing, such as 3R regeneration, on the signal at a particular NE. In [RFC6163] we discussed three types of processing not currently covered by GMPLS:

(A) Regeneration (possibly different types)

(B) Fault and Performance Monitoring

(C) Attribute Conversion

The extensions here MUST provide for the configuration of these types of processing at nodes along an LSP.

3.3. Bi-Directional WSON LSPs

WSON signaling MAY support LSP setup consistent with the wavelength continuity constraint for bi-directional connections. The following cases MAY be separately supported:
(a) Where the same wavelength is used for both upstream and downstream directions

(b) Where different wavelengths can be used for both upstream and downstream directions.

(Editor’s Note: an evaluation of current GMPLS bidirectional solutions should be evaluated if they would fit to the current WSON needs.)

3.4. Distributed Wavelength Assignment Support

WSON signaling MAY support the selection of a specific distributed wavelength assignment method.

This method is beneficial in cases of equipment failure, etc., where fast provisioning used in quick recovery is critical to protect carriers/users against system loss. This requires efficient signaling which supports distributed wavelength assignment, in particular when the centralized wavelength assignment capability is not available.

As discussed in the [RFC6163] different computational approaches for wavelength assignment are available. One method is the use of distributed wavelength assignment. This feature would allow the specification of a particular approach when more than one is implemented in the systems along the path.

3.5. Out of Scope

This draft does not address signaling information related to optical impairments.

4. WSON Signal Traffic Parameters, Attributes and Processing

As discussed in [RFC6163] single channel optical signals used in WSONs are called "optical tributary signals" and come in a number of classes characterized by modulation format and bit rate. Although WSONs are fairly transparent to the signals they carry, to ensure compatibility amongst various networks devices and end systems it can be important to include key lightpath characteristics as traffic parameters in signaling [RFC6163].

4.1. Traffic Parameters for Optical Tributary Signals

In [RFC3471] we see that the G-PID (client signal type) and bit rate (byte rate) of the signals are defined as parameters and in
4.2. Signal Attributes and Processing

Section 3.2. gave the requirements for signaling to indicate to a particular NE along an LSP what type of processing to perform on an optical signal or how to configure that NE to accept or transmit an optical signal with particular attributes.

One way of accomplishing this is via a new EXPLICIT_ROUTE subobject. Reference [RFC3209] defines the EXPLICIT_ROUTE object (ERO) and a number of subobjects, while reference [RFC5420] defines general mechanisms for dealing with additional LSP attributes. Although reference [RFC5420] defines a RECORD_ROUTE object (RRO) attributes subobject, it does not define an ERO subobject for LSP attributes.

Regardless of the exact coding for the ERO subobject conveying the input, output, or processing instructions. This new "processing" subobject would follow a subobject containing the IP address, or the interface identifier [RFC3477], associated with the link on which it is to be used along with any label subobjects [RFC3473].

The WSON Signal Processing object is defined as an LSP_ATTRIBUTES and extends the PATH message. It is defined as the following:

\[
\text{<WSON Processing>} ::= \text{<hop information>} \text{<Transmitter Capabilities>} \text{<Receiver Capabilities>} [\text{<RegenerationCapabilities}>]
\]

\[
\text{<Receiver Capabilities>} ::= \text{<ModulationTypeList>} \text{<FECTypeList>} \text{<BitRateRange>}
\]

\[
\text{<Transmitter Capabilities>} ::= \text{<ModulationTypeList>} \text{<FECTypeList>} \text{<BitRateRange>}
\]

Where:

\[
\text{<hop information>: Ipv4,Ipv6 address. Note: this not required if WSON Processing object become part of the ERO}
\]

\[
\text{<Transmitter Capabilities> is defined in [WSON-Encode]}
\]

\[
\text{<ReceiverCapabilities> is defined in [WSON-Encode]}
\]
<ModulationTypeList> is defined in [WSON-Encode]

<FECTypeList> is defined in [WSON-Encode]

<BitRateRange> is defined in [WSON-Encode]

<RegenerationCapabilities> is defined in [WSON-Encode]

<RegenerationCapabilities> are only applied in the intermediate nodes of the LSP. The head and tail nodes will ignore regeneration capability processing.

4.2.1. WSON Processing Object Encoding

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type               |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
˜                            Value                              ˜
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: to be defined by IANA

Value: sub-TLVS according to section 4.1.

5. Bidirectional Lightpath Setup

With the wavelength continuity constraint in CI-incapable [RFC3471] WSONs, where the nodes in the networks cannot support wavelength conversion, the same wavelength on each link along a unidirectional lightpath should be reserved. In addition to the wavelength
continuity constraint, requirement 3.2 gives us another constraint on wavelength usage in data plane, in particular, it requires the same wavelength to be used in both directions. [RFC6163] in section 6.1 reports on the implication to GMPLS signaling related to both bi-directionality and Distributed Wavelengths Assignment.

Current GMPLS solution defines a bidirectional LSP (as defined by [RFC3471]). The label distribution is based on Label_Set and Upstream_Label objects. In case of specific constraints such as the same wavelengths in both directions, it may require several signaling attempts using information from the Acceptable_Label_Set received from path error messages.

Some implementations may prefer using two unidirectional LSPs. This solution has been always available as per [RFC3209] however recent work introduces the association concept [RFC4872] and [ASSOC-Info]. Recent transport evolutions [ASSOC-ext] provide a way to associate two unidirectional LSPs as a bidirectional LSP. In line with this, a small extension can make this approach work for the WSON case.

6. RWA Related

6.1. Wavelength Assignment Method Selection

Routing + Distributed wavelength assignment (R+DWA) is one of the options defined by the [RFC6163]. The output from the routing function will be a path but the wavelength will be selected on a hop-by-hop basis.

Under this hypothesis the node initiating the signaling process needs to declare its own wavelength availability (through a label_set object). Each intermediate node may delete some labels due to connectivity constraints or its own assignment policy. At the end, the destination node has to make the final decision on the wavelength assignment among the ones received through the signaling process.

As discussed in [HZang00] a number of different wavelength assignment algorithms maybe employed. In addition as discussed in
[RFC6163] the wavelength assignment can be either for a unidirectional lightpath or for a bidirectional lightpath constrained to use the same lambda in both directions.

A simple TLV could be used to indication wavelength assignment directionality and wavelength assignment method. This would be placed in an LSP_REQUIRED_ATTRIBUTES object per [RFC5420]. The use of a TLV in the LSP required attributes object was pointed out in [Xu].

[TO DO: The directionality stuff needs to be reconciled with the earlier material]

Unique Wavelength: 0 same wavelength in both directions, 1 may use different wavelengths [TBD: shall we use only 1 bit]

Wavelength Assignment Method: 0 unspecified (any), 1 First-Fit, 2 Random, 3 Least-Loaded (multi-fiber). Others TBD.

```
+------------+------------+----------+
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Unique WL  |    WA Method  |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

7. Security Considerations

This document has no requirement for a change to the security models within GMPLS and associated protocols. That is the OSPF-TE, RSVP-TE, and PCEP security models could be operated unchanged.

However satisfying the requirements for RWA using the existing protocols may significantly affect the loading of those protocols. This makes the operation of the network more vulnerable to denial of service attacks. Therefore additional care maybe required to ensure that the protocols are secure in the WSON environment.

Furthermore the additional information distributed in order to address the RWA problem represents a disclosure of network capabilities that an operator may wish to keep private. Consideration should be given to securing this information.
8. IANA Considerations

TBD. Once finalized in our approach we will need identifiers for such things and modulation types, modulation parameters, wavelength assignment methods, etc...

9. Acknowledgments

Anyone who provide comments and helpful inputs
10. References

10.1. Normative References


10.2. Informative References


Author’s Addresses

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

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

Nicola Andriolli
Scuola Superiore Sant’Anna, Pisa, Italy
Email: nick@sssup.it

Alessio Giorgetti
Scuola Superiore Sant’Anna, Pisa, Italy
Email: a.giorgetti@sssup.it

Lin Guo
Key Laboratory of Optical Communication and Lightwave Technologies
Ministry of Education
P.O. Box 128, Beijing University of Posts and Telecommunications,
P.R.China
Email: guolintom@gmail.com

Hiroaki Harai
National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi, Koganei,
Tokyo, 184-8795 Japan

Phone: +81 42-327-5418
Email: harai@nict.go.jp

Yuefeng Ji
Key Laboratory of Optical Communication and Lightwave Technologies
Ministry of Education
P.O. Box 128, Beijing University of Posts and Telecommunications,
P.R.China
Email: jyf@bupt.edu.cn

Daniel King
Old Dog Consulting

Email: daniel@olddog.co.uk

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

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

Sugang Xu
National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi, Koganei,
Tokyo, 184-8795 Japan

Phone: +81 42-327-6927
Email: xsg@nict.go.jp

Giovanni Martinelli
Cisco
Via Philips 12
20052 Monza, IT

Phone: +39 039-209-2044
Email: giomarti@cisco.com

Intellectual Property Statement

The IETF Trust takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in any IETF Document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights.

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

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary
A framework for Management and Control of G.698.2 optical interface parameters
draft-kunze-g-698-2-management-control-framework-02

Abstract

This document provides a framework that describes a solution space for the control and management of optical interfaces parameters according to the Black Link approach as specified by ITU-T [ITU G.698.2]. In particular, it examines topological elements and related network management processes to operate this construct. This framework is scoped to address the Optical Channel (OCh)-layer covered by G.698.2. The focus is on enabling the wavelength provisioning process in a black link approach irrespective on how it is triggered i.e. by EMS, NMS or GMPLS. This document covers management as well as control plane considerations in different management cases of a single channel DWDM interface as defined by ITU-G.698.2. The purpose is to identify the necessary information elements and processes to be used by control or management devices for further processing. Hence wavelength routing and selection processes as defined e.g. in WSON are beyond the scope of this document.

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 10, 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
    1.1.  Requirements Language ............................... 4
2.  Terminology and Definitions ............................. 5
3.  Solution Space for optical interfaces using a DWDM Black
    Link ...................................................... 6
    3.1.  Comparison of approaches for transverse compatibility .. 7
        3.1.1.  Multivendor DWDM line system with transponders .. 7
        3.1.2.  Black Link Deployments ........................... 9
4.  Operational aspects using IUT-T G.698.2 specified single
    channel DWDM interfaces ............................... 10
    4.1.  Bringing into service .............................. 10
    4.2.  Configuration Management ............................ 11
    4.3.  In service (performance management) .................. 11
    4.4.  Fault Clearance ..................................... 11
5.  Solutions for managing and controlling the optical
    interface within Black Link scenarios ................. 11
    5.1.  BL Separate Operation and Management Approaches ....... 12
        5.1.1.  Direct connection to the management system ...... 13
        5.1.2.  Indirect connection to the DWDM management system .. 15
    5.2.  Control Plane Considerations ....................... 16
        5.2.1.  Considerations using GMPLS UNI .................. 17
6.  Requirements for BL deployments ........................ 18
    6.1.  Interoperability Aspects ........................... 18
7.  Acknowledgements ...................................... 19
8.  IANA Considerations ................................... 20
9.  Security Considerations ................................ 20
10. Contributors ........................................... 20
11. References ............................................ 21
    11.1.  Normative References .............................. 21
11.2. Informative References . . . . . . . . . . . . . . . . . . 22
1. Introduction

The usage of the Black Link approach in carrier applications (which include optical amplifiers) adds further networking option for operators enabling integration of G.698.2 optical interfaces into routers and other types of client devices.

Carriers deploy their networks today as a combination of transport and packet infrastructures to ensure high availability and flexible data transport. Both network technologies are usually managed by different operational units using different management concepts. This is the status quo in many carrier networks today. In the case of a black link deployment, where the optical transport interface moves into the client device (e.g., router), it is necessary to coordinate the management of the optical interface at the client domain with the optical transport domain. There are different levels of coordination, which are specified in this framework.

The objective of this document is to provide a framework that describes the solution space for the control and management of single channel interfaces as specified by the ITU-T Recommendation G.698.2 [ITU G.698.2]. In particular, it examines topological elements and related network management measures. From an architectural point of view, the black link is a set of pre-configured/qualified unidirectional, single-fiber, network connections between the G.698.2 reference points S and R. The optical transport network is managed and controlled in order to provide black links of the intended centre frequencies and the optical interfaces are managed and controlled to generate signals of the intended centre frequencies and further parameters as specified in ITU-T Recommendations G.698.2 and G.798.

Optical Routing and Wavelength assignment based on WSON is out of scope.

Furthermore, support for Fast Fault Detection, to e.g., trigger ODUk Protection Switching is out of scope of this work. Additionally, the wavelength ordering process and the process how to determine the demand for a new wavelength from A to Z is out of scope.

Note that the Control and Management Planes are two separate entities that are handling the same information in different ways. This document covers management as well as control plane considerations in different management cases of single channel DWDM interfaces.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
2. Terminology and Definitions

Current generation WDM networks are single vendor networks where the optical line system and the transponders are tightly integrated. The Black Link approach changes this situation by introducing a standardized interface at the level of OCh between the line system and transponders.

Black Link: The Black Link [ITU G.698.2] allows supporting an optical transmitter/receiver pair of a single vendor or from different vendors to provide a single optical channel interface and transport it over an optical network composed of amplifiers, filters, add-drop multiplexers which may be from a different vendor. Therefore the standard defines the ingress and egress parameters for the optical interfaces at the reference points Ss and Rs. In that case the optical connection between the two G.968.2 optical interfaces is referred to as a Black Link. G.698.2 provides an optical interface specification ensuring the realization of transversely compatible dense wavelength division multiplexing (DWDM) systems primarily intended for metro applications which include optical amplifiers and leads towards a multivendor DWDM optical transmission network.

Single Channel DWDM Interface: The single channel interfaces to DWDM systems defined in G.698.2, which currently include the following features: channel frequency spacing: 50 GHz and wider (defined in [ITU-T G.694.1]); bit rate of single channel: Up to 10 Gbit/s. Future revisions are expected to include application codes for bit rates up to 40 Gb/s. Single channel DWDM interfaces to/from other vendor(s): G.698.2 provides transverse compatibility at the single-channel point, using a direct wavelength-multiplexing configuration, for single channel DWDM interfaces to/from other vendors (but not at the multi-channel point).

Forward error correction (FEC): FEC is a way of improving the performance of high-capacity optical transmission systems. Employing FEC in optical transmission systems yields system designs that can accept relatively large BER (much more than 10-12) in the optical transmission line (before decoding).

Administrative domain [G.805]: For the purposes of this Recommendation an administrative domain represents the extent of resources which belong to a single player such as a network operator, a service provider or an end-user. Administrative domains of different players do not overlap amongst themselves.

Intra-domain interface (IaDI) [G.872]: A physical interface within an
administrative domain.

Inter-domain interface (IrDI) [G.872]: A physical interface that represents the boundary between two administrative domains.

Management Plane [G.8081]: The management plane performs management functions for the transport plane, the control plane and the system as a whole. It also provides coordination between all the planes. The following management functional areas are performed in the management plane: performance management; fault management; configuration management; accounting management and security management.

Control Plane [G.8081]: The control plane performs neighbour discovery, call control and connection control functions. Through signalling, the control plane sets up and releases connections, and may restore a connection in case of a failure. The control plane also performs other functions in support of call and connection control, such as neighbour discovery and routing information dissemination.

Transponder: A Transponder is a network element that performs O/E/O (Optical/Electrical/Optical) conversion. In this document it is referred only transponders with 3R (rather than 2R or 1R regeneration) as defined in [ITU.G.872].

3. Solution Space for optical interfaces using a DWDM Black Link

The management of optical interfaces using the Black Link approach deals with aspects related to the management of single-channel optical interface parameters of physical point-to-point and ring DWDM applications on single-mode optical fibres.

The Black Link approach allows the direct connection of a wide variety of equipments using a DWDM link, for example:

a. A digital cross-connect with multiple optical interfaces, supplied by a different vendor from the line system

b. Multiple optical client devices, each from a different vendor, supplying one channel each

c. A combination of the above

Table 1 provides a list of BL management tasks regarding the configuration of optical parameters.
<table>
<thead>
<tr>
<th>Task</th>
<th>Domain</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>determination of centre frequency</td>
<td>optical</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>configuration of centre frequency at</td>
<td>client</td>
<td>NR</td>
<td>NR</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>optical IF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>path computation of wavelength</td>
<td>optical</td>
<td>NR</td>
<td>NR</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>routing of wavelength</td>
<td>optical</td>
<td>NR</td>
<td>NR</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>wavelength setup across optical network</td>
<td>optical</td>
<td>?</td>
<td>?</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>detection of wavelength fault</td>
<td>client</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>fault isolation, identification of root failure</td>
<td>optical</td>
<td>NR</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>repair actions within optical network</td>
<td>optical</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>protection switching of wavelength</td>
<td>optical</td>
<td>NR</td>
<td>NR</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>restoration of wavelength</td>
<td>optical</td>
<td>NR</td>
<td>NR</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

Note: R = relevant, NR = not relevant

Table 1: List of tasks related to BL management

Furthermore the following deployment cases will be considered:

a. Passive WDM
b. P2P WDM systems
c. WDM systems with OADMs
d. Transparent optical networks supporting specific IPoWDM functions, interfaces, protocols etc.

Case a) is added for illustration only, since passive WDM is specified in ITU-T Recommendations G.695 and G.698.1.

Case b) and case c) are motivated by the usage of legacy equipment using the traditional connection as described in Figure 1 combined with the BL approach.

3.1. Comparison of approaches for transverse compatibility

3.1.1. Multivendor DWDM line system with transponders

As illustrated in Figure 1, for this approach interoperability is achieved via the use of optical transponders providing OEO (allowing conversion to appropriate parameters). The optical interfaces labelled "single channel non-DWDM interfaces from other vendor(s)"
and "Single channel non DWDM interfaces to/from other vendor(s)" can then be any short reach standardized optical interface that both vendors support, such as those found in [ITU-T G.957] [ITU-T G.691], [ITU-T G.693], [ITU-T G.959.1], etc.

![Diagram of Inter and Intra-Domain Interface Identification](image)

**Figure 1: Inter and Intra-Domain Interface Identification**

In the scenario of Figure 1 the administrative domain is defined by the Interdomain Interface (IrDI). This interface terminates the DWDM domain. The line side is characterized by the IaDI. This interface specifies the internal parameter set of the optical administrative domain. In the case of black link deployment this interface moves into the client devices and extends the optical and administrative domain towards the client node. ITU-T G.698.2 specifies the parameter set for a certain set of applications.

This document elaborates only the IaDI Interface as specified in ITU-T G.698.2 as transversely compatible and multi-vendor interface within one administrative domain controlled by the network operator. This administrative domain can contain several vendor domains (vendor A for the DWDM sub-network, and vendors B1 and B2 at the transmitter and receiver terminal side).
3.1.2. Black Link Deployments

In case of a Black Link deployment as shown in Figure 2, through the use of the single channel DWDM interfaces defined in [ITU G.698.2], multi-vendor interconnection can also be achieved while removing the need for one short reach transmitter and receiver pair per channel (eliminating the transponders).

Figure 2 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 optical multiplexer (OM) and an optical demultiplexer (OD) (which are used as a pair with the peer 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

Linear black link as per ITU-T G.698.2

Figure 2: Linear Black Link
In Figure 2, if the administrative domain consists of several domains (e.g. A for a DWDM network supporting the Black Link, B1 for the DWDM Tx, and B2 for the DWDM Rx), it is typical that there will be a separate Element Management Systems (EMS) will be used for each vendor domain (e.g. EMS-a for domain A, EMS-b1 for domain B1, and EMS-b2 for domain B2). Each EMS may have a common standard north bound management interface to a Network Management System (NMS), allowing consistent end-to-end management of the connection.

To facilitate consistent end-to-end network management, the north bound management interface from the EMS to the NMS should be consistent (from a management information point of view) with the standard protocol-neutral (or protocol-specific) information model used in the EMS south bound management interface to its subtending NEs (TX and/or RX). The [Black-Link-MIB] defines such a protocol-specific information using SNMP/SMI.

4. Operational aspects using IUT-T G.698.2 specified single channel DWDM interfaces

A Comparison of the Black Link with the traditional operation scenarios provides an insight of similarities and distinctions in operation and management. The following four use cases provide an overview about operation and maintenance processes.

4.1. Bringing into service

It is necessary to differentiate between two operational issues for setting up a light path (a Black Link connection is specific in having defined maximum impairments) within an operational network. The first step is the preparation of the connection if no optical signal is applied. Therefore it is necessary to define the path of the connection.

The second step is to setup the Black Link connection. This is done using the NMS of the optical transport network. From the operation point of view the task is similar in a Black Link scenario and in a traditional WDM environment. The Black Link connection is measured by using BER tester which use optical interfaces according to G.698.2. These measurements are carried out in accordance with ITU-T Recommendation M.xxxx. When needed further Black Link connections for resilience are brought into service in the same way.

If the optical interface moves into a client device some of changes from the operational point of view have to be considered. The centre frequency of the Black Link connections was determined by the setup process. The optical interfaces at both terminals are set to the centre frequency of the Black Link connection before interconnected...
with the dedicated ports of the WDM network. Optical monitoring is activated in the WDM network after the terminals are interconnected with the dedicated ports in order to monitor the status of the Black Link connection. The monitor functions of the optical interfaces at the terminals are also activated in order to monitor the end to end connection.

Furthermore it should be possible to automate this last step. After connecting the client device towards the first control plane managed transport node a control connection may e.g. be automatically established using LMP to exchange configuration information.

If tunable interfaces are used in the Black Link scenario it would be possible to define a series of backup wavelength routes for restoration that could be tested and stored in backup profile. In fault cases this wavelength routes can be used to recover the service.

4.2. Configuration Management

tbd.

4.3. In service (performance management)

tbd.

4.4. Fault Clearance

tbd.

5. Solutions for managing and controlling the optical interface within Black Link scenarios

Operation and management of WDM systems is traditionally seen as a homogenous group of tasks that could be carried out best when a single management system or an umbrella management system is used. Currently each WDM vendor provides an Element Management System (EMS) that also administers the wavelengths.

Therefore from the operational point of view in a pure Black Link or in a mixed setup with transponders there are the following approaches will be considered to manage and operate optical interfaces.

1. Separate operation and management of client device and the transport network

   a. Direct link between the client device and the management system of the optical network (e.g. EMS, NMS)
b. Indirect link to the management system of the optical network using a protocol between the client device and the directly connected WDM system node to exchange management information with the optical domain

2. Common operation and management of client device and the Transport network

The first option keeps the status quo in large carrier networks as mentioned above. In that case it must be ensured that the full FCAPS Management (Fault, Configuration, Accounting, Performance and Security) capabilities are supported. This means from the management staff point of view nothing changes. The transceiver/receiver optical interface will be part of the optical management domain and will be managed from the transport management staff.

The second solution addresses the case where underlying WDM transport network is mainly used to interconnect a homogeneous set of client nodes (e.g. IP routers or digital crossconnects). Since the service creation and restoration could be done by to higher layers (e.g. IP), this may lead to more efficient network operation and a higher level of integration.

5.1. BL Separate Operation and Management Approaches
5.1.1. Direct connection to the management system

As depicted in Figure 3 (case 1a) one possibility to manage the optical interface within the client domain is a direct connection to the management system of the optical domain. This ensures manageability as usual.

---

**Figure 3: Connecting G.698.2 optical interfaces to the Transport Management system**

The exchange of management information between client device and the management system assumes that some form of a direct management...
A communication link exists between the client device and the DWDM management system (e.g. EMS). This may be an Ethernet Link or a DCN connection (management communication channel MCC).

It must be ensured that the optical network interface can be managed in a standardized way to enable interoperable solutions between different optical interface vendors and vendors of the optical network management application. RFC 3591 [RFC3591] defines managed objects for the optical interface type but needs further extension to cover the optical parameters required by this framework document. Therefore an extension to this MIB for the optical interface has been drafted in [Black-Link-MIB]. In that case SNMP is used to exchange data between the client device and the management system of the WDM domain.

Note that a software update of the optical interface components of the client nodes must not lead obligatory to an update of the software of the EMS and vice versa.
5.1.2. Indirect connection to the DWDM management system

The alternative as shown in Figure 4 can be used in cases where a more integrated relationship between transport node (e.g. OM or OD) and client device is aspired. In that case a combination of control plane features and manual management will be used.

![Diagram](image)

CL = Client Device
/C = G.698.2 optical Interface
OM = Optical Mux
OD = Optical Demux
EMS = Element Management System
MI = Management Interface

Figure 4: Direct connection between peer node and first optical network node

For information exchange between the client node and the direct connected node of the optical transport network LMP as specified in RFC 4209 [RFC4209] can (should) be used. This extension of LMP may
be used between a peer node and an adjacent optical network node as depicted in Figure 4.

Recently LMP based on RFC 4209 does not yet support the transmission of configuration data (information). This functionality has to be added to the existing extensions of the protocol. The use of LMP-WDM assumes that some form of a control channel exists between the client node and the WDM equipment. This may be a dedicated lambda, an Ethernet Link, or other signaling communication channel (SCC).

5.2. Control Plane Considerations

The concept of black link equally applies to management and control plane mechanisms. The general GMPLS control Plane for wavelength switched optical networks is work under definition in the scope of WSON. One important aspect of the BL is the fact that it includes the wavelength that is supported by the given link. Thus a BL can logically be considered as a fiber that is transparent only for a single wavelength. In other words, the wavelength becomes a characteristic of the link itself. Nevertheless the procedure to light up the fiber may vary depending on the BL implementation. Since the implementation of the BL itself is unknown a priori, different sequences to light up wavelength need to be considered:

1. Transponders first, transponder tuning: The transmitter is switched on and the BL is immediately transparent to its wavelength. This requires the transmitter to carefully tune power and frequency not overload the line system or to create transients.

2. Transponder first, OLS tuning: The transmitter is switched on first and can immediately go to the max power allowed since the OLS performs the power tuning. This leads to an intermediate state where the receiver doesn't receive a valid signal while the transmitter is sending out one. Alarm suppression mechanisms shall be employed to overcome that condition.

3. OLS first, Transponder tuning: At first the OLS is tuned to be transparent for a given wavelength, then transponders need to be tuned up. Since the OLS in general requires the presence of a wavelength to fine-tune it is internal facilities there may be a period of time where a valid signal is transmitted but the receiver is unable to detect it. This equally need to be covered by alarm suppression mechanisms.

4. OLS first, OLS tuning: The OLS is programmed to be transparent for a given wavelength, then the transponders need to be switched on and further power tuning takes place. The sequencing of
enabling the link needs to be covered as well.

The preferred way to address these in a Control Plane enabled network is neighbour discovery including exchange of link characteristics and link property correlation. The general mechanisms are covered in RFC4209 [LMP-WDM] and RFC 4204[LMP] which provides the necessary protocol framework to exchange those characteristics between client and black link. LMP-WDM is not intended for exchanging routing or signalling information but covers:

- Control channel management
- Link property correlation
- Link verification
- Fault Manegement

Extensions to LMP/LMP-WDM covering the code points of the BL definition are needed. Additionally when client and server side are managed by different operational entities, Link state exchange is required to align the management systems.

5.2.1. Considerations using GMPLS UNI

The deployment of G.698.2 optical interfaces is leading to some functional changes related to the control plane models and has therefore some impact on the existing interfaces especially in the case of an overlay model where the edge node requests resources from the core node and the edges node do not participate in the routing protocol instance that runs among the core nodes. RFC 4208 [RFC4208] defines the GMPLS UNI that will be used between edge and core node. In case of a black link deployment additional functionalities are needed to setup a connection.

It is necessary to differentiate between topology/signalling information and configuration parameters that are needed to setup a wavelength path. RSVP-TE could be used for the signalling and the reservation of the wavelength path. But there are additional information needed before RSVP-TE can start the signalling process. There are three possibilities to proceed:

a. Using RSVP-TE only for the signalling and LMP as described above to exchange information to configure the optical interface within the edge node or

b. RSVP-TE will be used to transport additional information
c. Leaking IGP information instead of exchanging this information needed from the optical network to the edge node (overlay will be transformed to a border-peer model)

Furthermore following issues should be addressed:

a) The Communication between peering edge nodes using an out of band control channel. The two nodes have to exchange their optical capabilities. An extended version of LMP is needed to exchange FEC Modulation scheme, etc. that must be the same. It would be helpful to define some common profiles that will be supported. Only if the profiles match with both interface capabilities it is possible to start signalling.

b) Due to the bidirectional wavelength path that must be setup it is obligatory that the upstream edge node inserts a wavelength value into the path message for the wavelength path towards the upstream node itself. But in the case of an overlay model the client device may not have full information which wavelength must/should be selected and the information must be exchanged between the edge and the core node.

...additional points

6. Requirements for BL deployments

This section raises requirements from the carrier perspective and will be removed in a separate requirements draft if necessary.

6.1. Interoperability Aspects

For carrier network deployments, interoperability is a key requirement. Today it is state-of-the-art to interconnect e.g. client devices from different vendors via a WDM transport system using short-reach, grey interfaces. Applying the Black Link (BL) concept, client devices (e.g. routers) are now directly connected via transport interfaces which must be interoperable to each other.
A progressive approach addressing interoperability is shown in Figure 5. According to the concept of ITU-T G.698.2 the black link, the single channel (coloured) Tx and the Rx can be provided by different vendors. The single-channel reference points Ss and Rs indicate the demarcation between the Tx/Rx and the black link, and the set of optical parameters refers to these reference points according to G.698.2. However, G.698.2 does not give any insight into the client equipment (CL), e.g. routers or switches, containing the optical transmitters and receivers. This is a valid topic which is subject of other standards and multi-source agreement (MSAs) describing electrical interfaces, mechanical properties and environmental conditions. Such topics are out of the scope of this document.

**Figure 5: Interoperability aspects**

7. Acknowledgements

The author would like to thank Ulrich Drafz for the very good teamwork during the last years and the initial thoughts related to the packet optical integration. Furthermore the author would like to thank all people involved within Deutsche Telekom for the support and fruitful discussions.
8. IANA Considerations

This memo includes no request to IANA.

9. Security Considerations

The architecture and solution space in scope of this framework imposes no additional requirements to the security models already defined in RFC5920 for packet/optical networks using GMPLS, covering also Control Plane and Management interfaces. Respective security mechanisms of the components and protocols, e.g. LMP security models, can be applied unchanged.

As this framework is focusing on the single operator use case, the security concerns can be relaxed to a subset compared to a setup where information is exchanged between external parties and over external interfaces.

Concerning the access control to Management interfaces, security issues can be generally addressed by authentication techniques providing origin verification, integrity and confidentiality. Additionally, access to Management interfaces can be physically or logically isolated, by configuring them to be only accessible out-of-band, through a system that is physically or logically separated from the rest of the network infrastructure. In case where management interfaces are accessible in-band at the client device or within the optical transport network domain, filtering or firewalling techniques can be used to restrict unauthorized in-band traffic. Authentication techniques may be additionally used in all cases.

10. Contributors
11. References

11.1. Normative References


[RFC3591] Lam, H-K., Stewart, M., and A. Huynh, "Definitions of Managed Objects for the Optical Interface Type",


11.2. Informative References


Authors’ Addresses

Ruediger Kunze (editor)
Deutsche Telekom AG
Berlin, 10589
DE
Phone: +49 30 3497 3152
EMail: ruediger.kunze@telekom.de

Gert Grammel (editor)
Juniper Networks
dddddd
dddd, 1234
US
Phone: +1 45552
EMail: ggrammel@juniper.net

Gabriele Galimberti (editor)
Cisco
Via Philips,12
20052 - Monza
Italy
Phone: +390392091462
EMail: ggalimbe@cisco.com

Hans-Juergen Schmidtke (editor)
Juniper Networks
dddd, 1234
US
Phone: +1 45552
EMail: hschmidtke@juniper.net
Link Management Protocol Extensions for Grid Property Negotiation
draft-li-ccamp-grid-property-lmp-01

Abstract

The recent updated version of ITU-T [G.694.1] has introduced the flexible-grid DWDM technique, which provides a new tool that operators can implement to provide a higher degree of network optimization than is possible with fixed-grid systems. This document describes the extensions to the Link Management Protocol (LMP) to negotiate link grid property between the adjacent DWDM nodes before the link is brought up.

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 9, 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. Conventions Used in This Document ..................... 3
2. Terminology .................................................. 3
3. Problem Statement ............................................. 4
   3.1. Flexi-fixed Grid Nodes Interworking .................... 4
   3.2. Flexible-Grid Capability Negotiation .................... 5
   3.3. Problem Summary ......................................... 6
4. LMP extensions ............................................... 6
   4.1. Grid Property Subobject ................................ 6
5. Messages Exchange Procedure .................................. 8
   5.1. Flexi-fixed Grid Nodes Messages Exchange ............... 8
   5.2. Flexible Nodes Messages Exchange ....................... 9
6. IANA Considerations ......................................... 10
7. Security Considerations ..................................... 10
8. References .................................................... 10
   8.1. Normative references .................................... 10
   8.2. Informative References .................................. 11
Authors’ Addresses ............................................. 11
1. Introduction

The recent updated version of ITU-T [G.694.1] has introduced the flexible-grid DWDM technique, which provides a new tool that operators can implement to provide a higher degree of network optimization than is possible with fixed-grid systems. A flexible-grid network supports allocating an arbitrary spectral slot to a channel. Mixed bitrate transmission systems can allocate their channels with different spectral bandwidths/slot widths so that they can be optimized for the bandwidth requirements of the particular bitrate and modulation scheme of the individual channels. This technique is regarded to be a promising way to improve the spectrum utilization efficiency and fundamentally reduce the cost of the core network.

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. Therefore, this document describes the extensions to the Link Management Protocol (LMP) to negotiate a link grid property between two adjacent DWDM nodes before the link is brought up.

1.1. Conventions Used in This Document

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

2. Terminology

For the flexible DWDM grid, the allowed frequency slots have a nominal central frequency (in THz) defined by:

193.1 + n * 0.00625

where n is a positive or negative integer including 0 and 0.00625 is the nominal central frequency granularity in THz

and a slot width defined by:

12.5 * m

where m is a positive integer and 12.5 is the slot width granularity in GHz. Any combination of frequency slots is allowed as long as no two slots overlap.
In this contribution, some other definitions are listed below:

Grid granularity: Grid granularity includes two elements: nominal central frequency granularity and slot width granularity. The value of slot width granularity is always configured to be twice of the central frequency granularity, so that the spectral resources can be allocated without leaving any gaps. Therefore, when grid granularity appears alone, we just refer to the nominal central frequency granularity.

Tuning range: In this draft we just refer to the tuning range of the spectral bandwidth or slot width.

Channel spacing: In traditional fixed-grid network, the adjacent channel spacing is constant. While for the flexible-grid network, the adjacent channel spacing is determined by the two central frequencies.

3. Problem Statement

3.1. Flexi-fixed Grid Nodes Interworking

---+         ---+         ---+         ---+        ---+
| A |---------| B |=========| C |=========| D +--------+ E |
---+         ---+         ---+         ---+        ---+

Figure 1

öttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttöttottenham

n=-1 0 1 2

Fixed channel spacing of 50 GHz (Node C)

öttöttöttöttöttöttöttottenham

n=-8, m=4 n=0, m=4 n=8, m=4

n=-16 -14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14 16

Flexi-grid (Nodes B,D) 6.25 GHz
Central frequency granularity=6.25 GHz
Slot width granularity=12.5 GHz
Figure 1 shows an example of interworking between flexible and fixed-grid nodes. Nodes A, B, D, E support flexible-grid. All these nodes can support frequency slots with a central frequency granularity of 6.25 GHz and slot width granularity of 12.5 GHz. Given the flexibility in flexible-grid nodes, it is possible to configure the nodes in such a way that the central frequencies and slot width parameters are backwards compatible with the fixed DWDM grids (adjacent flexible frequency slots with channel spacing of 8*6.25 and slot width of 4*12.5 GHz is equivalent to fixed DWDM grids with channel spacing of 50 GHz).

As node C can only support the fixed-grid DWDM property with channel spacing of 50 GHz, to establish a LSP through node B,C,D, the links between B to C and C to D must set to align with the fixed-grid values. This link grid property must be negotiated before establishing the LSP.

3.2. Flexible-Grid Capability Negotiation

```
+---+            +---+
| F +------------| G |
+---+            +---+
     ^              ^
     | Tuning range |  \\
     | [12.5, 100]  |  \\
     | [25, 200]     |  \\
```

Figure 3

The updated version of ITU-T [G.694.1] has defined the flexible-grid with 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 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. For example, in figure 3, node F can only support slot width with tuning change from 12.5 to 100 GHz,
while node G supports tuning range from 25 GHz to 200 GHz. The link property of slot width tuning range for the link between F and G should be chosen as the range intersection, resulting in a range from 25 GHz to 100 GHz.

3.3. Problem Summary

In summary, in a DWDM Link between two nodes, the following properties can 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.

4. LMP extensions

4.1. Grid Property Subobject

According to [RFC4204], the LinkSummary message is used to verify the consistency of the link property on both sides of the link before it is brought up. The LinkSummary message contains negotiable and non-negotiable DATA_LINK objects, carrying a series of variable-length data items called subobjects, which illustrate the detailed link properties. The subobjects are defined in Section 12.12.1 in [RFC4204].

To solve the problems stated in section 3, this draft extends the LMP protocol by introducing a new DATA_LINK subobject called "Grid property", allowing the grid property correlation between adjacent nodes. The encoding format of this new subobject is as follows:

```
+-----------------+-----------------+-----------------+-----------------+-----------------+
|     Type       |     Length      |            Reserved            |
+-----------------+-----------------+-----------------+-----------------+-----------------+
| Grid |  C.S. |    Reserved    |      Min      |      Max      |
+-----------------+-----------------+-----------------+-----------------+-----------------+
```

Type=TBD, Grid property type.

Grid:
The value is used to represent which grid the node/interface supports. Values defined in [RFC6205] identify DWDM [G.694.1] and CWDM [G.694.2]. The value defined in [I-D.farrkingel-ccamp-flexigrid-lambda-label] identifies flexible DWDM.

<table>
<thead>
<tr>
<th>Grid</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>0</td>
</tr>
<tr>
<td>ITU-T DWDM</td>
<td>1</td>
</tr>
<tr>
<td>ITU-T CWDM</td>
<td>2</td>
</tr>
<tr>
<td>Flexible DWDM</td>
<td>3</td>
</tr>
<tr>
<td>Future use</td>
<td>4-7</td>
</tr>
</tbody>
</table>

C.S.:

For a fixed-grid node/interface, the C.S. value is used to represent the channel spacing, as the spacing between adjacent channels is constant. For a flexible-grid node/interface, this field should be used to represent the central frequency granularity.

<table>
<thead>
<tr>
<th>C.S. (GHz)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>12.5</td>
<td>4</td>
</tr>
<tr>
<td>6.25</td>
<td>5</td>
</tr>
<tr>
<td>Future use</td>
<td>6-15</td>
</tr>
</tbody>
</table>

Min & Max:

The slot width tuning range the interface supports (indicated by the
m value defined in section 2). For example, for slot width tuning range from 25 GHz to 100 GHz (with regarding to a node with slot width granularity of 12.5 GHz), the values of Min and Max should be 2 and 8 respectively. For fixed-grid nodes, these two fields are meaningless and should be set to zeros.

5. Messages Exchange Procedure

5.1. Flexi-fixed Grid Nodes Messages Exchange

To demonstrate the procedure of grid property correlation, the model shown in Figure 1 is reused. Node B starts sending messages.

- After inspecting its own node/interface property, node B sends node C a LinkSummary message including the MESSAGE ID, TE_LINK ID and DATA_LINK objects. The setting and negotiating of MESSAGE ID and TE_link ID can be referenced to [RFC4204]. As node B supports flexible-grid property, the Grid and C.S. values in the grid property subobject are set to be 3 and 5 respectively. The slot width tuning range is from 12.5 GHz to 200 GHz. Meanwhile, the N bit of the DATA_LINK object is set to 1, indicating that the property is negotiable.

- When node C receives the LinkSummary message from B, it checks the Grid, C.S., Min and Max values in the grid property subobject. Node C can only support fixed-grid DWDM and realizes that the flexible-grid property is not acceptable for the link. Since the receiving N bit in the DATA_LINK object is set, indicating that the Grid property of B is negotiable, node C responds to B with a LinkSummaryNack containing a new Error_code object and state that the property needs further negotiation. Meanwhile, an accepted grid property subobject (Grid=2, C.S.=2, fixed DWDM with channel spacing of 50 GHz) is carried in LinkSummaryNack message. At this moment, the N bit in the DATA_LINK object is set to 0, indicating that the grid property subobject is non-negotiable.

- As the channel spacing and slot width of node B can be configured to be any integral multiples of 6.25 GHz and 12.5 GHz respectively, node B supports the fixed DWDM values announced by node C. Consequently, node B will resend the LinkSummary message carrying the grid property subobject with values of Grid=2 and C.S.=2.

- Once received the LinkSummary message from node B, node C replies with a LinkSummaryACK message. After the message exchange, the link between node B and C is brought up with a fixed channel spacing of 50 GHz.
In the above mentioned grid property correlation scenario, the node supporting a flexible-grid is the one that starts sending LMP messages. The procedure where the initiator is the fixed-grid node is as follows:

- After inspecting its own interface property, Node C sends B a LinkSummary message containing a grid property subobject with Grid=2, C.S.=2. The N bit in the DATA_LINK object is set to 0, indicating that it is non-negotiable.

- As the channel spacing and slot width of node B can be configured to be any integral multiples of 6.25 GHz and 12.5 GHz respectively, node B is able to support the fixed DWDM parameters. Then, node B will make appropriate configuration and reply node C the LinkSummaryACK message.

- After the message exchange, the link between node B and C is brought up with a fixed channel spacing of 50 GHz.

5.2. Flexible Nodes Messages Exchange

To demonstrate the procedure of grid property correlation between to flexi-grid capable nodes, the model shown in figure 3 is reused. The procedure of grid property correlation (negotiating the grid granularity and slot width tuning range) is similar to the scenarios mentioned above.

- The Grid, C.S., Min and Max values in the grid property subobject sent from node F to G are set to be 3,5,1,8 respectively. Meanwhile, the N bit of the DATA_LINK object is set to 1, indicating that the grid property is negotiable.

- When node G has received the LinkSummary message from F, it will analyze the Grid, C.S., Min and Max values in the Grid property subobject. But node G can only support grid granularity of 12.5 GHz and a slot width tuning range from 25 GHz to 200 GHz. Considering the property of node F, node G then will respond F a LinkSummaryNack containing a new Error_code object and state that the property need further negotiation. Meanwhile, an accepted grid property subobject (Grid=3, C.S.=4, Min=1, Max=4, the slot width tuning range is set to the intersection of Node F and G) is carried in LinkSummaryNack message. Meanwhile, the N bit in the DATA_LINK object is set to 1, indicating that the grid property subobject is non-negotiable.
As the channel spacing and slot width of node F can be configured to be any integral multiples of 6.25 GHz and 12.5 GHz respectively, node F can support the larger granularity. The suggested slot width tuning range is acceptable for node F. In consequence, node F will resend the LinkSummary message carrying the grid subobject with values of Grid=3, C.S.=4, Min=1 and Max=4.

Once received the LinkSummary message from node F, node G replies with a LinkSummaryACK message. After the message exchange, the link between node F and G is brought up supporting central frequency granularity of 12.5 GHz and slot width tuning range from 25 GHz to 100 GHz.

From the perspective of the control plane, once the links have been brought up, wavelength constraint information can be advertised and the wavelength label can be assigned hop-by-hop when establishing a LSP based on the link grid property.

6. IANA Considerations
   TBD

7. Security Considerations
   TBD

8. References

8.1. Normative references


8.2. Informative References

[I-D.farrkingel-ccamp-flexigrid-lambda-label]
King, D., Farrel, A., Li, Y., Zhang, F., and R. Casellas,
"Generalized Labels for the Flexi-Grid in Lambda-Switch-
Capable (LSC) Label Switching Routers",
draft-farrkingel-ccamp-flexigrid-lambda-label-01 (work in
progress), October 2011.

Authors’ Addresses

Yao Li
ZTE
Email: li.yao3@zte.com.cn

Ramon Casellas
CTTC
Email: ramon.casellas@cttc.es

Yu Wang
China Academy of Telecom Research, MIIT
No.52 Huayuan Beilu, Haidian District, Beijing, P.R. China, 100083
Email: wangyu@mail.ritt.com.cn
Expressing Label Set in ERO

draft-margaria-ccamp-label-set-ero-00

Abstract

The paths chosen by Generalized MPLS (GMPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) can be constrained using the Explicit Route (ERO) object and related sub-objects. Standard ERO sub-objects can specify the Autonomous System (AS), LSR Node Ids, Numbered or unnumbered TE links, downstream and upstream labels, and PCE path keys thus restricting which resources are to be used by a TE-LSP.

The Explicit Label Control (ELC) in the explicit route object (ERO) allows both terminating an LSP on a particular outgoing port and label of an egress node, as well as restricting which label to use on any hop along the path determined by the route. However, currently, it’s not allowed to specify more than 2 labels (downstream and upstream label), and it is not possible to specify, for a given section or segment of a TE-LSP path, a set of labels to restrict which label to be allocated from a Set of candidate labels.

This memo provides extensions to the RSVP-TE and PCEP protocols to support Label Sets in the form of ERO sub-objects, being applicable to ERO and ERO-like (IRO, RRO, XRO) sub-objects, extending the ELC concept to a set of candidate labels.

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
# Table of Contents

1. Introduction ........................................... 4  
   1.1. Contributing Authors ............................ 5  
   1.2. Requirements Language ........................... 5  
2. Solution overview ...................................... 6  
3. Multiple consecutive Label ERO subobjects ............... 7  
4. Label Set ERO subobject ................................ 8  
   4.1. Procedures ....................................... 8  
5. LSP_ATTRIBUTE extensions ............................... 10  
   5.1. TARGETED_LSP_ATTRIBUTE TLV .................... 10  
   5.2. PCEP Extensions ................................. 13  
6. Evaluation of proposed solution alternatives ............ 14  
7. IANA Considerations .................................... 15  
   7.1. Label Set ERO subobject .......................... 15  
   7.2. LSP_ATTRIBUTE ................................... 15  
       7.2.1. Attribute Flags ............................. 15  
       7.2.2. Service ID TLV .............................. 15  
       7.2.3. Targeted LSP attribute TLV ................. 16  
8. Security Considerations ................................ 17  
9. Contributing Authors ................................... 18  
10. Acknowledgments ...................................... 19  
11. References ........................................... 20  
   11.1. Normative References ............................ 20  
   11.2. Informative References .......................... 20  
Authors’ Addresses ....................................... 21
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]. In general, ERO sub-objects identify the resources to be used by a GMPLS, and can be used to restrict, exclude (IRO/XRO), define (ERO) or report (RRO) such resources.

The Explicit Label Control (ELC) in the explicit route object (ERO) allows both terminating an LSP on a particular outgoing port and label of an egress node, as well as restricting which label to use on any hop along the path determined by the route. However, currently, it is not allowed to specify more than 2 labels (downstream and upstream label), and it is not possible to specify, for a given section or segment of a TE-LSP path, a set of labels to restrict which label to be allocated from a Set of candidate labels.

On the other hand, [RFC3473] defines the RSVP-TE LABEL_SET object, which can be used in a Path Message to restrict the choice of the generalized label, allocated by the downstream node to a set of labels.

Extending the semantics of the Explicit Label Control to a label set and restricting / limiting the choice of label within a given Label Set, while maintaining the applicability of ERO and ERO-like RSVP-TE and PCEP objects can be beneficial in the following cases:

- To constrain and restrict the choice of a Label at a given port (including egress port) to be selected from a set of labels but without strictly enforcing a single value (for example, when conveying a set of available labels due to hardware limitations such as tunable wavelengths).

- Due to known label switching constraint on some section of the TE-LSP path: explicitly specify the label constraint on a specific link by requesting a Label Set to limit the choice of the label.

- To constraint a distributed wavelength assignment (D-WA) for a TE-LSP DWDM transparent optical section

- To allow a PCE or any other centralized entity to indicate a set of labels to be used in signaling, not only in the initial Label set but in any hop along the path.

- To allow a Path Computation Client (PCC) to indicate, as an input constraint when requesting a combined R&WA computation to a PCE,
which set of wavelengths are acceptable at a given TE link, transparent segment or end-to-end path, depending on the label (wavelength) continuity restrictions of the underlying data plane.

- To exclude (i.e., in a XRO object) a set of Labels from being considered in a label allocation, reusing the efficient encoding that has been proposed for Label sets and Label set ERO sub-objects.

- To control resource sharing for pre-planned protecting LSP, where one can indicate which labels should be shared in addition to the PPRO disjointness constraints.

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.
2. Solution overview

In order to support specifying several labels as a potential set of labels to be used when allocating a label, several solutions are applicable and described in this document:

- Allowing several consecutive Label ERO subobjects in an ERO object.
- Defining a Label Set ERO subobject to be used in an ERO (and similar) object.
- Extending the LSP_ATTRIBUTES object with a new TLV targeting attributes at a given node.

A short overview of each solution is provided in the next sections and an evaluation of each one on is provided afterwards.
3. Multiple consecutive Label ERO subobjects

This approach would require relaxing the rules that define how Label sub-objects are used within an ERO/XRO/RRO object, and notably in what concerns Explicit Label Control. In particular, the procedure described in [RFC3473] section 5.1.1 is modified as follows:

The following SHOULD NOT result in a "Bad EXPLICIT_ROUTE object" errors:

For there to be two label subobjects with the same U-bit values

It is allowed to have several consecutive Label subobjects with the same U-bit values, which become equally valid alternatives for the downstream label.

To support the label subobject, a node must check to see if one or more sub-objects following its associated address/interface sub-objects are label subobjects. If this is the case, the sub-objects are examined: a RSVP-TE LABEL_SET object (Type 36) is constructed with the values of the labels that have the U-bit cleared. This LABEL_SET object MUST be included in the outgoing Path message and MAY be splitted into several LABEL_SET objects (LABEL_SET for the downstream direction). Note that this LABEL_SET does not replace the existing LABEL_SET and MAY be merged with it.

The new Label_Set objects included in the Path message do not replace existing Label_Set object.

If the U-bit of the subobject being examined is set (1), then the set of value of the Label subobject with U bit set should be used to restrict the choice of the upstream label associated with the bidirectional LSP. If this label is not acceptable, a "Bad EXPLICIT_ROUTE object" error SHOULD be generated. If the label is acceptable, the assigned label is copied into a new Upstream_Label object. This Upstream_Label object MUST be included on the corresponding outgoing Path message.
4. Label Set ERO subobject

In this solution a new Label Set subobject is defined.

The Label Set ERO 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |U|   Reserved  |   C-Type(1)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Label Set                        |
|                              ...                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

See [RFC3471] for a description of L,U and Label Set parameters.

Type  x Label Set

Length  The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always divisible by 4.

C-Type  The C-Type of the included Label Object. Copied from the Label Object.

4.1. Procedures

The Label Set subobject follows a similar procedure to the aforementioned one that is used when several Label sub-objects are allowed. The Label Set subobject must follow a subobject identifying the link where the Label Set is applicable (a subobject containing a IP address or an interface identifier) or a previous Label Set subobject. Several Label Set subobject may be present. The following SHOULD result in "bad EXPLICIT_ROUTE object" errors:

If the first Label Set subobject is not preceded by a subobject containing an IP address or an interface identifier associated with an output link

On unidirectional LSP setup, for there to be a label set subobject with the U-bit set

It is not allowed to mix Label Set and Label subobject.

The following text is adapted from the Label subobject procedure described in [RFC3473]
To support the label set subobject, a node must check to see if the subobject following its associate address/interface is a label set subobject. If it is, the following subobjects are examined. If the U-bit of the subobject being examined is clear (0), then value of the label set is copied into a RSVP-TE LABEL_SET object (Type 36). One LABEL_SET object MUST be included in the outgoing Path message. The LABEL-SET object MAY be splitted into several LABEL_SET objects or MAY be merged with the existing LABEL-SET objects of this LSP.

If the U-bit of the subobject being examined is set (1), then value of the label set is used to choose the label to be used for upstream traffic associated with the bidirectional LSP. If this label is not acceptable, a "Bad EXPLICIT_ROUTE object" error SHOULD be generated. If the label set is acceptable and a label assigned, the label is copied into a new Upstream_Label object. This Upstream_Label object MUST be included on the corresponding outgoing Path message.

After processing, the label set subobjects are removed from the ERO.

Note an implication of the above procedures is that the label set subobject should never be the first subobject in a newly received message. If the label subobject is the the first subobject an a received ERO, then it SHOULD be treated as a "Bad strict node" error.

Procedures by which an LSR at the head-end of an LSP obtains the information needed to construct the Label Set subobject are outside the scope of this document.
5. LSP_ATTRIBUTE extensions

In order to indicate, at a given hop or interface within the ERO, a set of candidate labels to be used when selecting the generalized label, it is also possible to use LSP_ATTRIBUTE extensions [RFC5240]. To this end, the procedure to generate the outgoing RSVP-TE Path message LABEL_SET object from the information contained in the LSP_ATTRIBUTE is similar conceptually to the previous ones. The following new TLVs are required for this solution:

- a TLV indicating attributes for a node/interface (one per node/interface)
- This TLV will contain sub-TLV, here:
  - Attribute Flag TLV
  - A Label Set TLV
  - Any Attribute TLV which are applicable to a specific Node/interface

A new flag should be defined for the Targeted LSP attribute, requiring this will then depend in which object this is added. The Label Set TLV also required a new flag, but it SHOULD NOT appear directly in the LSP_ATTRIBUTE, this solution add TLVs that can be scoped only to specific interface or node. The RRO subobject attribute processing is not modified, a Node MAY report all the bits from the Attribute flag TLV in LSP_ATTRIBUTES or/and LSP_REQUIRED_ATTRIBUTES or/and the Attribute flag TLV in the TARGETED_LSP_ATTRIBUTE TLV.

5.1. TARGETED_LSP_ATTRIBUTE TLV

A new TLV is introduced, the TARGETED_LSP_ATTRIBUTE TLV, which is valid on Path message only in LSP_ATTRIBUTE and LSP_REQUIRED_ATTRIBUTES Object.
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type | Length | Identifier Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Identifier (Variable) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
// TLVs //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Type x

Length  The Length contains the total length of the subobject in bytes, including the Type and Length fields. This length must be a multiple of four and must be at least eight.

Identifier Type  The type of identifier used, currently the following types are defined:

0  IPv4 address
1  IPv6 address
2  Unnumbered address

Identifier  Depending on the Identifier type this field contains:

IPv4 address  A 32 bit IPv4 address
IPv6 address  A 128 bit IPv6 Address
Unnumbered address  A 64 bit field containing a 32 bit Node Id and 32 bit unnumbered address

TLVs  A list of TLVs

An IPv4 Identifier type
An IPv4 Identifier type

```
+-----------------+-------+-------------------+
|     Type        |  Length | Identifier Type=0 (IPv4) |
|-----------------+-------+-------------------|
| IPv4 address    |       |                   |
```

```
//                       TLVs                                  |
//                       //                                    |
```

An IPv6 Identifier type

```
+-----------------+-------+-------------------+
|     Type        |  Length | Identifier Type=1 (IPv6) |
|-----------------+-------+-------------------|
| IPv6 address    |       |                   |
```

```
//                       TLVs                                  |
//                       //                                    |
```

An Unnumbered interface identifier

```
+-----------------+-------+-------------------+
|     Type        |  Length | Identifier Type=2 (Unnumbered) |
|-----------------+-------+-------------------|
| Node Id         |       |                   |
```

```
<table>
<thead>
<tr>
<th>Unnumbered interface id</th>
</tr>
</thead>
</table>
```

```
//                       TLVs                                  |
//                       //                                    |
```
5.2. PCEP Extensions

This solution does not fit to existing PCEP object. One possible solution would be to map the RSVP LSP_ATTRIBUTE logic to PCEP LSPA object and define a set of LSPA TLVs carrying relevant LSP_ATTRIBUTE TLVs. This is for further study.
6. Evaluation of proposed solution alternatives

The First two solutions would easily allow a centralized entity such as a PCE or a NMS to add this per-hop constraints information but would imply a greater impact to existing deployments. Let us note that, in general, a PCE currently uses the existing ERO sub-objects (with different semantics) in the following PCEP sub-objects.

- ERO - to indicate the result of a Path Computation given one or more requests.
- IRO - to specify which elements, resources, etc. must be used in a path computation.
- XRO - to specify which elements, resources, etc. must be excluded in a path computation.
- RRO - to report of existing Paths

Making use of the LSP_ATTRIBUTES would reduce the impact on existing deployment yet allow to mandate the support of the attribute if desired, but will introduce several source for Label information.
7. IANA Considerations

7.1. Label Set ERO subobject

IANA is requested to make the following subobject allocations from the "EXPLICIT_ROUTE Subobject Type" registry.

<table>
<thead>
<tr>
<th>Sub-object type</th>
<th>TBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Label Set</td>
</tr>
<tr>
<td>Reference</td>
<td>This document</td>
</tr>
</tbody>
</table>

7.2. LSP_ATTRIBUTE

IANA is request to add the following information for each TLV in the RSVP TLV type identifier registry.

- Whether allowed on TARGETED_LSP_ATTRIBUTE TLV

The existing registry is modified for existing TLVs.

7.2.1. Attribute Flags

The new TLV type definition is as follow

- TLV Type = 1
- TLV Name = Attribute Flags TLV
- Allowed on LSP_ATTRIBUTES object
- Allowed on LSP_REQUIRED_ATTRIBUTES object
- Allowed on TARGETED_LSP_ATTRIBUTE TLV

7.2.2. Service ID TLV

The new TLV type definition is as follow

- TLV Type = 1
- TLV Name = Attribute Flags TLV
- Allowed on LSP_ATTRIBUTES object
7.2.3. Targeted LSP attribute TLV

IANA is requested to make the following allocations from the RSVP Attribute TLV Space registry.

- TLV Type = n
- TLV Name = Targeted LSP attribute TLV
- Allowed on LSP_ATTRIBUTES object
- Allowed on LSP_REQUIRED_ATTRIBUTES object
- Not allowed on TARGETED_LSP_ATTRIBUTE TLV

IANA is requested to make the following allocation from the RSVP Attribute Flags registry:

<table>
<thead>
<tr>
<th>Bit No</th>
<th>Name</th>
<th>Attribute Flags</th>
<th>Attribute Flags</th>
<th>RRO Reference</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Targeted LSP</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>This document</td>
</tr>
<tr>
<td></td>
<td>Attribute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Label Set</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>This document</td>
</tr>
</tbody>
</table>

8. Security Considerations

None.
9. Contributing Authors
10. Acknowledgments

The research leading to these results has received funding from the European Community’s Seventh Framework Program FP7/2007-2013 under grant agreement no 247674.
11. References

11.1. Normative References


11.2. Informative References

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

Ramon Casellas
CTTC
Av. Carl Friedrich Gauss n.7
Castelldefels, Barcelona
Spain
Phone: +34 93 645 29 00
Email: ramon.casellas@cttc.es

Oscar Gonzalez de Dios
Telefonica Investigacion y Desarrollo
C/ Emilio Vargas 6
Madrid, 28043
Spain
Phone: +34 91 3374013
Email: ogondo@tid.es
WSON Optical Interface Class
draft-martinelli-wson-interface-class-02

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 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
   1.1. Requirements Language .............................. 3
2. Existing WSON 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 ........................... 7
5. Acknowledgements .......................................... 8
6. IANA Considerations ....................................... 8
7. Security Considerations ................................... 8
8. References ................................................ 9
   8.1. Normative References ............................... 9
   8.2. Informative References ............................. 9
Appendix A. Encoding example ................................ 10
Authors’ Addresses ............................................ 12
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. Regarding how to evaluate, there are a list of parameters identified according to ITU specification [ITU-G.698.1] and [ITU-G.698.2]. In particular the following set of parameters has been chosen: signal bit rate, modulation format, forward error correction.

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 [I-D.ietf-ccamp-wson-impairments]. 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 extend 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:

- Modulation Format. Only NRZ currently defined.
- FEC, according to G.709 and G.975.
- 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.

- 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.
- 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).
- 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 enpoints 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.

```
+---+    +----+   +----+     +-----+     +----+     +---+
|   |----| N1 |---| N2 |-----| REG |-----| N3 |-----| E |
+---+    +----+   +----+     +-----+     +----+     +---+
   cl1 <--------S1---------> cl1    cl1
   cl2                       cl2                     cl2
   cl3    cl3 <----S2----> cl3

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].

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S|   Reserved                |    OI Code Points             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Optical Interface Class                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Optical Interface Class  (Cont.)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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. If set to 1, code points to indentify ITU.T application codes. If set 0 identify a not ITU code points.

S=0, OI Code Points:

0: reserved

1: Enterprise Specific Optical Interface Class. In this case the first 32 bits of the 64 bits for the Optical Interface class shall match the IANA SMI enterprise number.
As an example, Figure 3 shows how the encoding looks like in case of S bit equals to 0 and Code point equals to 1.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|     Reserved                |      OI Code Point = 1        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         IANA SMI Enterprise Number                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Vendor Specific Optical Interface Class               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

**Figure 3: Vendor Specific Optical Interface Class**

S=1, OI Code Points:

0: reserved

1: [ITU-G.698.1] 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. This mapping is out of the scope of this document.

However 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 analysis. 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 estimation, 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, [draft-galimbe-kunze-g698-2-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.

8. References

8.1. Normative References

[I-D.ietf-ccamp-rwa-wson-encode]

[I-D.ietf-ccamp-wson-signal-compatibility-ospf]

[I-D.ietf-ccamp-wson-signaling]

[ITU-G.698.1]

[ITU-G.698.2]


8.2. Informative References

[I-D.ietf-ccamp-wson-impairments]
Appendix A. 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 length 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 format 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 4.
With the following notes:

- 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.
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.

Optical interface classes have a fixed size while basic encoding blocks of [I-D.ietf-ccamp-rwa-wson-encode] have sizes that vary 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: giomart@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
Munchen, 81541
Germany

Phone: +49-89-5159-16934
Email: cyril.margaria@nsn.com
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/1id-abstracts.txt

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

This Internet-Draft will expire on September 5, 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 Internet-Draft proposes requirements for combinations of rerouting and protection schemes that are of interest to carrier networks.

1. Introduction and Problem Statement

Combination of protection and rerouting mechanisms allow carriers to:

- Perform maintenance activities on the protected/protection LSPs while maintaining the protection.

- Offer services with higher availability by automatically combining restoration and protection schemes.

This document defines use cases for the combination of protection and rerouting mechanisms that are of interest to carriers that could be candidates for support in GMPLS.

2. Maintenance operations of protected services

The first area to consider is the combination of maintenance activities with the protection and restoration schemes. For example, it is possible to perform maintenance operation on one of the legs of a 1+1 connection but the operator may want to maintain the 1+1 protection during the maintenance activity. Temporarily or permanently rerouting the traffic of one of the legs for maintenance purposes requires proper support in the signaling procedures.

The requirement is to support up to 3 related LSPs that have resources reserved but of which at most 2 are instantiated end-to-end in the data plane.

Consider the following network topology:
The detailed signaling scenario for rerouting is as follows:

- Preconditions: The working (A-B-C-D-Z) and protecting (A-E-F-Z) LSPs are both established and neither of them has a failure condition.

- A third LSP (A-G-H-I) is established with resources reserved and established on nodes G, H and I but reserved only at nodes A and Z.

- The LSP that is put under maintenance is de-activated by keeping the resources reserved but no established at nodes A and Z while maintaining resources reserved and established at nodes B, C and D if maintenance is applied to working LSP or nodes E and F if maintenance is applied to protecting LSP.

- The maintenance LSP is activated by establishing the resources for the maintenance LSP at nodes A and Z.

- At this stage, if this is a permanent reroute operation, the original working or protecting LSP is torn down. Otherwise, the LSP is maintained for reversion at a later stage.

The reversion stage:

- The maintenance LSP is deactivated by de-establishing the resources for the maintenance LSP at nodes A and Z.

- The original LSP (working or protecting) is activated at nodes A and Z by establishing resources (working or protecting) at nodes A and Z.

- The maintenance LSP is torn down.
3. Combined restoration and protection schemes

In order to provide higher reliability, some service levels may combine restoration and protection. Two combinations that are useful to operators are included here. These combinations may require more than two LSPs to be associated together in case of make-before-break or when reversion is desired. Signaling extensions that support combined protection and restoration are required by identifying the type of recovery as a combination of protection (e.g. 1+1 bi-directional) and restoration types (e.g. full rerouting). The signaling extensions should also provide an indication of the relationship between the two mechanisms to distinguish the following scenarios:

- Second level restoration: offers protection against dual failures in the case of protected services. It offers the option to restore the LSP if both working and protection LSPs fail.
- Always on protection: offers the assurance of fast protection even after a failure by restoring the failed leg of a protected service.

3.1. Second Level Restoration

The requirement is to support up to 3 related LSPs that have resources reserved but of which at most 2 are instantiated end-to-end in the data plane. When both the working and protecting LSPs are under failure condition, this triggers restoration.

Consider the following network topology:

```
B --- C --- D
/               \   
A --- E --- F --- Z
\               /  
G --- H --- I
```

The detailed signaling scenario for rerouting is as follows:

- Preconditions: The working (A-B-C-D-Z) and protecting (A-E-F-Z) LSPs are both established and both are under failure condition.
- One of the original LSPs (working or protecting), as determined by the head-end local policy is deactivated at the A and Z nodes while maintaining the resources reserved.

- A restoration LSP (A-G-H-I) is established with resources reserved and established.

- At this stage, if this is non-revertive restoration, the original working or protecting LSP is torn down. Otherwise, the LSP is maintained for reversion at a later stage.

The reversion stage:

- The restoration LSP is deactivated by de-establishing the resources for the restoration LSP at nodes A and Z.

- The original LSP (working or protecting) is activated at nodes A and Z by establishing resources (working or protecting) at nodes A and Z.

- The restoration LSP is torn down.

3.2. Always on protection

The requirement is to support up to 4 related LSPs that have resources reserved but of which at most 2 are instantiated end-to-end in the data plane. When one of the working or protecting LSPs is under failure condition, this triggers restoration of that LSP.

Consider the following network topology:

```
B --- C --- D  
   / 
A --- E --- F --- Z  
  /  
G --- H --- I  
  
J ------- L
```

The detailed signaling scenario for rerouting is as follows:
- Preconditions: The working (A-B-C-D-Z) and protecting (A-E-F-Z) LSPs are both established and at least one is under failure condition. For the purpose of this example, let's assume working has failed.

- The failed LSP (working in this example), is deactivated at the A and Z nodes while maintaining the resources reserved.

- A restoration working LSP (A-G-H-I) is established with resources reserved and established.

- At this stage, we have 3 LSPs that are associated but only 2 are established end-to-end in the data plane. If this is a non-revertive restoration, the working LSP (A-B-C-D-Z) is torn down. Otherwise, it is maintained until reversion.

- If the second original LSP also fails (protecting LSP in this example), it is also deactivated at the A and Z nodes while maintaining the resources reserved and a restoration protecting LSP (A-J-L-Z) is established with resources reserved and established. Similarly to the previous bullet, if this is a non-revertive restoration, the protecting LSP is torn down. Otherwise, it is maintained until reversion.

The reversion stage:

- The working or protecting restoration LSP is deactivated by de-establishing the resources for the restoration LSP at nodes A and Z.

- The original LSP (working or protecting) is activated at nodes A and Z by establishing resources (working or protecting) at nodes A and Z.

- The working or protecting restoration LSP is torn down.

4. Security Considerations

None identified at this time

5. IANA Considerations

None identified at this time
6. Acknowledgments

The authors would like to thank members of the OIF Network & Operations WG and the CCAMP co-chairs for their comments and suggestions to this document.
Authors’ Addresses

Evelyne Roch
Ciena
Email: eroch@ciena.com

Lyndon Ong
Ciena
Email: lyong@ciena.com
Revertive Recovery Requirements

draft-roch-ccamp-full-reroute-reversion-00.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September 5, 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

Roch

Expires September 5, 2012
This draft identifies requirements for support of full rerouting recovery scheme in a revertive manner.

Table of Contents

1. Introduction and Problem Statement............................................2
2. Basic Requirements.........................................................................2
3. Analysis of Current Text..............................................................2
4. Example Scenario.............................................................................3
5. Security Considerations.................................................................4
6. IANA Considerations.........................................................................4
7. References..........................................................................................4
   7.1. Normative References...............................................................4
   7.2. Informative References...............................................................4
8. Acknowledgments................................................................................4

1. Introduction and Problem Statement

Carriers have expressed interest in supporting full-rerouting with reversion capability. Additional clarification of how this can be supported within GMPLS is needed.

2. Basic Requirements

We would like to be able to support of a version of rerouting that has the following characteristics:

- First the original LSP is established on demand.

- Secondly failure of the original LSP causes an alternate LSP to be created without any prior reservation of backup resources, potentially sharing some resources

- Finally, traffic reverts to the original LSP path once the failure is repaired.

3. Analysis of Current Text

The authors would like to get clarifications on how the combination of reversion and rerouting is supported in GMPLS as current specifications are not very explicit about this.

In our reading of [RFC4872], section 11, full LSP rerouting (protection type 0x01) involves tearing down the original LSP, although it says make-before-break can be used to establish the
alternate LSP before tearing down the original, so that some of the resources can be reused.

Rerouting without extra traffic (protection type 0x02) involves pre-establishment of the alternate LSP using a disjoint path, with no sharing of resources.

We also found that section 12 on reversion describes how reversion is supported for 1+1 bidirectional protection, 1:n protection with extra traffic, and rerouting without extra traffic, but has no description for reversion for the full LSP rerouting case.

It has been suggested that the decision to maintain the original LSP in the case of full LSP rerouting is made by the head-end despite the make-before-break terminology but we are concerned that this may lead to undefined behavior at the tail-end in case of multiple failures.

4. Example Scenario

For example, let’s consider the following network topology:

```
  B --- C --- D
     /               \
    A --- E --- F --- Z
      \               /
    G --- H --- I
```

If we start with an "original" connection A-B-C-D-Z that fails and is restored using "restoration 1" A-E-F-Z that later also fails and is restored to "restoration 2" A-G-H-I-Z. In networks where the SCN is congruent with the data path, the teardown of "original" and "restoration 1" connections may not reach the Z node until much later.

If Z receives a request to establish "restoration 2" LSP A-G-H-I-Z, it may still have "original" LSP A-B-C-D-Z and "restoration 1" LSP A-E-F-Z in place if the teardown of either or both of them failed. If Z cannot determine how to setup the bridge/selector, it cannot accept the request. If Z was informed or the revertiveness of the LSP, it could make that determination based on the type. If it is revertive, bridge with "original", otherwise no bridge needed.
It may be beneficial for this scenario and potentially other scenarios to distinguish between revertive and non-revertive full rerouting.

5. Security Considerations

None identified at this time.

6. IANA Considerations

None identified at this time.

7. Normative References


8. Informative References

9. Acknowledgements

The authors would like to thank members of the OIF Network & Operations WG and the CCAMP co-chairs for their comments and suggestions to this document.
Authors' Addresses

Evelyne Roch
Ciena
Email: eroch@ciena.com

Lyndon Ong
Ciena
Email: lyong@ciena.com
A Framework for control of Flex Grid Networks
draft-syed-ccamp-flexgrid-framework-ext-00.txt

Abstract

This document provides a framework for applying GMPLS architecture and protocols to Flex Grid.

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/1id-abstracts.txt

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

This Internet-Draft will expire on April 5, 2012.
Table of Contents

1. Introduction...........................................................................................................2
2. Terminology.............................................................................................................3
3. Acronyms................................................................................................................3
4. Requirements and constraints.................................................................................4
5. Use cases................................................................................................................6
6. Protocol Implications..............................................................................................9
7. Security Considerations........................................................................................9
8. IANA Considerations.............................................................................................9
9. References.............................................................................................................9
  9.1. Normative References.......................................................................................9
  9.2. Informative References.....................................................................................9
10. Acknowledgments ..............................................................................................10

1. Introduction

To enable scaling of existing transport systems to ultrahigh data rates of 1 Tbps and beyond, next generation systems providing super-channel switching capability are currently being developed. To allow efficient allocation of optical spectral bandwidth for such high bit rate systems, International Telecommunication Union Telecommunication Standardization Sector (ITU-T) is extending the G.694.1 grid standard (termed ’’Fixed-Grid’’) to include flexible grid (termed ’’Flex-Grid’’) support.
2. Terminology

A. Frequency Slot:
   The frequency range allocated to a slot [FLEX-GRID]
   It is a contiguous portion of the spectrum available for
   optical passband filter.

B. Spectral Slice:
   Refers to a minimum granularity of a frequency slot (e.g.
   12.5GHz).

C. Slot width:
   The full width of a frequency slot in a flexible grid[FLEX-
   GRID].
   The slot width is equal to number of spectral slices in the
   slot times the width of spectral slice.

D. Super-channel:
   Super-channel is a collection of one or more frequency slots
   to be treated as unified entity for management and control
   plane (Ref to figure-1).

E. Contiguous Spectrum Super-channel:
   Contiguous spectrum super-channel is a super-channel with a
   single frequency slot (Ref to figure-1).

F. Split-Spectrum super-channel:
   Split-Spectrum super-channel is a super-channel with multiple
   frequency slots.
   Each frequency slot will be allocated an independent passband
   filter, irrespective of whether frequency slots are adjacent
   or not.

Figure 1 Super-Channel (Refer to pdf version [5] of this draft for figures)

3. Acronyms

OCG: Optical Carrier Group

SCH: Super Channel

OCH: Optical Channel
OCC: Optical Channel Carrier
OTU$k$: Optical channel Transport Unit level $k$
ODU$k$: Optical channel Data Unit Level $k$
ODU$j$: Optical channel Data Unit Level $j$

4. Requirements and constraints

This section covers the high level requirements for the support of super-channels over Flex-Grid Infrastructure. Specifically, the scope of requirements and constraints listed in this section covers the functionality that shall be supported by the control plane sub-system. The Features are listed as list of Requirements Tagged as Rn, for better traceability and coverage in other related drafts and/or for references by other related standards across other standard bodies.

R1: Flexible size of super-channel

The protocol shall allow the super-channels on the Flex-Grid to be of different size/width. The number of slices and the granularity of each slice shall be flexible.

R2: Flexible mapping of super-channel

The super-channels shall be allowed to be mapped to any spectrum location in the ITU Grid.

The frequency slots allocation of super-channels on the ITU-Grid shall confirm to [FLEX GRID]

R3: Contiguous Spectrum and Split Spectrum super-channel

The protocol shall allow the use of super-channels which can be contiguous or non-contiguous.

Example: consider a system supporting 500GHz super-channel.

In case of contiguous spectrum, the super-channel is allocated with 40 slices of 12.5GHz granularity. This super-channel is placed directly on the Flex-Grid at any location.

In case of split spectrum, the super-channel is divided into multiple members. Considering the same example scenario, the 500GHz
super-channel can be divided into 2 member split spectrum channels. Each member is allocated a different flexible location on the Flex-Grid. Each frequency slot can be 250GHz, 20x12.5GHz slices allocated for frequency slot.

R4: Co-routing of split-spectrum super-channel

The protocol shall support the co-routing of frequency slots within the split-spectrum super-channels.

Please refer to the Figure 5 and Use Case 3, depicting the co-routing of split-spectrum super-channels.

R5: Flexible Modulation Formats for different super-channels on the same Flex-Grid

Each super-channel mapped on to the Flex-Grid system shall have the capability to carry mixed modulation signals.

R6: Fixed vs Flexible Grid super-channel interworking

The Control Plane protocol shall handle nodes which support flex-grid functionality in addition to nodes that only support fixed grid functionality.

This requirement is to enable introduction of flex-grid systems into existing fixed-grid network. This can also be used to deploy flex-grid system in certain segments of the network. Please also refer use case section of this document.

R7: Support for the CDC based super-channels over Flex-Grid

The super-channel over the Flex-Grid control plane frame work shall support CDC (Connectionless, Directionless and Contentionless) architecture. Further, flexibility of control shall be provided, such that, depending on deployment scenarios and application, a subset of CDC features are used on a given network segment. Hence, each type of ROADM shall be supported.

R8: Directionless/Contentionless super-channels

The protocol shall allow for routing the super-channels in different fiber directions/degrees, based on the following criteria:

a) Based on spectral slices
b) Based on fibers/nodes
The super-channels with the same frequency slot mapping are not allowed to be provisioned over a given fiber direction.

Please refer to the Figure 5 and Use Case 3, depicting the handling of same super-channel at a CDC node.

R9: Resizing of super-channel bandwidth

Depending on the spectral bandwidth changes, the protocol shall allow super-channels resizing.

R10: super-channel LSP restoration

The system shall support the super-channel based LSP restoration feature where the restored path is computed dynamically. During the restoration process, it shall be possible for the system to pick different frequency slots of super-channel, keeping the number and size of slices the same. Further, options for LSP restoration with pre-computed path (with or without resource reservation) shall be supported. Revertive and Non-Revertive restoration options shall be provided.

R11: Embedded Control Channel for super-channel routing and signaling

The system shall continue to use the standard mechanism for ECC defined in [ref: OSC based control channel], for OAM features required to be supported between network elements deploying super-channel over Flex-Grid.

R12: Management Plane and Control Plane feature interaction for super-channel

The system shall keep track of important bandwidth related parameters for the Flex-Grid based system. Important parameters include (but not limited to):

a) Available Spectral Slices
b) Provisioned super-channels along with provisioned spectral-slices

5. Use cases

The use cases described in this section are for information only. The OTN hierarchy described in this section is sure to be discussed in ITU SG-15 Q6 & Q14. Within the scope of this frame-work document, the main focus is super-channel entity. The remaining layers are described to illustrate the relationship with the digital layers.
With respect to the mapping hierarchy in the OTN layers, multiple OCHs are mapped to the SCH, and multiple OCCs (Optical Channel Carriers) are mapped to an OCH. This hierarchy is depicted in Figure 2 below. Specifically, the following flexibility of number of instances that are mapped between the layers shall be supported.

X number of OCC mapped to OCH
Y number of OCH mapped to SCH
Z number of SCH mapped to OCG

Figure 2 Super-Channel mapping to OTN hierarchy (Refer to pdf version [5] of this draft for figures)

Example Use Case 1: Super-Channel with multiple OCHs and multiple carriers per OCHs.

The following Figure 3 gives an example use case where multiple OCH are carrier over a single SCH. Please note that this is an example use case only. In general, the system shall be capable of supporting flexible mapping where there is flexible number of carriers mapped into an OCH and a flexible number of OCHs mapped to a single Super-Channel.

Figure 3 Super-Channel use case showing multiple OCH and multiple carriers per OCH (Refer to pdf version [5] of this draft for figures)

Example Use Case 2:

The following Figure 4 shows the case where multiple OCHs are carried over separate super-channels. These separate super-channel use case is used to realize the split-spectrum super-channel implementations. Further, these split-spectrum based super-channels can be co-routed together or can be diversely routed in the network.

Figure 4: Split-Spectrum Super-Channel use case showing multiple OCH and multiple carriers per OCH (Refer to pdf version [5] of this draft for figures)
Example Use Case 3: Network Level Use Case of super-channel

A network level diagram to illustrate the use of CDC based super-channel (contiguous spectrum and split-spectrum) is shown in Figure 5 below. In this scenario, N1 and N2 are digital/TDM nodes, where the client services originate. N2, N3, N4 and N5 are Optical/WDM nodes on which the super-channels are provisioned. Node N2 is CDC ROADM and Nodes N3, N4 and N5 are Colorless ROADMs only.

Four super-channels are provisioned in this example network. Super-Channels S1 are contiguous spectrum super-channels, both using the same frequency slots, and are added/dropped at Node N2. The contention for the same super-channel (with exactly the same frequency slot mapping) is avoided by routing these super-channels in different degrees of the network. Alternatively, if these super-channels have to go through the same fiber path, then the frequency slots occupied on the Flex-Grid shall be different.

Super-channels S2-1 and S2-2 illustrates the split-spectrum super-channel that is co-routed over the same fibers in the network. Super-channels S3-1 and S3-2 illustrates the split-spectrum super-channel that is diversely routed through Node N3 and Node N4.

Figure 5: Super-Channel Network Level use case (Refer to pdf version [5] of this draft for figures)

Example Use Case 4: Fixed and Flexible Grid Interworking

- In Figure 6:
  - The Nodes N2 and N3 are Flex-Grid and Fixed grid capable nodes
  - The Nodes N1 and N4 are fixed grid capable nodes.
- Fixed and Flexible support on the same interface
  - In Figure 6, this is represented by Link L3
- BW advertisement that include both fixed and flexible grid by Flex Grid capable nodes
- Signaling support for both fixed and flex-grid.

Figure 6: Use case for fixed and flex-grid interworking (Refer to pdf version [5] of this draft for figures)
6. Protocol Implications

Support GMPLS Routing extensions to satisfy requirements in Section 3.0.

Support GMPLS Signaling extensions to satisfy requirements in section 3.0.

7. Security Considerations

<Add any security considerations>

8. IANA Considerations

IANA needs to assign a new Grid field value to represent ITU-T Flex-Grid.

9. References

9.1. Normative References


9.2. Informative References


[3]  [RFC 6163] Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)


10. Acknowledgments
Authors’ Addresses

Sharfuddin Syed
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: ssyed@infinera.com

Rajan Rao
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: rrao@infinera.com

Marco Sosa
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: msosa@infinera.com

Biao Lu
Infinera
140 Caspian Ct., Sunnyvale, CA 94089
Email: blu@infinera.com

Contributor’s List

Radhakrishna Valiveti
Email: rvaliveti@infinera.com

Iftekhar Hussain
Email: IHussain@infinera.com

Abinder Dhillon
Email: ADhillon@infinera.com

Mike VanLeeuwen
Email: MVanleeuwen@infinera.com

Ping Pan
Email: ppan@infinera.com
OSPF extensions for support wavelength range allocation in flexible grid supported network
draft-wang-ccamp-flexigrid-wavelength-range-ospf-00.txt

Abstract

This document addresses the requirements and routing protocol extension of wavelength range allocation in flexible grid supported network in order to help spectrum utilization in the process of path computation.

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 1, 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
4. Extension of routing protocol .................. 5
   4.1. Relationship with WSON ............... 5
   4.2. Extensions of OSPF Protocol to Support Wavelength Range Allocation .......... 6
       4.2.1. Wavelength Range Allocation by Bitrates .......... 6
5. Security Considerations ..................... 8
6. IANA considerations ........................ 8
7. References ................................ 8
   7.1. Normative References .................. 8
   7.2. Informative References ................ 8
Authors’ Addresses .......................... 9
1. Introduction

In the latest version of ITU-T G.694.1, a new flexible grid technology is defined. The flexible grid technology is also a dense wavelength division multiplexing and is different from traditional fixed grid technology. Difference between fixed grid technology and flexible grid technology can be found in [G.694.1], the flexible grid technology allows mixed bit rates or mixed modulation formats transmission system to allocate frequency slots with different slot widths so that they can be optimized for the bandwidth requirements of a particular bit rate and modulation scheme of the individual channels. As shown in Figure 1, there are three different channel spacing exist in one system.

![Figure 1: flexible grid wavelength utilization](image)

Traditional fixed grid doesn’t allow system to allocate frequency slots with different widths, whereas flexible grid technology allows this. The situation that different slot widths exist in the flexible grid network brings much disorder. Spectrum fragments that can’t be used anyone will appear in the network in the process of allocation of wavelength for a channel. And this will result in the decline of spectrum utilization rate. So in order to raise the spectrum utilization rate, spectrum utilization scheme such as scheme of wavelength assignment and use of slot position, wavelength range, can be done in advance, and this will help raise the spectrum utilization rate. More advantage can be seen in section 2.

This document mainly addresses the routing protocol extension support the advertisement of wavelength range (or spectrum block) allocation information, such as the wavelength range positions and specific restriction information which are related to this wavelength range and so on. Certain policy would be involved here to help allocation wavelength range. One wavelength range can only be used for path setup with specific attributes, for example, with specific bitrates and/or modulation format.

When control plane uses path computation element to setup an end-to-
end path in flexible grid network, wavelength available information and restriction information should be considered in order to compute a suitable end-to-end path. The restriction information can be derived by path computation element through routing advertisement, and the wavelength restriction information depicted in this document should also be considered in the flexible network.

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

Concerns were expressed in ITU-T document about mixing and matching a variety of bitrates and modulation formats on a single fiber completely freedom. If flexible grid network allows completely freedom, indiscriminate positioning of the various channels with respect to each other, it is likely to lead to dramatically impaired system performance compared to a more careful choice of slot positioning / wavelength assignment. Effects like XPM (Cross-phase modulation) will appear if network allows completely freedom, indiscriminate positioning of various channels. Solution expressed by the ITU-T expert is that grouping of wavelength with the same bitrates and modulation formats is a sensible one. This document addresses the requirements and protocol extension of the grouping of wavelength with the same attributes such as bitrates and/or modulation formats. As signals of the same bitrates usually use the same modulation formation on a specific link, this document mainly pays attention to grouping of wavelength with the same bitrates and routing protocol extension is described in the next chapter.

Some other advantages are also brought if operator split supported wavelength into several wavelength ranges and one wavelength range can be only used for path with specific attributes in the flexible grid network except what have been described above. Grouping of wavelength of the same bitrates would help reduce fragments. Channels in the same wavelength range with the same bitrates looks almost like fixed grid technology, and they won’t generate much fragments in the path setup and release because every channel use the same slot width. It will also be convenient for operator to manage the flexible network if the operator groups the wavelength of the same bitrates.

The whole wavelength that subsystems support can be partitioned into
several wavelength ranges, and one wavelength range can only be used for paths setup with the specific bit rate. For example, a wavelength range which has a range of \((m_1, m_2)\) can only be used for signal with 400Gbit/s to setup path, and \(m_1\) represents the start label of the wavelength range, \(m_2\) represents the end label of the wavelength range. Also the wavelength can be partitioned according to the bitrates and/or modulation format if it is needed in some complex scenarios, that is to say a wavelength range can only be used for signal with the specific bitrates and/or modulation format to setup path.

When control plane uses path computation element to setup an end-to-end path through flexible grid network, wavelength available information and restriction information should be taken into consideration in order to compute a suitable end-to-end path. The wavelength range allocation information needs to be advertised by routing protocol in order to help the path computation. Section 4 describes the extension of OSPF routing protocol to advertise these wavelength range information in order to help with path computation.

4. Extension of routing protocol

4.1. Relationship with WSON

As described in the introduction section, flexible grid is a new DWDM technology which is different from traditional fixed grid technology. It is described in newest version of ITU-T [G.694.1]. Flexible grid can make use of current WSON control plane technology to setup path, for example, routing protocol and signaling protocol. Additional extensions may be needed because of new features introduced by flexible grid. This Section addresses the routing extension of the features which is described above base on the current WSON routing extension in IETF CCAMP.

[draft-ietf-ccamp-general-constraint-encode] defines a new link sub-TLV called Port Label Restrictions sub-TLV. Descriptions about Port Label Restrictions sub-TLV in this draft are referenced here: "Port label restrictions describe the label restrictions that the network element (node) and link may impose on a port. These restrictions represent what labels may or may not be used on a link and are intended to be relatively static. More dynamic information is contained in the information on available labels. Port label restrictions are specified relative to the port in general or to a specific connectivity matrix for increased modeling flexibility" and "For example, Port Label Restrictions describes the wavelength restrictions that the link and various optical devices such as OXCs, ROADMs, and waveband multiplexers may impose on a port in WSON."
These restrictions represent what wavelength may or may not be used on a link and are relatively static.

The wavelength range allocation information described in this document can be regarded as label restrictions which are imposed by network element (node) on a port, and the network element include various optical devices such as OXCs, ROADMs and waveband multiplexers and so on. These wavelength range allocation restrictions represent what labels may or may not be used on a link or what labels can only be used for channel with specific bitrates and/or modulation format. The restrictions described in this document can be seen as relative static and can be included in Port Label Restrictions sub-TLV.

4.2. Extensions of OSPF Protocol to Support Wavelength Range Allocation

Wavelength range allocation information should be known by path computation element if operators want to compute an end-to-end flexible grid network path. As described in the previous section, Port Label Restrictions sub-TLV can be used to carry this wavelength range allocation restriction information. Figure 1 is the format of Port Label Restrictions sub-TLV which is described in [draft-ietf-ccamp-general-constraint-encode] and definition of the parameters included in this sub-TLV can be found in this document.

```
<table>
<thead>
<tr>
<th>MatrixID</th>
<th>RestrictionType</th>
<th>Reserved/Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional Restriction Parameters per RestrictionType</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Figure 2: Port Label Restrictions sub-TLV

4.2.1. Wavelength Range Allocation by Bitrates

As described in section 3, signals on a single fiber with the same bitrates usually use the same modulation format, especially when the equipments come from one vendor. So here operator can allocation wavelength into several wavelength ranges by bitrates. In this section, OSPF protocol is extended to cover the wavelength range allocation information by bitrates.
The wavelength range allocation information by bitrates is needed in the process of path computation if an end-to-end path needs to be computed by path computation element and this information SHOULD be advertised by routing protocols. Figure 3 gives a new kinds of Port Label Restrictions sub-TLV which mainly extend the Additional Restriction Parameters field to cover the wavelength ranges allocation information. The parameters in the Additional Restriction Parameters field include Bit Rate which indicates the bitrates of the specific wavelength range and (Start Label, End Label) tuple which indicates the wavelength range.

```
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MatrixID | RestrictionType | F | Reserved |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bit Rate |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Start Label |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| End Label |
+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3

Definition of MatrixID and RestrictionType field can be found in the document [draft-ietf-ccamp-general-constraint-encode]. Value of RestrictionType needs to be assigned by IANA.

When fixed grid network and flexible grid network exist on a single fiber, and an indication is needed to distinguish both of them. "F" bit is introduced as an indication to distinguish fixed grid (F = 0) and flexible grid (F = 1).

"Bit Rate" field indicates the bitrates of the specific wavelength range.

The tuple (Start Label, End Label) indicates the wavelength range, and the label encoding format can be found in [RFC6205] when used to represent fixed grid, and in [draft-farrkingel-ccamp-flexigrid-lambda-label] when used to represent flexible grid. Note: label encoding format in [RFC6205] can also be used by flexible grid, but it’s an unresolved problem and needs further discussion.
In some situation, modulation format information may also be needed
to help allocation wavelength range, as signals with the same
bitrates on a single fiber can use different modulation format. In
this case, modulation formats information is needed to be carried in
Port Label Restrictions sub-TLV. Wavelength that is supported by
subsystems can be partitioned to service traditional fixed grid
technology.

5. Security Considerations

TBD

6. IANA considerations

TBD.

7. References

7.1. Normative References

[RFC2119]  Bradner, S., "Key words for use in RFCs to Indicate

GMPLS and Path Computation Element (PCE) Control of
Wavelength Switched Optical Networks (WSONs)", RFC 6163,
April 2011.

7.2. Informative References

[G.694.1 v1]  International Telecommunications Union, "Draft revised
G.694.1 version 1.3".

[flexible-grid-ospf-ext]  Fatai Zhang, Xiaobing Zi, Ramon Casellas, O. Gonzalez de
Dios, and D. Ceccarelli, "GMPLS OSPF-TE Extensions in
support of Flexible-Grid in DWDM Networks",
draft-zhang-ccamp-flexible-grid-ospf-ext-00.txt.

[flexigrid-lambda-label]  D. King, A. Farrel, Y. Li, F. Zhang, and R. Casellas,
"Generalized Labels for the Flexi-Grid in Lambda-Switch-
Capable (LSC) Label Switching Routers",
draft-farrkingel-ccamp-flexigrid-lambda-label-01.txt.

Authors’ Addresses

Qilei Wang
ZTE Corporation

Email: wang.qilei@zte.com.cn

Xihua Fu
ZTE Corporation
ZTE Plaza, No.10, Tangyan South Road, Gaoxin District
Xi’an 210012
P.R.China

Email: fu.xihua@zte.com.cn
Framework for GMPLS Control of Flexible Grid Network

draft-wang-ccamp-gmpls-flexigrid-framework-01.txt

Abstract

This document provides a framework for applying Generalized Multi-Protocol Label Switching (GMPLS) and the Path Computation Element (PCE) architecture to control the flexible grid network base on the Wavelength Switched Optical Networks (WSONs). GMPLS control of WSON which is addressed in RFC6163 is out of the scope of this document.

This document focuses on the topological elements changes and new path selection constraints that flexible grid technology takes. Impairments related technology is not covered in this document.

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 12, 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
Table of Contents

1. Introduction .................................................... 3
   1.1. Conventions used in this document ....................... 3
2. Terminology ...................................................... 4
3. Flexible Grid Networks .......................................... 4
   3.1. Flexible Grid Network ................................... 4
   3.2. WDM Links ................................................. 5
   3.3. Optical Transmitters and Receivers ..................... 5
   3.4. Optical Signals in Flexible Grid Network ............... 6
      3.4.1. Optical Tributary Signals ......................... 7
      3.4.2. WSON Signal Characteristics ....................... 7
   3.5. ROADMs, OXCs, Splitters, Combiners, and FOADMs ......... 7
      3.5.1. Reconfigurable Optical Add/Drop Multiplexers, OXCs
             and FOADM ........................................... 8
      3.5.2. Splitters and Combiners ................................ 9
      3.6. Electro-Optical Systems ................................ 9
4. Routing and wavelength Assignment in flexible grid network .... 10
5. GMPLS and PCE Control ........................................... 11
   5.1. Extension to GMPLS Signaling ............................ 11
   5.2. Extension to GMPLS Routing ................................ 11
      5.2.1. Available Wavelength Range ......................... 12
      5.2.2. Port Label Restriction .............................. 12
   5.3. Optical Path Computation and Implications for PCE ....... 13
      5.3.1. Optical Path Constraints and Electro-Optical
             Element Signal Compatibility .................... 13
      5.3.2. Discovery of RWA-Capable PCEs .................... 14
      5.3.3. Use of GCO ........................................ 14
6. Security Considerations ......................................... 14
7. References ....................................................... 14
   7.1. Normative References .................................... 14
   7.2. Informative References .................................. 14
Authors’ Addresses .................................................. 15
1. Introduction

Flexible grid is a new DWDM application which is defined in the newest version of [G.694.1]. Compared to traditional fixed grid network, a flexible grid network can select its data channels with arbitrary slot width, and mainly be used to setup path with higher bitrates (e.g., 100G or 400G or higher). Whereas traditional fixed grid DWDM technology always uses fixed slot width and is mainly used to setup path with lower bitrates signals. Flexible grid network is also a WDM-based optical network in which switching is performed selectively based on the center wavelength of optical channels, which means flexible grid channels can be represented as a lambda capable switching LSP by center wavelength and slot width from the control plane perspective.

Wavelength Switched Optical Network (WSON) which is addressed in [RFC6163] is the application of Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] operation to traditional fixed grid WDM network. As flexible grid network is a new WDM network which evolves from traditional fixed grid network, GMPLS also can be used to operate flexible grid network. Similar to fixed grid network, 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, which have flexible grid characteristics. WSON specific descriptions are addressed in [RFC6163] and are out of the scope of this document. People who are interested in this document are supposed to be familiar with [RFC6163].

This document provides a framework for applying the GMPLS architecture and protocols [RFC3945] and the PCE architecture [RFC4655] to the control and operation of flexible grid networks. In order to help GMPLS and PCE use for flexible grid network, this document first focuses on the subsystems and characteristics information that flexible grid network brings and then modeled the characteristics information by GMPLS and PCE. This work will help facilitate the development of protocol solution models and protocol extensions within the GMPLS and PCE protocol families.

1.1. Conventions used in this document

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

- Flexible Grid: a new WDM technology different from traditional fixed grid DWDM technology defined with the aim of allowing flexible optical spectrum management, in which the Slot Width of the wavelength ranges allocated to different channels are flexible (variable sized).

- Wavelength Range: [RFC6163] gives a description of this terminology. Wavelength range given a mapping between labels and the ITU-T grids, each range could be expressed in terms of a tuple, (lambda1, lambda2) or (freq1, freq2), where the lambdas or frequencies can be represented by 32-bit integers.

- Frequency slot: The definition in [G.694.1] is shown here. The frequency range allocated to a channel and unavailable to other channels within a flexible grid. A frequency slot is defined by its nominal central frequency and its slot width.

- Slot width: The full width of a frequency slot in a flexible grid.

3. Flexible Grid Networks

Wavelength Switched Optical Network (WSON) related documents cover the constraints information that needs to be considered in the process of path computation. Emergence of flexible grid DWDM technology raises some new characteristics and these new characteristics should be modeled by GMPLS and PCE from the perspective of control plane in order to help path computation. This document mainly focus on the flexible grid subsystems’ characteristics information and constraints information that impact the flexible grid path selection process (i.e. wavelength selection). Subsequent sections review and model flexible grid characteristics that need to be emphasized by control plane and these sections follow the sequence of the section addressed in [RFC6163].

3.1. Flexible Grid Network

As described in the newest version of [G.694.1], flexible DWDM grid allows frequency slots have a nominal central frequency (in THz) defined by: 193.1 + n x 0.00625 where n is a positive or negative integer including 0 and 0.00625 is the nominal central frequency granularity in THz and a slot width defined by: 12.5 x m where m is a positive integer and 12.5 is the slot width granularity in GHz. Any combination of frequency slots is allowed as long as no two frequency slots overlap.
3.2. WDM Links

According to the review of the newest version of [G.694.1], the nominal central frequencies for the flexible grid network are defined with a granularity of 6.25 GHz and the frequency slot widths are defined as a multiple of 12.5 GHz. A label representation which includes the information of central frequency and slot width is needed to provide a common label format to be used in signaling optical paths. The flexible grid labels can also be used to describe WDM links, ROADM ports, and wavelength converters for the purposes of path selection.

As described in section 3.1 of [RFC 6163], putting WDM over different types of fiber require significant engineering and a fairly limited range of wavelengths. Parameters that include wavelength range and channel spacing is needed to perform basic, impairment-unaware modeling of a WDM link.

- Wavelength range: wavelength range can be used to give a mapping between labels and the flexible grid and each range could be expressed in terms of a tuple, (\( \lambda_1, \lambda_2 \)) or (freq1, freq2). Maybe new label representation is needed to describe wavelength range.

- Channel Spacing: since flexible grid can provide a granularity of 6.25GHz, this new channel spacing value needs to be added.

In addition to the wavelength range and channel spacing, indication SHOULD also be added to indicate the link support flexible grid DWDM technology.

As indicated in [RFC6163], this information is relatively statically for a particular link as changes to these properties generally require hardware upgrades. Such information may be used locally during wavelength assignment via signaling.

3.3. Optical Transmitters and Receivers

Similar to WSON, flexible grid WDM optical systems make use of coupled optical transmitters and receivers to setup LSC LSP. In the case of an optical network without wavelength converters, an optical path needs to be routed from source transmitter to sink receiver and must use a single wavelength. Flexible grid brings some new characteristics to transmitters and receivers compare to traditional fixed grid characteristics like "Tunable", "Tuning range", "Tuning time" and "Spectral characteristics and stability" which are addressed in [RFC6163] for fixed grid. This section examines the new
characteristics that would impact optical transmitters and receivers in the process of control plane path computation. Modeling parameters for flexible grid optical transmitters and receivers from the control plane perspective are:

- **Tuning range**: As described in [RFC6163], this is the frequency or wavelength range over which the optics can be tuned. \((\text{lambda}_1, \text{lambda}_2)\) or \((\text{freq}_1, \text{freq}_2)\) can be used to represent the wavelength range, where \(\text{lambda}_1\) and \(\text{lambda}_2\) or \(\text{freq}_1\) and \(\text{freq}_2\) are the labels representing the lower and upper bounds in wavelength. As nominal central frequencies can’t be figured out before the path setup in flexible grid network and flexible grid label may be different from fixed grid label, "Tuning range" may be encode with some different format from traditional fixed grid technology.

- **Slot width**: this parameter indicates slot width needed by a transmitter or receiver and SHOULD be considered in the process of path computation.

When an end-to-end LSC LSP needs to be setup, operator first sends a path setup command which convey some characteristics information of the LSP, such as bitrates, to the source node. Path setup request is sent to path computation element to computes an end-to-end LSC LSP with specific slot width information, which bases on the bitrates and modulation format that transceiver and receiver support.

### 3.4. Optical Signals in Flexible Grid Network

Similar to the fixed grid switching (e.g., WSON), the fundamental unit of switching in flexible grid is also a "wavelength". The transmitters and receivers in these networks will deal with one wavelength at a time, while the switching systems themselves can deal with multiple wavelengths at a time. Key non-impairment-related parameters which are listed in [RFC6163] are shown below:

- **(a) Minimum channel spacing (GHz)**
- **(b) Minimum and maximum central frequency**
- **(c) Bitrates/Line coding (modulation) of optical tributary signals**

As described in [RFC6163], (a) and (b) are considered properties of the link and restrictions on the GMPLS Labels while (c) is a property of the "signal". For the purposes of modeling the flexible grid, new parameters which are related to the properties of the link and...
restrictions and property of "signal" SHOULD be considered:

- (d) Minimum and Maximum Slot Width
- (e) Slot Width

(d) is considered properties of the link and restrictions on the GMPLS Labels, and description can be found in the following section. (e) is a property of the "signal" and this property is determined by the transmitter and may be changed if signal traverse an OEO.

3.4.1. Optical Tributary Signals

In [RFC6163], "optical tributary signal classes" are characterized by a modulation format and bitrates range and both of them are key parameters in characterizing the optical tributary signal. Note that, with advances in technology, optical tributary signal classes that support flexible grid would be added.

For optical tributary signals in flexible grid, bitrates range and modulation format are still two key parameters, as a single wavelength with central frequency and slot width used by a signal sent from transmitter can be deduced from these two parameters base on the available wavelength and slot width range from the source to the destination.

3.4.2. WSON Signal Characteristics

Description about WSON signal characteristics in [RFC6163] also can be applied to this document. Fundamental unit of switching in flexible grid network is also "wavelength". WSON signal characteristics like optical tributary signal class (modulation format), forward error correction (FEC), central frequency (wavelength), bitrates and general protocol identifier (G-PID) are still used in flexible grid network in the process of path computation and some more modulation formats and FECs may be added to describe flexible grid network signal characteristics.

Except the parameter that have been included in [RFC6163], the parameter slot width is also needed here to specify the slot width that signal occupies.

3.5. ROADMs, OXCs, Splitters, Combiners, and FOADMs

This section mainly focuses on optical devices such as ROADMs, Optical Cross-Connects (OXCs), splitters, combiners, and Fixed
Optical Add/Drop Multiplexers (FOADMs) which can be used in flexible grid network and examines their parameters of these devices that can be used in the process of control plane path computation.

3.5.1. Reconfigurable Optical Add/Drop Multiplexers, OXCs and FOADM

A picture is shown here to facilitate the description of ROADM. ROADM is composed of WSSes (wavelength selective switch) and splitters which are used massively in current WDM network. WSS can be used to select the wavelength on the line side output port and splitter can be used on the line side input port to split the income wavelength.

Switched connectivity matrix is needed to show whether a wavelength on input port can be connected to an output port internal.

Besides the switched connectivity matrix which is applied to line side port and tributary side port included in [RFC6163], new wavelength restriction of the line side port on a ROADM which are brought by flexible grid are considered below:
Available wavelength range:

This parameter indicates the available wavelength that can be allocated to a LSP. (\(\lambda_1, \lambda_2\)) or \((\text{freq}_1, \text{freq}_2)\) can be used to represent the available wavelength range.

Maximum/Minimum slot width that a port support

This is an inherent attribution of the network subsystems, like WSS, and can be treated as port label restriction. Requirements and descriptions about the restrictions information can be found in [draft-wangl-ccamp-ospf-ext-constraint-flexi-grid]. For flexible grid subsystem’s ports, the possible values of slot width are within the range \([\text{Minimum Slot Width}, \text{Maximum Slot Width}]\) and with the slot width granularity of \(2 \times \text{C.S.}\) (Channel Spacing). The combination of C.S. and \([\text{Minimum Slot Width}, \text{Maximum Slot Width}]\) can represent any slot width that ROADM support.

Wavelength Range allocation

The whole wavelength that ROADM support can be partitioned into several wavelength ranges, and one wavelength range can only be used for paths setup with the specific bit rate and/or modulation format. The advertisement of this restrictions information will help reduce fragments in flexible grid network. Requirements related description can be found in [draft-wang-ccamp-flexible-grid-wavelength-range-ospf-te]. This is an optional requirement.

These restrictions information can also be applied to fixed optical Add/Drop Multiplexers.

3.5.2. Splitters and Combiners

Nothing is new except switched connectivity matrix and this has been addressed in [RFC6163].

3.6. Electro-Optical Systems

Some words can be found in [RFC6163]. OEO switches, wavelength converters, and regenerators all share a similar property: they can be more or less "transparent" to an "optical signal" depending on their functionality and/or implementation. Properties can be applied to flexible grid, and these properties can satisfy path computation without taking any new characteristics into consideration. Modeling of OEO switches, wavelength converters and regenerators can also be applied to flexible grid.
Regenerator can be used to restore signal quality. Bitrates range and modulation formats that the regenerator support need to be taken into consideration to help path computation, whereas slot width do not (May be someone will talk about slot width). If one regenerator is designed to handle signal with specific bitrates and modulation formats, then it would support the corresponding slot width because slot width can be derived by modulation format and bitrates. Even if the slot width is changed by the electro-optical systems due to the change of modulation format, the slot width that has already changed may not be explicitly specified because bitrates and modulation format are explicitly specified.

4. Routing and wavelength Assignment in flexible grid network

This section briefly describes the constraints information of routing and wavelength assignment in the flexible grid network. Similar to WSON, the input to basic RWA in flexible grid network are the requested optical path’s source and destination, the network topology, the locations and capabilities of any wavelength converters, the wavelengths available on each optical link and port label constraints information such as slot width range that a port support and wavelength range partition information by bitrates and/or modulation formats. The output that provided by RWA in flexible grid network are an explicit route through ROADMs, a wavelength for optical transmitter, the slot width that this wavelength occupies, and a set of locations (generally associated with ROADMs or switches) where wavelength conversion is to occur and the new wavelength to be used on each component link after that point in the route. Similar to WSON, an optical flexible grid path that from source to destination also must use a single wavelength that is available along that path without “colliding” with a wavelength used by any other optical path that may share an optical fiber.

In [RFC6163], three different ways of performing RWA in conjunction with the control plane are shown here:

1) Combined RWA

2) Separated R and WA (R + WA)

3) Routing and Distributed WA (R + DWA)

These ways can also be applied to flexible grid control plane path computation. Related description about these three architectures can be found in section 4.1 of [RFC6163].
5. GMPLS and PCE Control

Flexible grid brings some new characteristics to WDM network, and consequently WSON would add some extensions or change in order to control the flexible grid network. Extensions to GMPLS signaling, routing and PCE are described in this section.

5.1. Extension to GMPLS Signaling

Support for WSON signaling exists in [RFC3471], [RFC4328] and [draft-ietf-ccamp-wson-signaling]. However, a number of practical issues arise in the identification of wavelengths and signals in wavelength assignment in flexible grid.

A mapping between label and wavelength is needed to simplify the characterization of WDM links and WSON devices. The mapping like the one described in [draft-farrkingel-ccamp-flexigrid-lambda-label] provides label and wavelength mapping for communication between PCE and WSON PCCs. Different LSP may occupy different slot width if paths have different bitrates and modulation format in flexible grid network. So in the flexible grid network, not only central frequency is needed, but also slot width SHOULD be included to identify a channel in the process of path setup in flexible grid network.

GMPLS Signaling should be able to convey the central frequency and slot width information that a LSC LSP occupies. If the slot width is changed due to the change of modulation format, signaling should also be able to express this. Except methods that are specified in [draft-farrkingel-ccamp-flexigrid-lambda-label], [draft-hussain-ccamp-super-channel-label] and [draft-zhang-ccamp-flexible-grid-rsvp-te-ext] also provide methods to carry central frequency and slot width information in the process of signaling.

Note: extension to GMPLS signaling SHOULD be compatible with current signaling protocol.

5.2. Extension to GMPLS Routing

The following subsystem’s properties are needed by IGP to minimally characterize WSON, also these properties are needed to characterize flexible grid control plane. This section addresses the constraints information needed to model flexible grid from the control plane perspective base on the Wavelength Switched Optical Network (WSON).
1) WDM link properties (allowed wavelengths)

2) Optical transmitters (wavelength range)

3) ROADM/FOADM properties (connectivity matrix, port wavelength restrictions)

4) Wavelength converter properties (per network element, may change if a common limited shared pool is used)

Here 1, 2 and 3 are re-considered in the flexible grid network.

5.2.1. Available Wavelength Range

Wavelengths available on WDM link and port of optical transmitters are advertised through routing protocol, the wavelengths available information can be used by path computation element to compute a suitable end-to-end LSP. As different flexible grid channels always have different slot widths and channels’ central frequency position and slot width can’t be decided in advance, so mapping between label and wavelength may not be able to use the representation similar to [RFC6205] to represent every channel. Maybe new label formats and representation of wavelength available are needed in routing protocol to transfer IGP information between nodes and PCEs. Extensions to label set field SHOULD be able to represent the wavelength available validly in flexible grid network. Allowed wavelengths on WDM link and wavelength range on optical transmitters need to adapt to this change.[draft-dhillon-ccamp-super-channel-ospfte-ext], [draft-wangl-ccamp-ospf-ext-constraint-flexi-grid] and [draft-zhang-ccamp-flexible-grid-ospf-ext] give some different methods to represent the available wavelengths.

5.2.2. Port Label Restriction

Some new ROADM/FOADM properties brought by flexible grid need to be advertised by routing protocol in order to help path computation. In the section 3, properties of ROADM/FOADM are described as the port label restrictions information.

The first one, maximum/minimum slot width supported on one port need to be advertised. This slot width constraint information of a port (i.e., available slot width range of a WSS) SHOULD be known by path computation element in order to compute a suitable path. According to [draft-wangl-ccamp-ospf-ext-constraint-flexi-grid], combination of C.S. and [Minimum Slot Width, Maximum Slot Width] can represent any slot width that ROADM support. LMP can be run between two neighbor nodes to negotiate these attributes and related extension can be found in [draft-li-ccamp-grid-property-lmp]. This is optional
because routing protocol can also be used to deal with it.

The second one, wavelength range allocation information of ROADM/FOADM needs to be advertised through routing protocol. Grouping of wavelength of the same bitrates and/or modulation formats would help reduce fragments. Channels in the same wavelength range with the same bitrates looks almost like fixed grid technology, and they won’t generate much fragment in the path setup and release because every channel use the same slot width. Requirements of wavelength range allocation and protocol extensions can be found in [draft-wang-ccamp-flexible-grid-wavelength-range-ospf-te].

5.3. Optical Path Computation and Implications for PCE

Extensions to PCEP can be found in [draft-lee-pce-wson-rwa-ext] base on Wavelength Switched Optical Network. Emergence of flexible grid brings some extension to current draft. PCEP SHOULD be able to support flexible grid path computation.

5.3.1. Optical Path Constraints and Electro-Optical Element Signal Compatibility

Flexible grid may not change the computation architectures of WSON, but new constraints information SHOULD be taken into consideration in the process of path computation. According to the description in [RFC6163], when requesting a path computation to PCE, the PCC should be able to indicate:

1) The G-PID type of an LSP
2) The signal attributes at the transmitter and receiver.

And the PCE should be able to respond to the PCC with the following:

1) The conformity of the requested optical characteristics associated with the resulting LSP with the source, sink, and NE along the LSP.
2) Additional LSP attributes modified along the path.
3) Slot width of the LSP. This should be respond to the PCC as flexible grid channels always have different slot widths. Slot width information may be contained in the wavelength object which is carried in PCRep message from PCE to PCC.
5.3.2. Discovery of RWA-Capable PCEs

Not all PCEs within a domain would necessarily need the capability of flexible grid path computation. Therefore, it would be useful to indicate that a PCE has the ability to deal with flexible grid via the discovery mechanisms being established for PCE discovery in [RFC5088]. Extensions to [RFC5088] are needed to achieve this goal.

5.3.3. Use of GCO

Though GCO is able to reduce the fragment of the wavelength or spectrum, it is hard to be implemented in the network, because GCO would involve massive LSPs and disturb current service. As fragment can be reduced through early wavelength or spectrum allocation planning, GCO maybe avoided.

6. Security Considerations

TBD

7. References

7.1. Normative References


7.2. Informative References

[G.694.1 v1] International Telecommunications Union, "Draft revised G.694.1 version 1.3".


[flexible-grid-requirements] Fatai Zhang, Xiaobing Zi, O. Gonzalez de Dios, and Ramon Casellas, "Requirements for GMPLS Control of Flexible Grids".

Qilei Wang & Xihua Fu Expires September 12, 2012 [Page 14]
draft-zhang-ccamp-flexible-grid-requirements-01.txt

[flexible-grid-rsvp-te]
Fatai Zhang, O. Gonzalez de Dios, and D. Ceccarelli,
"RSVP-TE Signaling Extensions in support of Flexible Grid",
draft-zhang-ccamp-flexible-grid-rsvp-te-ext-00.txt

[flexigrid-lambda-label]
D. King, A. Farrel, Y. Li, F. Zhang, and R. Casellas,
"Generalized Labels for the Flexi-Grid in Lambda-Switch-Capable (LSC) Label Switching Routers",
draft-farrkingel-ccamp-flexigrid-lambda-label-01.txt

[ospf-ext-constraint-flexi-grid]
L Wang, Y Li, "OSPF Extensions for Routing Constraint Encoding in Flexible-Grid Networks",
draft-wangl-ccamp-ospf-ext-constraint-flexi-grid-00.txt

[super-channel-label]
Iftekhar Hussain, Abinder Dhillon, Zhong Pan, Marco Sosa and Bert Basch, Steve Liu, Andrew G. Malis, "Generalized Label for Super-Channel Assignment on Flexible Grid",
draft-hussain-ccamp-super-channel-label-02.txt

[super-channel-ospfte]
Abinder Dhillon, Iftekhar Hussain, Rajan Rao, Marco Sosa, "OSPFTE extension to support GMPLS for Flex Grid",
draft-dhillon-ccamp-super-channel-ospfte-ext-02.txt

Authors’ Addresses

Qilei Wang
ZTE Corporation

Email: wang.qilei@zte.com.cn

Xihua Fu
ZTE Corporation
ZTE Plaza, No.10, Tangyan South Road, Gaoxin District
Xi’an
P.R.China

Email: fu.xihua@zte.com.cn
OSPF Extensions for Routing Constraint Encoding in Flexible-Grid Networks
draft-wangl-ccamp-ospf-ext-constraint-flexi-grid-01

Abstract

In Flexible-Grid networks, network elements and links may impose additional routing constraints, which cannot be ignored in Routing and Spectrum Assignment (RSA) process. This document describes the requirements of such constraints, and then provides efficient encodings to specify how the information is carried.

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 10, 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 ....................... 3
3. Terminologies ............................................. 3
4. Requirements of Routing Constraint for RSA in Flexible-Grid Networks ............................................. 4
   4.1. Label set ............................................... 8
   4.2. Flexible-Grid Ability Constraint ..................... 8
   4.3. Optical Signal Compatibility Constraint ............. 9
5. Encoding ..................................................... 10
   5.1. Label Set ............................................... 10
   5.2. Flexible-Grid Ability Constraint ..................... 13
   5.3. Optical Signal Compatibility Constraint ............. 15
6. Encoding Example ............................................ 16
   6.1. Example of Label Set Encoding ....................... 16
   6.2. Example of Flexible-Grid Ability Constraint Encoding ............................................. 19
   6.3. Example of Signal Compatibility Encoding ........... 19
7. Security Considerations .................................... 20
8. IANA Considerations ....................................... 20
9. References ............................................... 20
   9.1. Normative References ................................ 20
   9.2. Informative References ............................... 20
Authors’ Addresses ........................................... 21
1. Introduction

Flexible-Grid technique breaks the rigid nature of traditional DWDM wavelength Grid, and enables flexible allocation of optical spectrum resources to accommodate ultra-high data rate traffic. Currently, there are several IETF draft addressing GMPLS routing and signaling extension to support Flexible-Grid DWDM Networks, such as [I-D.farrkingel-ccamp-flexigrid-lambda-label][I-D.li-ccamp-flexible-grid-label] [I-D.zhang-ccamp-flexible-grid-requirements][I-D.zhang-ccamp-flexible-grid-rsvp-te-ext]
[I-D.zhang-ccamp-flexible-grid-ospf-ext][I-D.hussain-ccamp-super-channel-label] [I-D.dhillon-ccamp-super-channel-ospf-te-ext][I-D.zhang-ccamp-flexi-grid-ospf-te-ext]. However, all these documents mainly focus on Label/Label-set extensions in Flexible-Grid Networks, and take spectral continuity and consecutivity into account, but ignore other aspects of RSA problem. In fact, Network elements (such as nodes and Optical-to-Electronic/Electronic-to-Optical sub-systems) and links may impose additional routing constraints such as flexible-grid ability/slot width range limitations on ports, asymmetric switch connectivity, and signal processing limitations of each OE/EO system. Without considering these constraints, it cannot be guaranteed to obtain available results in RSA process especially for network scenarios with various Flexible-Grid and Fixed-Grid elements, which leads to inefficient routing and high blocking probability of end-to-end paths.

This document describes the requirements of RSA, and then encodes the constraints imposed by network elements and links, which could be carried in OSPF Messages to flood to each node for efficient RSA. In addition, such information could be conveyed by other mechanisms to a Path Computation Element (PCE). Note that, impairment-related constraints are not considered here.

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. Terminologies

Center Frequency Granularity (CFG): The minimum step by which the center frequency of optical bandwidth can be increased or decreased.

Frequency slot: The frequency range allocated to a slot and
4. Requirements of Routing Constraint for RSA in Flexible-Grid Networks

In Flexible-Grid network, there is one key problem: how to route and allocate spectrum resources for each end-to-end optical channel, so to fulfill their requirements in an efficient way? To address this problem, some constraints must be taken into consideration, which are listed as follows.

- Spectrum availability constraint.
- Flexible-Grid ability constraint.
- Asymmetric switch connectivity constraint.
- Optical signal compatibility constraint.
- Other constraints.

The asymmetric switch connectivity constraint in Flexible-Grid network could be well addressed by Connectivity matrix sub-TLV used
in Wavelength Switched Optical Networks (WSON)
[I-D.ietf-ccamp-general-constraint-encode]. The spectrum
availability constraint is studied in several drafts
[I-D.lli-ccamp-flexible-grid-label]
[I-D.zhang-ccamp-flexible-grid-ospf-ext][I-D.dhillon-ccamp-super-channel-ospfte-ext], and could be represented by Label-set extensions. However, these extensions are not complete, so we reorganize the Flexible-Grid label-set according to WSON definition. In addition,
Flexible-Grid ability constraint (including grid type and slot width granularity/range) and optical signal compatibility constraint are also necessary for efficient RSA, but few document takes these into account. We will describe the requirements and encodings of such constraints in this draft.

Here a general scenario of Flexible-Grid Network is given in order to illustrate these requirements.

```
+----+A-E2    B-I1+----+B-E2    C-I1++----+
| A  |----------->| B  |----------->| C  |
   ------------   ------------
+----+A-I2    B-E1+----+B-I2    C-E1++----+
   |        O    |        O
   A-I1    A-E1    B-I3    B-E3    C-I2    C-E2
   D-E1    D-I1    E-E3    E-I3    F-E2    F-I2
   |        O    |        O
+----+D-E2    E-I1+----+E-E2    F-I1++----+
| D  |----------->| E  |----------->| F  |
   ------------   ------------
|        |<-------|        |
+----+D-I2    E-E1+----+E-I2    F-E1++----+
```

Figure 1. A sample network with both Fixed-Grid and Flexible-Grid elements
Figure 2 shows the network topology, while Figure 2 shows the architecture of nodes. The ROADM of Figure 2 is composed of WSSs and splitters. I1˜4/E1˜4 are line-side input/output ports, while I5˜8/E5˜8 are tributary-side add/drop ports to/from line-side 1˜4 respectively. The configuration of each line-side output port is shown as follows:
The granularity denotes the slot width granularity. The Min-width and Max-width denote the slot width range. There are three types of nodes: Node A, node D and node E are Flexible-Grid ROADMs, which only consist of Flexible-Grid elements; Node C is a Fixed-Grid ROADM, which only consists of Fixed-Grid elements; Node B and Node F are Mixed-Grid ROADMs, which consist of both Flexible-Grid and Fixed-Grid Elements. Both Flexible-Grid ROADM and Mixed-Grid ROADM can support Flexible-Grid LSPs to accommodate ultra-high data rate traffic such as beyond 100G. In addition, the Fixed-Grid ROADM can be smoothly updated to Mixed-Grid ROADM by adding Flexible-Grid ports. With appropriate RSA, the network is able to support both Fixed-Grid LSPs and Flexible-Grid LSPs in an efficient way.
4.1. Label set

In Flexible-Grid networks, the spectrum assignment is not a local matter due to spectral consecutiveness and continuity constraints, so it is needed to get the information of which slice may or may not be used on each link and node port along the path in RSA process. For example, in the network of Figure 1, when a LSP request from node A to node E with 150GHz slot width and route A->B->E arrives, the label restriction of input port A-I6, output port E-E7, switch port A-E2, B-I1, B-E3, E-I3 and spectrum availability of link AB, BE must be got for the spectrum assignment. All the information is described by the label set objects which is decided by the label format. The generalized label for the flexible grid can be referred to [I-D.farrkingel-ccamp-flexigrid-lambda-label] including central frequency and slot width information.

As specified in [I-D.li-ccamp-flexible-grid-label] in section 4.1, this kind of label format is backward compatible to support the traditional 5 ways of wavelength label set encoding [I-D.ietf-ccamp-general-constraint-encode].

- 1. Inclusive list
- 2. Exclusive list
- 3. Inclusive range
- 4. Exclusive range
- 5. Bitmap set

It can be seen that these 5 types of representations can be easily inherited by incorporating the new flexible label into the object. Note that in the procedure of flooding, any combination of the 5 types of label sets is feasible.

4.2. Flexible-Grid Ability Constraint

Flexible-Grid ability may include the grid type (Fixed-Grid or Flexible-Grid) and slot width granularity/range. This information can be seen as the attribution of network ports with relations to links or nodes. The RSA requirements of such fields are listed as follows:

Firstly, Flexible-Grid WSSs of different companies or product-types may have different slot width granularity and range, which may be a subset of possible values specified by ITU-T [G.694.1v2], so it should be taken into consideration in RSA process to avoid invalid
route selection. For example, in the network of Figure 1, when a Flexible-Grid LSP request from node A to node E with 250GHz slot width arrives, only the optical channel with a route A→D→E is able to carry the traffic due to the slot width range limitations on other ports.

In addition, the slot width granularity of network elements may impact the spectral efficiency. For example, when a Flexible-Grid LSP request from node A to node E with 87.5GHz slot width arrives, 100GHz Slot width must be assigned for the route A→D→E due to 25GHz slot width granularity, which performs poor in spectral efficiency.

Furthermore, although Flexible-Grid technology may offer full backwards compatibility with the standard ITU-T DWDM grids, it is a cost-efficient way to consider Flexible-Grid Ability constraints in RSA process for Fixed-Grid requirements. For example, in the network of figure 1, when a Fixed-Grid LSP request from node B to node F with 50GHz slot width arrives, it is a better route of B→C→F than the route B→E→F, because that flexible-Grid WSSs are more expensive than fixed-grid ones, and routing fixed-Grid requests on fixed-Grid elements could leave the Flexible-Grid elements and related spectrum resources to subsequent high data rate traffic.

4.3. Optical Signal Compatibility Constraint

Optical Signal Compatibility Constraint includes the signal processing ability (for example, data rate, FEC and modulation format) and modulation-related minimum slot width for each Optical-to-Electronic (OE)/Electronic-to-Optical (EO) subsystem. The RSA requirements of such fields are listed as follows:

Firstly, as described in [I-D.ietf-ccamp-rwa-wson-encode], OE/EO subsystems may be limited to process only certain types of optical signal in WSON or Flexible-Grid networks, so it is necessary to get sufficient information characterizing OE/EO elements in RSA process to determine the signal compatibility along the path. Examples of such subsystems include transponders, regenerators and so on.

In addition, for each Flexible-Grid Label Switch Path, the required slot width is determined by the attribution of optical signal. However, a client only requests "data rate" as its traffic parameter but do not care "slot width", so it is needed to establish the mapping relations between data-rate/modulation-format and slot width, which should be reflected in optical signal compatibility constraint. For example, in the network of Figure 1, when a LSP request from node A to Node E with 100Gbit/s data rate arrives, and both the transmitter of node A and the responder of node E support optical tributary signal class DP-QPSK 100G with the same FEC and
corresponding slot width 50GHz, the minimum slot width required by
this LSP should be 50GHz.

5. Encoding

5.1. Label Set

The general format for a label set is in accordance with that in
[I-D.ietf-ccamp-general-constraint-encode], with a new flag G (1bit)
representing the grid type of label sets (1~Flexible-Grid DWDM;
0~Fixed-Grid DWDM):

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|G| Act.|    Num Labels         |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         start Label                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         start Label(continue)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:       Additional fields as necessary per action          :
:                                                               :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

the label format is in accordance with that in
[I-D.farrkingel-ccamp-flexigrid-lambda-label].

In the case of Inclusive/Exclusive label lists (0/1), the label set
format is given 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1| 0or1| Num Labels (not used) |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         First Label                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         First Label(continue)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Last  Label                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Note that one label set may contain multiple labels. The lowest/highest frequency of the K-th label is calculated as follows:

Lowest frequency_k = (central frequency_k) - (slot width_k)/2
= (193.1 + n_k * C.S.) - (2 * C.S. * m_k)/2
= (193.1 + (n_k - m_k) * C.S.) THz;

Highest frequency_k = Lowest frequency_k + slot width_k
= (193.1 + (n_k + m_k) * C.S.) THz;

In the case of Inclusive/Exclusive label ranges (2/3), the label set format is given as follows:

```
0                   1                   2                   3                   
+------------------+------------------+------------------+------------------+------------------+
| 1 | 2or3 | Num Labels(not used) |             Length            |
+------------------+------------------+------------------+------------------+
|                    Start Label #1                             |
+------------------+------------------+------------------+
|                    Start Label #1(continue)                   |
+------------------+------------------+
|                     End Label #1                              |
+------------------+
|                     End Label #1(continue)                    |
+------------------+
```

Note that one label set may contain multiple label ranges. The value
of m in start/end label in meaningless on the label set, however, in order to keep the integrity of labels and avoid misunderstanding, it is set to default value: m = (slot width granularity)/12.5GHz.

The lowest/highest frequency of the K-th label range is calculated as follows:

Lowest frequency_k = (central frequency_kstart) - (slot width granularity)/2
= (193.1 + n_kstart * C.S.) - C.S.
= (193.1 + (n_kstart - 1) * C.S.) THz;

Highest frequency_k = (central frequency_kend) + (slot width granularity)/2
= (193.1 + n_kend * C.S.) + C.S.
= (193.1 + (n_kend + 1) * C.S.) THz;

In the case of bitmap (4), the label set format is given as follows:

```
  0  1  2  3  4  5  6  7  8  9  0  1  2  3  4  5  6  7  8  9  0  1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|  4  | Num Labels  | Length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Start Label                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Start Label(continue)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Bit Map Word #1 (Lowest numerical labels)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Bit Map Word #N (Highest numerical labels)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Based on [I-D.ietf-ccamp-general-constraint-encode], Num labels denote the number of slices represented by the bit map; where the slice denotes the basic slot unit, and the slot width of one slice is equal to the slot width granularity. As there may exist some situations that the unused bandwidth between two occupied bandwidth is odd times of the central frequency granularity (not integral times of the slot with granularity), two bits are needed to represent a
single slice. Each bit in the bit map represents a particular label of half a slice with a value of 1/0 indicating whether the part is in the set or not. Bit position zero and one represent the lowest slice and corresponds to the start label. The lowest/highest frequency of label range represented by bit position $K$ is calculated as follows:

$$\text{Lowest frequency}_k = (\text{central frequency}_\text{start}) + (K - 1) \times \frac{\text{slot width granularity}}{2}$$

$$= (193.1 + n_{\text{start}} \times C.S.) + (K - 1) \times C.S.$$ 

$$= 193.1 + (n_{\text{start}} + K -1) \times C.S.;$$

$$\text{Highest frequency}_k = \text{Low frequency}_k + C.S.$$ 

$$= 193.1 + (n_{\text{start}} + K) \times C.S.$$ 

The size of the bit map is $(2 \times \text{Num Label})$ 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) and beyond SHOULD be set to zero and MUST be ignored" [I-D.ietf-ccamp-general-constraint-encode].

5.2. Flexible-Grid Ability Constraint

To accommodate the feature of Flexible-Grid Ability constraint, we extend the Port Label Restriction sub-TLV defined in [I-D.ietf-ccamp-general-constraint-encode] for Flexible-Grid networks:

```
0                   1                   2                   3
+-----------------------------------------------+-----------------------------------------------+
| MatrixID | RstType = 5 |        Reserved               | Grid | C.S. | Reserved | Min-Width | Max-Width |
+-----------------------------------------------+-----------------------------------------------+
```

In WSON network, Matrix ID is used to represent "either the value in the corresponding Connectivity Matrix sub-TLV or takes the value OxFF to indicate the restriction applies to the port regardless of any Connectivity Matrix" [I-D.ietf-ccamp-general-constraint-encode]. RstType is used to represent the restriction type. This document defines a new RstType value to express the port Flexible-Grid Supporting Ability constraint in Flexible-Grid networks:
5: GRID_ABILITY.

The meaning of Grid and C.S. is defined in [I-D.farrkingel-ccamp-flexigrid-lambda-label], which is shown as follows:

| Grid          | Value |
|---------------+-------|
| Reserved      | 0     |
| ITU-T DWDM    | 1     |
| ITU-T CWDM    | 2     |
| Flexible DWDM | 3     |
| Any           | 4(TBA) |
| Future use    | 5-7   |

| C.S. (GHz)    | Value |
|---------------+-------|
| Reserved      | 0     |
| 100           | 1     |
| 50            | 2     |
| 25            | 3     |
| 12.5          | 4     |
| 6.25          | 5 (TBA) |
| Future use    | 6 ~ 15 |

A new Grid type "Any" is defined. the reason is explained later in this document.
"Within the fixed grid network, the C.S. value is used to represent the channel spacing, as the spacing between adjacent channels is constant. While for flexible grid situation, this field should be used to represent central frequency granularity."

[1-D.farrkingel-ccamp-flexigrid-lambda-label] Accordingly the slot width granularity is twice of the C.S..

Min-Width/Max-Width: 8bits, unsigned integer. Min-Width/Max-Width denotes the minimum/maximum slot width that the ROADM port supports, which is an inherent attribution of the network elements. The formula is shown as follows:

Minimum Slot Width (GHz) = 12.5GHz * Min-Width;
Maximum Slot Width (GHz) = 12.5GHz * Max-Width;

For flexible-Grid ports (Grid = 3), the possible values of slot width are within the range [Minimum Slot Width, Maximum Slot Width] and with the slot width granularity of 2 * C.S.; for Fixed-Grid ports (Grid = 1 or 2), Min-Width/Max-Width is meaningless and padded with 0. For any port with Grid type "any", it means that the port support any Grid type, any slot width granularity and any slot width range, so C.S. and Min-Width/Max-Width are meaningless and padded with 0. One example of such port is A-I1, which is comprised of optical splitter.

Note that, the similar field of Min-Width/Max-Width is also included in object "BW sub-TLV" proposed by [1-D.dhillon-ccamp-super-channel-ospfte-ext]. However, BW sub-TLV is mainly used to present the available label set, so it belongs to dynamic information according to [RFC6163] and should be flooded frequently whenever the link state changes (for example, after the setup/teardown of the path traversing the link). In this document, the Port Label Restriction sub-TLV with GRID_ABILITY type is regarded as relatively static information, as changes to these properties such as Grid, C.S. and Min-Width/Max-Width require hardware upgrades. It is more suitable to carry such information separated from available label set in order to alleviate unnecessary flooding.

Note that, the similar field of Min-Width/Max-Width is also included in object "BW sub-TLV" proposed by [1-D.dhillon-ccamp-super-channel-ospfte-ext]. However, BW sub-TLV is mainly used to present the available label set, so it belongs to dynamic information according to [RFC6163] and should be flooded frequently whenever the link state changes (for example, after the setup/teardown of the path traversing the link). In this document, the Port Label Restriction sub-TLV with GRID_ABILITY type is regarded as relatively static information, as changes to these properties such as Grid, C.S. and Min-Width/Max-Width require hardware upgrades. It is more suitable to carry such information separated from available label set in order to alleviate unnecessary flooding.

Other port label restrictions have no difference with that in [1-D.ietf-ccamp-general-constraint-encode].

5.3. Optical Signal Compatibility Constraint

To accommodate the feature of Optical Signal Compatibility Constraint, we extend the Modulation Type sub-TLV defined in [1-D.ietf-ccamp-rwa-wson-encode] for Flexible-Grid networks:
The meaning of S, I and Modulation ID is in accordance with that of [I-D.ietf-ccamp-rwa-wson-encode].

This document adds a new field "m" (8bit) to represent the minimum slot width requirement for corresponding Modulation ID:

Minimum Slot Width = 12.5GHz * m.

Note that the modulation type sub-TLV may contain multiple modulation IDs, which means the transmitter/responder/transponder/regennerator support multiple data rate/modulation format.

This sub-TLV establishes mapping relations between data rate/modulation format (Modulation ID) and slot width. In addition, it also provides the signal processing ability for each OE/EO element in the network. However, FEC may impact the value of m, but it is not discussed here and leaved for further study. New values of Modulation ID should be defined for ultra-high speed transmission, but it depends on transmission technique and not specified in this document.

Other signal compatibility constraints have no difference with that in [I-D.ietf-ccamp-rwa-wson-encode].

6. Encoding Example

6.1. Example of Label Set Encoding

Taking the network of figure 1 as an example, the available spectral resource of link AB is shown in figure 3.
Figure 3. Spectral resource state of link AB

In figure 3, the spectral resource is from 193.1 THz - 16 * 6.25 GHz to 193.1 THz + 10 * 6.25 GHz. For label list type, the label set format is given as follows:

```
<table>
<thead>
<tr>
<th>3</th>
<th>C.S.(5)</th>
<th>Identifier</th>
<th>n(-15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>m(1)</td>
<td>Reserved</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

For label range type, the label set format is given as follows:

```
<table>
<thead>
<tr>
<th>3</th>
<th>C.S.(5)</th>
<th>Identifier</th>
<th>n(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>m(4)</td>
<td>Reserved</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

For bitmap type, the label set format is given as follows:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------------------------------------------|
| 1 | 4 | Num Labels(26) | Length(16) |
|---------------------------------------------|
| 3 | C.S.(5) | Identifier | n(-15) |
|---------------------------------------------|
| m(1) | Reserved |
|---------------------------------------------|
| 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
+---------------------------------------------+
```
6.2. Example of Flexible-Grid Ability Constraint Encoding

Taking the network of figure 1 as an example, the Flexible-Grid ability constraint of A-E1 can be encoded as follows:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| MatrixID(0xff) | RstType(5)    |        Reserved               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| 3  |C.S.(5)      |    Reserved     |  Min-Width(4) | Max-Width(16) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

The Flexible-Grid ability constraint of A-E2 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| MatrixID(0xff) | RstType(5)    |        Reserved               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| 3  |C.S.(4)      |    Reserved     |  Min-Width(4) | Max-Width(24) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

The Flexible-Grid ability constraint of B-E2 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| MatrixID(0xff) | RstType(5)    |        Reserved               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
| 1  |C.S.(2)      |    Reserved     |  Min-Width(0) | Max-Width(0)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

6.3. Example of Signal Compatibility Encoding

Assuming an optical transmitter can support the following modulation types: optical tributary signal class DP-QPSK 100G (minimum slot width: 50GHz); optical tributary signal class DP-BPSK 100G (minimum slot width: 100GHz). The Modulation Type sub-TLV is given as follows:

```
```
7. Security Considerations

8. IANA Considerations

TBD.

9. References

9.1. Normative References


9.2. Informative References


Bernstein, G., Lee, Y., Li, D., Imajuku, W., and JR. Han, "General Network Element Constraint Encoding for GMPLS Controlled Networks", May 2011.

Bernstein, G., Lee, Y., Li, D., Imajuku, W., and JR. Han, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", October 2011.


Zhang, FT., Zi, XB., Casellas, R., Gonzales-de-Dios, O., and D. Ceccarelli, "GMPLS OSPF-TE Extensions in support of Flexible-Grid in DWDM Networks", October 2011.

Zhang, FT., Zi, XB., Gonzales-de-Dios, O., and R. Casellas, "Requirements for GMPLS Control of Flexible Grids", October 2011.

Zhang, FT., Gonzales-de-Dios, O., and D. Ceccarelli, "RSVP-TE Signaling Extensions in support of Flexible Grid", October 2011.

Authors’ Addresses

Lei Wang
ZTE
No.19, Huayuan East Road, Haidian District
Beijing  100191
P.R.China

Phone: +86 13811440067
Email: wang.lei131@zte.com.cn (hechen0001@gmail.com)
URI:   http://www.zte.com.cn/

Yao Li
ZTE
P.R.China

Phone: +86 025 52871109
Email: li.yao3@zte.com.cn
URI:   http://www.zte.com.cn/

Guoying Zhang
China Academy of Telecom Research, MIIT
No.52 Huayuan Beilu, Haidian District
Beijing  100083
P.R.China

Email: zhangguoying@mail.ritt.com.cn
GMPLS OSPF-TE Extensions in support of Flexible Grid in DWDM Networks

draft-zhang-ccamp-flexible-grid-ospf-ext-01.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/1id-abstracts.txt.

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

This Internet-Draft will expire on September 12, 2012.

Abstract

Zhang
This memo describes the OSPF-TE extensions in support of GMPLS control for flexi-grid in DWDM networks.

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 ................................................. 2
2. Terminology .................................................. 3
3. Requirements for Flexi-grid Routing .......................... 3
   3.1. Available Frequency Ranges .............................. 3
   3.2. Application Compliance Considerations ................... 5
   3.3. Comparison with Fix-grid DWDM Links ..................... 5
4. Extensions ................................................... 6
   4.1. Available Labels Set sub-TLV ............................ 6
     4.1.1. Inclusive/Exclusive Label Range .................... 7
     4.1.2. Inclusive/Exclusive Label Lists ........................ 7
     4.1.3. Bitmap ............................................. 8
   4.2. Extensions to Port Label Restriction sub-TLV ............ 8
   4.3. Examples for Label Set .................................. 9
5. IANA Considerations ......................................... 10
6. Security Considerations ..................................... 10
7. References .................................................. 10
   7.1. Normative References ................................... 10
   7.2. Informative References ................................ 11
8. Authors’ Addresses .......................................... 12

1. Introduction

[G.694.1v1] defines the Dense Wavelength Division Multiplexing (DWDM) frequency grids for WDM applications. A frequency grid is a reference set of frequencies used to denote allowed nominal central frequencies that may be used for defining applications. The channel spacing, i.e. the frequency spacing between two allowed nominal central frequencies could be 12.5 GHz, 25 GHz, 50 GHz, 100 GHz and integer multiples of 100 GHz as defined in [G.694.1v1]. All of the wavelengths on a fiber should use different central frequencies and occupy a fixed bandwidth of frequency.
[G.FLEXGRID], an updated version of [G.694.1v1] has been consented in December 2011 in support of flexi-grids. The terms "frequency slot" (The frequency range allocated to a channel and unavailable to other channels within a flexi-grid) and "slot width" (the full width of a frequency slot in a flexi-grid) are introduced to address flexi-grids. A channel is represented as a LSC (Lambda Switching Capable) LSP in the control plane, i.e. a LSC LSP should occupy a frequency slot on each fiber it traverses. In the case of flexi-grid, different LSC LSPs may have different slot width on a fiber, i.e. the slot width is flexible on a fiber.

[WSON-OSPF] defines the OSPF-TE extensions for WSON networks, which focuses on the fixed grids of DWDM. [GEN-OSPF] defines OSPF-TE extensions in support of the general network element constraints under the control of GMPLS. This document describes the additional requirements and extensions of routing protocol brought by flexi-grid.

This document uses the DWDM link model which is shown in [SSON-FWK] to describe the requirement and extensions for routing. The flexi-grid related terminologies can also refer to [SSON-FWK].

2. Terminology

Flexi-grid: See [SSON-FWK].

Slot Width: See [SSON-FWK].

Frequency Range: See [SSON-FWK].

SSON: Spectrum-Switched Optical Networks; See [SSON-FWK].

Flexi-LSP: See [SSON-FWK].

3. Requirements for Flexi-grid Routing

As described in [SSON-FWK], the main changes for routing brought by flexible grid are related to the DWDM links.

3.1. Available Frequency Ranges

In the case of flexi-grids, the central frequency steps from 193.1 THz with 6.25 GHz granularity. The central frequency is calculated as follows:

Central Frequency = 193.1 THz + n * 0.00625 THz
Where \( n \) is a positive or negative integer including 0.

Different flexi-LSPs could occupy frequency slots with different slot width. The frequency slot width of a flexi-LSP is defined as follows:

\[
\text{Slot width} = 0.0125 \text{ THz} \times m
\]

Where \( m \) is a positive integer.

The frequency slot of a flexi-LSP can be determined by the slot width and central frequency as follows.

Lowest frequency = (central frequency) - (slot width)/2
\[
= (193.1 + n \times 0.00625) - (0.0125 \times m)/2
\]

Highest frequency = (central frequency) + (slot width)/2
\[
= (193.1 + n \times 0.00625) + (0.0125 \times m)/2
\]

On a DWDM link, the frequency slots must not overlap with each other. However, the border frequencies of two frequency slots may be the same frequency, i.e. the highest frequency of a frequency slot may be the lowest frequency of the next frequency slot.

![Frequency Slot Diagram](image-url)

**Figure 1 - Two Frequency Slots on a link**

Figure 1 shows two adjacent frequency slots on a link. The highest frequency of frequency slot 1 denoted by \( n=2 \) is the lowest frequency of frequency slot 2.
of slot 2. In this example, it means that the frequency range from n=-2 to n=10 is occupied and is unavailable to other flexi-LSPs.

Hence, the available frequency ranges should be advertised for the flexi-grid DWDM links. A set of non-overlapping available frequency ranges SHOULD be disseminated in order to allow efficient resource management of Flexi-grid DWDM links and RSA procedures which are described in section 4 of [SSON-FWK].

3.2. Application Compliance Considerations

As described in [G.FLEXIGRID], devices or applications that make use of the flexi-grid may not have to 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).

Hence, the following information should be advertised for a flexi-grid DWDM link:

- Central frequency granularity: a multiple of 6.25 GHz.
- Slot width granularity: a multiple of 12.5 GHz.
- Slot width range: the minimal and maximal slot width supported by a port.

The combination of slot width range and slot width granularity can be used to determine the slot widths set supported by a port.

3.3. Comparison with Fix-grid DWDM Links

In case of fix-grid DWDM links, each wavelength has a pre-defined central frequency and all the wavelengths occupy the same frequency range (channel spacing). Hence all the wavelengths in the DWDM links can be identified uniquely and the status (available or not) of the wavelengths can be advertised through routing protocol.
Figure 2 - A Link supports Fixed Wavelengths with 50 GHz Channel Spacing

Figure 2 shows a link that supports fix-grid with 50 GHz channel spacing. The central frequencies of the wavelengths are pre-defined by ‘n’ and each wavelength occupies a fixed 50 GHz frequency range as described in \([G.694.1v1]\).

Different from the fix-grid DWDM links, the slot width of the wavelengths are flexible on a flexi-grid DWDM link as described in section 2.1, i.e., the value of m in the formula is uncertain before a frequency slot is allocated. So, the available frequency ranges instead of the specific "wavelengths" should be advertised for a flexi-grid DWDM link.

4. Extensions

As described in \([SSON-FWK]\), the network connectivity topology constructed by the links/nodes and node capabilities are the same as WSON which can be advertised by GMPLS routing protocol (refer to section 6.2 of \([RFC6163]\). In case of flexi-grid, the available frequency ranges instead of the specific "wavelengths" should be advertised for the link, which is different from the fixed grid DWDM. This section defines the GMPLS OSPF-TE extensions in support of advertising the available frequency ranges for the flexi-grid DWDM links.

4.1. Available Labels Set sub-TLV

As described in section 2.1, the available frequency ranges other than the available frequency slots should be advertised for the flexi-grid DWDM links. The Available Labels Set sub-TLV defined in \([GEN-OSPF]\) can be re-used to advertise the available frequency ranges for the flexi-grid DWDM links.

The label format defined in \([FLEXIBLE-SIG]\) shown below MUST be used to encode the Label fields in Available Labels Set sub-TLV:
In case of Grid=1 (ITU-T DWDM), a new value of C.S. is defined for flexible grid.

If the C.S. is set to "Flexible grid" in an Available Labels Set sub-TLV, it means that the corresponding link supports flexible grid.

Note that according to the label format defined in [FLEXIBLE-SIG], for the case where the channel spacing value is set to "Flexible grid", a channel spacing of 6.25 GHz should be used in the central frequency computation formula.

4.1.1. Inclusive/Exclusive Label Range

The inclusive/exclusive label ranges format of Available Labels Set sub-TLV defined in [GEN-OSPF] can be used for specifying the frequency ranges of the flexi-grid DWDM links.

Note that it needs multiple Available Labels Set sub-TLVs if there are multiple discontinuous frequency ranges on a link.

4.1.2. Inclusive/Exclusive Label Lists

The inclusive/exclusive label lists format of Available Labels Set sub-TLV defined in [GEN-OSPF] can be used for specifying the available central frequencies of the flexi-grid DWDM links.
4.1.3. Bitmap

The bitmap format of Available Labels Set sub-TLV defined in [GEN-OSPF] can be used for specifying the available central frequencies of the flexi-grid DWDM links.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  4    |   Num Labels          |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Base Label                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Bit Map Word #1 (Lowest numerical labels)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Bit Map Word #N (Highest numerical labels)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

In this case, the Base Label specifies the lowest available central frequency.

Each bit in the bit map represents a particular central frequency with a value of 1/0 indicating whether the central frequency is in the set or not. Bit position zero represents the lowest central frequency and corresponds to the base label, while each succeeding bit position represents the next central frequency logically above the previous.

4.2. Extensions to Port Label Restriction sub-TLV

As described in Section 3.2, there are some restrictions on a port to support flexi-grid. Port Label Restriction sub-TLV is defined in [GEN-OSPF] that can be used to describe the label restrictions on a
port. A new restriction type, i.e. flexi-grid Restriction Type is defined here to specify the restrictions on a port to support flexi-grid.

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MatrixID      | RstType = TBA |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     C.F.G     |    S.W.G      |   Min Width   |   Max Width   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

C.F.G (Central Frequency Granularity, 8 bits): A positive integer. Its value indicates the multiple of 6.25 GHz in terms of central frequency granularity.

S.W.G (Slot Width Granularity, 8 bits): A positive integer. Its value indicates the multiple of 12.5 GHz in terms of slot width granularity.

Min Width (8 bits): A positive integer. Its value indicates the multiple of 12.5 GHz in terms of the supported minimal slot width.

Max Width (8 bits): A positive integer. Its value indicates the multiple of 12.5 GHz in terms of the supported maximal slot width.

4.3. Examples for Label Set

Figure 3 shows an example of available frequency range of a flexi-grid DWDM link.

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11
...+-----------------------------------------------+
|--Available Frequency Range--|

Figure 3 - Flexi-grid DWDM Link

The symbol ‘+’ represents the allowed nominal central frequency. The symbol ‘--’ represents a 6.25 GHz frequency unit. The number on the top of the line represents the ‘n’ in the frequency calculation formula (193.1 + n * 0.00625). The nominal central frequency is 193.1 THz when n equals zero.

Assume that the central frequency granularity is 6.25GHz, the label set can be encoded as follows:

Inclusive Label Range:
o  Start Label = -2;
  o  End Label = 8.

The available central frequencies (-1, 0, 1, 2, 3, 4, 5, 6, 7) can be deduced by the Inclusive Label Range, because the Central Frequency Granularity is 6.25 GHz.

Inclusive Label Lists:

  o  Label 1 = -1;
  o  Label 2 = 0;
  o  Label 3 = 1;
  o  Label 4 = 2;
  o  Label 5 = 3;
  o  Label 6 = 4;
  o  Label 7 = 5;
  o  Label 8 = 6;
  o  Label 9 = 7.

Bitmap:

  o  Base Label = -1;
  o  Bitmap = 111111111 (padded out to a full multiple of 32 bits)

5. IANA Considerations

This document introduces a new Restriction Type for the Port Label Restriction sub-TLV defined in [GEN-OSPF]:

Restriction Type: TBA (flexi-grid)

6. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC3630], [RFC4203].

7. References

7.1. Normative References


7.2. Informative References

8. 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  
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

Ramon Casellas, Ph.D.  
CTTC  
Spain  
Phone: +34 936452916  
Email: ramon.casellas@cttc.es

Oscar Gonzalez de Dios  
Telefonica Investigacion y Desarrollo  
Emilio Vargas 6  
Madrid, 28045  
Spain  
Phone: +34 913374013  
Email: ogondoio@tid.es

Daniele Ceccarelli  
Ericsson  
Via A. Negrone 1/A  
Genova - Sestri Ponente  
Italy  
Email: daniele.ceccarelli@ericsson.com

Xiaobing Zi  
Huawei Technologies  
F3-5-B R&D Center, Huawei Base
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.
RSVP-TE Signaling Extensions in support of Flexible Grid

draft-zhang-ccamp-flexible-grid-rsvp-te-ext-01.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/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at

This Internet-Draft will expire on September 12, 2012.

Abstract

This memo describes the signaling extensions of GMPLS control of
flexible grid network.
1. Introduction

[G.694.1v1] defines the DWDM frequency grids for WDM applications. A frequency grid is a reference set of frequencies used to denote allowed nominal central frequencies that may be used for defining applications. The channel spacing, i.e. the frequency spacing between two allowed nominal central frequencies can be 12.5 GHz, 25 GHz, 50 GHz, 100 GHz and integer multiples of 100 GHz as defined in [G.694.1v1]. All of the wavelengths on a fiber SHALL use different central frequencies and occupy a fixed bandwidth of frequency.

[G.FLEXIGRID], an updated version of [G.694.1v1] has been consented in December 2011 in support of flexible grids. The terms "frequency
slot (i.e. the frequency range allocated to a specific channel and unavailable to other channels within a flexible grid)” and “slot width” (i.e. the full width of a frequency slot in a flexible grid) are introduced to define a flexible grid. A channel is represented as an LSC (Lambda Switching Capable) LSP in the control plane and occupies a frequency slot on each fiber it traverses. In the case of flexible grid, the different flexi-LSPs may have different slot widths on a given fiber, referring to [SSON-FWK].

[WSON-SIG] describes the requirements and extensions for WSON signaling. It focuses on the control of optical networks using a fixed DWDM grid. This document describes the additional requirements and extensions for signaling of LSPs using the flexi-grid capabilities.

2. Terminology

Flexi-grid: See [SSON-FWK].
Slot Width: See [SSON-FWK].
Frequency Range: See [SSON-FWK].
SSON: Spectrum-Switched Optical Networks; See [SSON-FWK].
flexi-LSP: See [SSON-FWK].
RSA: See [SSON-FWK].

3. Requirements for Flexible Grid Signaling

A flexi-LSP SHOULD occupy a frequency slot, i.e. a range of frequencies. The process of computing a route and the allocation of a frequency slot is referred to as RSA (Routing and Spectrum Assignment).

[SSON-FWK] describes three types of architecture approaches to RSA, which are: combined RSA, separated RSA and distributed SA. The first two approaches among them could be called "centralized SA", since both routing and spectrum (frequency slot) assignment are performed by centralized entity before the signaling procedure.

In the case of centralized SA, the assigned frequency slot SHOULD be specified in the Path message. In the case of distributed SA, the slot width of the flexi-LSP SHOULD be specified in the Path message, allowing the involved network elements (e.g., the egress node) to perform such distributed assignment.
Similar to a fixed grid network, if the capability of shifting or converting the whole optical spectrum allocated to a flexi-LSP is not available, the flexi-LSP is subject to the Optical "Spectrum Continuity Constraint", as described in [SSON-FWK].

3.1. Slot Width

The slot width is an end-to-end parameter representing how much frequency resource is requested for a flexi-LSP. Since different LSPs may request different amounts of frequency resource in flexible grid networks, the slot width SHOULD be carried in the signaling message, so that all the nodes along the LSP can know how much frequency resource (including both central frequency and slot width) will be allocated for the LSP.

3.2. Frequency Slot

The frequency slot information represents which part of the frequency resource is allocated on each link for a flexi-LSP. This information SHOULD be carried hop-by-hop in signaling message so that each node can indicate its neighbor the resource reservation on the link between them.

The frequency slot can be represented by the two parameters: central frequency and slot width, as follows:

\[
\text{Frequency slot} = [(\text{central frequency}) - (\text{slot width})/2] ~ [\text{central frequency} + (\text{slot width})/2]
\]

Since the slot width information is carried in the signaling message (as described in Section 2.1), also the central frequency parameter SHOULD be carried in the signaling message for frequency slot determination.

As described in [G.FLEXIGRID], for the flexible DWDM grid, the allowed frequency slots have a nominal central frequency (in THz) defined by:

\[
193.1 + n * 0.00625, \text{ where } n \text{ is a positive or negative integer including 0, and 0.00625 is the nominal central frequency granularity in THz.}
\]

and a slot width defined by:

\[
12.5 * m, \text{ where } m \text{ is a positive integer and 12.5 is the slot width granularity in GHz.}
\]
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).

Figure 1 shows an example of two flexi-LSPs traversing a link and illustrates how to determine the frequency slot based on the central frequency and slot width information.

<table>
<thead>
<tr>
<th>Frequency Slot 1</th>
<th>Frequency Slot 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11</td>
<td></td>
</tr>
</tbody>
</table>

...+------------------+-+------------------+-+------------------+-+------------------+-+------------------+-+

^                     ^
Central F = 193.1 THz  Central F = 193.14375 THz
Slot width = 25 GHz    Slot width = 37.5 GHz

Figure 1 - Two flexi-LSPs traverse a Link

The two wavelengths shown in figure 1 have the following meaning:

flexi-LSP 1: central frequency = 193.1 THz, slot width = 25 GHz. It means the frequency slot [193.0875 THz, 193.1125 THz] is assigned to this flexi-LSP.

flexi-LSP 2: central frequency = 193.14375 THz, slot width = 37.5 GHz. It means the frequency slot [193.125 THz, 193.1625 THz] is assigned to this flexi-LSP.

Note that the frequency slots of two flexi-LSPs on a fiber MUST NOT overlap with each other.

4. Extensions

This section defines the extensions of signaling for flexible grid.

4.1. SSON Traffic Parameters

As described in Section 2, the slot width represents how much frequency resource is requested for a flexi-LSP, i.e., it describes the end-to-end traffic profile of the LSP. Therefore, the slot width SHOULD be regarded as a traffic parameter for a flexi-LSP.
The SSON traffic parameters are organized as follows:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     m         |                      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Reserved                 |                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

m (8 bits): the slot width is specified by \( m \times 12.5 \text{ GHz} \).

Note that the slot width of a fixed grid defined in [G.694.1v1] can also be specified by \( m \) because the defined channel spacings (12.5 GHz, 25 GHz, 50 GHz, 100 GHz and integer multiples of 100 GHz) are also the multiple of 12.5 GHz. Therefore, the traffic parameters are general for SSON including both fixed grid (i.e. WSON) and flexible grid.

The SSON traffic parameters are carried in the `SENDER_TSPEC` object within a Path message and in the `FLOWSPEC` object within a Resv message:

SSON `SENDER_TSPEC`: Class = 12, C-Type = to be assigned by IANA, preferred 8.

SSON `FLOWSPEC`: Class = 9, C-Type = to be assigned by IANA, preferred 8.

### 4.2. Generalized Label

In the case of a flexible grid link, the allocated central frequency is calculated as follows:

\[
\text{Central Frequency} = (193.1 + n \times 0.00625) \text{ THz}
\]

Where \( n \) can be a positive or negative integer, or 0.

The Generalized Label object is used to indicate the resource reserved on a link. In Flexible Grid networks, it is used to indicate which frequency slot is allocated on a link for the given flexi-LSP.

Since the frequency slot assigned to a flexi-LSP can be determined by the combination of [central frequency, slot width], while the slot width of a flexi-LSP is specified in the traffic parameters, the Label object just needs to carry the assigned central frequency.
Therefore, the wavelength label format defined in [RFC6205] can be reused to specify the central frequency of a flexi-LSP, without any change on the label format.

<table>
<thead>
<tr>
<th>Grid</th>
<th>C.S.</th>
<th>Identifier</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

The meaning of Grid, Identifier and n fields are not changed. The usage of the label format is also not changed.

According to [G.FLEXIGRID], flexible grid still belongs to DWDM, so there is no need to introduce a new type of Grid, i.e., Grid=1 (ITU-T DWDM) SHOULD be used for flexible grid.

In case of Grid=1 (ITU-T DWDM), according to [G.697v2.1], a new value of C.S. is defined for flexible 6.25 GHz grid. The C.S.(Channel Spacing) field is defined as follows:

<table>
<thead>
<tr>
<th>C.S. (GHz)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserved</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>12.5</td>
<td>4</td>
</tr>
<tr>
<td>Flexible grid</td>
<td>5 (TBA)</td>
</tr>
<tr>
<td>Future use</td>
<td>6-15</td>
</tr>
</tbody>
</table>

The frequency is calculated as such in [G.FLEXIGRID]:

Frequency (THz) = 193.1 THz + n * channel spacing (THz)
For the case where the channel spacing value is set to "Flexible grid", a channel spacing of 6.25 GHz MUST be used in the above formula.

4.3. Signaling Procedures

This section describes the signaling procedures for distributed SA and centralized SA (See [SSON-FWK]).

4.3.1. Distributed SA

In this case, only the route is provided by a PCE or ingress node before the signaling procedure. The available central frequencies SHALL be collected hop by hop and the egress node SHOULD select a proper central frequency for the LSP.

After the route is computed, the ingress node SHOULD find out the available central frequencies for the LSP on the next link of the route. If the frequency slot does not overlap with the existing flexi-LSPs, the central frequency is considered to be available for the requesting flexi-LSP.

Then a Path message is sent to the next node on the route. The Path message MUST contain a SSON SENDER_TSPEC object to specify the slot width of the flexi-LSP. A LABEL_SET object SHALL be added to the Path message, which contains the candidate central frequencies for the LSP on the next link.

When an intermediate node receives a Path message, it can get the slot width from the SSON SENDER_TSPEC object. Then it SHOULD find the available central frequencies for the LSP on the next link of the route similar to the ingress node. The common part of the two available central frequency sets, i.e. the set received from the Path message and the set of the next link, SHALL be selected as the new available central frequency set for the LSP. If the new set is null, the Path message SHALL be rejected by a PathErr message. Otherwise, the LABEL_SET object in the Path message SHALL be updated according to the new set and the Path message is forwarded to the next node on the route.

When an egress node receives a Path message, it SHOULD select an available central frequency from the LABEL SET object based on local policy and determine the frequency slot based on the slot width and the selected central frequency (See section 2.2). Then a Resv message is responded so that the nodes along the LSP can establish the optical cross-connect based on the frequency slot determined by
the slot width in the traffic parameters and the central frequency in the label.

4.3.2. Centralized SA

In this case, both of the route and the frequency slot are provided by the PCE or ingress node. When signaling the LSP, the slot width is carried in the traffic parameters, and the assigned central frequency is carried in the Label ERO. When the nodes along the LSP receive the Path message carrying this information, they can determine the frequency slot by the slot width and the central frequency, so that they can establish the optical cross-connect based on the central frequency. The procedures of ERO and Label ERO are the same as described in [RFC3209] and [RFC3473].

5. Example

An example is provided as below. In this example, assume that there are two links and three nodes for the network topology and a flexi-LSP is assumed to be created from Node N1 to Node N3.

```
+-----+ link1 +-----+ link2 +-----+
| N1  +----------+ N2 +----------+ N3 |
+-----+ +-----+ +-----+
```

Frequency resources on link1 (central frequency granularity = 12.5 GHz):

```
-9 -8 -7 -6 -5 -4 -3 -2 -1  0  1  2  3  4  5  6  7  8  9 10 11
...+------------+------------+------------+------------+------------+
|--Available Frequency Range--|
```

Frequency resources on link2 (central frequency granularity = 12.5 GHz):

```
-8 -6 -4 -2  0  2  4  6  8 10
...+------------+------------+------------+------------+------------+
|--Available Frequency Range--|
```

The symbol '+' represents the allowed nominal central frequency. The symbol "--" represents a 6.25 GHz frequency unit. The number on the top of the line represents the 'n' in the frequency calculation formula \(193.1 + n \times 0.00625\). The nominal central frequency is 193.1 THz when n equals zero.

A flexi-LSP establishment request:
o Source node: N1
o Sink node: N3
o Slot width: 25 GHz

The usable central frequencies set for the flexi-LSP is
[n=0,1,2,3,4,5,6] on link1. But on link2, because the central
frequency granularity is 12.5 GHz, The usable central frequencies
set for the flexi-LSP is [n=0,2,4].

In the case of Centralized SA, PCE or ingress node (N1) could
allocate an available frequency slot to the flexi-LSP, e.g. n=2 and
slot width=50 Ghz. During the LSP setup procedures, the slot width
(50 GHz, i.e. m=4) should be specified in the traffic parameters
objects and the central frequency (n=2) should be specified in the
label objects.

6. IANA Considerations

6.1. RSVP Objects Class Types

This document introduces two new Class Types for existing RSVP
objects. IANA is requested to make allocations from the "Resource
ReSerVation Protocol (RSVP) Parameters" registry using the "Class
Names, Class Numbers, and Class Types" sub-registry.

<table>
<thead>
<tr>
<th>Class Number</th>
<th>Class Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>FLOWSPEC</td>
<td>[RFC2205]</td>
</tr>
</tbody>
</table>

Class Type (C-Type):

(TBA) SSON FLOWSPEC  [This.I-D]

<table>
<thead>
<tr>
<th>Class Number</th>
<th>Class Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>SENDER_TSPEC</td>
<td>[RFC2205]</td>
</tr>
</tbody>
</table>

Class Type (C-Type):

(TBA) SSON SENDER_TSPEC  [This.I-D]

6.2. DWDM Channel Spacing

The IANA has created a registry and manages the space of DWDM
Channel Spacing as described in section 5.2 of [RFC6205]. It is
requested that the IANA makes assignments from the DWDM Channel
Spacing.
### 6.3. PCEP Object

This document introduces a new Object-Type for existing PCEP objects. It is requested that the IANA makes an assignment from the object-type of GENERALIZED-BANDWIDTH.

<table>
<thead>
<tr>
<th>Object-Class</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBA</td>
<td>GENERALIZED-BANDWIDTH</td>
<td>[GMPLS-PCE]</td>
</tr>
</tbody>
</table>

Object-Type:

(TBA) SSON

[This.I-D]

### 7. Security Considerations

This document introduces no new security considerations to [RFC3473].

### 8. References

#### 8.1. Normative References


8.2. Informative References


9. Contributors’ Address

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

Phone: 
Email: ramon.casellas@cttc.es
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

Oscar Gonzalez de Dios
Telefonica Investigacion y Desarrollo
Emilio Vargas 6
Madrid, 28045
Spain

Phone: +34 913374013
Email: ogondio@tid.es

Felipe Jimenez Arribas
Telefonica Investigacion y Desarrollo
Emilio Vargas 6
Madrid, 28045
Spain
Email: felipej@tid.es

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

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

Phone: +86-755-28973229
Email: zixiaobing@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) 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.
RSVP-TE Extensions to Exchange MPLS-TP LSP Identifiers

draft-zhang-ccamp-mpls-tp-rsvpte-ext-tunnel-num-02

Abstract

The MPLS Transport Profile (MPLS-TP) identifiers document [RFC6370] specifies an initial set of identifiers, such as local assigned tunnel number and Global_ID, which can be used to form Maintenance Entity Point Identifier (MEP_ID). As to some Operation, Administration and Maintenance (OAM) functions, such as Connectivity Verification (CV) [RFC6428], source MEP_ID must be inserted in the OAM packets, so that the peer endpoint can compare the received and expected MEP_IDs to judge whether there is a mismatch [RFC6371], which means that the two MEP nodes need to pre-store each other’s MEP_IDs.

This document defines the signaling extensions to exchange the Label Switched Path (LSP) identifiers.

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 8, 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
1. Introduction .................................................. 3
2. Conventions used in this document ............................. 3
3. Operation ........................................................ 3
   3.1. Co-routed Bidirectional LSP ............................... 3
   3.2. Associated Bidirectional LSP .............................. 4
4. RSVP-TE Extensions ............................................. 5
   4.1. Association Type ............................................ 5
5. IANA Considerations ............................................ 5
6. Security Considerations ........................................ 5
7. Acknowledgement ................................................ 6
8. References ....................................................... 6
   8.1. Normative references ...................................... 6
   8.2. Informative References .................................... 6
Author’s Address .................................................. 7
1.  Introduction

The MPLS Transport Profile (MPLS-TP) identifiers document [RFC6370] specifies a initial set of identifiers, such as local assigned tunnel number (Tunnel_Num) and Global_ID, which can be used to form Maintenance Entity Point Identifier (MEP_ID). The MPLS-TP LSP_MEP_ID is Node_ID::Tunnel_Num::LSP_Num, and in situations where global uniqueness is required, this becomes: Global_ID::Node_ID::Tunnel_Num::LSP_Num. In order to realize some Operation, Administration and Maintenance (OAM) functions, such as Connectivity Verification (CV) [RFC6428], source MEP-ID MUST be inserted in the OAM packets, in this way the peer endpoint can compare the received and expected MEP-IDs to judge whether there is a mismatch [RFC6371]. Hence, the two MEP nodes must pre-store each other’s MEP-IDs before sending the CV packets.

Obviously, the exchange of MEP_IDs can be accomplished by the Network Management System (NMS), but it is complex when the LSPs cross different administration domains, which involves the cooperation of NMSs. When the LSPs are set up by control plane, Resource ReserVation Protocol Traffic Engineering (RSVP-TE) messages will be more suitable to realize the exchange of MEP_IDs.

Since the LSP identifiers can be carried in an Extended ASSOCIATION object, which may also be used in a single session [I-D.ietf-ccamp-assoc-ext], it is naturally to define the signaling extensions of co-routed and associated bidirectional LSP to exchange the LSP identifiers based on 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.  Operation

3.1.  Co-routed Bidirectional LSP

Consider that LSP1 is across different administration domains, which is initialized at A1 node with an Extended ASSOCIATION object inserted in Path message. Association Type is set to "LSP Identifiers", Association ID set to A1-Tunnel_Num, Association Source set to A1-Node_ID, Global Association Source set to A1-Global_ID, and the Extended Association ID field is omitted. Upon receipt of the Extended Association Object, the terminating node Z9 checks the
Association Type field. If it is "LSP Identifiers" and an Upstream_Label exists in Path message, the Extended ASSOCIATION object must be carried in the Resv message also. Similarly, Association Type is set to "LSP Identifiers", Association ID set to Z9-Tunnel_Num, Association Source set to Z9-Node_ID, Global Association Source set to Z9-Global_ID, and the Extended Association ID field is omitted.

3.2. Associated Bidirectional LSP

The document [I-D.ietf-ccamp-mpls-tp-rsvpte-ext-associated-lsp] discusses the provisioning models and signaling procedures of associated bidirectional LSPs. Consider the example provided there, when LSP1 and LSP2 are bound together to be an associated bidirectional LSP which is across several administration domains, the global ID filled in the Extended Association objects with Association Type set to "Double Sided Associated Bidirectional LSP" or "Single Sided Associated Bidirectional LSP" is A-Global_ID or B-Global_ID. If it is A-Global_ID, node A still does know the global ID of node B in case that LSP1 and LSP2 are across several administration domains. Since multiple Association objects have always been supported in Path messages, an Extended Association object with Association Type "LSP Identifiers" can be inserted in the Path messages of associated bidirectional LSPs to let the terminating nodes exchange each others LSP identifiers.

If double sided provisioning model is used, the values of an Extended Association object in LSP1’s Path message are set as below:
Association Type set to "LSP Identifiers", Association ID set to A-Tunnel_Num, Association Source set to A-Node_ID, Global Association Source set to A-Global_ID, and the Extended Association ID field omitted; the object in LSP2’s Path message are set similarly:
Association Type is set to "LSP Identifiers", Association ID set to B-Tunnel_Num, Association Source set to B-Node_ID, Global Association Source set to B-Global_ID, and the Extended Association ID field omitted. While in case that single sided provisioning model is adopted, in the initialized LSP1’s Path message, the values of an Extended Association object are set as following: Association Type set to "LSP Identifiers", Association ID set to A-Tunnel_Num, Association Source set to A-Node_ID, Global Association Source set to A-Global_ID, and the Extended Association ID field omitted. When node B receives this Path message, LSP2 is triggered to be established by the received Extended ASSOCIATION objects with the Association Type "Single Sided Associated Bidirectional LSPs" and "LSP Identifiers". The Extended Association Object with Association Type "LSP Identifiers" inserted in LSP2’s Path message is like: Association ID set to B-Tunnel_Num, Association Source set to B-Node_ID, Global Association Source set to B-Global_ID, and the
Extended Association ID field omitted.

4. RSVP-TE Extensions

4.1. Association Type

Within the current document, a new Association Type is defined in the Extended ASSOCIATION object.

<table>
<thead>
<tr>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 (TBD)</td>
<td>LSP Identifiers (L)</td>
</tr>
</tbody>
</table>

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

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. Since the Association Type "LSP Identifiers" used here is to carry LSP identifier, there is no need to associate Path state to Path state or Resv state to Resv state, one specific rule is added: when the Association Type is "LSP Identifiers", the Extended ASSOCIATION object can appear in Path or Resv message across sessions or in a single session, and the values can be different.

5. IANA Considerations

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

One bit ("LSP Identifiers") needs to be allocated in the Association Type Registry.

6. Security Considerations

A new Association Type is defined in this document, and except this, there are no security issues about the Extended ASSOCIATION object are introduced here. For Association object related security issues, see the documents [RFC4872], [RFC4873], and [I-D.ietf-ccamp-assoc-ext].
For a more comprehensive discussion on GMPLS security please see the Security Framework for MPLS and GMPLS Networks [RFC5920].

7. Acknowledgement

This document was prepared based on the discussion with George Swallow, valuable comments and input were also received from Berger Lou, Venkatesan Mahalingam, Jaihari Kalijanakiraman, Muliu Tao and Wenjuan He.

8. References

8.1. Normative references

[I-D.ietf-ccamp-assoc-ext]

[I-D.ietf-ccamp-mpls-tp-rsvpte-ext-associated-lsp]


8.2. Informative References


[RFC6371] Busi, I. and D. Allan, "Operations, Administration, and
Maintenance Framework for MPLS-Based Transport Networks”, RFC 6371, September 2011.


Author’s Address

Fei Zhang
ZTE Corporation

Email: zhang.fei3@zte.com.cn

Xiao Bao
ZTE Corporation

Email: bao.xiao1@zte.com.cn
RSVP-TE Extensions for Collecting SRLG Information
draft-zhang-ccamp-srlg-fa-configuration-05

Abstract

This document provides extensions for the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) to support automatic collection of Shared Risk Link Group (SRLG) Information for the TE link formed by a LSP.

Status of this Memo

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

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

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

This Internet-Draft will expire on September 13, 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. RSVP-TE Requirements ................................. 3
   2.1. SRLG Collection Indication ...................... 3
   2.2. SRLG Collection ................................. 3
   2.3. SRLG Update ................................... 3
3. RSVP-TE Extensions (Encoding) ......................... 3
   3.1. SRLG Collection Flag ......................... 3
   3.2. SRLG sub-object ............................... 4
4. Signaling Procedures ...................................... 5
   4.1. SRLG Collection ................................. 5
   4.2. SRLG Update ................................... 6
5. Manageability Considerations .......................... 6
   5.1. Policy Configuration ............................. 6
   5.2. Coherent SRLG IDs ............................... 6
6. Security Considerations ................................... 6
7. IANA Considerations ...................................... 7
   7.1. RSVP Attribute Bit Flags ...................... 7
   7.2. ROUTE_RECORD Object ........................... 7
8. Acknowledgements .......................................... 7
9. Normative References ...................................... 7
Authors’ Addresses ........................................... 8
1. Introduction

It is important to understand which TE links in the network might be at risk from the same failures. In this sense, a set of links may constitute a 'shared risk link group' (SRLG) if they share a resource whose failure may affect all links in the set [RFC4202].

On the other hand, as described in [RFC4206] and [RFC6107], H-LSP (Hierarchical LSP) or S-LSP (stitched LSP) can be used for carrying one or more other LSPs. Both of the H-LSP and S-LSP can be formed as a TE link. In such cases, it is important to know the SRLG information of the LSPs that will be used to carry further LSPs.

This document provides an automatic mechanism to collect the SRLG for the TE link formed by a LSP. Note that how to use the collected SRLG information is out of scope of this document.

2. RSVP-TE Requirements

2.1. SRLG Collection Indication

The head nodes of the LSP must be capable of indicating whether the SRLG information of the LSP should be collected during the signaling procedure of setting up an LSP.

2.2. SRLG Collection

The SRLG information can be collected during the setup of an LSP. Then the endpoints of the LSP can get the SRLG information and use it for routing, sharing and TE link configuration purposes.

2.3. SRLG Update

When the SRLG information changes, the endpoints of the LSP need to be capable of updating the SRLG information of the path. It means that the signaling should be capable of updating the newly SRLG information to the endpoints.

3. RSVP-TE Extensions (Encoding)

3.1. SRLG Collection Flag

In order to indicate nodes that SRLG collection is desired, this document defines a new flag in the Attribute Flags TLV, which is carried in an LSP_REQUIRED_ATTRIBUTES Object.
The SRLG Collection flag is meaningful on a Path message. If the SRLG Collection flag is set to 1, it means that the SRLG information should be reported to the head and tail node along the setup of the LSP.

The rules of the processing of the Attribute Flags TLV are not changed.

3.2. SRLG sub-object

This document defines a new RRO sub-object (ROUTE_RECORD sub-object) to record the SRLG information of the LSP. Its format is modeled on the RRO sub-objects defined in RFC 3209 [RFC3209].

<table>
<thead>
<tr>
<th>Type</th>
<th>Length</th>
<th>Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRLG ID 1 (4 bytes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;.....&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRLG ID n (4 bytes)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Type

The type of the sub-object, to be assigned by IANA, which is recommended 34.

Length

The Length contains the total length of the sub-object in bytes, including the Type and Length fields. The Length depends on the number of SRLG IDs.

SRLG Id

The 32-bit identifier of the SRLG.

Reserved

This field is reserved. It SHOULD be set to zero on transmission and MUST be ignored on receipt.
The rules of the processing of the LSP_REQUIRED_ATTRIBUTES Object and ROUTE_RECORD Object are not changed.

4. Signaling Procedures

4.1. SRLG Collection

Typically, the head node gets the route information of an LSP by adding a RRO which contains the sender’s IP addresses in the Path message. If a head node also desires SRLG recording, it sets the SRLG Collection Flag in the Attribute Flags TLV which can be carried in an LSP_REQUIRED_ATTRIBUTES Object.

When a node receives a Path message which carries an LSP_REQUIRED_ATTRIBUTES Object and the SRLG Collection Flag is set, if local policy determines that the SRLG information should not be provided to the endpoints, it must return a PathErr message to reject the Path message. Otherwise, it must add an SRLG sub-object to the RRO to carry the local SRLG information. Then it forwards the Path message to the next node in the downstream direction.

Following the steps described above, the intermediate nodes of the LSP can collect the SRLG information in the RRO during the forwarding of the Path message hop by hop. When the Path message arrives at the tail node, the tail node can get the SRLG information from the RRO.

Before the Resv message is sent to the upstream node, the tail node adds an SRLG sub-object to the RRO. The collected SRLG information can be carried in the SRLG sub-object. Therefore, during the forwarding of the Resv message in the upstream direction, the SRLG information is not needed to be collected hop by hop.

Based on the above procedure, the endpoints can get the SRLG information automatically. Then the endpoints can for instance advertise it as a TE link to the routing instance based on the procedure described in [RFC6107] and configure the SRLG information of the FA automatically.

It is noted that a node (e.g. the edge node of a domain) may edit the RRO to remove the route information (e.g. node, interface identifier information) before forwarding it due to some reasons (e.g. confidentiality or reduce the size of RRO), but the SRLG information should be retained if it is desirable for the endpoints of the LSP.
4.2. SRLG Update

When the SRLG information of a link is changed, the LSPs using that link should be aware of the changes. The procedures defined in Section 4.4.3 of RFC 3209 [RFC3209] MUST be used to refresh the SRLG information.

5. Manageability Considerations

5.1. Policy Configuration

In a border node of inter-domain or inter-layer network, the following SRLG processing policy should be capable of being configured:

- Whether the SRLG IDs of the domain or specific layer network can be exposed to the nodes outside the domain or layer network.

- If the SRLG IDs must not be exposed to the nodes outside of the domain or specific layer network by policy, the border node must reject the Path message desiring SRLG recording and send a PathErr message with the defined error code 'Policy Control Failure'/’Inter-domain policy failure’.

5.2. Coherent SRLG IDs

In a multi-layer multi-domain scenario, SRLG ids may be configured by different management entities in each layer/domain. In such scenarios, maintaining a coherent set of SRLG IDs is a key requirement in order to be able to use the SRLG information properly. Thus, SRLG IDs must be unique. Note that current procedure is targeted towards a scenario where the different layers and domains belong to the same operator, or to several coordinated administrative groups.

Further scenarios, where coherence in the SRLG IDs cannot be guaranteed are out of the scope of the present document and are left for further study.

6. Security Considerations

TBD.
7. IANA Considerations

7.1. RSVP Attribute Bit Flags

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

This document introduces a new Attribute Bit Flag:

- Bit number: TBD (0)
- Defining RFC: this I-D
- Name of bit: SRLG Collection Flag
- The meaning of the Attribute Flags TLV on a Path is defined in this I-D

7.2. ROUTE_RECORD Object

IANA has made the following assignments in the "Class Names, Class Numbers, and Class Types" section of the "RSVP PARAMETERS" registry located at http://www.iana.org/assignments/rsvp-parameters. We request that IANA make assignments from the ROUTE_RECORD RFC 3209 [RFC3209] portions of this registry.

This document introduces a new RRO sub-object:

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD (34)</td>
<td>SRLG sub-object</td>
<td>This I-D</td>
</tr>
</tbody>
</table>

8. Acknowledgements

The authors would like to thank Igor Bryskin, Ramon Casellas and Lou Berger for their useful comments to the document.

9. Normative References


[RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in


Authors’ Addresses

Fatai Zhang (editor)
Huawei
F3-5-B RD Center
Bantian, Longgang District, Shenzhen 518129
P.R.China

Phone:
Email: zhangfatai@huawei.com

Dan Li
Huawei
F3-5-B RD Center
Bantian, Longgang District, Shenzhen 518129
P.R.China

Phone:
Email: danli@huawei.com

Oscar Gonzalez de Dios (editor)
Telefonica I+D
Don Ramon de la Cruz
Madrid, 28006
Spain

Phone: +34 913328832
Email: ogondio@tid.es
Framework for GMPLS and PCE Control of Spectrum Switched Optical Networks
draft-zhang-ccamp-sson-framework-00.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September 05, 2012.
Abstract

A new flexible grid of DWDM has been developed within the ITU-T Study Group 15 to allow a more efficient spectrum allocation. In such environment a data plane connection is switched based on the allocated variable width optical spectrum frequency slot. This new switching capability is referred to as Spectrum Switched Optical Networks (SSON). This draft describes the framework for the application of a GMPLS control plane to a SSON.

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. Terminology .................................................. 4
3. New characteristics of SSON .................................. 5
3.1. Overview of Flexible Grid .................................. 6
3.2. ROADM ...................................................... 7
3.3. Optical Transmitters and Receivers ......................... 8
4. Routing and Spectrum Assignment .............................. 9
4.1. Architectural Approaches to RSA ........................... 10
4.1.1. Combined RSA (R&SA) ..................................... 10
4.1.2. Separated RSA (R+SA) ................................... 11
4.1.3. Routing and Distributed SA (R+DSA) ...................... 11
5. Requirements for GMPLS Control Plane ......................... 11
5.1. Routing ................................................... 11
5.1.1. Available Frequency Ranges of DWDM Links ............... 12
5.1.2. Available Slot Width Ranges of DWDM Links ............... 12
5.1.3. Tunable Optical Transmitters and Receivers .............. 12
5.2. Signaling ................................................. 12
5.2.1. Slot Width Requirement ................................. 13
5.2.2. Frequency Slot Representation .......................... 13
5.3. PCE ....................................................... 13
5.3.1. RSA Computation Type ................................... 13
5.3.2. RSA path re-optimization request/reply .................. 14
5.3.3. Frequency Constraints ................................... 14
6. Security Considerations ..................................... 15
7. References .................................................. 15
7.1. Normative References ...................................... 15
7.2. Informative References .................................... 15
8. Authors’ Addresses .......................................... 16
1. Introduction

[G.694.1v1] defines the DWDM frequency grids for WDM applications. A frequency grid is a reference set of frequencies used to denote allowed nominal central frequencies that may be used for defining applications. The channel spacing, i.e. the frequency spacing between two allowed nominal central frequencies could be 12.5 GHz, 25 GHz, 50 GHz, 100 GHz or integer multiples of 100 GHz as defined in [G.694.1v1]. The frequency spacing of the channels on a fiber is fixed.

The data rate of optical signals becomes higher and higher with the advancement of the optical technology. In the near future, it is anticipated that high data rate signals (beyond 100 Gbit/s or even 400 Gbit/s) will be deployed in optical networks. These signals may not be accommodated in the channel spacing specified in old [G.694.1v1]. Moreover, "mixed rate" scenarios will be prevalent, and the optical signals with different rates may require different spectrum width. As a result, when the optical signals with different rates are mixed to be transmitted on the same fiber, the frequency allocation needs to be more flexible so as to improve spectral efficiency.

An updated version of [G.694.1v1], i.e., [G.FLEXIGRID] has been consented in December 2011 in support of flexible grids. The terms "frequency slot (the frequency range allocated to a channel and unavailable to other channels within a flexible grid)" and "slot width" (the full width of a frequency slot in a flexible grid) are introduced to address flexible grid extensions. A channel is represented as a LSC (Lambda Switching Capable) LSP in the control plane, and it means that a LSC LSP should occupy a frequency slot on each fiber it traverses. In the case of flexible grid, different LSC LSPs may have different slot widths on a fiber.

Thus the concept of Wavelength Switched Optical Network (WSON) is extended to Spectrum Switched Optical Network (SSON) which includes flexible capabilities (i.e. flexi-grid). In SSON, a data plane connection is switched based on an optical spectrum frequency slot of a variable (flexible) slot width, rather than based on a single wavelength within a fixed grid and with a fixed channel spacing as is the case for WSON. In this sense, a WSON can be seen as a particular case of a SSON in which all slot widths are equal and central frequencies depend on the used channel spacing.
WSON related documents are currently being developed with the focus of the GMPLS control of fixed grid optical networks. This document describes the new characteristics of SSON and provides the framework of GMPLS control for the new features of SSON beyond WSON.

Note that this document focuses on the general properties of SSON. Information related to optical impairments is out of its scope and will be addressed in a separate draft.

2. Terminology

Flexible Grid: a new WDM frequency grid defined with the aim of allowing flexible optical spectrum management, in which the Slot Width of the frequency ranges allocated to different channels are flexible (variable size).

Frequency Range: a frequency range is defined as the portion of frequency spectrum included between a lowest and a highest frequency.

Frequency Slot: the frequency range allocated to a slot and unavailable to other slots within a flexible grid. A frequency slot is defined by its nominal central frequency and its slot width.

Slot Width: the full width (in Hz) of a frequency slot. A slot width can be expressed as a multiple (m) of a basic slot width (e.g. 12.5 GHz).

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 (WSON) can be seen as a particular case of SSON in which all slot widths are equal and depend on the used channel spacing.

Flexi-LSP: a control plane construct that represents a data plane connection in which the switching involves a frequency slot of a variable (flexible) slot width. The mapped client signal is transported over the frequency slot, using spectrum efficient modulations such as Coherent Optical Orthogonal Frequency Division Multiplexing (CO-OFDM) and Forward Error Correction (FEC) techniques. Although still in the scope of LSC, the term flexi-LSP is used, when needed, to differentiate from regular WSON LSP in which switching is based on a nominal wavelength.

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. RSA is introduced to describe the routing and spectrum assignment procedures.

3. New characteristics of SSON

Wavelength Switched Optical Networks (WSONs) are 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. WSON framework is described in [RFC6163]. The introduced flexible grid brings some changes on GMPLS controlled optical networks.

The concept of WSON is extended to SSON, to highlight that such subsystems are extended with flexible capabilities (i.e. flexi-grid). Note that, when modeling SSONs, a WSON can be seen as a particular case of SSON where all LSC LSPs use a fixed (and equal) slot width which depends on the used channel spacing.

In WSON, the joint determination of an optical path (physical route) along with the nominal wavelength on a fiber is known as Routing and Wavelength Assignment (RWA). Wavelength Assignment (WA) is the determination of which wavelength can be used for a specific optical path. In analogy with WSON, in SSON, the determination of a path and a frequency slot (including both central frequency and slot width) is referred to as Routing and Spectrum Assignment (RSA). Spectrum Assignment (SA) is the process of determining the spectrum range that can be used for one specific flex-LSP given a physical route.

Compared to WSON, flexibility needs to be introduced in optical network devices such as ROADMs or optical transponders in order to fully benefit from SSON (flexible grid) improved spectrum management. Consequently, transceivers may be able to fully leverage flexible optical channels with advanced modulation formats, and ROADMs may need to be extended to allow flexible spectrum switching.

A flexible grid is defined for the DWDM system in [G.FLEXIGRID]. Compared to fixed grids a flexible grid has a smaller granularity for the central frequencies and the slot width may range from say, 12.5 GHz to hundreds of GHz, in order to accommodate different client data rates. The subsequent sections analyze the new characteristics of flexible grid based on standard [G.FLEXIGRID],
and then model ROADM s, and optical transponders in SSON with an emphasis on those aspects that are of relevance to the control plane.

3.1. Overview of Flexible Grid

- Central Frequency

According to the definition of flexible DWDM grid in [G.FLEXIGRID], the allowed nominal central frequencies are calculated as in the case of flexible grid:

\[
\text{Central Frequency} = 193.1 \text{ THz} + n \times 0.00625 \text{ THz}
\]

Where 193.1 THz is ITU-T ‘ anchor frequency’ for transmission over the C band, n is a positive or negative integer including 0 and 0.00625 THz is the nominal central frequency granularity.

- Slot Width

A slot width is defined by:

\[
12.5 \text{ GHz} \times m
\]

Where m is a positive integer and 12.5 GHz is the slot width granularity.

Note that, when flexi-grid is supported on a WDM link, the slot width of different flexi-LSPs may be different.

The WDM link for flexible grid can be represented as shown in figure 1.

-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11
...

^ 193.1THz

Figure 1 Fiber link model for flexible grid

The symbol ‘+’ represents the allowed nominal central frequencies. The symbol ‘--’ represents the basic 6.25 GHz frequency slot. The number on the top of the line represents the ‘n’ in the frequency calculation formula. The nominal central frequency is 193.1 THz when n equals zero.
As described in [G.FLEXIGRID], the flexible DWDM grid 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 have to be capable of supporting every possible slot width or central frequency granularity. For example, ROADM and transceivers in SSON may support subset of all possible slot width or position defined in [G.FLEXIGRID].

3.2. ROADM

To support flexi grid, a ROADM is a key device which allows spectrum-based optical switching. A classic degree-4 ROADM is shown in Figure 2.

```
+-----------------------+
|                      |
| Input (I1)           |
| Line side-1          |
|                      |
| Output (E1)          |
| Line side-3          |
| Input (I3)           |
|                      |
| Output (E3)          |
|                      |
| Line side-1          |
|                      |
| Input (I2)           |
| Line side-2          |
|                      |
| Output (E2)          |
+-----------------------+

Figure 2 Degree-4 Bidirectional ROADM
```

The key feature of ROADMs is their highly asymmetric switching capability which is described in [RFC6163] in detail. The asymmetric switching feature of flexible ROADM in SSON is similar to fixed ROADM in WSON. The ports on ROADM include line side port which is connected to WDM link, tributary side input/output port which is connected to transmitter/receiver. The main difference between ROADMs in SSON and WSON is the capability of ports on ROADM, which are characterized as follows:

From a SSON control plane perspective (in terms of path computation and resource allocation), ROADM line side ports are characterized by the following aspects:
Available frequency ranges: the set or union of frequency ranges that are not allocated (i.e. available or unused. The relative grouping and distribution of available frequency ranges in a fiber is usually referred to as "fragmentation").

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 defined in [G.FLEXIGRID] (i.e. 12.5GHz) or integer multiples of 12.5GHz.

These properties can be injected into the link parameters from the control plane perspective, which is described in Section 5.

Since the tributary side port is connected to a transmitter and receiver, the characterization of tributary side ports is described in the next section.

3.3. Optical Transmitters and Receivers

In WSON, the optical transmitter is the wavelength source and the optical receiver is the wavelength sink of the WDM system. In each direction, the wavelength used by the transmitter and receiver along a path shall be consistent if no wavelength converter is available. The central frequency used by a transmitter or receiver may be fixed or tunable.

In SSON the optical spectrum (frequency slot width) used by different flexi-LSPs may be variable. Optical transmitters/receivers may have different restriction on the following properties:

- Available central frequencies: The set of central frequencies which can be used by an optical transmitter/receiver.

- 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.
Similarly, other information on transmitters and receivers capabilities, in regard to signal processing is needed to perform efficient RSA, much like in WSON [WSON-ENCODE].

4. Routing and Spectrum Assignment

A LSC flexi-LSP occupies a frequency slot, i.e. a range of frequencies, on each link the LSP traverses.

Much like in WSON, in which if there is no (available) wavelength converter in an optical network an LSP is subject to the "wavelength continuity constraint" (see section 4 of [RFC6163]), in SSON if the capability of shifting or converting the whole optical spectrum allocated to a flexi-LSP is not available, the flexi-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 flexi-LSP. Note that the mapping between client signals data rates (10, 40, 100... Gbps) and optical slot widths (which are dependent on modulation formats and other physical layer parameters) is out of the scope of this document. The frequency slot can be deduced from the central frequency and slot width parameters as follows:

Lowest frequency = (central frequency) - (slot width)/2;

Highest frequency = (central frequency) + (slot width)/2.

Hence, when a route is computed the spectrum assignment process (SA) should determine the central frequency for a flexi-LSP 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.

Figure 2 shows two LSC LSPs that traverse a link.
The two wavelengths shown in figure 2 have the following meaning:

Flexi-LSP 1: central frequency = 193.1 THz, slot width = 25 GHz. It means the frequency slot [193.0875 THz, 193.1125 THz] is assigned to this LSC LSP.

Flexi-LSP 2: central frequency = 193.14375 THz, slot width = 37.5 GHz. It means the frequency slot [193.125 THz, 193.1625 THz] is assigned to this LSC LSP.

Note that the frequency slots of two LSC flexi-LSPs on a fiber do not overlap with each other, and a guard band may be considered to counteract inter-channel detrimental effects.

4.1. 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.

4.1.1. 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 capability, etc.

The computation entity could reside on the following elements, which depends on the implementation:

- PCE: PCE gets the detailed network information and implements the RSA algorithm for RSA requests from the PCCs.
4.1.2. 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.

4.1.3. 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 wavelength continuity constraint should be collected hop by hop along the route. This procedure can be implemented by the GMPLS signaling protocol.

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 central frequencies that meet the slot width requirement of the LSC LSP, i.e. the frequency slot which is determined by the central frequency and slot width MUST NOT overlap with the existing LSC LSPs.

5. Requirements for GMPLS Control Plane

According to the different architecture approaches to RSA some additional requirements have to be considered for the GMPLS control of SSONs.

5.1. Routing

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.

5.1.1. Available Frequency Ranges 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 LSC LSPs could make use of different slot widths on the same link. Hence, the available frequency ranges should be advertised.

5.1.2. Available Slot Width Ranges of DWDM Links

The available slot width ranges needs to be advertised in order to understand whether a LSC LSCP with a given slot width can be set up or not.

Whether a LSC LSP with a certain slot width can be set up or not is constrained by the available slot width ranges of flexible ROADM. So the available slot width ranges should be advertised.

5.1.3. Tunable Optical Transmitters and Receivers

The slot width of a LSC 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.

5.2. Signaling

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 LSC LSP. This brings the following changes to the GMPLS signaling.
5.2.1. Slot Width Requirement

In order to allocate a proper frequency slot for a LSC LSP, the signaling should specify the slot width requirement of a LSC LSP. Then 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 LSC LSP to assign the proper frequency resource. Hence, the slot width requirement should be specified in the signaling message when a LSC LSP is being set up.

5.2.2. Frequency Slot Representation

The frequency slot can be determined by the central frequency (n value) and slot width (m value) as described in section 5. Such parameters should be able to be specified by the signaling protocol.

5.3. PCE

[WSON-PCE] describes the architecture and requirements of PCE for WSON. In the case of flexible grid, RSA instead of RWA is used for routing and frequency slot assignment. Hence PCE should implement RSA for flexible grids. The architecture and requirements of PCE for flexible grids are similar to what is described in [WSON-PCE]. This section describes the changes brought by flexible grids.

5.3.1. RSA Computation Type

A PCEP request within a PCReq message MUST be able to specify the computation type of the request:

- Combined RSA: Both the route and frequency slot should be provided by PCE.
- Routing Only: Only the route is requested to be provided by PCE.

The PCEP response within a PCRep Message MUST be able to specify the route and the frequency slot assigned to the route.

RSA in SSON MAY include the check of signal processing capabilities, which MAY be provided by the IGP. A PCC should be able to indicate additional restrictions for such signal compatibility, either on the endpoint or any given link (such as regeneration points).
A PCC MUST be able to specify whether the PCE MUST also assign a Modulation list and/or a FEC list, as defined in [WSON-ENCODE] and [WSON-PCE].

A PCC MUST be able to specify whether the PCE MUST or SHOULD include or exclude specific modulation formats and FEC mechanisms.

In the case where a valid path is not found, the response MUST be able to specify the reason (e.g., no route, spectrum not found, etc.)

5.3.2. RSA path re-optimization request/reply

For a re-optimization request, the PCEP request MUST provide the path to be re-optimized and include the following options:

- Re-optimize the path keeping the same frequency slot.
- Re-optimize spectrum keeping the same path.
- Re-optimize allowing both frequency slot and the path to change.

The corresponding PCEP response for the re-optimized request MUST provide the Re-optimized path and frequency slot.

In case a path is not found, the response MUST include the reason (e.g., no route, frequency slot not found, both of route and frequency slot not found, etc.)

5.3.3. Frequency Constraints

A PCE should consider the following constraints brought by the transmitters and receivers:

- Available central frequencies: The set of central frequencies that can be used by an optical transmitter or receiver.
- Slot width: The slot width needed by a transmitter or receiver.

These constraints may be provided by the requestor (PCC) in the PCEP request or reside within the PCE’s TEDB which stores the transponder’s capabilities.

A PCC may also specify the frequency constraints for policy reasons. In this case, the constraints should be specified in the request sent to the PCE. In any case, the PCE will compute the route and assign the frequency slot to meet the constraints specified in
the aforementioned request and it will then return the result of the path computation to the PCC in the corresponding response.

6. Security Considerations

This document does not introduce any further security issues other than those described in [RFC6163] and [RFC5920].

7. References

7.1. Normative References


7.2. Informative References


8. 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
1700 Alma Drive, Suite 100
Plano, TX  75075
US
Phone: +1 972 509 5599 x2240
Fax: +1 469 229 5397
EMail: ylee@huawei.com

Oscar Gonzalez de Dios
Telefonica Investigacion y Desarrollo
Emilio Vargas 6
Madrid,   28045
Spain
Phone: +34 913374013
Email: ogondio@tid.es

Ramon Casellas
CTTC
Av. Carl Friedrich Gauss, 7
Castelldefels, 08860, Spain
Phone: +34 936452900
Email: ramon.casellas@cttc.es

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

Zhang                  Expires September 2012                [Page 16]
Email: daniele.ceccarelli@ericsson.com

Xiaobing Zi
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28973229
Email: zixiaobing@huawei.com

Jianrui Han
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28973229
Email: hanjianrui@huawei.com

Felipe Jimenez Arribas
Telefonica Investigacion y Desarrollo
Emilio Vargas 6
Madrid, 28045
Spain
Email: felipej@tid.es

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.