

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 25, 2011

D. Ceccarelli, Ed.
D. Caviglia
Ericsson
F. Zhang
D. Li
Huawei Technologies
S. Belotti
P. Grandi
Alcatel-Lucent
R. Rao
K. Pithewan
Infinera Corporation
J. Drake
Juniper
June 23, 2011

Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS)
Control of Evolving G.709 OTN Networks
draft-ceccarelli-ccamp-gmpls-ospf-g709-06

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 December 25, 2011.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
1.1. Terminology	3
2. OSPF-TE Extensions	3
3. TE-Link Representation	4
4. ISCD format extensions	5
4.1. Switch Capability Specific Information	7
5. Examples	12
5.1. Example of T and S bits utilization	13
5.2. Example of ODUFlex advertisement	13
5.3. Example of single stage muxing	15
5.4. Example of multi stage muxing - Unbundled link	16
5.5. Example of multi stage muxing - Bundled links	18
6. Compatibility	20
7. Security Considerations	21
8. IANA Considerations	21
9. Contributors	21
10. Acknowledgements	23
11. References	23
11.1. Normative References	23
11.2. Informative References	24
Authors' Addresses	24

1. Introduction

G.709 OTN [G709-V3] includes new fixed and flexible ODU containers, two types of Tributary Slots (i.e., 1.25Gbps and 2.5Gbps), and supports various multiplexing relationships (e.g., ODUj multiplexed into ODUk ($j < k$)), two different tributary slots for ODUk ($K=1, 2, 3$) and ODUflex service type, which is being standardized in ITU-T. In order to present this information in the routing process, this document provides OTN technology specific encoding for OSPF-TE.

For a short overview of OTN evolution and implications of OTN requirements on GMPLS routing please refer to [OTN-FWK]. The information model and an evaluation against the current solution are provided in [OTN-INFO].

The routing information for Optical Channel Layer (OCh) (i.e., wavelength) is out of the scope of this document. Please refer to [WSO-Frame] for further information.

1.1. Terminology

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

2. OSPF-TE Extensions

In terms of GMPLS based OTN networks, each OTUk can be viewed as a component link, and each component link can carry one or more types of ODUj ($j < k$).

Each TE LSA can carry a top-level link TLV with several nested sub-TLVs to describe different attributes of a TE link. Two top-level TLVs are defined in [RFC 3630]. (1) The Router Address TLV (referred to as the Node TLV) and (2) the TE link TLV. One or more sub-TLVs can be nested into the two top-level TLVs. The sub-TLV set for the two top-level TLVs are also defined in [RFC 3630] and [RFC 4203].

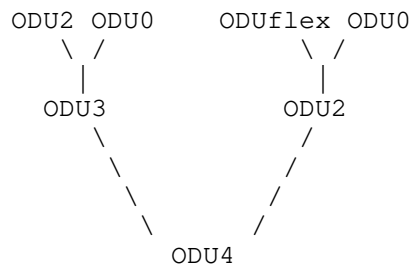
As discussed in [OTN-FWK] and [OTN-INFO], the OSPF-TE must be extended so to be able to advertise the termination and switching capabilities related to each different ODUj and ODUk/OTUk and the advertisement of related multiplexing capabilities. This leads to the need of defining a new Switching Capability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

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 with switching capability only, and the ODUj->ODUk format is used to indicate the ODUj into ODUk multiplexing capability.

This notation can be iterated dependently from 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 multiplexed container is represented with increasing granularity till the leaf of the tree. The tree can be structured with more than one branch if the server ODUk/OTUk supports more than one hierarchy.

If for example a multiplexing hierarchy like the following one is considered:



The ODU4 is the root of the muxing tree, ODU3 and ODU2 are containers directly multiplexed into the server and then ODU2, ODU0 are the leaves of 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 ODUk-TE-Link. The ODUk-TE-Link advertises ODUj Switching Capacity. The advertised capacity could include ODUk switching capacity. Figure-1 below provides an illustration of one hop ODUk TE-links.

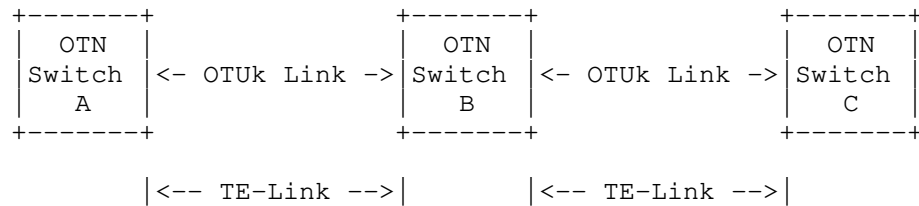


Figure 1: ODUk TE-Link

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

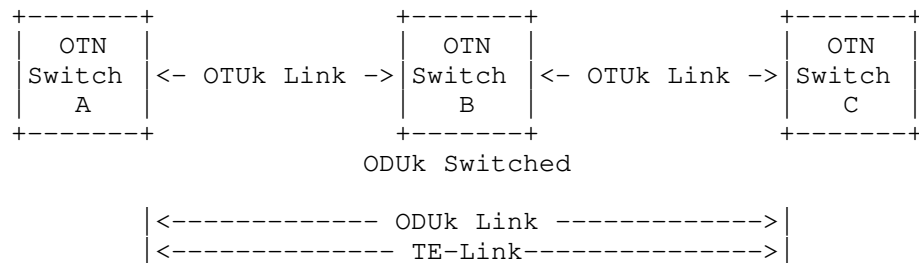


Figure 2: Multiple hops TE-Link

4. ISCD format extensions

The Interface Switching Capability Descriptor describes switching capability of an interface [RFC 4202]. This document defines a new Switching Capacity value for OTN [G.709-v3] as follows:

Value -----	Type -----
101 (TBA by IANA)	OTN-TDM capable (OTN-TDM)

Switching Capability and Encoding values MUST be used as follows:

Switching Capability = OTN-TDM
 Encoding Type = G.709 ODUk (Digital Path) [as defined in RFC4328]

Both fixed and flexible ODUs use the same switching type and encoding values. When Switching Capability and Encoding fields are set to values as stated above, the Interface Switching Capability Descriptor should be interpreted as follows:

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
Switching Cap	Encoding	Reserved	
Max LSP Bandwidth at priority 0			
Max LSP Bandwidth at priority 1			
Max LSP Bandwidth at priority 2			
Max LSP Bandwidth at priority 3			
Max LSP Bandwidth at priority 4			
Max LSP Bandwidth at priority 5			
Max LSP Bandwidth at priority 6			
Max LSP Bandwidth at priority 7			
Switch Capability Specific Information (variable length)			

Maximum LSP Bandwidth

The MAX LSP bandwidth field is used accordingly to RFC4204: i.e. $0 \leq \text{Max LSP Bandwidth} \leq \text{ODUk/OTUk}$ and intermediate values are those on

the branch of OTN switching hierarchy supported by the interface. E.g. in the OTU4 link it could be possible to have ODU4 as MAX LSP Bandwidth for some priorities, ODU3 for others, ODU2 for some others etc. The bandwidth unit is in bytes per second and the encoding is in IEEE floating point format. The discrete values for various ODUs is shown in the table below.

ODU Type	ODU nominal bit rate	Value in Byte/Sec
ODU0	1 244 160 kbits/s	0x4D1450C0
ODU1	239/238 x 2 488 320 kbit/s	0x4D94F048
ODU2	239/237 x 9 953 280 kbit/s	0x4E959129
ODU3	239/236 x 39 813 120 kbit/s	0X4F963367
ODU4	239/227 x 99 532 800 kbit/s	0x504331E3
ODU2e	239/237 x 10 312 500 kbit/s	0x4E9AF70A
ODUflex for CBR Client signals	239/238 x client signal bit rate	MAX LSP BANDWIDTH
ODUflex for GFP-F Mapped client signal	Configured bit rate	MAX LSP BANDWIDTH
ODU flex resizable	Configured bit rate	MAX LSP BANDWIDTH

The ISCD includes a variable number of SCSI TLVs as described in the following sections. A single ISCD TLV MAY be used for the advertisement of unbundled or bundled links also with different server layers. A different SCSI TLV MUST be used for each different muxing hierarchy (muxing tree in the following examples).

The Maximum LSP Bandwidth at priority 'p' field MUST be set accordingly to [RFC4204]. It MUST be set to zero for non supported priorities.

E.g. if 3 OTU3 and 4 OTU2 interfaces are bundled together, a single ISCD TLV may be advertised with a different SCSI for each muxing hierarchy.

4.1. Switch Capability Specific Information

The technology specific part of the ISCD can include a variable number of SCSI TLVs. The definitio of a SCSI TLV allows the encoding

being future proof and easily extensible. This document defines SCSI TLV type 1 (TBA by IANA), which is used to describe a tree of the OTN muxing hierarchy. The muxing hierarchy tree is encoded as an order independent list of TLVs called Bandwidth TLVs. Two types of Bandwidth TLV are defined (TBA by IANA):

- Type 1 - Used for fixed containers
- Type 2 - Used for flexible containers

The format of the SCSI TLV is depicted in the following figure:

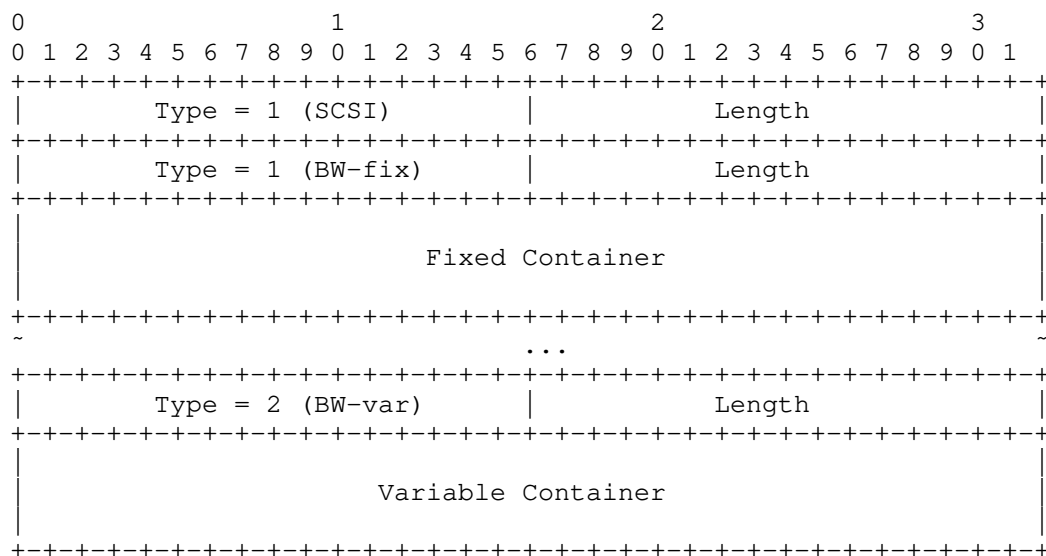


Figure 3: SCSI TLV

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

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

Figure 4: Bandwidth TLV - Type 1 -

The values of the fields shown in figure 4 are explained after figure 5.

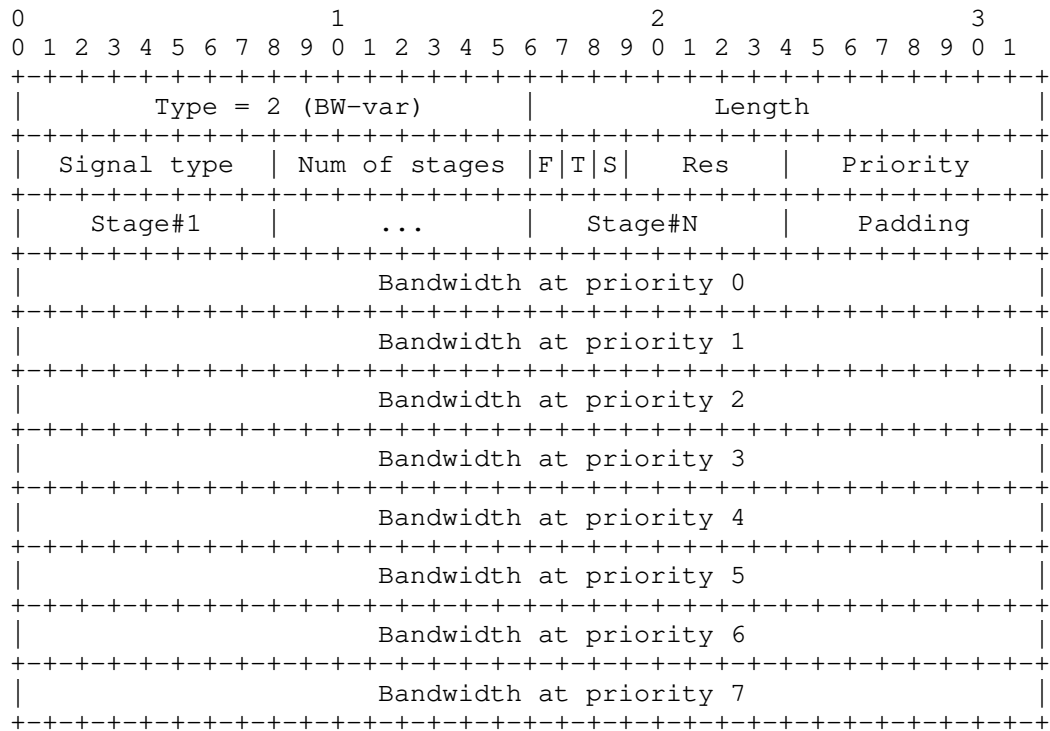


Figure 5: Bandwidth TLV - Type 2 -

- Signal Type: Indicates the ODU type being advertised

Value	Type
-----	-----
1	ODU1
2	ODU2
3	ODU3
4	ODU4
10	ODU0
11	ODU2e
20	ODUflex CBR
21	ODUflex GFP-F resizable
22	ODUflex GFP-F non resizable
60000-65535	Experimental

- Number of stages: Indicates the number of multiplexing stages level. It is 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:
 - F flag : This flag defines the meaning of the Bandwidth being advertised. When the F bit is cleared, the type of bandwidth being advertised is the Unreserved Bandwidth of the given signal type. On the other side, when the F bit is set, the Bandwidth fields represent the MAX LSP bandwidth.
 - 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 is not permitted.

- Priority : 8 bits field with 1 flag for each priority. Bit set indicates priority supported, bit cleared 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 is always indicating the server ODU container (ODUk/OTUk). The values of the Stage fields are the same ones defined for the Signal Type field.
- Padding: Given that the number of Stages is variable, a padding to 32 bits field might be needed.
- Unreserved Bandwidth/Max LSP BW : In case of fixed containers the Bandwidth field is 16 bits long and indicates the Unreserved Bandwidth in number of available containers, while in case of variable container the Bandwidth field (both in case of Unreserved or MAX LSP) is 32 bits long and expressed in IEEE floating point format. Only Unreserved/MAX LSP bandwidth for supported priorities MUST be advertised.

[EDITOR NOTE]: TO BE MOVED TO THE INFO MODEL DRAFT 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 ODU1.

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).

5. Examples

The examples in the following pages are not normative and are not intended to infer or mandate any specific implementation.

5.1. Example of T and S bits utilization

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

Figure 6: Example 1 - T and S bits utilization

5.2. 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 but in some cases also information about the Unreserved bandwidth could be useful. The F flag is used to distinguish between the two cases.

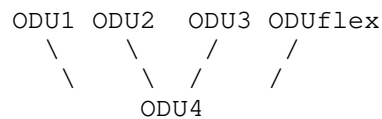
										1											2											3		
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1			
----- -----																																		

Type = 2 (BW-var)				Length			
S. type=ODUflex	#stages= 1	F=0	T S	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							
Type = 2 (BW-var)				Length			
S. type=ODUflex	#stages= 1	F=1	T S	Res	Priority		
Stage#1=ODU3	Padding						
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 7: Example 2 - ODUflex advertisement

5.3. Example of single stage muxing

Supposing to have 1 OTU4 component link supporting single stage muxing of ODU1, ODU2, ODU3 and ODUflex, the supported hierarchy can be summarized in a tree as in the following figure. For sake of simplicity we assume that also in this case only priorities 0 and 3 are supported.



and the related SCSIs as follows:

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type = 1 (SCSI)										Length																													
Type = 1 (BW-fix)										Length																													
Sig type=ODU4										#stages= 0										F=0 T S Res 1 0 0 1 0 0 0 0																			
Unres ODUk at Prio 0 =1										Unres ODUk at Prio 3 =1																													
Type = 1 (BW-fix)										Length																													
Sig type=ODU1										#stages= 1										F=0 T S Res 1 0 0 1 0 0 0 0																			
Stage#1=ODU4										Padding																													
Unres ODUk at Prio 0 =40										Unres ODUk at Prio 3 =40																													
Type = 1 (BW-fix)										Length																													
Sig type=ODU2										#stages= 1										F=0 T S Res 1 0 0 1 0 0 0 0																			
Stage#1=ODU4										Padding																													
Unres ODUk at Prio 0 =10										Unres ODUk at Prio 3 =10																													
Type = 1 (BW-fix)										Length																													

```

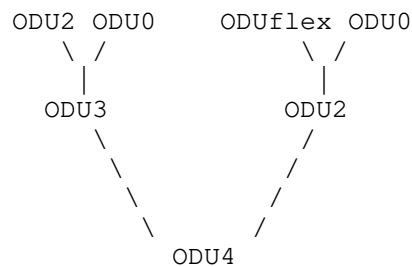
+++++
|Sig type=ODU3 | #stages= 1 |F=0|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU4 | Padding
+++++
| Unres ODUk at Prio 0 =2 | Unres ODUk at Prio 3 =2 |
+++++
| Type = 2 (BW-var) | Length
+++++
|S. type=ODUflex| #stages= 1 |F=1|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU4 | Padding
+++++
| MAX LSP Bandwidth at priority 0 =100Gbps
+++++
| MAX LSP Bandwidth at priority 3 =100Gbps
+++++
|S. type=ODUflex| #stages= 1 |F=0|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU4 | Padding
+++++
| Unreserved Bandwidth at priority 0 =100Gbps
+++++
| Unreserved Bandwidth at priority 3 =100Gbps
+++++

```

Figure 8: Example 3 - Single stage muxing

5.4. Example of multi stage muxing - Unbundled link

Supposing to have 1 OTU4 component link with muxing capabilities as show in the following figure:



and supported priorities 0 and 3, the advertisement is composed by the following SCSI:

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (SCSI)          |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU4 | #stages= 0 | F | T | S | Res | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =1 | Unres ODUk at Prio 3 =1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU3 | #stages= 1 | F | T | S | Res | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 | Padding |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =2 | Unres ODUk at Prio 3 =2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU2 | #stages= 1 | F | T | S | Res | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 | Padding |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =10 | Unres ODUk at Prio 3 =10 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU2 | #stages= 2 | F | T | S | Res | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU3 | Stage#2=ODU4 | Padding |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =10 | Unres ODUk at Prio 3 =10 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU0 | #stages= 2 | F | T | S | Res | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU3 | Stage#2=ODU4 | Padding |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =80 | Unres ODUk at Prio 3 =80 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Type = 1 (BW-fix)         |          Length          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

```

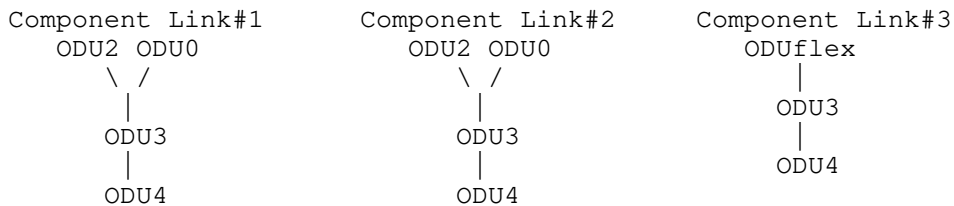
+++++
|Sig type=ODU0 | #stages= 2 |F|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU2 | Stage#2=ODU4 | Padding |
+++++
| Unres ODUk at Prio 0 =80 | Unres ODUk at Prio 3 =80 |
+++++
| Type = 2 (BW-var) | Length |
+++++
|S.type=ODUflex | #stages= 2 |F=0|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU2 | Stage#2=ODU4 | Padding |
+++++
| Unreserved Bandwidth at priority 0 =100Gbps |
+++++
| Unreserved Bandwidth at priority 3 =100Gbps |
+++++
| Type = 2 (BW-var) | Length |
+++++
|S.type=ODUflex | #stages= 2 |F=1|T|S| Res |1|0|0|1|0|0|0|0|
+++++
| Stage#1=ODU2 | Stage#2=ODU4 | Padding |
+++++
| MAX LSP Bandwidth at priority 0 =10Gbps |
+++++
| MAX LSP Bandwidth at priority 3 =10Gbps |
+++++

```

Figure 9: Example 4 - Multi stage muxing - Unbundled link

5.5. Example of multi stage muxing - Bundled links

In this example 3 OTU4 component links with the following muxing capabilities trees are considered



Considering only supported priorities 0 and 3, the advertisement is

composed by a single ISCD with 2 SCSI TLVs, one for the advertisement of Component Link#1 and #2 and the second one for Component Link#3:

- SCSI 1 -

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Type = 1 (SCSI)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Type = 1 (BW-fix)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU4 | #stages= 0 | F|T|S| Res | 1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =2 | Unres ODUk at Prio 3 =2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Type = 1 (BW-fix)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU3 | #stages= 1 | F|T|S| Res | 1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU4 |                               Padding                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =4 | Unres ODUk at Prio 3 =4 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Type = 1 (BW-fix)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU2 | #stages= 2 | F|T|S| Res | 1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU3 | Stage#2=ODU4 |                               Padding                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =20 | Unres ODUk at Prio 3 =20 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Type = 1 (BW-fix)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Sig type=ODU0 | #stages= 2 | F|T|S| Res | 1|0|0|1|0|0|0|0|
+-----+-----+-----+-----+-----+-----+-----+-----+
| Stage#1=ODU2 | Stage#2=ODU4 |                               Padding                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Unres ODUk at Prio 0 =160 | Unres ODUk at Prio 3 =160 |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

SCSI - 2 -

```

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

```

	Type = 1 (SCSI)		Length	
+	+	+	+	+
	Type = 1 (BW-fix)		Length	
+	+	+	+	+
	Sig type=ODU4		#stages= 0	
+	+	+	+	+
	Unres ODUk at Prio 0 =1		Unres ODUk at Prio 3 =1	
+	+	+	+	+
	Type = 1 (BW-fix)		Length	
+	+	+	+	+
	Sig type=ODU3		#stages= 1	
+	+	+	+	+
	Stage#1=ODU4		Padding	
+	+	+	+	+
	Unres ODUk at Prio 0 =2		Unres ODUk at Prio 3 =2	
+	+	+	+	+
	Type = 2 (BW-var)		Length	
+	+	+	+	+
	S. type=ODUflex		#stages= 2	
+	+	+	+	+
	Stage#1=ODU3		Stage#1=ODU4	
+	+	+	+	+
	MAX LSP Bandwidth at priority 0 =40Gbps			
+	MAX LSP Bandwidth at priority 3 =40Gbps			
+	+	+	+	+
	S. type=ODUflex		#stages= 2	
+	+	+	+	+
	Stage#1=ODU3		Stage#1=ODU4	
+	+	+	+	+
	Unreserved Bandwidth at priority 0 =80Gbps			
+	Unreserved Bandwidth at priority 3 =80Gbps			
+	+	+	+	+

Figure 10: Example 5 - Multi stage muxing - Bundled lilnks

6. Compatibility

Backwards compatibility with implementations based on [RFC4328] can be achieved advertising the [RFC4328] based ISCDs in addition to the ISCD defined in this document.

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 Kunjidhapatham

akunjidhapatham@infinera.com

Snigdho Bardalai

sbardalai@infinera.com

Steve Balls

Steve.Balls@metaswitch.com

Jonathan Hardwick

Jonathan.Hardwick@metaswitch.com

Xihua Fu

fu.xihua@zte.com.cn

Cyril Margaria

cyril.margaria@nsn.com

10. Acknowledgements

11. References

11.1. Normative References

- [MLN-EXT] D.Papadimitriou, M.Vigoureux, K.Shiomoto, D.Brungard, J.Le Roux, "Generalized Multi-Protocol Extensions for Multi-Layer and Multi-Region Network (MLN/MRN)", February 2010.
- [OTN-FWK] F.Zhang, D.Li, H.Li, S.Belotti, D.Ceccarelli, "Framework for GMPLS and PCE Control of G.709 Optical Transport networks, work in progress draft-ietf-ccamp-gmpls-g709-framework-04", March 2011.
- [OTN-INFO] S.Belotti, P.Grandi, D.Ceccarelli, D.Caviglia, F.Zhang, D.Li, "Information model for G.709 Optical Transport Networks (OTN), work in progress draft-ietf-ccamp-otn-g709-info-model-00", April 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2154] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", RFC 2154, June 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC2370] Coltun, R., "The OSPF Opaque LSA Option", RFC 2370, July 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

11.2. Informative References

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

[G.709-v3] ITU-T, "Draft revised G.709, version 3", consented by ITU-T on Oct 2009.

[Gsup43] ITU-T, "Proposed revision of G.sup43 (for agreement)", December 2008.

Authors' Addresses

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

Email: daniele.ceccarelli@ericsson.com

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

Email: diego.caviglia@ericsson.com

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Shenzhen 518129 P.R.China Bantian, Longgang District
Phone: +86-755-28972912

Email: zhangfatai@huawei.com

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

Email: danli@huawei.com

Sergio Belotti
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

Pietro Vittorio Grandi
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: pietro_vittorio.grandi@alcatel-lucent.com

Rajan Rao
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA

Email: rrao@infinera.com

Khuzema Pithewan
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA

Email: kpithewan@infinera.com

John E Drake
Juniper

Email: jdrake@juniper.net

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2012

X. Fu
Q. Wang
Y. Bao
ZTE Corporation
R. Jing
X. Huo
China Telecom
July 8, 2011

RSVP-TE Extension for MRN/MLN Application
draft-fuxh-ccamp-boundary-explicit-control-ext-03

Abstract

[RFC5212] defines a Multi-Region and Multi-Layer Networks (MRN/MLN). [RFC4206] introduces a region boundary determination algorithm and a Hierarchy LSP (H-LSP) creation method. However, in some scenarios, there must be some additional information to facilitate hierarchy LSP creation. This document extends RSVP-TE to meet this requirement.

Status of this Memo

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

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

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

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

Copyright Notice

Copyright (c) 2011 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. Requirement Identification	3
2.1. Indication of Server Layer	3
2.2. Requirement in OTN Multi-Layer Network	4
2.2.1. Indication of ODUk Signal Type	4
2.2.2. Indication of Multi Stages Multiplexing Hierarchy	4
3. Mechanism and Protocol Extensions	5
3.1. Controlling FA-LSPs Boundaries	5
3.1.1. Boundaries Determination	6
3.1.2. Example	6
3.2. Explicit Route Boundary Object (ERBO)	7
3.2.1. Switching Capability subobject	8
3.2.2. Encoding Type subobject	8
3.2.3. Signal Type subobject	9
3.2.4. Multiplexing Hierarchy subobject	10
3.2.5. Signaling Procedure	11
3.3. Exclude Route Object (XRO)	12
3.3.1. Encoding Type subobject	12
3.3.2. Signal Type subobject	13
3.3.3. Multiplexing Hierarchy subobject	13
4. Security Considerations	14
5. IANA Considerations	14
6. References	14
6.1. Normative References	14
6.2. Informative References	15
Authors' Addresses	15

1. Introduction

This document describes some requirements of explicitly control Multi-Region and Multi-Layer Network. It extends mechanisms and protocols defined in [RFC4206] and [RFC6001] to meet these requirement.

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. Requirement Identification

2.1. Indication of Server Layer

[RFC4206] describes a region boundary determination algorithm and a hierarchical LSP creation method. It is well applied to multi-region network. However it isn't fully applied to multi-layer network within the same switching capability.

In the following figure, three LSPs belong to the same TDM region and different latyers, but boundary node (e.g., B) could not determine that STM-N FA-LSP should be triggered according to the region boundary determination algorithm defined in [RFC4206]. The solution MUST support to explicitly indicate which server layer must be triggered.

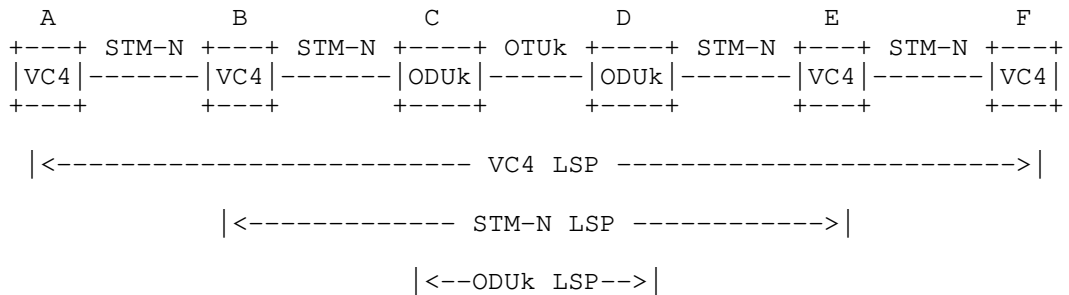


Figure 1: Example of Server Layer Indication

2.2. Requirement in OTN Multi-Layer Network

2.2.1. Indication of ODUk Signal Type

In Figure 2, node B and C in the OTN network are connected to 2.5G TS network by two OTU3 links. They can support flexible multi stages multiplexing hierarchies. There are two multi stages multiplexing hierarchies for ODU0 being mapped into OTU3 link in B and C (i.e., ODU0-ODU1-ODU3 and ODU0-ODU2-ODU3). But boundary node (e.g., B) could not determine which kind of ODUk FA-LSP (ODU1, ODU2 or ODU3) should be triggered during one e2e ODU0 connection signaling according to the region boundary determination algorithm defined in [RFC4206].

If path computation entity select the ODU0-ODU2-ODU3 multi stages multiplexing hierarchy in Node B and C for one end-to-end ODU0 service from A to Z, there has to be an ODU2 or ODU3 FA-LSP between B and C. The solution MUST support to explicitly indicate which type of ODUk FA-LSP must be triggered for ODUj (k>j).

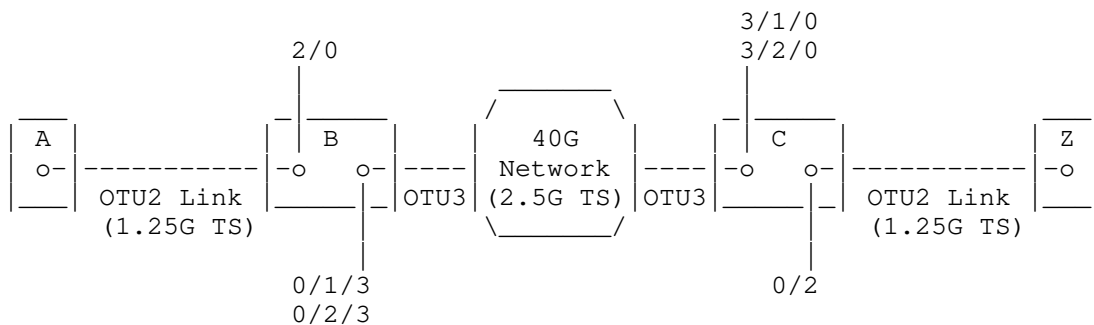


Figure 2 Example of ODUk Signal Type Indication

2.2.2. Indication of Multi Stages Multiplexing Hierarchy

In figure 2, if ODU3 FA-LSP will be triggered between B and C to directly support one end-to-end ODU0 service from A to Z, B should be informed which multi stages multiplexing hierarchy should be used for ODU0 mapping into ODU3. So the solution MUST support to explicitly indicate which multi stages multiplexing hierarchy must be applied to a special interface.

3. Mechanism and Protocol Extensions

This section defines protocol mechanisms and extensions to achieve the requirement described in the previous section.

- o A generic boundaries determination mechanism is introduced first. Path computation entity or interim LSR along one end-to-end LSP which traverses multi-layer can rely on this mechanism to determine the boundary nodes of FA-LSP.
- o Path computation entity can determine regions' boundaries. After PCE compute an end-to-end paths across multi-layer, the boundary nodes and some limitation about how to create FA-LSP must be inform to interim nodes during signaling.

A new object, Explicit Route Boundary Object (ERBO), is introduced to explicitly indicate a pair of FA-LSP boundary nodes and some attributes which indicates how to create FA-LSPs.

This document also introduces some new subobjects as part of the XRO that explicitly indicate which Signal Type, Multiplexing Hierarchy and Encoding Type have to be excluded before initiating FA-LSP creation.

3.1. Controlling FA-LSPs Boundaries

The boundary determination mechanism in [RFC4206] depends on the comparing of interface switching capabilities. For multi-layer network within the same TDM switching capability, the comparing of interface switching capabilities relies on the max LSP bandwidth of interface. But one interface in OTN network could support several ODUk signal type, the max LSP bandwidth makes no any sense to path computation entity. The mechanism in [RFC4206] isn't well applied to OTN multi-layer network. The solution MUST support the boundaries determination of ODUk FA-LSP.

This document introduces a generic mechanism to determine the boundaries of FA-LSPs by using termination and switching capability from IGP database. It can be applied to multi-layer network within same switching capability (e.g, OTN network) and multi-region network. So this mechanism is compatible with the one in [RFC4206]. The switching and termination capability could be induced by IACD [RFC6001] in multi-region network. In OTN multi-layer network, the switching and termination Capability [OTNv3-OSPF] is advertised by using SCSI (Switch Capability Specific Information) within ISCD.

3.1.1. Boundaries Determination

Suppose an LSP's path is as follows: node-0, link-1, node-1, link-2, node-2, ..., link-n, node-n. Moreover, for link-i denote by [link-i, node-(i-1)] the interface that connects link-i to node-(i-1), and by [link-i, node-i] the interface that connects link-i to node-i.

Suppose interface [link-(i+1), node-i] supports switching capability of one signal type ST-x and termination capability of one signal type ST-y. Interface [link-(i+1), node-(i+1)] supports switching capability of ST-y. Switching capability of ST-y (e.g., LSC) is larger than ST-x (e.g., TDM/G.709) or ST-x (e.g., ODUj) could be mapped into ST-y (e.g., ODUk (k>j)). So we say that the LSP has crossed a region boundary at node-i. The 'other edge' of the region with respect to the LSP path is node-k, where k is the smallest number greater than i such that interface [link-k, node-(k-1)] supports switching capability of ST-y and interface [link-k, node-k] supports switching capability of ST-x and termination capability of ST-y.

3.1.2. Example

A multi-layer OTN network is illustrated in figure 3. Node B and D support ODUj being mapping into ODUk (k>j). Interface IF-B and IF-D support ODUj switching capability (ODUj(S)) and ODUk termination capability (ODUk(T)). Interface within C only supports ODUk switching capability. So Node B and D could be boundaries of ODUk FA-LSP for ODUj LSP.

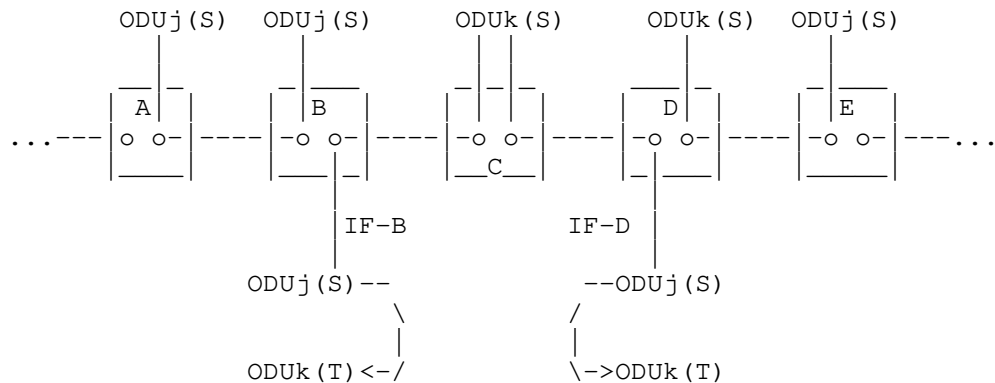


Figure 3 Example of Controlling ODUk FA-LSPs Boundaries

A multi-region network is illustrated in figure 4. Node B and D which are hybrid nodes support PSC being mapping into ODUk (e.g., by GFP-F). Interface IF-B and IF-D support PSC switching capability (PSC(S)) and ODUk termination capability (ODUk(T)). Interface within C only supports ODUk switching capability. So Node B and D could be boundaries of ODUk FA-LSP for PSC LSP.

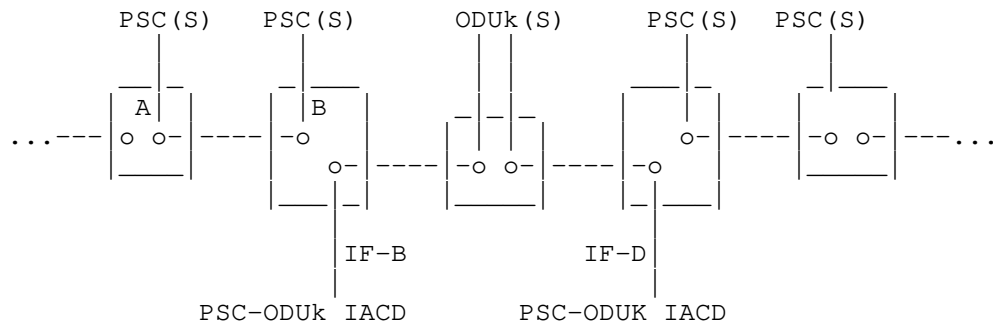


Figure 4 Example of Controlling ODUk FA-LSPs Boundaries

3.2. Explicit Route Boundary Object (ERBO)

In order to explicitly control hierarchy LSP creation, this document introduce a new object (ERBO-Explicit Route Boundary Object) carried in Path message. The format of ERBO object is the same as ERO. It looks more like the SERO defined in [RFC4873].

One or more ERBOs may be carried in Path message. Multiple ERBOs could support cascading of FA easy. An ERBO must contain at least two subobjects. The first and final one indicate the source and sink node of a FA-LSP or Composite Link [CL-REQ] which will be passed by one e2e LSP. Other subobjects may be inserted into ERBO between source and sink node to indicates how to select the FA/Component Link or create them.

The purpose is not to extend ERO and to limit the modifications to existing RSVP-TE procedures. ERBO is a top object and parsed easy. Many attributes could be inserted into ERBO in the future for other requirements.

This document defines four subobjects (i.e., Switching Cap, Encoding Type, Signal Type and Multiplexing Hierarchy) in ERBO. These subobjects may be inserted into ERBO between source and sink node to indicates how to select the FA/Component Link or create them. It is very convenient to use these subobjects independently or combine them.

For example, Signal Type and Multiplexing Hierarchy subobject are enough for OTN multi-layer network application.

3.2.1. Switching Capability subobject

A new subobject, called the switching capability subobject, is defined for use in the ERBO. The format of the switching capability subobject is defined as follows:

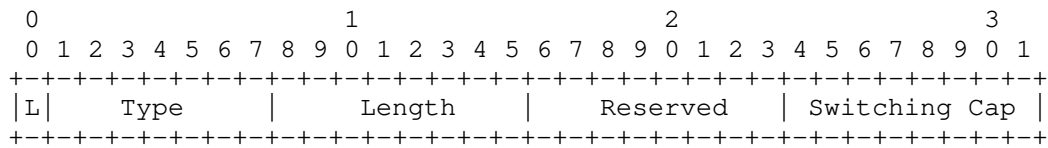


Figure 5 Switching Capability subobject in ERBO

- o L-bit: 0 indicates that the attribute specified MUST be included. 1 indicates that the attribute specified SHOULD be included.
- o Type: To be defined.
- o Length: It is always 4.
- o Switching Capability (SC): Indicates which corresponding server layer should be triggered by the boundary node. The value of switching capability is the same as the one in [RFC3471].

3.2.2. Encoding Type subobject

A new subobject, called the encoding type subobject, is defined for use in the ERBO. The format of the encoding type subobject is defined as follows:

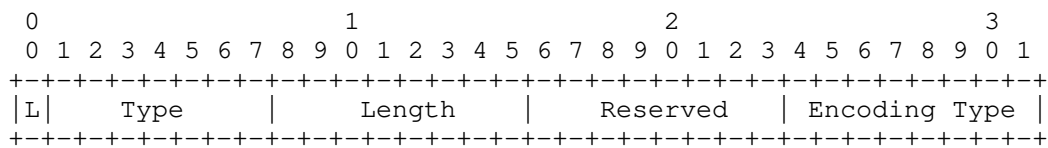


Figure 6 Encoding Type subobject in ERBO

- o L-bit: 0 indicates that the attribute specified MUST be included. 1 indicates that the attribute specified SHOULD be included.

- o Type: To be defined.
- o Length: It is always 4.
- o Encoding Type: It may need to further indicate which encoding type (e.g., SDH/SONET or G.709 in TDM) should be triggered. It is the same as the one in [RFC3471].

3.2.3. Signal Type subobject

A new subobject, called the signal type subobject, is defined for use in the ERBO. The format of the encoding type subobject is defined as follows:

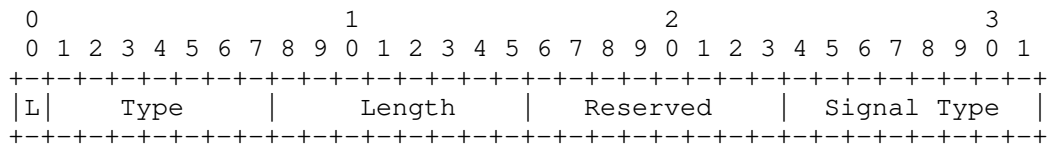


Figure 7 Signal Type subobject in ERBO

- o L-bit: 0 indicates that the attribute specified MUST be included.
1 indicates that the attribute specified SHOULD be included.
- o Type: To be defined.
- o Length: It is always 4.
- o Signal Type: If there are several sub-layers within one server layer, it can further indicates which sub-layer should be triggered by the boundary node. Following is the signal type in OTN.

Value	Type
-----	-----
0	Not significant
1	ODU1
2	ODU2
3	ODU3
4	ODU4
5	ODU0
6	ODUflex
7	ODUflex(G.hao)
8	ODU2e
9	STM-1
10	STM-4
11	STM-16
12	STM-64
13-255	Reserved (for future use)

3.2.4. Multiplexing Hierarchy subobject

A new subobject, called the Multiplexing Hierarchy (MH) subobject, is defined for use in the ERBO. The format of the multiplexing hierarchy subobject is defined as follows:

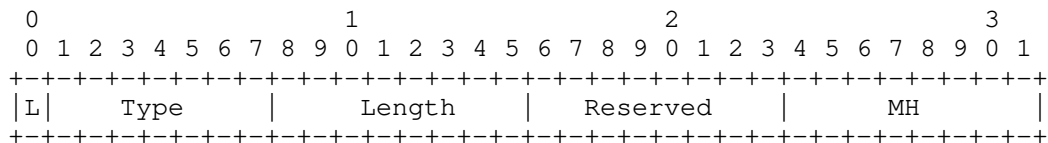


Figure 8 Multiplexing Hierarchy subobject in ERBO

- o L-bit: 0 indicates that the attribute specified MUST be included.
1 indicates that the attribute specified SHOULD be included.
- o Type: To be defined.
- o Length: It is always 4.
- o Multiplexing Hierarchy (MH): It explicitly indicates the multiplexing hierarchy used for boundary node to configure it to the data plane and trigger one specific corresponding tunnel creation. Following is the multiplexing hierarchy in current OTN.

Value	Type
-----	-----
0	ODU1-ODU0
1	ODU2-ODU0
2	ODU2-ODU1
3	ODU2-ODU1-ODU0
4	ODU2-ODUflex
5	ODU3-ODU0
6	ODU3-ODU1
7	ODU3-ODU1-ODU0
8	ODU3-ODU2
9	ODU3-ODU2-ODU0
10	ODU3-ODU2-ODU1
11	ODU3-ODU2-ODU1-ODU0
12	ODU3-ODU2-ODUflex
13	ODU3-ODUflex
14	ODU3-ODU2e
15	ODU4-ODU0
16	ODU4-ODU1
17	ODU4-ODU1-ODU0
18	ODU4-ODU2
19	ODU4-ODU2-ODU0
20	ODU4-ODU2-ODU1
21	ODU4-ODU2-ODU1-ODU0
22	ODU4-ODU2-ODUflex
23	ODU4-ODU3
24	ODU4-ODU3-ODU0
25	ODU4-ODU3-ODU1
26	ODU4-ODU3-ODU1-ODU0
27	ODU4-ODU3-ODU2
28	ODU4-ODU3-ODU2-ODU0
29	ODU4-ODU3-ODU2-ODU1
30	ODU4-ODU3-ODU2-ODU1-ODU0
31	ODU4-ODU3-ODU2-ODUflex
32	ODU4-ODU3-ODUflex
33	ODU4-ODU3-ODU2e
34	ODU4-ODUflex
35	ODU4-ODU2e

3.2.5. Signaling Procedure

In order to signal an end-to-end LSP across multi layer, the LSP source node sends the RSVP-TE PATH message with ERO which indicates LSP route and ERBO which indicates the LSP route boundary. If there are cascading FAs need to be created, there must be multiple associated ERBOs. There must be nesting routing information in ERO. The first and final address of node in ERBO SHOULD also be listed in the ERO. This ensures that they are along the LSP path. When a

interim node receives a PATH message, it will check ERBO to see if it is the layer boundary node. If a interim node isn't a layer boundary, it will process the PATH message as the normal one of single layer LSP. If a interim node finds its address is in ERBO, it is a layer boundary node. So it will directly extract another boundary egress node from ERBO. If it is necessary, it must also extract the server layer/sub-layer routing information from ERO based on a pair of boundary node. Then the layer boundary node holds the PATH message and selects or creates a server layer/sub-layer LSP based on the detailed information of subobject carried in ERBO.

3.3. Exclude Route Object (XRO)

[RFC6001] introduce SC (Switching Capability) subobjects into XRO [RFC4874] which enables (when desired) the explicit identification of at least one switching capability to be excluded from the resource selection process described multi-region signaling. This document adds more subobjects into the XRO to make multi-region and multi-layer signaling more flexible.

- o Encoding Type: explicitly indicates the encoding type should be excluded (e.g., SONET/SDH or G.709 in TDM).
- o Signal Type (ST) : explicitly indicates at least one ODUk signal type have to be excluded from the resource selection.
- o Multiplexing Hierarchy (MH): explicitly indicates at least one MH have to be excluded from the resource selection.

L bit and Attribute is the same as the Switching Capability (SC) subobject defined in [RFC6001].

3.3.1. Encoding Type subobject

A new subobject, called the encoding type subobject, is defined for use in the XRO. The format of the encoding type subobject is defined as follows:

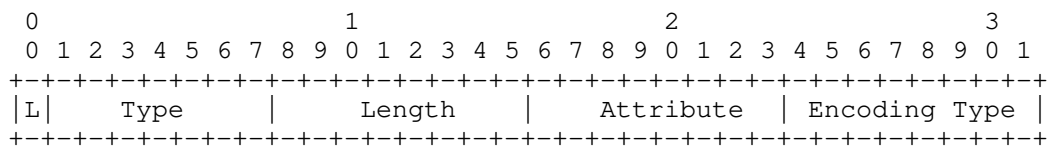


Figure 9 Encoding Type subobject in XRO

- o L-bit: 0 indicates that the attribute specified MUST be excluded. 1 indicates that the attribute specified SHOULD be avoided.
- o Type: To be defined.
- o Length: It is always 4.
- o Attribute: 0 reserved value. 1 indicates that the specified encoding type SHOULD be excluded or avoided with respect to the preceding numbered or unnumbered interface subobject.
- o Encoding Type: It indicates which Encoding Type has to be excluded. It is the same as the one in [RFC3471].

3.3.2. Signal Type subobject

A new subobject, called the signal type subobject, is defined for use in the XRO. The format of the encoding type subobject is defined as follows:

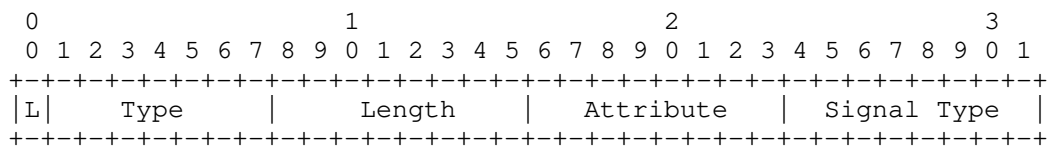


Figure 10 Signal Type subobject in XRO

- o L-bit: 0 indicates that the attribute specified MUST be excluded.
1 indicates that the attribute specified SHOULD be avoided.
- o Type: To be defined.
- o Length: It is always 4.
- o Attribute: 0 reserved value. 1 indicates that the specified signal type SHOULD be excluded or avoided with respect to the preceding numbered or unnumbered interface subobject.
- o Signal Type: It indicates which Signal Type has to be excluded.
The value of ST is the same as the one in ERBO.

3.3.3. Multiplexing Hierarchy subobject

A new subobject, called the Multiplexing Hierarchy (MH) subobject, is defined for use in the XRO. The format of the multiplexing hierarchy subobject is defined as follows:

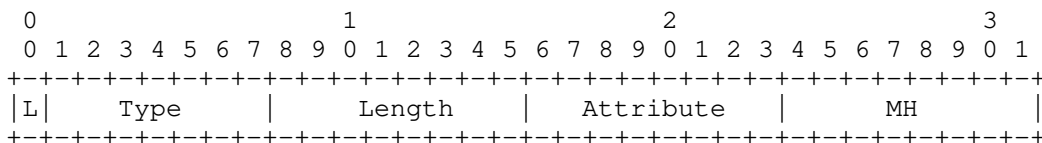


Figure 11 Multiplexing Hierarchy subobject in XRO

- o L-bit: 0 indicates that the attribute specified MUST be excluded.
1 indicates that the attribute specified SHOULD be avoided.
- o Type: To be defined.
- o Length: It is always 4.
- o Attribute: 0 reserved value. 1 indicates that the specified multiplexing hierarchy SHOULD be excluded or avoided with respect to the preceding numbered or unnumbered interface subobject.
- o Multiplexing Hierarchy (MH): It explicitly indicates which MH has to be excluded over a specified TE link, The value of multiplexing hierarchy is the same as the one in ERBO.

4. Security Considerations

This document does not introduce any new security considerations from the ones already detailed in [RFC5920] that describes the MPLS and GMPLS security threats, the related defensive techniques, and the mechanisms for detection and reporting.

5. IANA Considerations

TBD

6. References

6.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC5212] Shiimoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, July 2008.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

6.2. Informative References

- [CL-REQ] C. Villamizar, "Requirements for MPLS Over a Composite Link", draft-ietf-rtgwg-cl-requirement-04 .
- [OTNv3-OSPF]
D. Ceccarelli, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolving G.709 OTN Networks", draft-ceccarelli-ccamp-gmpls-ospf-g709-06 .

Authors' Addresses

Xihua Fu
ZTE Corporation

Email: fu.xihua@zte.com.cn

Qilei Wang
ZTE Corporation

Email: wang.qilei@zte.com.cn

Yuanlin Bao
ZTE Corporation

Email: bao.yuanlin@zte.com.cn

Ruiquan Jing
China Telecom

Email: jingrq@ctbri.com.cn

Xiaoli Huo
China Telecom

Email: huoxl@ctbri.com.cn

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 2, 2012

GMG. G.Galimberti, Ed.
Cisco
RK. R.Kunze, Ed.
Deutsche Telekom
July 1, 2011

A SNMP MIB to manage the optical colored interfaces of a DWDM network
draft-galimbe-kunze-g-698-2-snmp-mib-00

Abstract

This memo defines a portion of the Management Information Base (MIB) used by Simple Network Management Protocol (SNMP) in TCP/IP- based internets. In particular, it defines objects for managing Optical Interfaces 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. [ITU.G698.2]

The MIB module defined in this memo can be used for Optical Parameters monitoring and/or configuration of such optical interface.

Copyright Notice

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

Status of This Memo

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

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

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

This Internet-Draft will expire on January 2, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the

document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. The Internet-Standard Management Framework	4
3. Conventions	4
4. Overview	4
4.1. Optical Parameters Description	6
4.1.1. General	6
4.1.2. Parameters at Ss	7
4.1.3. Optical path from point Ss to Rs	8
4.1.4. Interface at point Rs	9
4.1.5. Alarms and Threshold definition	9
4.1.6. Performance Monitoring (PM) description	11
4.1.7. Generic Parameter description	12
4.2. Use of ifTable	13
5. Structure of the MIB Module	13
5.1. The optIfOTMn group	14
5.1.1. optIfOTMnTable	14
5.2. The optIfOTSn groups	14
5.2.1. optIfOTSn Configuration group	14
5.3. The [TEMPLATE TODO] Subtree	14
5.4. The Notifications Subtree	15
6. Object Definitions	15
7. Relationship to Other MIB Modules	17
7.1. Relationship to the [TEMPLATE TODO] MIB	17
7.2. MIB modules required for IMPORTS	17
8. Definitions	17
9. Security Considerations	17
10. IANA Considerations	18
11. Contributors	20
12. References	20
12.1. Normative References	20
12.2. Informative References	23
Appendix A. Change Log	24
Appendix B. Open Issues	24

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 Interfaces 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. Whereas the standardization of black link for 2.5 and 10G is settled for 40G and 100G interfaces and Black Link extensions are still in progress. For carrier network deployments, interoperability is a key requirement. Today it is state-of-the-art to interconnect IP Routers from different vendors and WDM transport systems using short-reach, grey interfaces. Applying the Black Link (BL) concept, routers now get directly connected to each via transport interfaces which must be interoperable to each other.

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 using optical interface parameters at the single-channel connection points between optical transmitters and the optical multiplexer, as well as between optical receivers and the optical demultiplexer in the DWDM system. This Recommendation uses a methodology which does not specify the details of the optical link, e.g. the maximum fibre length, explicitly. The Recommendation currently includes unidirectional DWDM applications at 2.5 and 10 Gbit/s with 100 GHz channel frequency spacing and may be extended to 40 and 100 Gbit/s channels with a lower channel frequency spacing.

The Building a SNMP MIB describing the optical parameters defined in G.698 [ITU.G698.2] allow the different vendors and operator to retrieve, provision and exchange information related to Optical Networks in a standardized way. This ensures interworking in case of using optical interfaces from different vendors at the end of the link. Decoupling DWDM layer from the optical layer The Optical Parameters and their values characterize the features and the performances of the Network optical components and allow a reliable network design in case of Multivendor Optical Networks.

Although RFC 3591 [RFC3591] describe and define 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.

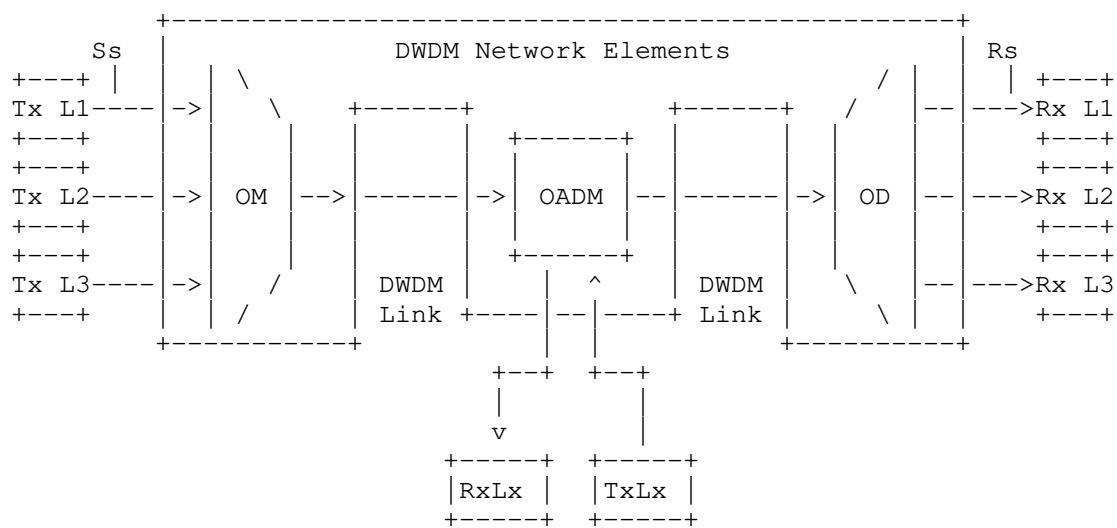
4. Overview

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 refer to [ITU.G698.2] for the Black Link and DWDM parameter description.

The optical objects will be managed using the MIB II ifTable and ifStackTable. Additional tables will also be supported to monitor layer specific status and provide performance monitoring data. In the tables, some entries are required for OTN systems only. A Configuration (Config) table, Current Performance Monitoring (PM) table, and Interval PM table will be maintained for the OTSn, OMSn, OChGroup, and OCh layers on a source and sink trail termination basis. These tables will be linked to the ifTable by using the ifIndex that is associated with that layer.

An Alarm (Aalarm) table will be maintained for the OTSn, OMSn, OChGroup, and OCh layers on a source and sink trail termination basis. These tables will be linked to the ifTable by using the ifIndex that is associated with that layer.

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]

These objects are used when the particular media being used to

realize an interface is an Optical Transport interface. At present, this applies to these values of the ifType variable in the Internet-standard MIB:

opticalChannel (195), opticalChannelGroup (219), opticalTransport (196).

The definitions contained herein are based on the OTN specifications in ITU-T G.872 [ITU.G872], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1].

4.1. Optical Parameters Description

The terminology used in this document describes the optical parameters, the states and the Alarms at the points Ss, Rs and DWDM depicted in fig.1. The terms are defined in ITU-T Recommendations G.698.2 [ITU.G698.2]. Those definitions are made to increase the readability of the document.

4.1.1. General

Minimum channel spacing:

This is the minimum nominal difference in frequency between two adjacent channels (G).

Bit rate/line coding of optical tributary signals:

Optical tributary signal class NRZ 2.5G or NRZ 10G nominally 2.4 Gbit/s to nominally 10.71 Gbit/s. 40Gbit/s and 100Gbit/s are under definition (G, S).

Channel Modulation Format:

This parameter indicate what kind of modulation format is used at Ss (G).

FEC Coding:

This parameter indicate what Forward Error Correction (FEC) code is used at Ss and Rs (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, 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.

single-channel application codes (see G.698.2):

This parameter indicates the transceiver allocation 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**Maximum and minimum mean channel output power:**

The mean launched power at Ss is the average power of a pseudo-random data sequence coupled into the DWDM link. It is defined the change (Max and Min) of the parameter (G, S)

Minimum and maximum central frequency:

The central frequency is the nominal single-channel frequency 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 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]

Lowest OSNR at Ss with worst case (residual) dispersion. Lowest OSNR at Ss with no dispersion (G)

Electrical Signal Framing:

This is the indication of what framing (GE, Sonet/SDH, OTN) the Ss and Rs ports are set (G, S)

4.1.3. Optical path from point Ss to Rs**Maximum and minimum (residual) chromatic dispersion:**

These parameters define the maximum and minimum value of the optical path "end to end chromatic dispersion" that the system shall be able to tolerate. (G)

Minimum optical return loss at Ss:

These parameter defines minimum optical return loss 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 (G)

4.1.4. Interface at point Rs**Maximum and minimum mean input power:**

The maximum and minimum values of the average received power 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)

Minimum maximum Chromatic Dispersion (CD) :

This parameter defines the CD range a Receiver (Rs) can tolerate in order to decode the received signal (G)

Maximum Polarization Mode Dispersion (PMD) :

This parameter defines the maximum PMD value a Receiver (Rs) can tolerate in order to decode the received signal (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

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.872 [ITU.G872], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1

[ITU.G874.1]. The SNMP MIB of the above list is already defined and specified by the RFC3591

OTN alarms defined in RFC3591:

Threshold Crossing Alert (TCA Alarm)

LOW-TXPOWER

HIGH-TXPOWER

LOW-RXPOWER

HIGH-RXPOWER

OTUk-LOF or more generic LOF

Backward Defect Indication (BDI)

Trace Identifier Mismatch (tim)

Signal Degrade (sd)

Server Signal Failure (SSF)

Alarm Indication Signal (AIS)

Loss of Multiframe (lom)

OTN Thresholds (for TCA) defined in RFC3591

LOW-TXPOWER

HIGH-TXPOWER

LOW-RXPOWER

HIGH-RXPOWER

The list below reports the new Alarms and Thresholds not managed in RFC3591

Laser Bias Current:

This parameter report the Bias current of the Laser Transmitter (G)

Laser Bias Current Threshold:

This parameter is to set the Bias current Threshold of the Laser Transmitter used to rise the related Alarm (G, S)

Forward Defect Indication (FDI):

This parameter indicates a notification to the receiver that a failure occurred in the network (G)

Backward Error Indication (BEI):

This parameter indicates the number of Errors occurred in the opposite line direction (G)

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 = 1×10^{-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 (bber) :

The ratio of Background Block Errors (BBE) to total blocks in available time during a fixed measurement interval. The count of total blocks excludes all blocks during SESs. (G)

FEC corrected Bit Error (FECcorrErr):

The number of bits corrected by the FEC are counted over one second (G)

FEC un-corrected Bit Error :

The number of bits un-corrected by the FEC are counted over one second (G)

Pre-FEC Bit Error :

The number of Errored bits at receiving side before the FEC function counted over one second (G)

OTN Valid Intervals :

The number of contiguous 15 minute intervals for which valid OTN performance monitoring data is available for the particular interface (G)

FEC Valid Intervals :

The number of contiguous 15 minute intervals for which valid FEC PM data is available for the particular interface. (G)

4.1.7. Generic Parameter description

This section describes the Generic Parameters at Ss and Rs points according to ITU-T Recommendations G.872 [ITU.G872], G.709 [ITU.G709], G.798 [ITU.G798], G.874 [ITU.G874], and G.874.1 [ITU.G874.1].

Interface Admin Status :

The Administrative Status of an Interface: Up/Down - In Service/Out of Service (can be Automatic in Service) (G/S)

Interface Operational Status :

The Operational Status of an Interface: Up/Down - In Service/Out of Service (G)

Loopbacks :

The Interface loopbacks used for maintenance purposes, they are Terminal or Line (may be with send AIS) (G/S)

TDC (mode/status/settings) :

Tuneable Dispersion Compensation settings (G/S)

Pre-FEC BER (Mantissa + Exponent) :

Bit Error Rate at the Rs interface before error correction (G/S)

Q factor :

(G)

Q margin :

(G)

4.2. Use of ifTable

This section specifies how the MIB II interfaces group, as defined in RFC 2863 [RFC2863], is used for optical interfaces. As described in the RFC 3591 figure 1 [RFC3591] Only the ifGeneralInformationGroup will be supported for the ifTable and the ifStackTable to maintain the relationship between the various layers. The OTN layers are managed in the ifTable using IfEntries that correlate to the layers depicted in Figure 1. For example, a DWDM device with an Optical Network Node Interface (ONNI) will have an Optical Transmission Section (OTS) physical layer, an Optical Multiplex Section (OMS) layer (transports multiple optical channels), and an Optical Channel (OCh) layer. There is a one to one relationship between the OMS and OTS layers. The OMS layer has fixed connectivity via the OTS and thus no connectivity flexibility at the OMS layer is supported. This draft extend the RFC 3591 [RFC3591] as far as the OMSn and OTSn are concerned. The sections 2.5 and 2.6 of RFC 3591 [RFC3591] must be considered as a reference for the ifStackTable use and Optical Network Terminology.

5. Structure of the MIB Module

The managed Optical Networking interface objects are arranged into the following groups of tables:

The optIfOTMn group handles the OTM information structure of an optical interface.

optIfOTMnTable

The optIfPerfMon group handles the current 15-minute and 24-hour interval elapsed time, as well as the number of 15-minute intervals

for all layers

optIfPerfMonIntervalTable

The optIfOTSn groups handle the configuration and performance monitoring information for OTS layers.

optIfOTSnConfigTable

optIfOTSnSinkCurrentTable

optIfOTSnSinkIntervalTable

optIfOTSnSinkCurDayTable

optIfOTSnSinkPrevDayTable

optIfOTSnSrcCurrentTable

optIfOTSnSrcIntervalTable

optIfOTSnSrcCurDayTable

optIfOTSnSrcPrevDayTable

5.1. The optIfOTMn group

5.1.1. optIfOTMnTable

This table contains the OTM structure information of an optical interface.

5.2. The optIfOTSn groups

5.2.1. optIfOTSn Configuration group

5.2.1.1. optIfOTSn Configuration Table

This table contains information on configuration of optIfOTSn interfaces, in addition to the information on such interfaces contained in the ifTable.

5.3. The [TEMPLATE TODO] Subtree

5.4. The Notifications Subtree

6. Object Definitions

```
OPT-IF-MIB DEFINITIONS ::= BEGIN
```

```
IMPORTS
```

```
    MODULE-IDENTITY, OBJECT-TYPE, Gauge32, Integer32,  
        Unsigned32, transmission  
        FROM SNMPv2-SMI  
    TEXTUAL-CONVENTION, RowPointer, RowStatus, TruthValue  
        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 OTN Interface objects.

optIfMibModule MODULE-IDENTITY
    LAST-UPDATED "200308130000Z"
    ORGANIZATION "IETF ATOM MIB Working Group"
    CONTACT-INFO
        "WG charter:
         http://www.ietf.org/html.charters/atommib-charter.html

        Mailing Lists:
            General Discussion: atommib@research.telcordia.com
            To Subscribe: atommib-request@research.telcordia.com
        Editor: Hing-Kam Lam
        Postal: Lucent Technologies, Room 4C-616
                101 Crawfords Corner Road
                Holmdel, NJ 07733
                Tel: +1 732 949 8338
                Email: hklam@lucent.com"
    DESCRIPTION
        "The MIB module to describe pre-OTN and OTN interfaces.

        Copyright (C) The Internet Society (2003). This version
        of this MIB module is part of RFC 3591; see the RFC
        itself for full legal notices."
    REVISION "200308130000Z"
    DESCRIPTION
        "Initial version, published as RFC 3591."
    ::= { transmission 133 }
```

```
OptIfBitRateK ::= TEXTUAL-CONVENTION
    STATUS current
    DESCRIPTION
        "Indicates the index 'k' that is used to
        represent a supported bit rate and the different
        versions of OPuk, ODUk and OTUk.
        Allowed values of k are defined in ITU-T G.709.
        Currently allowed values in G.709 are:
            k=1 represents an approximate bit rate of 2.5 Gbit/s,
            k=2 represents an approximate bit rate of 10 Gbit/s,
            k=3 represents an approximate bit rate of 40 Gbit/s."
    SYNTAX Integer32
```

`optIfOTMnBitRates OBJECT-TYPE``SYNTAX BITS { bitRateK1(0), bitRateK2(1), bitRateK3(2) }``MAX-ACCESS read-only``STATUS current``DESCRIPTION`

"This attribute is a bit map representing the bit rate or set of bit rates supported on the interface.

The meaning of each bit position is as follows:

bitRateK1(0) is set if the 2.5 Gbit/s rate is supported

bitRateK2(1) is set if the 10 Gbit/s rate is supported

bitRateK3(2) is set if the 40 Gbit/s rate is supported

Note that each bit position corresponds to one possible value of the type OptIfBitRateK.

The default value of this attribute is system specific."

::= { optIfOTMnEntry 3 }

7. Relationship to Other MIB Modules

7.1. Relationship to the [TEMPLATE TODO] MIB

7.2. MIB modules required for IMPORTS

8. Definitions

[TEMPLATE TODO]: put your valid MIB module here.

A list of tools that can help automate the process of checking MIB definitions can be found at <http://www.ops.ietf.org/mib-review-tools.html>

9. Security Considerations

There are a number of management objects defined in this MIB module with a MAX-ACCESS clause of read-write and/or read-create. Such objects may be considered sensitive or vulnerable in some network environments. The support for SET operations in a non-secure environment without proper protection can have a negative effect on network operations. These are the tables and objects and their sensitivity/vulnerability:

o

There are no management objects defined in this MIB module that have a MAX-ACCESS clause of read-write and/or read-create. So, if this MIB module is implemented correctly, then there is no risk that an intruder can alter or create any management objects of this MIB module via direct SNMP SET operations.

Some of the readable objects in this MIB module (i.e., objects with a MAX-ACCESS other than not-accessible) may be considered sensitive or vulnerable in some network environments. It is thus important to control even GET and/or NOTIFY access to these objects and possibly to even encrypt the values of these objects when sending them over the network via SNMP.

SNMP versions prior to SNMPv3 did not include adequate security. Even if the network itself is secure (for example by using IPsec), even then, there is no control as to who on the secure network is allowed to access and GET/SET (read/change/create/delete) the objects in this MIB module.

It is RECOMMENDED that implementers consider the security features as provided by the SNMPv3 framework (see [RFC3410], section 8), including full support for the SNMPv3 cryptographic mechanisms (for authentication and privacy).

Further, deployment of SNMP versions prior to SNMPv3 is NOT RECOMMENDED. Instead, it is RECOMMENDED to deploy SNMPv3 and to enable cryptographic security. It is then a customer/operator responsibility to ensure that the SNMP entity giving access to an instance of this MIB module is properly configured to give access to the objects only to those principals (users) that have legitimate rights to indeed GET or SET (change/create/delete) them.

10. IANA Considerations

Option #1:

The MIB module in this document uses the following IANA-assigned OBJECT IDENTIFIER values recorded in the SMI Numbers registry:

Descriptor	OBJECT IDENTIFIER value
-----	-----
sampleMIB	{ mib-2 XXX }

Option #2:

Editor's Note (to be removed prior to publication): the IANA is requested to assign a value for "XXX" under the 'mib-2' subtree and to record the assignment in the SMI Numbers registry. When the assignment has been made, the RFC Editor is asked to replace "XXX" (here and in the MIB module) with the assigned value and to remove this note.

Note well: prior to official assignment by the IANA, an internet draft MUST use placeholders (such as "XXX" above) rather than actual numbers. See RFC4181 Section 4.5 for an example of how this is done in an internet draft MIB module.

Option #3:

This memo includes no request to IANA.

11. Contributors

Arnold Mattheus
Deutsche Telekom
Darmstadt
Germany
email a.mattheus@telekom.de

Manuel Paul
Deutsche Telekom
Berlin
Germany
email Manuel.Paul@telekom.de

Frank Luennemann
Deutsche Telekom
Munster
Germany
email Frank.Luennemann@telekom.de

Najam Saquib
Cisco
Ludwig-Erhard-Strasse 3
ESCHBORN, HESSEN 65760
GERMANY
email nasaquib@cisco.com

Walid Wakim
Cisco
9501 Technology Blvd
ROSEMONT, ILLINOIS 60018
UNITED STATES
email wwakim@cisco.com

Ori Gerstel
Cisco
32 HaMelacha St., (HaSharon Bldg)
SOUTH NETANYA, HAMERKAZ 42504
ISRAEL
email ogerstel@cisco.com

12. References

12.1. Normative References

[RFC2863]

McCloghrie, K. and F.
Kastenholz, "The
Interfaces Group MIB",

- RFC 2863, June 2000.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, April 1999.
- [RFC2580] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Conformance Statements for SMIv2", STD 58, RFC 2580, April 1999.
- [RFC3591] Lam, H-K., Stewart, M., and A. Huynh, "Definitions of Managed Objects for the Optical Interface Type", RFC 3591, September 2003.
- [ITU.G698.2] International Telecommunications Union, "Amplified multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation

G.698.2, November 2009.

[ITU.G709]

International
Telecommunications
Union, "Interface for
the Optical Transport
Network (OTN)", ITU-
T Recommendation G.709,
March 2003.

[ITU.G872]

International
Telecommunications
Union, "Architecture of
optical transport
networks", ITU-
T Recommendation G.872,
November 2001.

[ITU.G798]

International
Telecommunications
Union, "Characteristics
of optical transport
network hierarchy
equipment functional
blocks", ITU-
T Recommendation G.798,
October 2010.

[ITU.G874]

International
Telecommunications
Union, "Management
aspects of optical
transport network
elements", ITU-
T Recommendation G.874,
July 2010.

[ITU.G874.1]

International
Telecommunications
Union, "Optical
transport network (OTN):
Protocol-neutral
management information
model for the network
element view", ITU-
T Recommendation
G.874.1, January 2002.

- | | |
|--------------|---|
| [ITU.G959.1] | International
Telecommunications
Union, "Optical
transport network
physical layer
interfaces", ITU-
T Recommendation
G.959.1, November 2009. |
| [ITU.G826] | International
Telecommunications
Union, "End-to-end error
performance parameters
and objectives for
international, constant
bit-rate digital paths
and connections", ITU-
T Recommendation G.826,
November 2009. |
| [ITU.G8201] | International
Telecommunications
Union, "Error
performance parameters
and objectives for
multi-operator
international paths
within the Optical
Transport Network
(OTN)", ITU-
T Recommendation G.8201,
September 2003. |
| [ITU.G694.1] | International
Telecommunications
Union, "Spectral grids
for WDM applications:
DWDM frequency grid",
ITU-T Recommendation
G.694.1, June 2002. |

12.2. Informative References

- [RFC3410] Case, J., Mundy, R., Partain, D., and B. Stewart, "Introduction and Applicability Statements for Internet-

- [RFC2629] Standard Management Framework", RFC 3410, December 2002.
- [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.
- [RFC4181] Heard, C., "Guidelines for Authors and Reviewers of MIB Documents", BCP 111, RFC 4181, September 2005.
- [I-D.kunze-black-link-management-framework] Kunze, R., "A framework for Black Link Management and Control", draft-kunze-black-link-management-framework-00 (work in progress), March 2011.
- [RFC4054] Strand, J. and A. Chiu, "Impairments and Other Constraints on Optical Layer Routing", RFC 4054, May 2005.

Appendix A. Change Log

This optional section should be removed before the internet draft is submitted to the IESG for publication as an RFC.

Note to RFC Editor: please remove this appendix before publication as an RFC.

Appendix B. Open Issues

Note to RFC Editor: please remove this appendix before publication as an RFC.

Authors' Addresses

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

Phone: +390392091462
EMail: ggalimbe@cisco.com

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

Phone: +49xxxxxxxxxxx
EMail: RKunze@telekom.de

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 5, 2012

WJ. He, Ed.
F. Zhang
ZTE
July 4, 2011

RSVP-TE Extensions to Notification for Shared Mesh Protection
draft-he-ccamp-notification-shared-mesh-protection-00

Abstract

The shared mesh protection(SMP) mechanism enables multiple protection paths within a shared mesh protection domain to share protection resources, which allows only one of the n working paths to be protected at the same time. This document extends RSVP-TE to notify the state of shared resources in MPLS Transport Profile (MPLS-TP) mesh topology.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 5, 2012.

Copyright Notice

Copyright (c) 2011 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. Shared Mesh Protection	3
3.1. Shared Mesh Protection Planning	4
3.2. Signaling Protection LSPs	4
3.3. Processing	4
3.3.1. Basic Operation	5
3.3.2. Rerouting	5
3.3.3. Preemption	5
4. IANA Considerations	6
5. Security Considerations	6
6. Acknowledgement	6
7. Informative References	6
Authors' Addresses	7

1. Introduction

In mesh protection, network resources may be shared to provide protection for working paths that do not share the same endpoints. This form of protection can make very efficient use of network resources, but requires careful synchronization to ensure that only one set of traffic is switched to the protection resources at any time.

[RFC4872] defines the shared mesh restoration schemes based on control plane extensions, but does not cover the shared mesh protection scenarios.

In order to coordinate the use of protection resources, this document specifies the notification schemes to notify the endpoint of the protecting LSP the state of the shared resource.

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. Shared Mesh Protection

Consider the following topology:

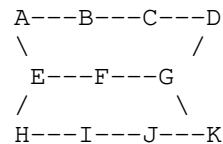


Figure 1 : A Shared Mesh Protection Topology

The working LSPs W1[via nodes A,B,C,D] and W2[via nodes H,I,J,K] could be protected by P1[via nodes A,E,F,G,D] and P2[via nodes H,E,F,G,K] respectively. For both cases, 1:1 protection may be used.

Thus, it is possible for the network resources on the segment EFG to be shared by the two protection paths. In this way, shared mesh protection can substantially reduce the amount of network resources that have to be reserved.

If there are no failures affecting either of the two working paths, the network segment EFG may carry no traffic. In the event of only one failure, the segment EFG carries traffic from the working path that experiences the failure.

3.1. Shared Mesh Protection Planning

Shared mesh protection will typically be subject to careful network planning. Operator plans the shared mesh protection group (SMPG) which includes the protected paths and protecting paths. Different SPMGs do not share protection resources and are protected independently.

In Figure 1, the working LSP W1,W2 and protecting LSPs P1,P2 consist of a shared mesh protection group, in which the protecting LSPs P1 and P2 share the segment FEG although they belong to different sessions. In order to achieve this, the "Resource Sharing" Association type that defined in [RFC4873] and [I-D.ietf-ccamp-assoc-ext] can be used here. When operators plan shared mesh protection group, they will assign a group ID and a virtual address for the shared mesh protection group. The protecting LSP will be signaled with the "Resource Sharing" type ASSOCIATION object, the Association ID is set to the group ID, and the Association source is set to the group virtual address.

3.2. Signaling Protection LSPs

When the protecting LSPs are signaled, the PROTECTION object, Notify Request object and "Recovery" type ASSOCIATION object are included in the Path message. Furthermore, the "Resource Sharing" type ASSOCIATION object SHOULD be inserted, the Association ID set to the associated protection group ID and the Association source set to the protection group virtual address.

Any node processing a Path message for which the node does not have a matching state, and which contains an ASSOCIATION object with a "Resource Sharing" type, examines existing LSPs for matching Association Type, Association Source, and Association ID values. If any match is found, then [RFC3209] style resource sharing SHOULD be provided between the new and old LSPs.

3.3. Processing

This section illustrates the realization of the shared mesh protection for the topology shown in Figure 1.

3.3.1. Basic Operation

If a failure occurs on the W2, the shared mesh protection will operate as follows:

- a. Node H detects the signal failure, switches the traffic to P2, and sends Path message to node E with O bit of protection object setting to 1.
- b. Upon receipt of the Path message with the O bit of protection object from 0 to 1, node E compares the protection switching priority of P2 and P1, then send notify message with the new error code/sub-code "notify error/resource occupied by the high priority" or "notify error/resource occupied by the low priority" to P1's ingress node.

When the fault of the working LSP disappears, the shared mesh protection will operate as follows:

- a. The ingress node H will switch traffic to the working LSP, and refresh the Path message with the O bit of protection object setting to 0.
- b. Upon receiving the Path message with the O bit of protection object from 1 to 0, sharing start endpoint (node E) will send notify message with new error code/sub-code "notify error/resource available" to the other protecting LSP's ingress node.

3.3.2. Rerouting

If the ingress of the protecting LSP receives notify message with "notify error/resource occupied by the high priority", the node should reroute the protecting LSP. Because that the traffic of higher priority LSP may also be lost during the preemption, the node may also reroute for a more optimized path according to the local policy, when the node receives notify message with "notify error/resource occupied by the low priority". If the protecting LSP reroutes, the new LSP will exclude the shared segment which was occupied by the other LSP.

3.3.3. Preemption

During the sharing resource was occupied by one of the protecting LSPs, the other working LSP may also experiences some fault. In this case, these resource MUST be preempted by the high priority LSP.

In Figure 1, if a failure occurs on the W1 while the W2 is still in failure state, the shared mesh protection will operate as follows:

- o The node A will not switch the traffic, if it has received the notification that the resource has been occupied by the high priority.
- o The node A has not received the notification that the resource has been occupied by the high priority LSP. The operation is as follows:
 1. Node A switches the traffic to P1, and sends Path message to node E with O bit of protection object setting to 1.
 2. Once Node E (sharing start endpoint) receives the Path message with the O bit of protection object from 0 to 1, it compares the protection switching priority of P2 and P1. The node will send notify message with the new error code/sub-code "notify error/resource occupied by the high priority" to P2's ingress node.
 3. Upon receipt of the notify message that the resource has been occupied by the high priority, node H will switch the traffic from P2 to W2, and sends Path message to node E with O bit of protection object setting to 0.

4. IANA Considerations

TBD

5. Security Considerations

TBD

6. Acknowledgement

TBD

7. Informative References

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

Berger, L., Faucheur, F., and A. Narayanan, "RSVP Association Object Extensions", draft-ietf-ccamp-assoc-ext-00 (work in progress), May 2011.

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

[RFC4872] Lang, J., Rekhter, Y., and D. Papadimitriou, "RSVP-TE

Extensions in Support of End-to-End Generalized Multi-
Protocol Label Switching (GMPLS) Recovery", RFC 4872,
May 2007.

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

Authors' Addresses

Wenjuan He (editor)
ZTE

Email: he.wenjuan1@zte.com.cn

Fei Zhang
ZTE

Email: zhang.fei3@zte.com.cn

Internet Draft
Updates: 2205, 3209, 3473
Category: Standards Track
Expiration Date: November 1, 2011

Lou Berger (LabN)
Francois Le Faucheur (Cisco)
Ashok Narayanan (Cisco)

May 1, 2011

RSVP Association Object Extensions

draft-ietf-ccamp-assoc-ext-00.txt

Abstract

The RSVP ASSOCIATION object was defined in the context of GMPLS (Generalized Multi-Protocol Label Switching) controlled label switched paths (LSPs). In this 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 this document defines how the ASSOCIATION object can be more generally applied. This document also defines extended ASSOCIATION objects which, in particular, can be used in the context of Transport Profile of Multiprotocol Label Switching (MPLS-TP). This document updates RFC 2205, RFC 3209, and RFC 3473.

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on November 1, 2011

Copyright and License Notice

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

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

Table of Contents

1	Introduction	3
1.1	Conventions Used In This Document	4
2	Non-GMPLS Recovery Usage	4
2.1	Upstream Initiated Association	4
2.1.1	Path Message Format	5
2.1.2	Path Message Processing	5
2.2	Downstream Initiated Association	6
2.2.1	Resv Message Format	7
2.2.2	Resv Message Processing	7
2.3	Association Types	8
2.3.1	Resource Sharing Association Type	8
3	IPv4 and IPv6 Extended ASSOCIATION Objects	9
3.1	IPv4 and IPv6 Extended ASSOCIATION Object Format	10
3.2	Processing	11
4	Security Considerations	13
5	IANA Considerations	13
5.1	IPv4 and IPv6 Extended ASSOCIATION Objects	13
5.2	Resource Sharing Association Type	14
6	Acknowledgments	14
7	References	14
7.1	Normative References	14
7.2	Informative References	15
8	Authors' Addresses	15

1. Introduction

End-to-end and segment recovery are defined for GMPLS (Generalized Multi-Protocol Label Switching) controlled label switched paths (LSPs) in [RFC4872] and [RFC4873] respectively. Both definitions use the ASSOCIATION object to associate recovery LSPs with the LSP they are protecting. Additional narrative on how such associations are to be identified is also provided in [ASSOC-INFO].

This document expands the possible usage of the ASSOCIATION object to non-GMPLS recovery contexts. This document reviews how association should be made in the case where the object is carried in a Path message and defines usage with Resv messages. This section also discusses usage of the ASSOCIATION object outside the context of GMPLS LSPs.

Some examples of non-LSP association in order to enable resource sharing are:

- o Voice Call-Waiting:
A bidirectional voice call between two endpoints A and B is signaled using two separate unidirectional RSVP reservations for the flows A->B and B->A. If endpoint A wishes to put the A-B call on hold and join a separate A-C call, it is desirable that network resources on common links be shared between the A-B and A-C calls. The B->A and C->A subflows of the call can share resources using existing RSVP sharing mechanisms, but only if they use the same destination IP addresses and ports. However, there is no way in RSVP today to share the resources between the A->B and A->C subflows of the call since by definition the RSVP reservations for these subflows must have different IP addresses in the SESSION objects.
- o Voice Shared Line:
A single number that rings multiple endpoints (which may be geographically diverse), such as phone lines on a manager's desk and their assistant. A VoIP system that models these calls as multiple P2P unicast pre-ring reservations would result in significantly over-counting bandwidth on shared links, since today unicast reservations to different endpoints cannot share bandwidth.
- o Symmetric NAT:
RSVP permits sharing of resources between multiple flows addressed to the same destination D, even from different senders S1 and S2. However, if D is behind a NAT operating in symmetric mode [RFC5389], it is possible that the destination port of the flows S1->D and S2->D may be different outside the NAT. In this case, these flows cannot share resources using RSVP today, since the SESSION objects for these two flows outside the NAT would have different ports.

This document also defines the extended ASSOCIATION objects which can be used in the context of Transport Profile of Multiprotocol Label Switching (MPLS-TP). Although, the scope of the extended ASSOCIATION objects is not limited to MPLS-TP.

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. Non-GMPLS Recovery Usage

While the ASSOCIATION object, [RFC4872], is defined in the context of GMPLS Recovery, the object can have wider application. [RFC4872] defines the object to be used to "associate LSPs with each other", and then defines an Association Type field to identify the type of association being identified. It also defines that the Association Type field is to be considered when determining association, i.e., there may be type-specific association rules. As discussed above, this is the case for Recovery type association objects. The text above, notably the text related to resource sharing types, can also be used as the foundation for a generic method for associating LSPs when there is no type-specific association defined.

The remainder of this section defines the general rules to be followed when processing ASSOCIATION objects. Object usage in both Path and Resv messages is discussed. The usage applies equally to GMPLS LSPs [RFC3473], MPLS LSPs [RFC3209] and non-LSP RSVP sessions [RFC2205], [RFC2207], [RFC3175] and [RFC4860]. As described below, association is always done based on matching either Path state to Path state, or Resv state to Resv state, but not Path state to Resv State. This section applies to the ASSOCIATION objects defined in [RFC4872].

2.1. Upstream Initiated Association

Upstream initiated association is represented in ASSOCIATION objects carried in Path messages and can be used to associate RSVP Path state across MPLS Tunnels / RSVP sessions. (Note, per [RFC3209] an MPLS tunnel is represented by a RSVP SESSION object, and multiple LSPs may be represented within a single tunnel.) Cross-session association based on Path state is defined in [RFC4872]. This definition is extended by this section, which defined generic association rules and usage for non-LSP uses. This section does not modify processing required to support [RFC4872] and [RFC4873], and which is reviewed above in Section 3 of [ASSOC-INFO].

2.1.1. Path Message Format

This section provides the Backus-Naur Form (BNF), see [RFC5511], for Path messages containing ASSOCIATION objects. BNF is provided for both MPLS and for non-LSP session usage. Unmodified RSVP message formats and some optional objects are not listed.

The format for MPLS and GMPLS sessions is unmodified from [RFC4872], and can be represented based on the BNF in [RFC3209] as:

```
<Path Message> ::= <Common Header> [ <INTEGRITY> ]
                   <SESSION> <RSVP_HOP>
                   <TIME_VALUES>
                   [ <EXPLICIT_ROUTE> ]
                   <LABEL_REQUEST>
                   [ <SESSION_ATTRIBUTE> ]
                   [ <ASSOCIATION> ... ]
                   [ <POLICY_DATA> ... ]
                   <sender descriptor>
```

The format for non-LSP sessions as based on the BNF in [RFC2205] is:

```
<Path Message> ::= <Common Header> [ <INTEGRITY> ]
                   <SESSION> <RSVP_HOP>
                   <TIME_VALUES>
                   [ <ASSOCIATION> ... ]
                   [ <POLICY_DATA> ... ]
                   [ <sender descriptor> ]
```

In general, relative ordering of ASSOCIATION objects with respect to each other as well as with respect to other objects is not significant. Relative ordering of ASSOCIATION objects of the same type SHOULD be preserved by transit nodes. Association type specific ordering requirements MAY be defined in the future.

2.1.2. Path Message Processing

This section is based on the processing rules described in [RFC4872] and [RFC4873], and which is reviewed in [ASSOC-INFO]. These procedures apply equally to GMPLS LSPs, MPLS LSPs and non-LSP session state.

A node that wishes to allow downstream nodes to associate Path state across RSVP sessions MUST include an ASSOCIATION object in the outgoing Path messages corresponding to the RSVP sessions to be associated. In the absence of Association Type-specific rules for identifying association, the included ASSOCIATION objects MUST be identical. When there is an Association Type-specific definition of association rules, the definition SHOULD allow for association based on identical ASSOCIATION objects. This document does not define any

Association Type-specific rules. (See Section 3 for a discussion of an example of Association Type-specific rules which are derived from [RFC4872].)

When creating an ASSOCIATION object, the originator MUST format the object as defined in Section 16.1 of [RFC4872]. The originator MUST set the Association Type field based on the type of association being identified. The Association ID field MUST be set to a value that uniquely identifies the sessions to be associated within the context of the Association Source field. The Association Source field MUST be set to a unique address assigned to the node originating the association.

A downstream node can identify an upstream initiated association by performing the following checks. When a node receives a Path message it MUST check each ASSOCIATION object received in the Path message to see if it contains an Association Type field value supported by the node. For each ASSOCIATION object containing a supported association type, the node MUST then check to see if the object matches an ASSOCIATION object received in any other Path message. To perform this matching, a node MUST examine the Path state of all other sessions and compare the fields contained in the newly received ASSOCIATION object with the fields contained in the Path state's ASSOCIATION objects. An association is deemed to exist when the same values are carried in all fields of the ASSOCIATION objects being compared. Processing once an association is identified is type specific and is outside the scope of this document.

Note that as more than one association may exist, the described matching MUST continue after a match is identified, and MUST be performed against all local Path state.

Unless there are type-specific processing rules, downstream nodes MUST forward all ASSOCIATION objects received in a Path message in any corresponding outgoing Path messages.

2.2. Downstream Initiated Association

Downstream initiated association is represented in ASSOCIATION objects carried in Resv messages and can be used to associate RSVP Resv state across MPLS Tunnels / RSVP sessions. Cross-session association based on Path state is defined in [RFC4872]. This section defines cross-session association based on Resv state. This section places no additional requirements on implementations supporting [RFC4872] and [RFC4873].

2.2.1. Resv Message Format

This section provides the Backus-Naur Form (BNF), see [RFC5511], for Resv messages containing ASSOCIATION objects. BNF is provided for both MPLS and for non-LSP session usage. Unmodified RSVP message formats and some optional objects are not listed.

The format for MPLS, GMPLS and non-LSP sessions are identical, and is represented based on the BNF in [RFC2205] and [RFC3209]:

```
<Resv Message> ::= <Common Header> [ <INTEGRITY> ]  
                  <SESSION> <RSVP_HOP>  
                  <TIME_VALUES>  
                  [ <RESV_CONFIRM> ] [ <SCOPE> ]  
                  [ <ASSOCIATION> ... ]  
                  [ <POLICY_DATA> ... ]  
                  <STYLE> <flow descriptor list>
```

Relative ordering of ASSOCIATION objects with respect to each other as well as with respect to other objects is not currently significant. Relative ordering of ASSOCIATION objects of the same type MUST be preserved by transit nodes. Association type specific ordering requirements MAY be defined in the future.

2.2.2. Resv Message Processing

This section apply equally to GMPLS LSPs, MPLS LSPs and non-LSP session state.

A node that wishes to allow upstream nodes to associate Resv state across RSVP sessions MUST include an ASSOCIATION object in the outgoing Resv messages corresponding to the RSVP sessions to be associated. In the absence of Association Type-specific rules for identifying association, the included ASSOCIATION objects MUST be identical. When there is an Association Type-specific definition of association rules, the definition SHOULD allow for association based on identical ASSOCIATION objects. This document does not define any Association Type-specific rules.

When creating an ASSOCIATION object, the originator MUST format the object as defined in Section 16.1 of [RFC4872]. The originator MUST set the Association Type field based on the type of association being identified. The Association ID field MUST be set to a value that uniquely identifies the sessions to be associated within the context of the Association Source field. The Association Source field MUST be set to a unique address assigned to the node originating the association.

An upstream node can identify a downstream initiated association by performing the following checks. When a node receives a Resv message

it MUST check each ASSOCIATION object received in the Resv message to see if it contains an Association Type field value supported by the node. For each ASSOCIATION object containing a supported association type, the node MUST then check to see if the object matches an ASSOCIATION object received in any other Resv message. To perform this matching, a node MUST examine the Resv state of all other sessions and compare the fields contained in the newly received ASSOCIATION object with the fields contained in the Resv state's ASSOCIATION objects. An association is deemed to exist when the same values are carried in all fields of the ASSOCIATION objects being compared. Processing once an association is identified is type specific and is outside the scope of this document.

Note that as more than one association may exist, the described matching MUST continue after a match is identified, and MUST be performed against all local Resv state.

Unless there are type-specific processing rules, upstream nodes MUST forward all ASSOCIATION objects received in a Resv message in any corresponding outgoing Resv messages.

2.3. Association Types

Two association types are currently defined: recovery and resource sharing. Recovery type association is only applicable within the context of recovery, [RFC4872] and [RFC4873]. Resource sharing is generally useful and its general use is defined in this section.

2.3.1. Resource Sharing Association Type

The resource sharing association type was defined in [RFC4873] and was defined within the context of GMPLS and upstream initiated association. This section presents a definition of the resource sharing association that allows for its use with any RSVP session type and in both Path and Resv messages. This definition is consistent with the definition of the resource sharing association type in [RFC4873] and no changes are required by this section in order to support [RFC4873]. The Resource Sharing Association Type MUST be supported by any implementation compliant with this document.

The Resource Sharing Association Type is used to enable resource sharing across RSVP sessions. Per [RFC4873], Resource Sharing uses the Association Type field value of 2. ASSOCIATION objects with an Association Type with the value Resource Sharing MAY be carried in Path and Resv messages. Association for the Resource Sharing type MUST follow the procedures defined in Section 4.1.2 for upstream (Path message) initiated association and Section 4.2.1 for downstream (Resv message) initiated association. There are no type-specific association rules, processing rules, or ordering requirements. Note

that as is always the case with association as enabled by this document, no associations are made across Path and Resv state.

Once an association is identified, resources SHOULD be shared across the identified sessions. Resource sharing is discussed in general in [RFC2205] and within the context of LSPs in [RFC3209].

3. IPv4 and IPv6 Extended ASSOCIATION Objects

[RFC4872] defines the IPv4 ASSOCIATION object and the IPv6 ASSOCIATION object. As defined, these objects each contain an Association Source field and a 16-bit Association ID field. The combination of the Association Source and the Association ID uniquely identifies the association. Because the association-ID field is a 16-bit field, an association source can allocate up to 65536 different associations and no more. There are scenarios where this number is insufficient. (For example where the association identification is best known and identified by a fairly centralized entity, which therefore may be involved in a large number of associations.)

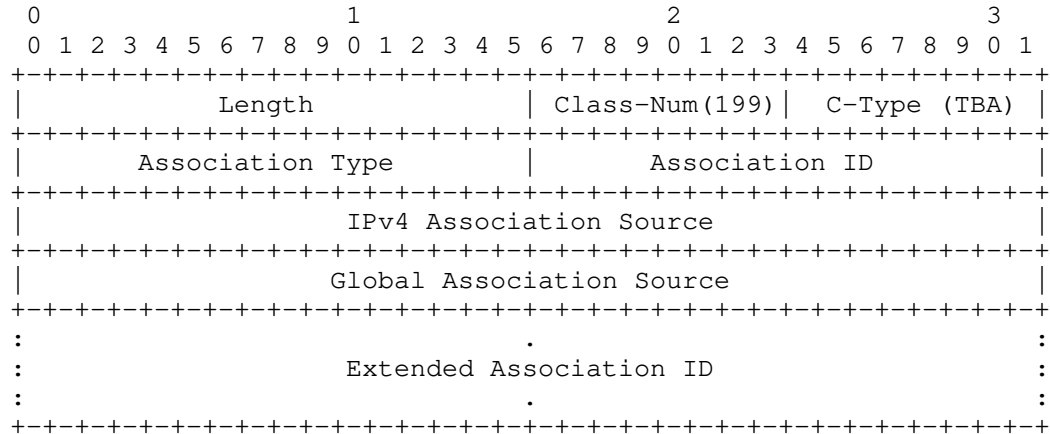
Furthermore, per [TP-IDENTIFIERS], MPLS-TP LSPs can be identified in two forms that cannot be supported using the existing ASSOCIATION objects. The first form is a global identifier and the second uses an ITU Carrier Code (ICC). The [TP-IDENTIFIERS] defined "global identifier", or Global_ID, is based on [RFC5003] and includes the operator's Autonomous System Number (ASN). [TP-IDENTIFIERS] identifies the ICC as "a string of one to six characters, each character being either alphabetic (i.e. A-Z) or numeric (i.e. 0-9) characters. Alphabetic characters in the ICC SHOULD be represented with upper case letters."

This sections defines new ASSOCIATION objects to support extended identification in order to address the limitations described above. Specifically, the IPv4 Extended ASSOCIATION object and IPv6 Extended ASSOCIATION object are defined below. Both new objects include the fields necessary to enable identification of a larger number of associations, as well as MPLS-TP required identification.

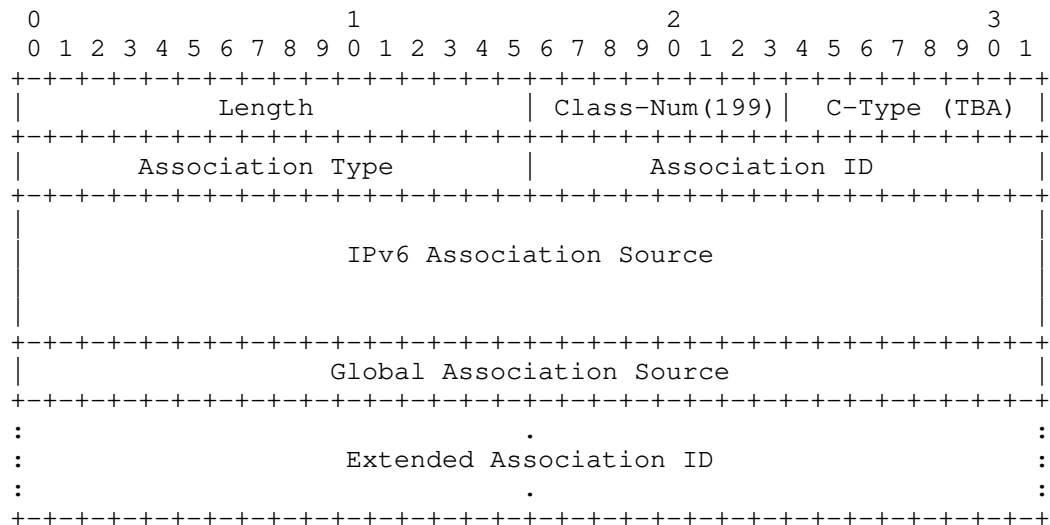
The IPv4 Extended ASSOCIATION object and IPv6 Extended ASSOCIATION object SHOULD be supported by an implementation compliant with this document. The processing rules for the IPv4 and IPv6 Extended ASSOCIATION object are described below, and are based on the rules for the IPv4 and IPv6 ASSOCIATION objects as described above.

3.1. IPv4 and IPv6 Extended ASSOCIATION Object Format

The IPv4 Extended ASSOCIATION object (Class-Num of the form 11bbbbbb with value = 199, C-Type = TBA) has the format:



The IPv6 Extended ASSOCIATION object (Class-Num of the form 11bbbbbb with value = 199, C-Type = TBA) has the format:



Association Type: 16 bits

Same as for IPv4 and IPv6 ASSOCIATION objects, see [RFC4872].

Association ID: 16 bits

Same as for IPv4 and IPv6 ASSOCIATION objects, see [RFC4872].

Association Source: 4 or 16 bytes

Same as for IPv4 and IPv6 ASSOCIATION objects, see [RFC4872].

Global Association Source: 4 bytes

This field contains a value that is a unique global identifier. This field MAY contain the 2-octet or 4-octet value of the provider's Autonomous System Number (ASN). It is expected that the global identifier will be derived from the globally unique ASN of the autonomous system hosting the Association Source. The special value of zero (0) indicates that no global identifier is present. Note that a Global Association Source of zero SHOULD be limited to entities contained within a single operator.

If the Global Association Source field value is derived from a 2-octet AS number, then the two high-order octets of this 4-octet field MUST be set to zero.

Please note that, as stated in [TP-IDENTIFIERS], the use of the provider's ASN as a global identifier DOES NOT have anything at all to do with the use of the ASN in protocols such as BGP.

This field is based on the definition of Global_ID defined in [RFC5003] and used by [TP-IDENTIFIERS].

Extended Association ID: variable, 4-byte aligned

This field contains data that is additional information to support unique identification. The length and contents of this field is determined by the Association Source. This field MAY be omitted, i.e., have a zero length. This field MUST be padded with zeros (0s) to ensure 32-bit alignment.

3.2. Processing

The processing of a IPv4 or IPv6 Extended ASSOCIATION object MUST identical to the processing of a IPv4 or IPv6 ASSOCIATION object as described above except as extended by this section. This section applies to both upstream-initiated (Path message) and downstream-initiated (Resv message) association.

The following are the modified procedures for Extended ASSOCIATION object processing:

- o When creating an Extended ASSOCIATION object, the originator MUST format the object as defined in this document.

- o The originator MUST set the Association Type, Association ID and Association Source fields as described in Section 4.
- o When ASN-based global identification of the Association Source is desired, the originator MUST set the Global Association Source field. When ASN-based global identification is not desired, the originator MUST set the Global Association Source field to zero (0).
- o The Extended ASSOCIATION object originator MAY include the Extended Association ID field. The field is included based on local policy. The field MUST be included when the Association ID field is insufficient to uniquely identify association within the scope of the source of the association. When included, this field MUST be set to a value that, when taken together with the other fields in the object, uniquely identifies the sessions to be associated.

When used in support of ICC identified (MPLS-TP) LSPs, this field MUST be at least eight (8) bytes long, and MAY be longer; the first six (6) bytes MUST be set to the ICC as defined in Section 3.2 of [TP-IDENTIFIERS] and the next two bytes MUST be set to zero (0). For non-ICC identified MPLS-TP LSPs, this field MUST either be omitted, or MUST have the first 6 bytes set to all zeros (0s).

- o The object Length field is set based on the length of the Extended Association ID field. When the Extended Association ID field is omitted, the object Length field MUST be set to 16 or 28 for the IPv4 and IPv6 ASSOCIATION objects, respectively. When the Extended Association ID field is present, the object Length field MUST be set to indicate the additional bytes carried in the Extended Association ID field, including pad bytes.

Note: per [RFC2205], the object Length field is set to the total object length in bytes, and is always a multiple of 4, and at least 4.

Identification of association is not modified by this section. It is important to note that Section 4 defines association identification based on ASSOCIATION object matching, and that such matching is based on the comparison of all fields in a ASSOCIATION object (unless type-specific comparison rules are defined). This applies equally to ASSOCIATION objects and Extended ASSOCIATION objects.

4. Security Considerations

A portion of this document reviews procedures defined in [RFC4872] and [RFC4873] and does not define any new procedures. As such, no new security considerations are introduced in this portion.

Section 4 defines broader usage of the ASSOCIATION object, but does not fundamentally expand on the association function that was previously defined in [RFC4872] and [RFC4873]. Section 5 increases the number of bits that are carried in an ASSOCIATION object (by 32), and similarly does not expand on the association function that was previously defined. This broader definition does allow for additional information to be conveyed, but this information is not fundamentally different from the information that is already carried in RSVP. Therefore there are no new risks or security considerations introduced by this document.

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

5. IANA Considerations

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

5.1. IPv4 and IPv6 Extended ASSOCIATION Objects

Upon approval of this document, IANA will make the assignment of two new C-Types (which are defined in section 5.1) for the existing ASSOCIATION object in the "Class Names, Class Numbers, and Class Types" section of the "Resource Reservation Protocol (RSVP) Parameters" registry located at <http://www.iana.org/assignments/rsvp-parameters>:

199 ASSOCIATION [RFC4872]

Class Types or C-Types

- | | | |
|---|----------------------------------|-----------------|
| 3 | Type 3 IPv4 Extended Association | [this document] |
| 4 | Type 4 IPv6 Extended Association | [this document] |

5.2. Resource Sharing Association Type

This document also broadens the potential usage of the Resource Sharing Association Type defined in [RFC4873]. As such, IANA is requested to change the Reference of the Resource Sharing Association Type included in the associate registry. This document also directs IANA to correct the duplicate usage of '(R)' in this Registry. In particular, the Association Type registry found at <http://www.iana.org/assignments/gmpls-sig-parameters/> should be updated as follows:

OLD:		
2	Resource Sharing (R)	[RFC4873]
NEW		
2	Resource Sharing (S)	[RFC4873][this-document]

There are no other IANA considerations introduced by this document.

6. Acknowledgments

Valuable comments and input was received from Dimitri Papadimitriou. We thank Subha Dhesikan for her contribution to the early work on sharing of resources across RSVP reservations.

7. References

7.1. Normative References

- [RFC2205] Braden, R., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification", RFC 2205, September 1997.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4872] Lang, J., Rekhter, Y., and Papadimitriou, D., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., Farrel, A., "GMPLS Segment Recovery", RFC 4873, May 2007.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax Used to Form Encoding Rules in Various Routing Protocol Specifications", RFC 5511, April 2009

7.2. Informative References

- [ASSOC-INFO] Berger, L., Faucheur, F., Narayanan, A., "Usage of The RSVP Association Object", work in progress, draft-ietf-ccamp-assoc-info.
- [RFC2207] Berger., L., O'Malley., T., "RSVP Extensions for IPSEC RSVP Extensions for IPSEC Data Flows", RFC 2207, September 1997.
- [RFC3175] Baker, F., Iturralde, C., Le, F., Davie, B., "Aggregation of RSVP for IPv4 and IPv6 Reservations", RFC 3175, September 2001.
- [RFC4860] Le, F., Davie, B., Bose, P., Christou, C., Davenport, M., "Generic Aggregate Resource ReSerVation Protocol (RSVP) Reservations", RFC 4860, May 2007.
- [RFC5003] Metz, C., Martini, L., Balus, F., Sugimoto, J., "Attachment Individual Identifier (AII) Types for Aggregation", RFC 5003, September 2007.
- [RFC5389] Rosenberg, J., Mahy, R., Matthews, P., Wing, D., "Session Traversal Utilities for NAT (STUN)", RFC 5389, October 2008.
- [RFC5920] Fang, L., et al, "Security Framework for MPLS and GMPLS Networks", work in progress, RFC 5920, July 2010.
- [TP-IDENTIFIERS] Bocci, M., Swallow, G., Gray, E., "MPLS-TP Identifiers", work in progress, draft-ietf-mpls-tp-identifiers.

8. Authors' Addresses

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

Francois Le Faucheur
Cisco Systems
Greenside, 400 Avenue de Roumanille
Sophia Antipolis 06410
France
Email: flefauch@cisco.com

Ashok Narayanan
Cisco Systems
300 Beaver Brook Road
Boxborough, MA 01719
United States
Email: ashokn@cisco.com

Generated on: Mon, May 02, 2011 10:17:58 AM

Internet Draft
Category: Informational
Expiration Date: November 1, 2011

Lou Berger (LabN)

May 1, 2011

Usage of The RSVP Association Object

draft-ietf-ccamp-assoc-info-02.txt

Abstract

The RSVP ASSOCIATION object was defined in the context of GMPLS (Generalized Multi-Protocol Label Switching) controlled label switched paths (LSPs). In this context, the object is used to associate recovery LSPs with the LSP they are protecting. This document reviews how association is to be provided in the context of GMPLS recovery. No new procedures or mechanisms are defined by this document and it is strictly informative in nature.

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on November 1, 2011

Copyright and License Notice

Copyright (c) 2011 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	Background	3
2.1	LSP Association	3
2.2	End-to-End Recovery LSP Association	5
2.3	Segment Recovery LSP Association	8
2.4	Resource Sharing LSP Association	8
3	Association of GMPLS Recovery LSPs	9
4	Security Considerations	10
5	IANA Considerations	10
6	Acknowledgments	10
7	References	10
7.1	Normative References	10
7.2	Informative References	11
8	Author's Addresses	11

1. Introduction

End-to-end and segment recovery are defined for GMPLS (Generalized Multi-Protocol Label Switching) controlled label switched paths (LSPs) in [RFC4872] and [RFC4873] respectively. Both definitions use the ASSOCIATION object to associate recovery LSPs with the LSP they are protecting. This document provides additional narrative on how such associations are to be identified. This document does not define any new procedures or mechanisms and is strictly informative in nature.

It may not be immediately obvious to the informed reader why this document is necessary, however questions were repeatedly raised in the Common Control and Measurement Plane (CCAMP) working group on the proper interpretation of the ASSOCIATION object in the context of end-to-end and segment recovery, and the working group agreed that this document should be produced in order to close the matter. This document formalizes the explanation provided in an e-mail to the working group authored by Adrian Farrel, see [AF-EMAIL]. This document in no way modifies the normative definitions of end-to-end and segment recovery, see [RFC4872] or [RFC4873].

2. Background

This section reviews the definition of LSP association in the contexts of end-to-end and segment recovery as defined in [RFC4872] and [RFC4873]. This section merely reiterates what has been defined, if differences exist between this text and [RFC4872] or [RFC4873], the earlier RFCs provide the authoritative text.

2.1. LSP Association

[RFC4872] introduces the concept and mechanisms to support the association of one LSP to another LSP across different RSVP-TE sessions. Such association is enabled via the introduction of the ASSOCIATION object. The ASSOCIATION object is defined in Section 16 of [RFC4872]. It is explicitly defined as having both general application and specific use within the context of recovery. End-to-end recovery usage is defined in [RFC4872] and is covered in Section 2.2. Segment recovery usage is defined in [RFC4873] and is covered in Section 2.3. Resource sharing LSP association is also defined in [RFC4873], while strictly speaking such association is beyond the scope of this document, for completeness it is covered in Section 2.4. The remainder of this section covers generic usage of the ASSOCIATION object.

In general, LSP association using the ASSOCIATION object can take place based on the values carried in the ASSOCIATION object. This means that association between LSPs can take place independent from

and across different sessions. This is a significant enhancement from the association of LSPs that is possible in base MPLS [RFC3209] and GMPLS [RFC3473].

When using ASSOCIATION object, LSP association is always initiated by an upstream node that inserts appropriate ASSOCIATION objects in the Path message of LSPs that are to be associated. Downstream nodes then correlate LSPs based on received ASSOCIATION objects. Multiple types of LSP association is supported by the ASSOCIATION object, and downstream correlation is made based on the type.

[RFC4872] defines C-Types 1 and 2 of the ASSOCIATION object. Both objects have essentially the same semantics, only differing in the type of address carried (IPv4 and IPv6). The defined objects carry multiple fields. The fields, taken together, enable the identification of which LSPs are association with one another. The [RFC4872] defined fields are:

- o Association Type:
This field identifies the usage, or application, of the association object. The currently defined values are Recovery [RFC4872] and Resource Sharing [RFC4873]. This field also scopes the interpretation of the object. In other words, the type field is included when matching LSPs (i.e., the type fields must match), and the way associations are identified may be type dependent.
- o Association Source:
This field is used to provide global scope (within the address space) to the identified association. There are no specific rules in the general case for which address should be used by a node creating an ASSOCIATION object beyond that the address is "associated to the node that originated the association", see [RFC4872].
- o Association ID:
This field provides an "identifier" that further scopes an association. Again, this field is combined with the other ASSOCIATION object fields to support identification of associated LSPs. The generic definition does not provide any specific rules on how matching is to be done, so such rules are governed by the Association Type. Note that the definition permits the association of an arbitrary number of LSPs.

As defined, the ASSOCIATION object may only be carried in a Path message, so LSP association takes place based on Path state. The definition permits one or more objects to be present. The support for multiple objects enables an LSP to be associated with other LSPs in more than one way at a time. For example, an LSP may carry one ASSOCIATION object to associate the LSP with another LSP for end-to-end recovery, and at the same time carry a second ASSOCIATION object

to associate the LSP with another LSP for segment recovery, and at the same time carry a third ASSOCIATION object to associate the LSP with yet another LSP for resource sharing.

2.2. End-to-End Recovery LSP Association

The association of LSPs in support of end-to-end LSP recovery is defined in Section 16.2 of [RFC4872]. There are also several additional related conformance statements (i.e., use of [RFC2119] defined key words) in Sections 7.3, 8.3, 9.3, 11.1. When analyzing the definition, as with any Standards Track RFC, it is critical to note and differentiate which statements are made using [RFC2119] defined key words, which relate to conformance, and which statements are made without such key words, which are only informative in nature.

As defined in Section 16.2, end-to-end recovery related LSP association may take place in two distinct forms:

- a. Between multiple (one or more) working LSPs and a single shared (associated) recovery LSP. This form essentially matches the shared 1:N ($N \geq 1$) recovery type described in the other sections of [RFC4872].
- b. Between a single working LSP and multiple (one or more) recovery LSPs. This form essentially matches all other recovery types described in [RFC4872].

Both forms share the same Association Type (Recovery) and the same Association Source (the working LSP's tunnel sender address). They also share the same definition of the Association ID, which is (quoting [RFC4872]):

"The Association ID MUST be set to the LSP ID of the LSP being protected by this LSP or the LSP protecting this LSP. If unknown, this value is set to its own signaled LSP ID value (default). Also, the value of the Association ID MAY change during the lifetime of the LSP."

The interpretation of the above is fairly straightforward. The Association ID carries one of 3 values:

- The LSP ID of the LSP being protected.
- The LSP ID of the LSP protecting an LSP.
- In the case where the matching LSP is not yet known (i.e., initiated), the LSP ID value of the LSP itself.

The text also explicitly allows for changing the Association ID during the lifetime of an LSP. But this is only an option, and is neither required (i.e., "MUST") nor recommended (i.e., "SHOULD"). It should be noted that the document does not describe when such a

change should be initiated, or the procedures for such a change. Clearly care needs to be taken when changing the Association ID to ensure that the old association is not lost during the transition to a new association.

The text does not preclude, and it is therefore assumed, that one or more ASSOCIATION objects may also be added to an LSP that was originated without any ASSOCIATION objects. Again this is a case that is not explicitly discussed in [RFC4872].

From the above, this means that the following combinations may occur:

- Case 1. When the ASSOCIATION object of the LSP being protected is initialized before the ASSOCIATION objects of any recovery LSPs are initialized, the Association ID in the LSP being protected and any recovery LSPs will carry the same value and this value will be the LSP ID value of the LSP being protected.
- Case 2. When the ASSOCIATION object of a recovery LSP is initialized before the ASSOCIATION object of any protected LSP is initialized, the Association ID in the recovery LSP and any LSPs being protected by that LSP will carry the same value and this value will be the LSP ID value of the recovery LSP.
- Case 3. When the ASSOCIATION objects of both the LSP being protected and the recovery LSP are concurrently initialized, the value of the Association ID carried in the LSP being protected is the LSP ID value of the recovery LSP, and the value of the Association ID carried in the recovery LSP is the LSP ID value of the LSP being protected. As this case can only be applied to LSPs with matching tunnel sender addresses, the scope of this case is limited to end-to-end recovery. Note that this is implicit in [RFC4872] as its scope is limited to end-to-end recovery.

In practical terms, case 2 will only occur when using the shared 1:N ($N \geq 1$) end-to-end recovery type and case 1 will occur with all other end-to-end recovery types. Case 3 is allowed, and it is subject to interpretation how often it will occur. Some believe that this case is the common case and, furthermore, that working and recovery LSPs will often first be initiated without any ASSOCIATION objects and then case 3 objects will be added once the LSPs are established. Others believe that case 3 will rarely if ever occur. Such perspectives have little impact on interoperability as a [RFC4872] compliant implementation needs to properly handle (identify associations for) all three cases.

It is important to note that Section 16.2 of [RFC4872] provides no

further requirements on how or when the Association ID value is to be selected. The other sections of the document do provide further narrative and 3 additional requirements. In general, the narrative highlights case 3 identified above but does not preclude the other cases. The 3 additional requirements are, by [RFC4872] Section number:

- o Section 7.3 -- "The Association ID MUST be set by default to the LSP ID of the protected LSP corresponding to N = 1."

When considering this statement together with the 3 cases enumerated above, it can be seen that this statement clarifies which LSP ID value should be used when a single shared protection LSP is established simultaneously with (case 3), or after (case 2), more than one LSP to be protected.

- o Section 8.3 -- "Secondary protecting LSPs are signaled by setting in the new PROTECTION object the S bit and the P bit to 1, and in the ASSOCIATION object, the Association ID to the associated primary working LSP ID, which MUST be known before signaling of the secondary LSP."

This requirement clarifies that the Rerouting without Extra-Traffic type of recovery is required to follow either case 1 or 3, but not 2, as enumerated above.

- o Section 9.3 -- "Secondary protecting LSPs are signaled by setting in the new PROTECTION object the S bit and the P bit to 1, and in the ASSOCIATION object, the Association ID to the associated primary working LSP ID, which MUST be known before signaling of the secondary LSP."

This requirement clarifies that the Shared-Mesh Restoration type of recovery is required to follow either case 1 or 3, but not 2, as enumerated above.

- o Section 11.1 -- "In both cases, the Association ID of the ASSOCIATION object MUST be set to the LSP ID value of the signaled LSP."

This requirement clarifies that when using the LSP Rerouting type of recovery is required to follow either case 1 or 3, but not 2, as enumerated above.

2.3. Segment Recovery LSP Association

GMPLS segment recovery is defined in [RFC4873]. Segment recovery reuses the LSP association mechanisms, including the Association Type field value, defined in [RFC4872]. The primary text to this effect in [RFC4873] is:

3.2.1. Recovery Type Processing

Recovery type processing procedures are the same as those defined in [RFC4872], but processing and identification occur with respect to segment recovery LSPs. Note that this means that multiple ASSOCIATION objects of type recovery may be present on an LSP.

This statement means that case 2 as enumerated above is to be followed and furthermore that Association Source is set to the tunnel sender address of the segment recovery LSPs. The explicit exclusion of case 3 is not listed as its non-applicability was considered obvious to the informed reader. (Perhaps having this exclusion explicitly identified would have obviated the need for this document.)

2.4. Resource Sharing LSP Association

Section 3.2.2 of [RFC4873] defines an additional type of LSP association which is used for "Resource Sharing". Resource sharing enables the sharing of resources across LSPs with different SESSION objects. Without this object only sharing across LSPs with a shared SESSION object was possible, see [RFC3209].

Resource sharing is indicated using a new Association Type value. As the Association Type field value is not the same as is used in Recovery LSP association, the semantics used for the association of LSPs using an ASSOCIATION object containing the new type differs from Recovery LSP association.

Section 3.2.2 of [RFC4873] states the following rules for the construction of an ASSOCIATION object in support of resource sharing LSP association:

- o The Association Type value is set to "Resource Sharing".
- o Association Source is set to the originating node's router address.
- o The Association ID is set to a value that uniquely identifies the set of LSPs to be associated.

The setting of the Association ID value to the working LSP's LSP

ID value is mentioned, but using the "MAY" key word. Per [RFC2119], this translates to the use of LSP ID value as being completely optional and that the choice of Association ID is truly up to the originating node.

Additionally, the identical ASSOCIATION object is used for all LSPs that should be associated using Resource Sharing. This differs from recovery LSP association where it is possible for the LSPs to carry different Association ID fields and still be associated (see case 3 in Section 2.2).

3. Association of GMPLS Recovery LSPs

The previous section reviews the construction of an ASSOCIATION object, including the selection of the value used in the Association ID field, as defined in [RFC4872] and [RFC4873]. This section reviews how a downstream receiver identifies that one LSP is associated within another LSP based on ASSOCIATION objects. Note that this section in no way modifies the normative definitions of end-to-end and segment recovery, see [RFC4872] or [RFC4873].

As the ASSOCIATION object is only carried in Path messages, such identification only takes place based on Path state. In order to support the identification of the recovery type association between LSPs, a downstream receiver needs to be able to handle all three cases identified in Section 2.2. Cases 1 and 2 are simple as the associated LSPs will carry the identical ASSOCIATION object. This is also always true for resource sharing type LSP association, see Section 2.4. Case 3 is more complicated as it is possible for the LSPs to carry different Association ID fields and still be associated. The receiver also needs to allow for changes in the set of ASSOCIATION objects included in an LSP.

Based on the [RFC4872] and [RFC4873] definitions related to the ASSOCIATION object, the following behavior can be followed to ensure that a receiver always properly identifies the association between LSPs:

- o Covering cases 1 and 2 and resource sharing type LSP association:

For ASSOCIATION objects with the Association Type field values of "Recovery" (1) and "Resource Sharing" (2), the association between LSPs is identified by comparing all fields of each of the ASSOCIATION objects carried in the Path messages associated with each LSP. An association is deemed to exist when the same values are carried in all fields of an ASSOCIATION object carried in each LSP's Path message. As more than one association may exist (e.g., in support of different association types or end-to-end and segment recovery), all carried ASSOCIATION objects need to be examined.

- o Covering case 3:

Any ASSOCIATION object with the Association Type field value of "Recovery" (1) that does not yield an association in the prior comparison needs to be checked to see if a case 3 association is indicated. As this case only applies to end-to-end recovery, the first step is to locate any other LSPs with the identical SESSION object fields and the identical tunnel sender address fields as the LSP carrying the ASSOCIATION object. If such LSPs exist, a case 3 association is identified by comparing the value of the Association ID field with the LSP ID field of the other LSP. If the values are identical, then an end-to-end recovery association exists. As this behavior only applies to end-to-end recovery, this check need only be performed at the egress.

No additional behavior is needed in order to support changes in the set of ASSOCIATION objects included in an LSP, as long as the change represents either a new association or a change in identifiers made as described in Section 2.2.

4. Security Considerations

This document reviews procedures defined in [RFC4872] and [RFC4873] and does not define any new procedures. As such, no new security considerations are introduced in this document..

5. IANA Considerations

There are no new IANA considerations introduced by this document.

6. Acknowledgments

This document formalizes the explanation provided in an e-mail to the working group authored by Adrian Farrel, see [AF-EMAIL]. This document was written in response to questions raised in the CCAMP working group by Nic Neate <nhn@dataconnection.com>. Valuable comments and input was also received from Dimitri Papadimitriou, Francois Le Faucheur and Ashok Narayanan.

7. References

7.1. Normative References

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

- [RFC4872] Lang, J., Rekhter, Y., and Papadimitriou, D., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., Farrel, A., "GMPLS Segment Recovery", RFC 4873, May 2007.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

7.2. Informative References

- [AF-EMAIL] Farrel, A. "Re: Clearing up your misunderstanding of the Association ID", CCAMP working group mailing list, <http://www.ietf.org/mail-archive/web/ccamp/current/msg00644.html>, November 18, 2008.

8. Author's Addresses

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

Generated on: Mon, May 02, 2011 10:15:27 AM

Network Working Group
Internet Draft
Category: Informational

Fatai Zhang
Dan Li
Huawei
Han Li
CMCC
S. Belotti
Alcatel-Lucent
D. Ceccarelli
Ericsson
March 11, 2011

Expires: September 11, 2011

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

draft-ietf-ccamp-gmpls-g709-framework-04.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September 11, 2011.

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

1. Introduction

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

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

[RFC3471] and [RFC3473], routing and OSPF extensions are described in [RFC4202] and [RFC4203], and the Link Management Protocol (LMP) is described in [RFC4204].

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

This document reviews relevant aspects of OTN technology evolution that affect the GMPLS control plane protocols and examines why and how to update the mechanisms described in [RFC4328]. This document additionally provides a framework for the GMPLS control of OTN networks and includes a discussion of the implication for the use of the Path Computation Element (PCE) [RFC4655]. No additional Switching Type and LSP Encoding Type are required to support the control of the evolved OTN, because the Switching Type and LSP Encoding Type defined in [RFC4328] are still applicable.

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 [WSO-Frame] 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 ODUn tributary signals in its OPUn area.

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

3. G.709 Optical Transport Network (OTN)

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

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

3.1. OTN Layer Network

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

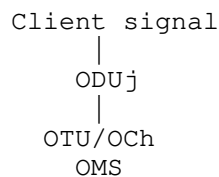


Figure 1 - Basic OTN signal hierarchy

Client signals are mapped into ODU_j containers. These ODU_j containers are multiplexed onto the OTU/OCh. The individual OTU/OCh signals are combined in the Optical Multiplex Section (OMS) using WDM multiplexing, and this aggregated signal provides the link between the nodes.

3.1.1. Client signal mapping

The client signals are mapped into a Low Order (LO) ODU_j. Appendix A gives more information about LO ODU.

The current values of *j* defined in [G709-V3] are: 0, 1, 2, 2e, 3, 4, Flex. The approximate bit rates of these signals are defined in [G709-V3] and are reproduced in Tables 1 and 2.

ODU Type	ODU nominal bit rate
ODU0	1 244 160 kbits/s
ODU1	239/238 x 2 488 320 kbit/s
ODU2	239/237 x 9 953 280 kbit/s
ODU3	239/236 x 39 813 120 kbit/s
ODU4	239/227 x 99 532 800 kbit/s
ODU2e	239/237 x 10 312 500 kbit/s
ODUflex for CBR Client signals	239/238 x client signal bit rate
ODUflex for GFP-F Mapped client signal	Configured bit rate

Table 1 - ODU types and bit rates

NOTE - The nominal ODU_k rates are approximately: 2 498 775.126 kbit/s (ODU1), 10 037 273.924 kbit/s (ODU2), 40 319 218.983 kbit/s (ODU3), 104 794 445.815 kbit/s (ODU4) and 10 399 525.316 kbit/s (ODU2e).

ODU Type	ODU bit-rate tolerance
ODU0	+/- 20 ppm
ODU1	+/- 20 ppm
ODU2	+/- 20 ppm
ODU3	+/- 20 ppm
ODU4	+/- 20 ppm
ODU2e	+/- 100 ppm
ODUflex for CBR Client signals	+/- 100 ppm
ODUflex for GFP-F Mapped client signal	+/- 100 ppm

Table 2 - ODU types and tolerance

One of two options is for mapping client signals into ODUflex depending on the client signal type:

- Circuit clients are proportionally wrapped. Thus the bit rate and tolerance are defined by the client signal.
- Packet clients are mapped using the Generic Framing Procedure (GFP). [G709-V3] recommends that the 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 ODUs are mapped into the TS of the OPUK. Note that in the case where $j=k$ the ODUj is mapped into the OTU/OCh without multiplexing.

The initial versions of G.709 [G709-V1] only provided a single TS granularity, nominally 2.5Gb/s. [G709-V3], approved in 2009, added an additional TS granularity, nominally 1.25Gb/s. The number and type of TSs provided by each of the currently identified OTUk is provided below:

	2.5Gb/s	1.25Gb/s	Nominal Bit rate
OTU1	1	2	2.5Gb/s
OTU2	4	8	10Gb/s
OTU3	16	32	40Gb/s
OTU4	--	80	100Gb/s

To maintain backwards compatibility while providing the ability to interconnect nodes that support 1.25Gb/s TS at one end of a link and 2.5Gb/s TS at the other, the 'new' equipment will fall back to the use of a 2.5Gb/s TS if connected to legacy equipment. This information is carried in band by the payload type.

The actual bit rate of the TS in an OTUk depends on the value of k. Thus the number of TS occupied by an ODUj may vary depending on the values of j and k. For example an ODU2e uses 9 TS in an OTU3 but only 8 in an OTU4. Examples of the number of TS used for various cases are provided below:

- ODU0 into ODU1, ODU2, ODU3 or ODU4 multiplexing with 1,25Gbps TS granularity
 - o ODU0 occupies 1 of the 2, 8, 32 or 80 TS for ODU1, ODU2, ODU3 or ODU4
- ODU1 into ODU2, ODU3 or ODU4 multiplexing with 1,25Gbps TS granularity
 - o ODU1 occupies 2 of the 8, 32 or 80 TS for ODU2, ODU3 or ODU4
- ODU1 into ODU2, ODU3 multiplexing with 2.5Gbps TS granularity
 - o ODU1 occupies 1 of the 4 or 16 TS for ODU2 or ODU3

- ODU2 into ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU2 occupies 8 of the 32 or 80 TS for ODU3 or ODU4
- ODU2 into ODU3 multiplexing with 2.5Gbps TS granularity
 - o ODU2 occupies 4 of the 16 TS for ODU3
- ODU3 into ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU3 occupies 31 of the 80 TS for ODU4
- ODUFlex into ODU2, ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODUFlex occupies n of the 8, 32 or 80 TS for ODU2, ODU3 or ODU4 (n <= Total TS numbers of ODUk)
- ODU2e into ODU3 or ODU4 multiplexing with 1.25Gbps TS granularity
 - o ODU2e occupies 9 of the 32 TS for ODU3 or 8 of the 80 TS for ODU4

In general the mapping of an ODU_j (including ODUFlex) into the OTU_k TSs is determined locally, and it can also be explicitly controlled by a specific entity (e.g., head end, NMS) through Explicit Label Control [RFC3473].

3.1.2.1. Structure of MSI information

When multiplexing an ODU_j into a HO ODU_k (k>j), G.709 specifies the information that has to be transported in-band in order to allow for correct demultiplexing. This information, known as Multiplex Structure Information (MSI), is transported in the OPU_k overhead and is local to each link. In case of bidirectional paths the association between TPN and TS MUST be the same in both directions.

The MSI information is organized as a set of entries, with one entry for each HO ODU_j TS. The information carried by each entry is:

- Payload Type: the type of the transported payload.
- Tributary Port Number (TPN): the port number of the ODU_j transported by the HO ODU_k. The TPN is the same for all the TSs assigned to the transport of the same ODU_j instance.

For example, an ODU2 carried by a HO ODU3 is described by 4 entries in the OPU3 overhead when the TS size is 2.5 Gbit/s, and by 8 entries when the TS size is 1.25 Gbit/s.

On each node and on every link, two MSI values have to be provisioned:

- The TxMSI information inserted in OPU (e.g., OPU3) overhead by the source of the HO 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 [WSON-FRAME].

While [G872-2001] considered the ODU as a set of layers in the same way as SDH has been modeled, recent ITU-T OTN architecture progress [G872-Am2] includes an agreement to model the ODU as a single layer network with the bit rate as a parameter of links and connections. This allows the links and nodes to be viewed in a single topology as a common set of resources that are available to provide ODU_j connections independent of the value of j. Note that when the bit rate of ODU_j is less than the server bit rate, ODU_j connections are supported by HO-ODU (which has a one-to-one relationship with the OTU).

From an ITU-T perspective, the ODU connection topology is represented by that of the OTU link layer, which has the same topology as that of the OCh layer (independent of whether the OTU supports HO-ODU, where multiplexing is utilized, or LO-ODU in the case of direct mapping).

Thus, the OTU and OCh layers should be visible in a single topological representation of the network, and from a logical perspective, the OTU and OCh may be considered as the same logical, switchable entity.

Note that the OTU link layer topology may be provided via various infrastructure alternatives, including point-to-point optical connections, flexible optical connections fully in the optical domain, flexible optical connections involving hybrid sub-lambda/lambda nodes involving 3R, etc.

The document will be updated to maintain consistency with G.872 progress when it is consented for publication.

4.1. Connection management of the ODU

LO ODU_j can be either mapped into the OTU_k signal ($j = k$), or multiplexed with other LO ODU_js into an OTU_k ($j < k$), and the OTU_k is mapped into an OCh. See Appendix A for more information.

From the perspective of control plane, there are two kinds of network topology to be considered.

(1) ODU layer

In this case, the ODU links are presented between adjacent OTN nodes, which is illustrated in Figure 2. In this layer there are ODU links with a variety of TSs available, and nodes that are ODXCs. Lo ODU connections can be setup based on the network topology.

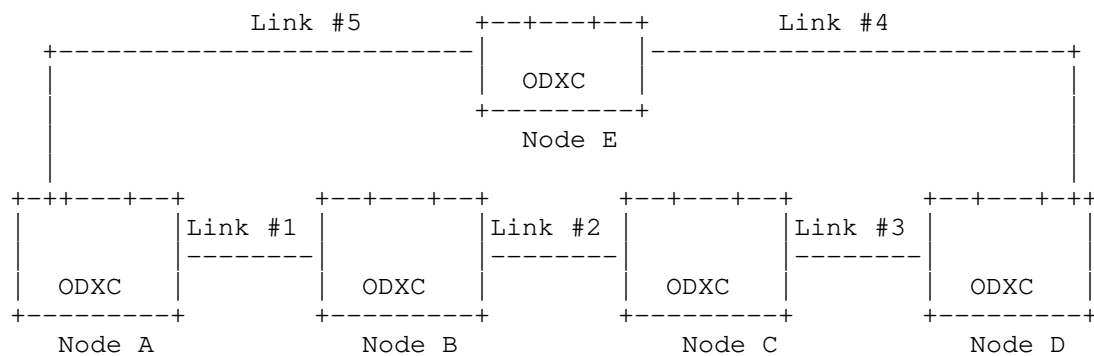


Figure 2 - Example Topology for LO ODU connection management

If an ODU_j connection is requested between Node C and Node E routing/path computation must select a path that has the required number of TS available and that offers the lowest cost. Signaling is then invoked to set up the path and to provide the information (e.g., selected TS) required by each transit node to allow the configuration of the ODU_j to OTU_k mapping ($j = k$) or multiplexing ($j < k$), and demapping ($j = k$) or demultiplexing ($j < k$).

(2) ODU layer with OCh switching capability

In this case, the OTN nodes interconnect with wavelength switched node (e.g., ROADM, OXC) that are capable of OCh switching, which is illustrated in Figure 3 and Figure 4. There are ODU layer and OCh layer, so it is simply a MLN. OCh connections may be created on demand, which is described in section 5.1.

In this case, an operator may choose to allow the underlined OCh layer to be visible to the ODU routing/path computation process in which case the topology would be as shown in Figure 4. In Figure 3 below, instead, a cloud representing OCH capable switching nodes is represented. In Figure 3, the operator choice is to hide the real RWA network topology.

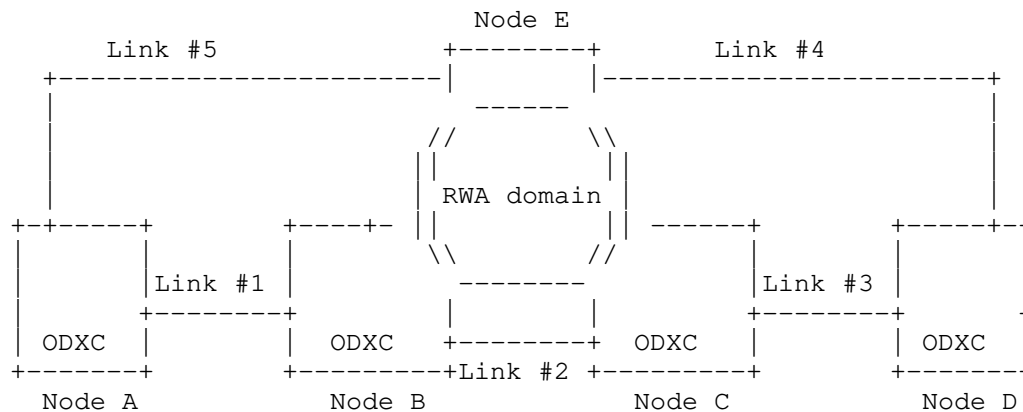


Figure 3 - RWA Hidden Topology for LO ODU connection management

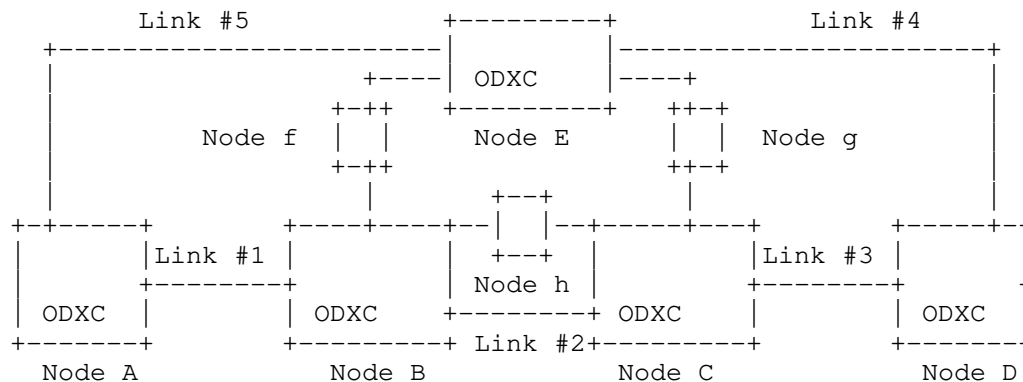


Figure 4 - RWA Visible Topology for LO ODUj connection management

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

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

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

5. GMPLS/PCE Implications

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

5.1. Implications for LSP Hierarchy with GMPLS TE

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

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

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

LSP hierarchy could 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

New label should be defined to identify the type of TS (i.e., the 2.5 Gbps TS granularity and the new 1.25 Gbps TS granularity).

- 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]). [VCAT] has a clear description of VCAT and LCAS control in SONET/SDH and OTN networks.

- Support for constraint signaling

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. In this case, an entire lambda of capacity is consumed in transporting the ODUk connection service. On the other hand, the operator might leverage sub-lambda 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 over associated OTU sections.

The identification of constraints and associated encoding in the signaling for differentiating full lambda LSP or sub lambda LSP is for further study.

- Support for Control of Hitless Adjustment of ODUFlex (GFP)

[G.HAO] 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.

Since the [G.HAO] is being developed currently, the control of HAO is for further study.

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 a ODU_j connection request. In order to compute the lowest cost path 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 other issues should also be considered which are discussed below.

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. So routing protocol should be extended when TDM is ODU type to support representation of ODU switching information, especially the following requirements should be considered:

- Support for carrying the link multiplexing capability

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

- Support any ODU and ODUflex

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

One LO ODU may need different number of TSs when multiplexed into different HO ODUs. For example, for ODU_{2e}, 9 TSs are needed when multiplexed into an ODU₃, while only 8 TSs are needed when

multiplexed into an ODU4. For ODUFlex, the total number of TSs to be reserved in a HO ODU equals the maximum of [bandwidth of ODUFlex / bandwidth of TS of the HO ODU].

Therefore, the routing protocol must be capable of carrying the necessary and sufficient link bandwidth information for performing accurate route computation for any of the fixed rate ODUs as well as ODUFlex.

- Support for differentiating between terminating and switching capability

Due to internal constraints and/or limitations, the type of signal being advertised by an interface could be just switched (i.e. forwarded to switching matrix without multiplexing/demultiplexing actions), just terminated (demuxed) or both of them. The capability advertised by an interface needs further distinction in order to separate termination and switching capabilities.

Therefore, to allow the required flexibility, the routing protocol should clearly distinguish the terminating and switching capability.

- Support 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 or different line rates, 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)

As described in Section 5.2, the routing requirements for supporting hitless adjustment of ODUFlex (GFP) (HAO) are for further study.

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 to allocate the TS over the TE link correctly.

- 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 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

The 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 $i \leq 4$) of node B are combined together. When creating an ODU1 service in this ODU2 link, node B reserves the TS#i and TS#i+4 with the granularity of 1.25Gbps. But in the label sent from B to A, it is indicated that the TS#i with the granularity of 2.5Gbps is reserved.

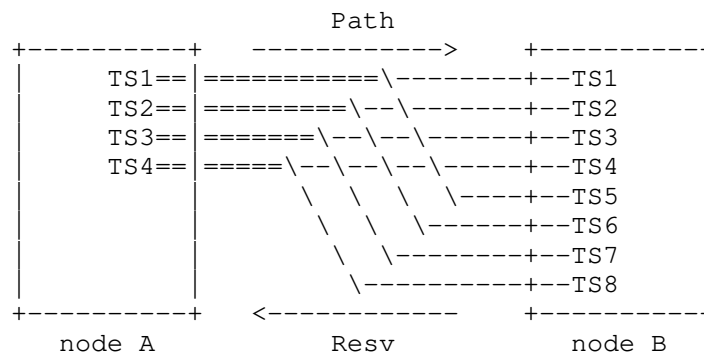


Figure 5 - Interworking between 1.25Gbps TS and 2.5Gbps TS

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 control 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

- [RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol LabelSwitching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, Jan 2006.
- [RFC3471] Berger, L., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] L. Berger, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4201] K. Kompella, Y. Rekhter, Ed., "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.
- [RFC4202] K. Kompella, Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] K. Kompella, Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4205] K. Kompella, Y. Rekhter, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4205, October 2005.
- [RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [RFC4206] K. Kompella, Y. Rekhter, Ed., " Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC6107] K. Shiimoto, A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC6107, February 2011.

- [RFC6001] Dimitri Papadimitriou et al, "Generalized Multi-Protocol Label Switching (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC6001, February 21, 2010.
- [RFC5440] JP. Vasseur, JL. Le Roux, Ed., " Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [VCAT] G. Bernstein et al, "Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS)", draft-ietf-ccamp-gmpls-vcat-lcas-11.txt, March 9, 2011.
- [G709-V3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709 Recommendation, December 2009.

10.2. Informative References

- [G709-V1] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation and Amendment1, November 2001.
- [G709-V2] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation, March 2003.
- [G7042] ITU-T G.7042/Y.1305, "Link capacity adjustment scheme (LCAS) for virtual concatenated signals", March 2006.
- [G872-2001] ITU-T, "Architecture of optical transport networks", November 2001 (11 2001).
- [G872-Am2] Draft Amendment 2, ITU-T, "Architecture of optical transport networks".
- [G.HAO] TD 382 (WP3/15), 31 May - 11 June 2010, Q15 Plenary Meeting in Geneva, Initial draft G.hao "Hitless Adjustment of ODUflex (HAO)".
- [HZang00] H. Zang, J. Jue and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks", Optical Networks Magazine, January 2000.

- [WSON-FRAME] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks (WSON)", draft-ietf-ccamp-rwa-wson-framework, work in progress.
- [PCE-APS] Tomohiro Otani, Kenichi Ogaki, Diego Caviglia, and Fatai Zhang, "Requirements for GMPLS applications of PCE", draft-ietf-pce-gmpls-aps-req-01.txt, July 2009.
- [GMPLS-SEC] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", Work in Progress, October 2009.

11. 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

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

Phone: +86-755-28973237
Email: huawei.danli@huawei.com

Han Li
China Mobile Communications Corporation
53 A Xibianmennei Ave. Xuanwu District
Beijing 100053 P.R. China

Phone: +86-10-66006688
Email: lihan@chinamobile.com

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

Email: sergio.belotti@alcatel-lucent.it

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

12. Contributors

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

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

Malcolm Betts
Huawei Technologies Co., Ltd.

Email: malcolm.betts@huawei.com

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

Email: pietro_vittorio.grandi@alcatel-lucent.it

Eve Varma
Alcatel-Lucent
1A-261, 600-700 Mountain Av
PO Box 636
Murray Hill, NJ 07974-0636
USA
Email: eve.varma@alcatel-lucent.com

APPENDIX A: ODU connection examples

This appendix provides a description of ODU terminology and connection examples. This section is not normative, and is just intended to facilitate understanding.

In order to transmit a client signal, an ODU connection 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_j client that is mapped into an OTU_k server. For example, if a STM-16 signal is encapsulated into ODU1, and then the ODU1 is mapped into OTU1, the ODU1 is a LO ODU (from a multiplexing perspective).

(2) An ODU_j client that is mapped into an ODU_k ($j < k$) server occupying several TSs. For example, if ODU1 is multiplexed into ODU2, and ODU2 is mapped into OTU2, the ODU1 is a LO ODU and the ODU2 is a HO ODU (from a multiplexing perspective).

Thus, a LO ODU_j represents the container transporting a client of the OTN that is either directly mapped into an OTU_k ($k = j$) or multiplexed into a server HO ODU_k ($k > j$) container. Consequently, the HO ODU_k represents the entity transporting a multiplex of LO ODU_j tributary signals in its OPU_k area.

In the case of LO ODU_j mapped into an OTU_k ($k = j$) directly, Figure 6 give an example of this kind of LO ODU connection.

In Figure 6, The LO ODU_j is switched at the intermediate ODXC node. OCh and OTU_k are associated with each other. From the viewpoint of connection management, the management of OTU_k is similar with OCh. LO ODU_j and OCh/OTU_k have client/server relationships.

For example, one LO ODU1 connection can be setup between Node A and Node C. This LO ODU1 connection is to be supported by OCh/OTU1

connections, which are to be set up between Node A and Node B and between Node B and Node C. LO ODU1 can be mapped into OTU1 at Node A, demapped from it in Node B, switched at Node B, and then mapped into the next OTU1 and demapped from this OTU1 at Node C.

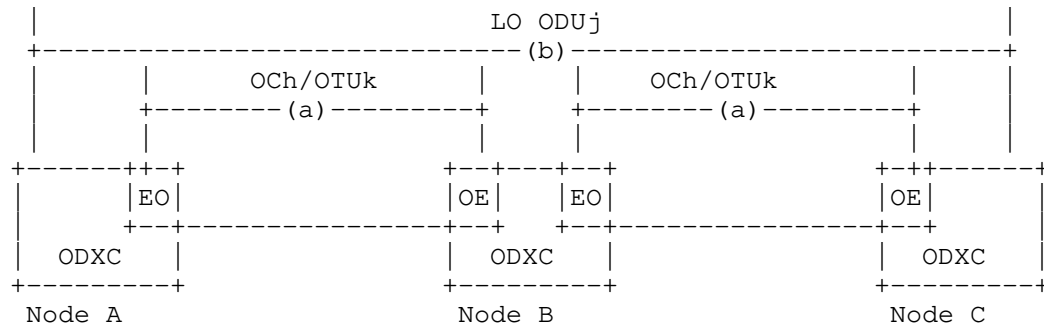


Figure 6 - Connection of LO ODUj (1)

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

In Figure 7, OCh, OTUk, HO ODUk are associated with each other. The LO ODUj is multiplexed/de-multiplexed into/from the HO ODU at each ODXC node and switched at each ODXC node (i.e. trib port to line port, line card to line port, line port to trib port). From the viewpoint of connection management, the management of these HO ODUk and OTUk are similar to OCh. LO ODUj and OCh/OTUk/HO ODUk have client/server relationships. When a LO ODU connection is setup, it will be using the existing HO ODUk (/OTUk/OCh) connections which have been set up. Those HO ODUk connections provide LO ODU links, of which the LO ODU connection manager requests a link connection to support the LO ODU connection.

For example, one HO ODU2 (/OTU2/OCh) connection can be setup between Node A and Node B, another HO ODU3 (/OTU3/OCh) connection can be setup between Node B and Node C. LO ODU1 can be generated at Node A, switched to one of the 10G line ports and multiplexed into a HO ODU2 at Node A, demultiplexed from the HO ODU2 at Node B, switched at Node B to one of the 40G line ports and multiplexed into HO ODU3 at Node B, demultiplexed from HO ODU3 at Node C and switched to its LO ODU1 terminating port at Node C.

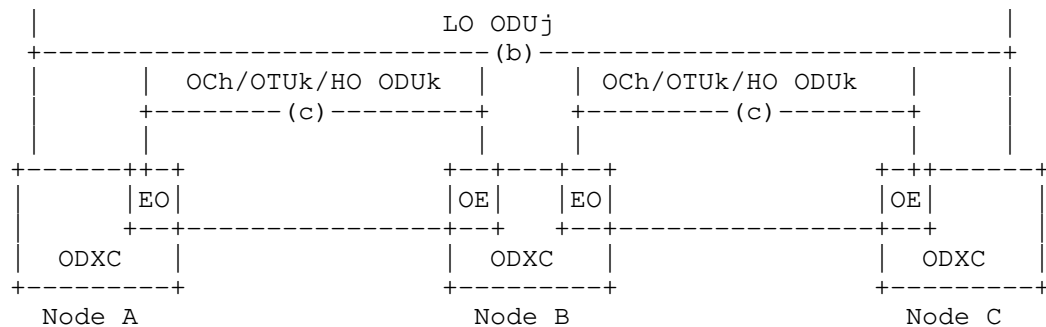


Figure 7 - Connection of LO ODUj (2)

Intellectual Property

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

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

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

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

those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

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

Disclaimer of Validity

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

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Network Working Group
Internet Draft
Updates: 4204, 4207, 4209, 5818
Category: Standards Track

Dan Li
Huawei
D. Ceccarelli
Ericsson
Lou Berger
LabN

Expires: December 2011

June 7, 2011

Link Management Protocol Behavior Negotiation and Configuration Modifications

draft-ietf-ccamp-lmp-behavior-negotiation-04.txt

Abstract

The Link Management Protocol (LMP) is used to coordinate the properties, use, and faults of data links in Generalized Multiprotocol Label Switching (GMPLS) networks. This document defines an extension to LMP to negotiate capabilities and indicate support for LMP extensions. The defined extension is compatible with non-supporting implementations.

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on December 7, 2011.

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.

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. LMP Message Modifications.....	4
2.1. Modified Message Formats.....	4
2.2. Processing	5
3. LMP Behavior Negotiation.....	6
3.1. BehaviorConfig C-Type Format.....	6
3.2. Processing	7
4. Backward Compatibility.....	7
5. Security Considerations.....	8
6. IANA Considerations	9
6.1. New LMP Class Type.....	9
6.2. New Capabilities Registry.....	9
7. Contributors	10
8. Acknowledgments	10
9. References	10
9.1. Normative References.....	10
9.2. Informative References.....	11
10. Authors' Addresses	11

1. Introduction

The Link Management Protocol (LMP) [RFC4204] has been successfully deployed in Generalized Multiprotocol Label Switching (GMPLS)-controlled networks.

New LMP behaviors and protocol extensions have been introduced in a number of IETF documents as set out later in this section. It is likely that future extensions will be made to support additional functions.

In a network, if one LMP node supports a new behavior or protocol extension but its adjacent node does not, it is beneficial to have a protocol mechanism to discover the capabilities of peer nodes so that the right protocol extensions can be selected and the correct features can be enabled. There are no such procedures defined in the base LMP specification [RFC4204]. [RFC4209] defined a specific mechanism to identify support for the functions defined in that document. This document defines an LMP extension to support the identification of supported LMP functions in a generic fashion, as well as how a node supporting these extensions would communicate with legacy nodes.

In [RFC4204], the basic behaviors have been defined around the use of the standard LMP messages, which include Config, Hello, Verify, Test, LinkSummary, and ChannelStatus. Per [RFC4204], these behaviors MUST be supported when LMP is implemented, and the message types from 1 to 20 have been assigned by IANA for these messages. Support for all functions required by [RFC4204] is assumed by this document.

In [RFC4207], the SONET/SDH technology-specific behavior and information for LMP is defined. The Trace behavior is added to LMP, and the message types from 21 to 31 were assigned by IANA for the messages that provide the TRACE function. The Trace function has been extended for the support of OTNs (Optical Transport Networks) in [LMP-TEST].

In [RFC4209], extensions to LMP are defined to allow it to be used between a peer node and an adjacent Optical Line System (OLS). The LMP object class type and sub-object class name have been extended to support DWDM behavior.

In [RFC5818], the data channel consistency check behavior is defined, and the message types from 32 to 34 have been assigned by IANA for messages that provide this behavior.

It is likely that future extensions to LMP for other functions or technologies will require the definition of further LMP messages.

This document describes an LMP extension, which is referred to as behavior negotiation, which enables nodes at the ends of a link to identify the LMP messages and functions supported by the adjacent node. The extension makes use of a new CONFIG object. The use of this new object does not preclude the use of existing or yet to be defined CONFIG object.

This document also modifies the format of messages that carry CONFIG object to allow for multiple objects. Multiple CONFIG objects allow behavior negotiation concurrent with existing usage of the CONFIG object, i.e., HelloConfig C-Type defined in [RFC4204] and LMP_WDM_CONFIG C-Type defined in [RFC4209]. This document modifies the ConfigAck message to include CONFIG objects so that acceptable parameters are explicitly identified. It also describes how a node which supports the extensions defined in this document interacts with a legacy LMP node.

2. LMP Message Modifications

LMP Config, ConfigNack and ConfigAck messages are modified by this document to allow for the inclusion of multiple CONFIG objects. The Config and ConfigNack messages were only defined to carry on CONFIG object in [RFC4204]. The ConfigAck message, which was defined without carrying any CONFIG objects in [RFC4204], is modified to enable explicit identification of negotiated configuration parameters. The inclusion of CONFIG objects in ConfigAck messages is triggered by the use of the BehaviorConfig object (defined below) in a received Config message.

2.1. Modified Message Formats

The format of the Config message as updated by this document is as follows:

```
<Config Message> ::= <Common Header> <LOCAL_CCID> <MESSAGE_ID>  
                    <LOCAL_NODE_ID> <CONFIG> [ <CONFIG> ... ]
```

The format of the ConfigAck message as updated by this document is as follows:

```
<ConfigAck Message> ::= <Common Header> <LOCAL_CCID> <LOCAL_NODE_ID>  
                        <REMOTE_CCID> <MESSAGE_ID_ACK>  
                        <REMOTE_NODE_ID>[ <CONFIG> ... ]
```

The format of the ConfigNack message as updated by this document is as follows:

```
<ConfigNack Message> ::= <Common Header> <LOCAL_CCID>
                           <LOCAL_NODE_ID>  <REMOTE_CCID>
                           <MESSAGE_ID_ACK> <REMOTE_NODE_ID>
                           <CONFIG> [ <CONFIG> ... ]
```

2.2. Processing

Nodes which support the extensions defined in this document MAY include multiple CONFIG objects when sending a Config, ConfigAck and ConfigNack message. A maximum of a single object of any particular C-type SHALL be included. A node which receives a message with multiple CONFIG objects of the same C-type SHALL process the first object of a particular C-type and ignore any subsequent CONFIG objects of the same C-type. Unless specified as part of the CONFIG object definition, ordering of CONFIG objects is not significant.

Nodes which support the extensions defined in this document MUST include a BehaviorConfig type object when sending a Config message to a neighbor whose support for the extensions is either known or unknown. (But not when the neighbor is known to not support the extensions.) Inclusion of other CONFIG objects in a Config message is at the discretion of the message sender, and is based on the rules defined by as part of CONFIG object definition. Nodes MAY include, HelloConfig, LMP_WDM_CONFIG, BehaviorConfig object types in a single message.

Inclusion of multiple CONFIG objects in a ConfigNack message is based on the processing of a received Config message. Per [RFC4204] "Parameters where agreement was reached MUST NOT be included in the ConfigNack Message." As such, a ConfigNack message MUST NOT include CONFIG objects which are acceptable and MUST include any CONFIG objects which are not acceptable. When a CONFIG object is included in a ConfigNack message, per [RFC4204], the object is to include "acceptable alternate values for negotiable parameters".

When sending a ConfigAck message, nodes supporting the extensions defined in this document MUST include all CONFIG objects received in the corresponding Config message when that message includes a CONFIG object of type BehaviorConfig.

3. LMP Behavior Negotiation

The Config message is used in the control channel negotiation phase of LMP [RFC4204]. The LMP behavior negotiation procedure is defined in this document as an addition to this phase.

The Config message is defined in Section 12.3.1 of [RFC4204] and carries the CONFIG object (class name 6) as defined in Section 13.6 of [RFC4204].

Two class types have been defined:

- C-Type = 1, HelloConfig, defined in [RFC4204]
- C-Type = 2, LMP_WDM_CONFIG, defined in [RFC4209]

This document defines a third C-Type to report and negotiate LMP mechanisms and behaviors. Its usage indicates support for the extensions defined in this document.

3.1. BehaviorConfig C-Type Format

Class = 6

- C-Type = (To be assigned by IANA), BehaviorConfig

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
S D C										Must Be Zero (MBZ)																													

Flags:

S: 1 bit

This bit indicates support for the Trace behavior of SONET/SDH technology-specific defined in [RFC4207].

D: 1 bit

This bit indicates support for the DWDM behavior defined in [RFC4209].

C: 1 bit

This bit indicates support for the data channel consistency check behavior defined in [RFC5818].

Must Be Zero (MBZ): Variable length

The remaining bits in the flags field MUST be set to zero (0). The number of bits present is based on the Length field of the LMP object header and MUST include enough bits so the Length field MUST be at least 8, and MUST be a multiple of 4.

Other bits may be defined in future documents, in which case the number of MBZ bits field is expected to change.

3.2. Processing

The inclusion of a BehaviorConfig type object in a message is discussed above in Section 2.2.

When sending a BehaviorConfig type object, the N-bit (negotiable) in the LMP object header must be set (N=1) in the LMP object header.

When sending a BehaviorConfig type object in Config and ConfigAck messages, the flags field SHOULD be set based on the supported capabilities of the sending node. When sending a ConfigAck message, the flags field MUST be set to the value received in the corresponding Config message.

When receiving a BehaviorConfig type object, the node compares the flags field against its capacities. Any bit set in the MBZ portion of the flags field MUST be interpreted as unacceptable. Processing related to unacceptable values in CONFIG objects is defined in [RFC4204] and is not modified by this document.

4. Backward Compatibility

The required use of the BehaviorConfig type CONFIG object enables nodes which support the extensions defined in this document to explicitly identify when a neighboring node does not. When a non-supporting node receives a Config message with the BehaviorConfig type CONFIG object or multiple CONFIG objects its behavior is likely to be one of the following behaviors:

- a) Reject the Config message because of the unknown BehaviorConfig object type and send a ConfigAck message which includes the unsupported C-type.

- b) Reject the message because of multiple CONFIG objects and send a ConfigNack message which includes all but one of the CONFIG objects.
- c) Silently ignore the one or more of the CONFIG object, and respond with a ConfigAck message that does not include any CONFIG objects.
- d) Treat the message as malformed, and discard it without any response.

Behaviors (a) and (b) result in ConfigNack messages with a BehaviorConfig type object whose contents are identical to what was sent in the Config message. Behavior (c) results in a ConfigAck message without a BehaviorConfig type CONFIG object. In each of these cases, the node SHOULD explicitly identify that the LMP neighbor does not support the extensions defined in this document.

Behavior (d) results in no response at all. When the node reaches the, [RFC4204] defined, "retry limit", the node SHOULD infer that the LMP neighbor does not support the extensions defined in this document.

Once a node identifies a neighbor as not supporting the extensions defined in this document, the node SHOULD follow previously defined Config message usage.

5. Security Considerations

[RFC4204] describes how LMP messages between peers can be secured, and these measures are equally applicable to messages carrying the new CONFIG object defined in this document.

The procedures described in this document do not of itself constitute a security risk since they do not cause any change in network state. It would be possible, if the messages were intercepted or spoofed to cause bogus alerts in the management plane, or to cause LMP peers to consider that they could or could not operate protocol extensions, and so the use of the LMP security measures are RECOMMENDED.

Note, however, that [RFC4204] refers to [RFC2401], which has been replaced by [RFC4301]. Also, the reference to IKEv2 in [RFC4301] is out of date, and the current reference for IKEv2 is [RFC5996].

7. Contributors

Diego Caviglia
Ericsson
Via A. Negrone 1/A 16153
Genoa Italy
Phone: +39 010 600 3736
Email: diego.caviglia@ericsson.com

8. Acknowledgments

Thanks to Adrian Farrel and Richard Graveman for their useful comments.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.
- [RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005
- [RFC5996] C. Kaufman, P. Hoffman, Y. Nir, P. Eronen, "Internet Key Exchange Protocol: IKEv2", RFC 5996, September 2010.
- [RFC4204] J. Lang, Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.
- [RFC4207] J. Lang, Ed., "Synchronous Optical Network (SONET)/ Synchronous Digital Hierarchy (SDH) Encoding for Link Management Protocol (LMP) Test Messages", RFC 4207, October 2005.
- [RFC4209] A. Fredette, Ed., "Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems", RFC 4209, October 2005.
- [RFC5818] D. Li, Ed., "Data Channel Status Confirmation Extensions for the Link Management Protocol", RFC 5818, April 2010.

9.2. Informative References

[LMP TEST] D. Ceccarelli, Ed., "Link Management Protocol (LMP) Test Messages Extensions for Evolutive Optical Transport Networks (OTN)" draft-ceccarelli-ccamp-gmpls-g709-lmp-test-02.txt, May, 2010.

10. Authors' Addresses

Dan Li
Huawei Technologies
F3-5-B R&D Center, Huawei Industrial Base,
Shenzhen 518129 China
Phone: +86 755-289-70230
Email: danli@huawei.com

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

Lou Berger
LabN Consulting, L.L.C.
Email: lberger@labn.net

MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 3, 2011

F. Zhang, Ed.
ZTE
R. Jing
China Telecom
June 01, 2011

RSVP-TE Extensions to Establish Associated Bidirectional LSP
draft-ietf-ccamp-mpls-tp-rsvpte-ext-associated-lsp-01

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 using a new Association Type 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 December 3, 2011.

Copyright Notice

Copyright (c) 2011 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. Association of Two Reverse Unidirectional LSPs	4
3.1. Provisioning Model	4
3.2. Signaling Procedure	4
3.2.1. Single Sided Provisioning Model	5
3.2.2. Double Sided Provisioning Model	6
3.2.3. Asymmetric Bandwidth LSPs	8
3.2.3.1. Error Handling	9
3.3. Recovery Considerations	10
4. Extensions to the Extended ASSOCIATION object	10
5. REVERSE_TSPEC Object	13
6. IANA Considerations	13
6.1. Association Type	13
6.2. REVERSE_TSPEC Object	14
7. Security Considerations	14
8. Acknowledgement	14
9. References	15
9.1. Normative references	15
9.2. Informative References	15
Authors' Addresses	16

1. Introduction

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

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

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

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

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

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

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

The above requirements are also repeated in [I-D.ietf-ccamp-mpls-tp-cp-framework].

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

This document provides a method to bind two unidirectional Label Switched Paths (LSPs) into an associated bidirectional LSP. The association is achieved by using a new Association Type 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. Association of Two Reverse Unidirectional LSPs

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 west to east) and LSP2 [via nodes B,D,C,A] (from east to west) are being established or have been established. These LSPs can be bound together to form an associated bidirectional LSP.

LSP1 is uniquely identified [I-D.ietf-mpls-tp-identifiers] by: West-Global_ID::West-Node_ID::West-Tunnel_Num::West-LSP_Num.

LSP2 is uniquely identified [I-D.ietf-mpls-tp-identifiers] by: East-Global_ID::East-Node_ID::East-Tunnel_Num::East-LSP_Num.

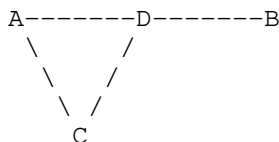


Figure 1: An example of associated bidirectional LSP

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, according to the scenarios described above, the following cases may occur:

1. Both LSPs do not exist.

LSP1 is initialized at node A with the Extended ASSOCIATION object inserted in the Path message, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to West-Node_ID, Global Association Source set to West-Global_ID, and Extended Association ID set to West-Tunnel_Num. Terminating node B is triggered to set up LSP2 by the received Extended ASSOCIATION object with the Association Type set to the value "Association of two reverse unidirectional LSPs", the Association Object inserted in LSP2's Path message is the same as in LSP1's Path message.

2. LSP1 exists, LSP2 needs to be established.

LSP1 is refreshed with the Extended ASSOCIATION object inserted in the Path message, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to West-Node_ID, Global Association Source set to West-Global_ID, and Extended Association ID set to West-Tunnel_Num. Terminating node B is triggered to set up LSP2 by the received Extended ASSOCIATION object with the Association Type set to the value "Association of two reverse unidirectional LSPs", the Association Object inserted in LSP2's Path message is the same as in LSP1's Path message.

3. LSP1 does not exist, LSP2 has been established.

LSP1 is initialized with the Extended ASSOCIATION object inserted in the Path message, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num. Terminating node B is triggered to refresh LSP2's Path message, with the received Extended ASSOCIATION object inserted.

4. Both LSP1 and LSP2 exist.

LSP1 is refreshed with the Extended ASSOCIATION object inserted in the Path message, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num. Terminating node B is triggered to refresh LSP2's Path message, with the received Extended ASSOCIATION object inserted.

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, Similarly, according to the scenarios described above, the following cases may occur:

1. LSP1 and LSP2 do not exist.

LSP1 and LSP2 are concurrently initialized with the Extended ASSOCIATION object inserted in the their Path messages, For LSP1, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to West-Node_ID, Global Association Source set to West-Global_ID, and Extended Association ID set to West-Tunnel_Num. For LSP2, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID is set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num. According to the general rules defined in [I-D.ietf-ccamp-assoc-ext], the two LSPs cannot be bound together to be an associated bidirectional LSP because of the different values. In this case, the two edge nodes firstly MUST compare their Global-Node_ID, then the bigger one sends Path refresh message, replacing the old Extended ASSOCIATION object with the new Extended ASSOCIATION object carried in the reverse LSP. Based on this Path refresh message, the two LSPs can be

bounded together to be an associated bidirectional LSP also.

2. LSP1 exists, LSP2 needs to be established.

For LSP1, Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to West-Node_ID, Global Association Source set to West-Global_ID, and Extended Association ID set to West-Tunnel_Num. For LSP2, Node B has known the existence of LSP1, so the Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to West-Node_ID, Global Association Source set to West-Global_ID, and Extended Association ID set to West-Tunnel_Num.

3. LSP1 does not exist, LSP2 has been established.

For LSP1, Node A has known the existence of LSP2. So the Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num. For LSP2, Node B does not know the existence of LSP1, so Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num.

4. Both LSP1 and LSP2 exist.

In this case, Both node A and Node B know the existence of the reverse LSPs. The two edge nodes firstly MUST compare their Global-Node_ID, then the bigger one sends Path refresh message, with the reverse LSP's identifier inserted in the Extended ASSOCIATION object, and the smaller one sends Path refresh message, with its own LSP's identifier inserted in the Extended ASSOCIATION object. For example, assuming that the node A has the bigger Global-Node_ID. For LSP1, the Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to East-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num. For LSP2, the Association Type is set to "Association of two reverse unidirectional LSPs", Association ID set to West-LSP_Num, Association Source set to East-Node_ID, Global Association Source set to East-Global_ID, and Extended Association ID set to East-Tunnel_Num.

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.

3.2.3. 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 [I-D.ietf-ccamp-asymm-bw-bidir-lsps-bis], which introduces three new objects named UPSTREAM_FLOWSPEC object, UPSTREAM_TSPEC object and UPSTREAM_ADSPEC object to represent the asymmetric upstream traffic flow. For the asymmetric bandwidth associated bidirectional LSPs, the existing SENDER_TSPEC, ADSPEC, and FLOWSPEC are complemented with the addition of a new REVERSE_TSPEC object, which is used exactly in the same fashion as the old SENDER_TSPEC object.

Consider the topology described in Figure 1 in the context of asymmetric associated bidirectional LSP, the following cases may occur:

1. LSP1 and LSP2 do not exist.

For the single sided provisioning, taking LSP2 triggered by LSP1 as an example. The REVERSE_TSPEC object MUST be carried in the LSP1's Path message together with the Extended ASSOCIATION object whose Association Type is "Association of two reverse unidirectional LSPs". The terminating node B is triggered to set up the reverse LSP2 with the corresponding asymmetric bandwidth, and the REVERSE_TSPEC object is converted to the SENDER_TSPEC object in the Path message.

For the double sided provisioning, the REVERSE_TSPEC object MUST be carried in the two LSPs' Path message together with the Extended ASSOCIATION object whose Association Type is "Association of two reverse unidirectional LSPs". Then the two terminating ends MUST compare the values of the SENDER_TSPEC and REVERSE_TSPEC objects in the two Path messages. If the values match, the end with the bigger Global-Node_ID sends Path refresh message, carrying the Extended ASSOCIATION object of the reverse LSP.

2. LSP1 exists, LSP2 needs to be established.

For the single sided provisioning, taking LSP2 triggered by LSP1 as an example. The REVERSE_TSPEC object MUST be carried in the LSP1's Path refresh message together with the Extended ASSOCIATION object whose Association Type is "Association of two reverse unidirectional LSPs". The terminating node B is triggered to set up the reverse LSP2 with the corresponding asymmetric bandwidth, and the REVERSE_TSPEC object is converted to the SENDER_TSPEC object in the Path message.

For the double sided provisioning, the REVERSE_TSPEC object MUST be carried in the LSP1's Path refresh message with the Extended ASSOCIATION object whose Association Type is "Association of two reverse unidirectional LSPs". There is no need to put the REVERSE_TSPEC object in LSP2's Path message, for the Extended ASSOCIATION object has indicated that LSP2 needs to be bound with LSP1.

3. LSP1 does not exist, LSP2 has been established.

For the single sided provisioning, taking LSP2 triggered by LSP1 as an example. There is no need to put the REVERSE_TSPEC object in LSP1's Path message, for the Extended ASSOCIATION object has indicates that LSP1 needs to be bound with LSP2.

For the double sided provisioning, just the same reason, the REVERSE_TSPEC object only needs to be carried in the LSP2's Path refresh message.

4. Both LSP1 and LSP2 exist.

For the single sided provisioning, taking LSP2 triggered by LSP1 as an example. There is no need to put the REVERSE_TSPEC object in LSP1's Path message also for the Extended ASSOCIATION object has indicates that LSP1 needs to be bound with LSP2.

As to the double sided provisioning, just the same reason, the REVERSE_TSPEC object does not need to be carried in the two LSPs' Path messages.

Based on the same values of the Association objects in the two LSPs' Path message, and the match of the REVERSE_TSPEC and SENDER_TSPEC objects in the two LSPs' Path message (if the REVERSE_TSPEC object exists), the two LSPs can be bound together to be an associated bidirectional LSP.

3.2.3.1. Error Handling

Nodes not supporting the new class number of the REVERSE_TSPEC object SHOULD respond with an "Unknown Object Class".

3.3. 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.

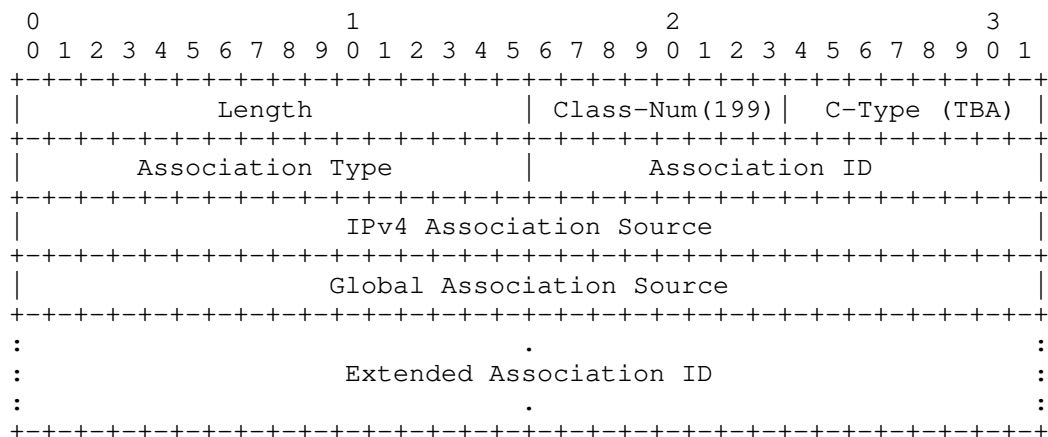
In the case that LSP3 is established after the failure occurs, the Extended ASSOCIATION object with LSP2's identifier SHOULD be inserted in LSP3's Path message since LSP2 has already existed. If LSP1 and LSP2 are associated together by the LSP1's identifier, LSP2's Path message is refreshed, an additional Extended ASSOCIATION object with LSP2's identifier are inserted. If LSP1 and LSP2 are bound together by the LSP2's identifier, there is no need to insert an additional Extended ASSOCIATION object in LSP2's Path message.

In the case that LSP3 is established before the failure occurs. For single sided provisioning, LSP3 is refreshed with the Extended ASSOCIATION object, its values are filled by LSP2's identifier. Then LSP2 is refreshed with this Extended ASSOCIATION object or not, see the description in the above paragraph. For double sided provisioning, if node A has the bigger Global-Node_ID than node B, LSP3 is refreshed with the Extended ASSOCIATION object whose values are filled by LSP2's identifier, and LSP2 is refreshed with this Extended ASSOCIATION object or not, see the description in the above paragraph. If node A has the smaller Global-Node_ID than node B, LSP3 is refreshed with the Extended ASSOCIATION object whose values are filled by LSP3's identifier, and LSP2 is refreshed with this Extended ASSOCIATION object.

4. Extensions to the Extended ASSOCIATION object

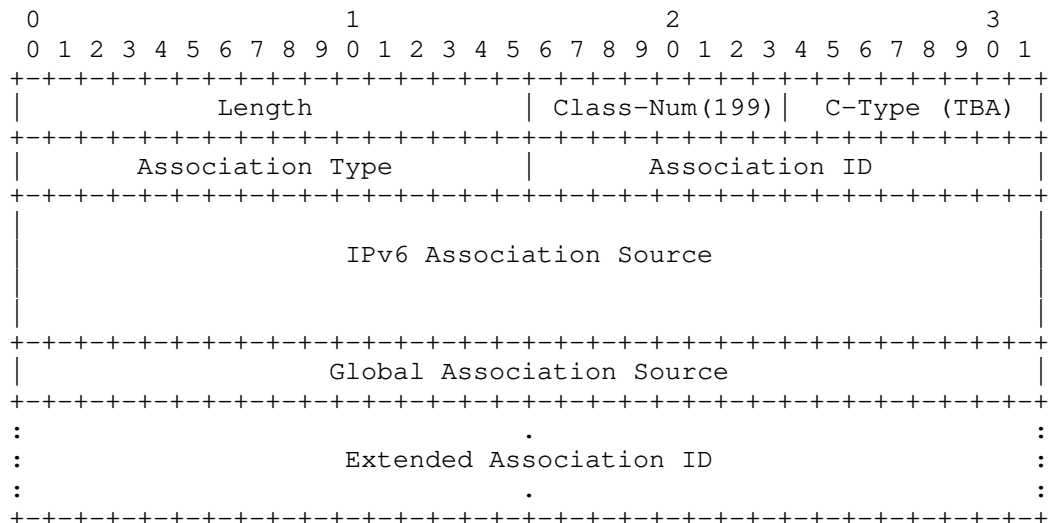
The Extended ASSOCIATION object is defined in [I-D.ietf-ccamp-assoc-ext], which enables MPLS-TP required identification.

The Extended IPv4 ASSOCIATION object (Class-Num of the form 11bbbbbb with value = 199, C-Type = TBA) has the format:



Extended IPv4 ASSOCIATION object

The Extended IPv6 ASSOCIATION object (Class-Num of the form 11bbbbbb with value = 199, C-Type = TBA) has the format:



Extended IPv6 ASSOCIATION object

- o Association Type:

In order to bind two reverse unidirectional LSPs to be an associated bidirectional LSP, this document defines a new Association Type:

Value	Type
-----	-----
4 (TBD)	Association of two reverse unidirectional LSPs (A)

If the downstream nodes are not aware of the Association Type, they MUST return a PathErr message with error code/sub-code "LSP Admission Failure/Bad Association Type".

Under the context of this Association Type, any node associating an associated bidirectional LSP MUST insert an ASSOCIATION object with the following setting:

- o Association ID:

The Association ID MUST be set to its own signaled LSP ID (default); if known, it MAY be set to the LSP ID of the associated reverse LSP.

- o Association Source:

The Association source MUST be set to the tunnel sender address of this LSP (default); if known, it May be set to the tunnel sender address of the peer node.

- o Global Association Source:

The format is described in [I-D.ietf-ccamp-assoc-ext].

- o Extended Association ID:

Because the two LSPs (one is from west to east, and the other is from east to west) are in different tunnels, the Association ID is insufficient to uniquely identify association for associated bidirectional LSP. Hence, this document adds specific rules: the first 16-bits MUST be set to its own tunnel ID (default); if known, it May be set to the tunnel ID of the the associated reverse tunnel.

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 type in this document is only defined for use in upstream initialized association. Thus it 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.

5. REVERSE_TSPEC Object

The REVERSE_TSPEC object is used in Path, PathTear, PathErr, and Notify message (via sender descriptor). This includes the definition of class type and format. It's class number is TBD (of the form 0bbbbbbb), and class type and format is the same as the SENDER_TSPEC object.

This object modifies the RSVP message-related formats defined in [RFC2205], [RFC3209] and [RFC3473]. See [RFC5511] for the syntax used by RSVP. The format of the sender description for asymmetric associated bidirectional LSPs is:

```
<sender descriptor>::= <SENDER_TEMPLATE> <SENDER_TSPEC>
                        [<ADSPEC>]
                        [<RCEORD_ROUTE>]
                        [<SUGGESTED_LABEL>]
                        [<RECOVERY_LABEL>]
                        <REVERSE_TSPEC>
```

6. IANA Considerations

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

6.1. Association Type

Within the current document, a new Association Type is defined in the Extended ASSOCIATION object.

Value -----	Type -----
4 (TBD)	Association of two reverse unidirectional LSPs (A)

6.2. REVERSE_TSPEC Object

A new class named REVERSE_TSPEC has been created in the 0bbbbbbb rang (123,TBD) with the following definition:

Class Types or C-types:

Same values as SENDER_TPSCE object (C-Num 12)

There are no other IANA considerations introduced by this document.

7. Security Considerations

This document introduces a new association type, and except this, there are no security issues about the Extended ASSOCIATION object are introduced here.

Furthermore, this document introduces the REVERSE_TSPEC object for use in GMPLS signaling [RFC3473], which is parallel the existing SENDER_TSPEC object. As such, any vulnerabilities that are due to the use of the old SENDER_TSPEC object now apply here also.

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].

8. 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, and Wenjuan He for the initial discussions.

9. References

9.1. Normative references

- [I-D.ietf-ccamp-assoc-ext]
Berger, L., Faucheur, F., and A. Narayanan, "RSVP Association Object Extensions", draft-ietf-ccamp-assoc-ext-00 (work in progress), May 2011.
- [I-D.ietf-mpls-tp-identifiers]
Bocci, M., Swallow, G., and E. Gray, "MPLS-TP Identifiers", draft-ietf-mpls-tp-identifiers-04 (work in progress), March 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4872] Lang, J., Rekhter, Y., and D. Papadimitriou, "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.
- [RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.

9.2. Informative References

- [I-D.ietf-ccamp-assoc-info]
Berger, L., "Usage of The RSVP Association Object", draft-ietf-ccamp-assoc-info-02 (work in progress), May 2011.
- [I-D.ietf-ccamp-asymm-bw-bidir-lsps-bis]
Takacs, A., Berger, L., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", draft-ietf-ccamp-asymm-bw-bidir-lsps-bis-01 (work in progress), January 2011.
- [I-D.ietf-ccamp-mpls-tp-cp-framework]
Andersson, L., Berger, L., Fang, L., Bitar, N., Gray, E., Takacs, A., Vigoureux, M., and E. Bellagamba, "MPLS-TP Control Plane Framework", draft-ietf-ccamp-mpls-tp-cp-framework-06 (work in progress), February 2011.

- [RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax Used to Form Encoding Rules in Various Routing Protocol Specifications", RFC 5511, April 2009.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

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

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2012

A. Takacs
Ericsson
D. Fedyk
Alcatel-Lucent
J. He
Huawei
July 11, 2011

GMPLS RSVP-TE extensions for OAM Configuration
draft-ietf-ccamp-oam-configuration-fwk-06

Abstract

OAM is an integral part of transport connections, hence it is required that OAM functions are activated/deactivated in sync with connection commissioning/decommissioning; avoiding spurious alarms and ensuring consistent operation. In certain technologies OAM entities are inherently established once the connection is set up, while other technologies require extra configuration to establish and configure OAM entities. This document specifies extensions to RSVP-TE to support the establishment and configuration of OAM entities along with LSP signaling.

Requirements Language

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

Status of this Memo

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

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

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

This Internet-Draft will expire on January 12, 2012.

Copyright Notice

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

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

Table of Contents

1. Introduction	4
2. Requirements	6
3. RSVP-TE based OAM Configuration	9
3.1. Establishment of OAM Entities and Functions	9
3.2. Adjustment of OAM Parameters	11
3.3. Deleting OAM Entities	11
4. RSVP-TE Extensions	13
4.1. LSP Attributes Flags	13
4.2. OAM Configuration TLV	14
4.2.1. OAM Function Flags Sub-TLV	15
4.2.2. Technology Specific sub-TLVs	16
4.3. Administrative Status Information	16
4.4. Handling OAM Configuration Errors	16
4.5. Considerations on Point-to-Multipoint OAM Configuration	17
5. IANA Considerations	19
6. Security Considerations	20
7. Acknowledgements	21
8. References	22
8.1. Normative References	22
8.2. Informative References	22
Authors' Addresses	24

1. Introduction

GMPLS is designed as an out-of-band control plane supporting dynamic connection provisioning for any suitable data plane technology; including spatial switching (e.g., incoming port or fiber to outgoing port or fiber), wavelength-division multiplexing (e.g., DWDM), time-division multiplexing (e.g., SONET/SDH, G.709), and Ethernet Provider Backbone Bridging -- Traffic Engineering (PBB-TE) and MPLS. In most of these technologies there are Operations, Administration and Maintenance (OAM) functions employed to monitor the health and performance of the connections and to trigger data plane (DP) recovery mechanisms. Similarly to connections, OAM functions follow general principles but also have some technology specific characteristics.

OAM is an integral part of transport connections, hence it is required that OAM functions are activated/deactivated in sync with connection commissioning/decommissioning; avoiding spurious alarms and ensuring consistent operation. In certain technologies OAM entities are inherently established once the connection is set up, while other technologies require extra configuration to establish and configure OAM entities. In some situations the use of OAM functions, like those of Fault- (FM) and Performance Management (PM), may be optional confirming to actual network management policies. Hence the network operator must be able to choose which kind of OAM functions to apply to specific connections and with what parameters the selected OAM functions should be configured and operated. To achieve this objective OAM entities and specific functions must be selectively configurable.

In general, it is required that the management plane and control plane connection establishment mechanisms are synchronized with OAM establishment and activation. In particular, if the GMPLS control plane is employed it is desirable to bind OAM setup and configuration to connection establishment signaling to avoid two separate management/configuration steps (connection setup followed by OAM configuration) which increases delay, processing and more importantly may be prone to misconfiguration errors. Once OAM entities are setup and configured, pro-active as well as on-demand OAM functions can be activated via the management plane. On the other hand, it should be possible to activate/deactivate pro-active OAM functions via the GMPLS control plane as well.

This document describes requirements on OAM configuration and control via RSVP-TE, and specifies extensions to the RSVP-TE protocol providing a framework to configure and control OAM entities along with the capability to carry technology specific information. Extensions can be grouped into generic elements that are applicable

to any OAM solution and technology specific elements that provide additional configuration parameters, which are only needed for a specific OAM technology. This document specifies the technology agnostic elements, which alone can be used to establish and control OAM entities in the case no technology specific information is needed, and specifies the way additional technology specific OAM parameters are provided.

This document addresses end-to-end OAM configuration, that is, the setup of OAM entities bound to an end-to-end LSP, and configuration and control of OAM functions running end-to-end in the LSP. Configuration of OAM entities for LSP segments and tandem connections are out of the scope of this document.

The mechanisms described in this document provide an additional option for bootstrapping OAM that is not intended to replace or deprecate the use of other technology specific OAM bootstrapping techniques; e.g., LSP Ping [RFC4379] for MPLS networks. The procedures specified in this document are intended only for use in environments where RSVP-TE signaling is already in use to set up the LSPs that are to be monitored using OAM.

2. Requirements

MPLS OAM requirements are described in [RFC4377], which provides requirements to create consistent OAM functionality for MPLS networks.

The following list is an excerpt of MPLS OAM requirements documented in [RFC4377]. Only a few requirements are discussed that bear a direct relevance to the discussion set forth in this document.

- o It is desired to support the automation of LSP defect detection. It is especially important in cases where large numbers of LSPs might be tested.
- o In particular some LSPs may require automated ingress-LSR to egress-LSR testing functionality, while others may not.
- o Mechanisms are required to coordinate network responses to defects. Such mechanisms may include alarm suppression, translating defect signals at technology boundaries, and synchronizing defect detection times by setting appropriately bounded detection timeframes.

MPLS-TP defines a profile of MPLS targeted at transport applications [RFC5921]. This profile specifies the specific MPLS characteristics and extensions required to meet transport requirements, including providing additional OAM, survivability and other maintenance functions not currently supported by MPLS. Specific OAM requirements for MPLS-TP are specified in [RFC5654] [RFC5860]. MPLS-TP poses requirements on the control plane to configure and control OAM entities:

- o From [RFC5860]: OAM functions MUST operate and be configurable even in the absence of a control plane. Conversely, it SHOULD be possible to configure as well as enable/disable the capability to operate OAM functions as part of connectivity management, and it SHOULD also be possible to configure as well as enable/disable the capability to operate OAM functions after connectivity has been established.
- o From [RFC5654]: The MPLS-TP control plane MUST support the configuration and modification of OAM maintenance points as well as the activation/ deactivation of OAM when the transport path or transport service is established or modified.

Ethernet Connectivity Fault Management (CFM) defines an adjunct connectivity monitoring OAM flow to check the liveness of Ethernet networks [IEEE-CFM]. With PBB-TE [IEEE-PBBTE] Ethernet networks

support explicitly-routed Ethernet connections. CFM can be used to track the liveness of PBB-TE connections and detect data plane failures. In IETF the GMPLS controlled Ethernet Label Switching (GELS) (see [RFC5828] and [RFC6060]) work extended the GMPLS control plane to support the establishment of PBB-TE data plane connections. Without control plane support separate management commands would be needed to configure and start CFM.

GMPLS based OAM configuration and control should be general to be applicable to a wide range of data plane technologies and OAM solutions. There are three typical data plane technologies used for transport application, which are wavelength based such as WSON, TDM based such as SDH/SONET, packet based such as MPLS-TP [RFC5921] and Ethernet PBB-TE [IEEE-PBBTE]. In all these data planes, the operator MUST be able to configure and control the following OAM functions.

- o It MUST be possible to explicitly request the setup of OAM entities for the signaled LSP and provide specific information for the setup if this is required by the technology.
- o Control of alarms is important to avoid false alarm indications and reporting to the management system. It MUST be possible to enable/disable alarms generated by OAM functions. In some cases selective alarm control may be desirable when, for instance, the operator is only concerned about critical alarms thus the non-service affecting alarms should be inhibited.
- o When periodic messages are used for liveness check (continuity check) of LSPs it MUST be possible to set the frequency of messages allowing proper configuration for fulfilling the requirements of the service and/or meeting the detection time boundaries posed by possible congruent connectivity check operations of higher layer applications. For a network operator to be able to balance the trade-off in fast failure detection and overhead it is beneficial to configure the frequency of continuity check messages on a per LSP basis.
- o Pro-active Performance Monitoring (PM) functions are continuously collecting information about specific characteristics of the connection. For consistent measurement of Service Level Agreements (SLAs) measurement points must use common probing rate to avoid measurement errors.
- o The extensions MUST allow the operator to use only a minimal set of OAM configuration and control features if the data plane technology, the OAM solution or network management policy allows. The extensions must be reusable as much as reasonably possible. That is generic OAM parameters and data plane or OAM technology

specific parameters must be separated.

3. RSVP-TE based OAM Configuration

In general, two types of Maintenance Points (MPs) can be distinguished: Maintenance End Points (MEPs) and Maintenance Intermediate Points (MIPs). MEPs reside at the ends of an LSP and are capable of initiating and terminating OAM messages for Fault Management (FM) and Performance Monitoring (PM). MIPs on the other hand are located at transit nodes of an LSP and are capable of reacting to some OAM messages but otherwise do not initiate messages. Maintenance Entity (ME) refers to an association of MEPs and MIPs that are provisioned to monitor an LSP. The ME association is achieved by configuring MPs to belong to the same ME.

When an LSP is signaled, forwarding association is established between endpoints and transit nodes via label bindings. This association creates a context for the OAM entities monitoring the LSP. On top of this association OAM entities may be configured to unambiguously identify MPs and MEs.

In addition to MP and ME identification parameters pro-active OAM functions (e.g., Continuity Check (CC), Performance Monitoring) may have specific parameters requiring configuration as well. In particular, the frequency of periodic CC packets and the measurement interval for loss and delay measurements may need to be configured.

In some cases all the above parameters may be either derived from some exiting information or pre-configured default values can be used. In the simplest case the control plane needs to provide information whether or not OAM entities need to be setup for the signaled LSP. If OAM entities are created signaling must provide means to activate/deactivate OAM message flows and associated alarms.

OAM identifiers as well as the configuration of OAM functions are technology specific, i.e., vary depending on the data plane technology and the chosen OAM solution. In addition, for any given data plane technology a set of OAM solutions may be applicable. The OAM configuration framework allows selecting a specific OAM solution to be used for the signaled LSP and provides technology specific TLVs to carry further detailed configuration information.

3.1. Establishment of OAM Entities and Functions

In order to avoid spurious alarms OAM functions should be setup and enabled in the appropriate order. When using the GMPLS control plane, establishment and enabling of OAM functions MUST be bound to RSVP-TE message exchanges.

An LSP can be signaled and established without OAM configuration

first, and OAM entities can be added later with a subsequent re-signaling of the LSP. Alternatively, the LSP can be setup with OAM entities right with the first signaling of the LSP. The below procedures apply to both cases.

Before the initiator first sends a Path messages with OAM Configuration information, it MUST establish and configure the corresponding OAM entities locally, however OAM source functions MUST NOT start sending any OAM messages. In the case of bidirectional connections, the initiator node MUST setup the OAM sink function to be prepared to receive OAM messages but MUST suppress any OAM alarms (e.g., due to missing or unidentified OAM messages). The Path message MUST be sent with the "OAM Alarms Enabled" ADMIN_STATUS flag cleared, i.e, data plane OAM alarms are suppressed.

When the Path message arrives at the receiver, the remote end MUST establish and configure OAM entities according to the OAM information provided in the Path message. If this is not possible a PathErr SHOULD be sent and neither the OAM entities nor the LSP SHOULD be established. If OAM entities are established successfully, the OAM sink function MUST be prepared to receive OAM messages but MUST not generate any OAM alarms (e.g., due to missing or unidentified OAM messages). In the case of bidirectional connections, an OAM source function MUST be setup and, according to the requested configuration, the OAM source function MUST start sending OAM messages. Then a Resv message is sent back, including the OAM Configuration TLV that corresponds to the actually established and configured OAM entities and functions. Depending on the OAM technology, some elements of the OAM Configuration TLV MAY be updated/changed; i.e., if the remote end is not supporting a certain OAM configuration it may suggest an alternative setting, which may or may not be accepted by the initiator of the Path message. If it is accepted, the initiator will reconfigure its OAM functions according to the information received in the Resv message. If the alternate setting is not acceptable a ResvErr may be sent tearing down the LSP. Details of this operation are technology specific and should be described in accompanying technology specific documents.

When the initiating side receives the Resv message it completes any pending OAM configuration and enables the OAM source function to send OAM messages.

After this round, OAM entities are established and configured for the LSP and OAM messages are already exchanged. OAM alarms can now be enabled. The initiator, while still keeping OAM alarms disabled sends a Path message with "OAM Alarms Enabled" ADMIN_STATUS flag set. The receiving node enables the OAM alarms after processing the Path message. The initiator enables OAM alarms after it receives the Resv

message. Data plane OAM is now fully functional.

3.2. Adjustment of OAM Parameters

There may be a need to change the parameters of an already established and configured OAM function during the lifetime of the LSP. To do so the LSP needs to be re-signaled with the updated parameters. OAM parameters influence the content and timing of OAM messages and identify the way OAM defects and alarms are derived and generated. Hence, to avoid spurious alarms, it is important that both sides, OAM sink and source, are updated in a synchronized way. First, the alarms of the OAM sink function should be suppressed and only then should expected OAM parameters be adjusted. Subsequently, the parameters of the OAM source function can be updated. Finally, the alarms of the OAM sink side can be enabled again.

In accordance with the above operation, the LSP MUST first be re-signaled with "OAM Alarms Enabled" ADMIN_STATUS flag cleared and including the updated OAM Configuration TLV corresponding to the new parameter settings. The initiator MUST keep its OAM sink and source functions running unmodified, but it MUST suppress OAM alarms after the updated Path message is sent. The receiver MUST first disable all OAM alarms, then update the OAM parameters according to the information in the Path message and reply with a Resv message acknowledging the changes by including the OAM Configuration TLV. Note that the receiving side has the possibility to adjust the requested OAM configuration parameters and reply with an updated OAM Configuration TLV in the Resv message, reflecting the actually configured values. However, in order to avoid an extensive negotiation phase, in the case of adjusting already configured OAM functions, the receiving side SHOULD NOT update the parameters requested in the Path message to an extent that would provide lower performance than what has been configured previously.

The initiator MUST only update its OAM sink and source functions after it received the Resv message. After this Path/Resv message exchange (in both unidirectional and bidirectional LSP cases) the OAM parameters are updated and OAM is running according to the new parameter settings. However OAM alarms are still disabled. A subsequent Path/Resv message exchange with "OAM Alarms Enabled" ADMIN_STATUS flag set is needed to enable OAM alarms again.

3.3. Deleting OAM Entities

In some cases it may be useful to remove some or all OAM entities and functions from an LSP without actually tearing down the connection.

To avoid any spurious alarm, first the LSP SHOULD be re-signaled with

"OAM Alarms Enabled" ADMIN_STATUS flag cleared but unchanged OAM configuration. Subsequently, the LSP is re-signaled with "OAM MEP Entities desired" and "OAM MIP Entities desired" LSP ATTRIBUTES flags cleared, and without the OAM Configuration TLV, this MUST result in the deletion of all OAM entities associated with the LSP. All control and data plane resources in use by the OAM entities and functions SHOULD be freed up. Alternatively, if only some OAM functions need to be removed, the LSP is re-signalled with the updated OAM Configuration TLV. Changes between the contents of the previously signalled OAM Configuration TLV and the currently received TLV represent which functions SHOULD be removed/added.

First, OAM source functions SHOULD be deleted and only after that SHOULD the associated OAM sink functions be removed, this will ensure that OAM messages do not leak outside the LSP. To this end the initiator, before sending the Path message, SHOULD remove the OAM source, hence terminating the OAM message flow associated to the downstream direction. In the case of a bidirectional connection, it SHOULD leave in place the OAM sink functions associated to the upstream direction. The remote end, after receiving the Path message, SHOULD remove all associated OAM entities and functions and reply with a Resv message without an OAM Configuration TLV. The initiator completely removes OAM entities and functions after the Resv message arrived.

4. RSVP-TE Extensions

4.1. LSP Attributes Flags

In RSVP-TE the Flags field of the SESSION_ATTRIBUTE object is used to indicate options and attributes of the LSP. The Flags field has 8 bits and hence is limited to differentiate only 8 options. [RFC5420] defines new objects for RSVP-TE messages to allow the signaling of arbitrary attribute parameters making RSVP-TE easily extensible to support new applications. Furthermore, [RFC5420] allows options and attributes that do not need to be acted on by all Label Switched Routers (LSRs) along the path of the LSP. In particular, these options and attributes may apply only to key LSRs on the path such as the ingress LSR and egress LSR. Options and attributes can be signaled transparently, and only examined at those points that need to act on them. The LSP_ATTRIBUTES and the LSP_REQUIRED_ATTRIBUTES objects are defined in [RFC5420] to provide means to signal LSP attributes and options in the form of TLVs. Options and attributes signaled in the LSP_ATTRIBUTES object can be passed transparently through LSRs not supporting a particular option or attribute, while the contents of the LSP_REQUIRED_ATTRIBUTES object must be examined and processed by each LSR. One TLV is defined in [RFC5420]: the Attributes Flags TLV.

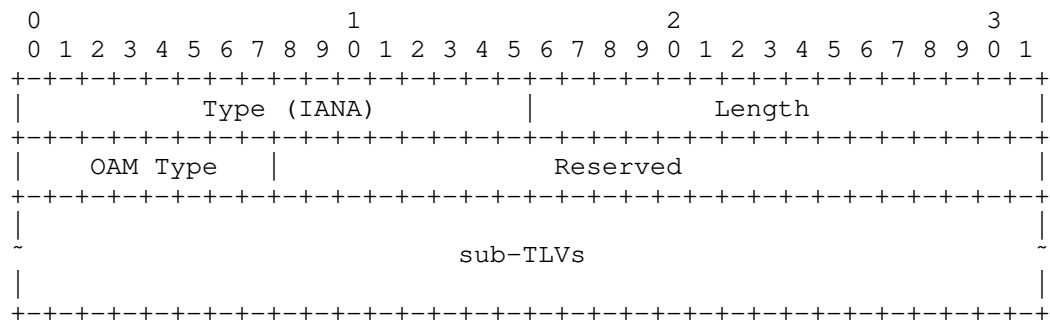
One bit (IANA to assign): "OAM MEP entities desired" is allocated in the LSP Attributes Flags TLV to be used in the LSP_ATTRIBUTES object. If the "OAM MEP entities desired" bit is set it is indicating that the establishment of OAM MEP entities are required at the endpoints of the signaled LSP. If the establishment of MEPs is not supported an error must be generated: "OAM Problem/MEP establishment not supported".

If the "OAM MEP entities desired" bit is set and additional parameters need to be configured, an OAM Configuration TLV MAY be included in the LSP_ATTRIBUTES Object.

One bit (IANA to assign): "OAM MIP entities desired" is allocated in the LSP Attributes Flags TLV to be used in the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES objects. This bit can only be set if the "OAM MEP entities desired" bit is set in. If the "OAM MIP entities desired" bit is set in the LSP_ATTRIBUTES Flags TLV in the LSP_REQUIRED_ATTRIBUTES Object, it is indicating that the establishment of OAM MIP entities is required at every transit node of the signalled LSP. If the establishment of a MIP is not supported an error MUST be generated: "OAM Problem/MIP establishment not supported".

4.2. OAM Configuration TLV

This TLV provides information about which OAM technology/method should be used and carries sub-TLVs for any additional OAM configuration information. The OAM Configuration TLV MAY be carried in the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES object in Path and Resv messages. When carried in the LSP_REQUIRED_ATTRIBUTES object it is indicating that intermediate nodes MUST recognize and eventually react on the OAM configuration onformation.



Type: indicates a new type: the OAM Configuration TLV (3) (IANA to assign).

OAM Type: specifies the technology specific OAM method. When carried in the LSP_REQUIRED_ATTRIBUTES Object, if the requested OAM method is not supported at a given node an error MUST be generated: "OAM Problem/Unsupported OAM Type". When carried in the LSP_ATTRIBUTES Object, intermediate nodes not supporting the OAM Type pass the object forward unchanged as specified in [RFC5420] only Label Edge Nodes MUST generate the error if the OAM Type is not supported.

OAM Type	Description
-----	-----
0-255	Reserved

This document defines no types. IANA is requested to maintain the values in a new "RSVP-TE OAM Configuration Registry".

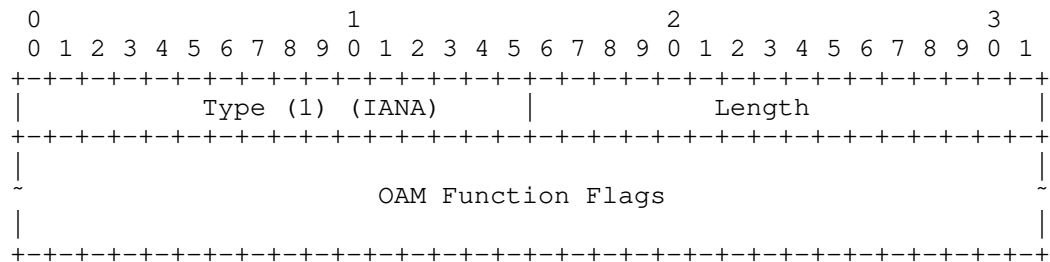
The receiving node based on the OAM Type will check if a corresponding technology specific OAM configuration sub-TLV is included. If the included technology specific OAM configuration sub-TLV is different than what is specified in the OAM Type an error MUST be generated: "OAM Problem/OAM Type Mismatch". IANA is requested to maintain the sub-TLV space in the new "RSVP-TE OAM Configuration

Registry".

Note that there is a hierarchical dependency in between the OAM configuration elements. First, the "OAM MEP (and MIP) entities desired" flag needs to be set. Only when that is set MAY an "OAM Configuration TLV" be included in the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Object. When this TLV is present, based on the "OAM Type" field, it MAY carry a technology specific OAM configuration sub-TLV. If this hierarchy is broken (e.g., "OAM MEP entities desired" flag is not set but an OAM Configuration TLV is present) an error MUST be generated: "OAM Problem/Configuration Error".

4.2.1. OAM Function Flags Sub-TLV

As the first sub-TLV the "OAM Function Flags sub-TLV" MUST be always included in the "OAM Configuration TLV". "OAM Function Flags" specifies which pro-active OAM functions (e.g., connectivity monitoring, loss and delay measurement) and which fault management signals MUST be established and configured. If the selected OAM Function(s) is(are) not supported, an error MUST be generated: "OAM Problem/Unsupported OAM Function".



OAM Function Flags is bitmap with extensible length based on the Length field of the TLV. Bits are numbered from left to right. IANA is requested to maintain the OAM Function Flags in the new "RSVP-TE OAM Configuration Registry". This document defines the following flags.

OAM Function Flag bit#	Description
0	Continuity Check (CC)
1	Connectivity Verification (CV)
2	Fault Monitoring Signal (FMS)
3	Performance Monitoring/Loss (PM/Loss)
4	Performance Monitoring/Delay (PM/Delay)
5	Performance Monitoring/Throughput Measurement (PM/Throughput)

4.2.2. Technology Specific sub-TLVs

One technology specific sub-TLV MAY be defined for each "OAM Type". This sub-TLV MUST contain any further OAM configuration information for that specific "OAM Type". The technology specific sub-TLV, when used, MUST be carried within the OAM Configuration TLV. IANA is requested to maintain the sub-TLV space in the new "RSVP-TE OAM Configuration Registry".

4.3. Administrative Status Information

Administrative Status Information is carried in the ADMIN_STATUS Object. The Administrative Status Information is described in [RFC3471], the ADMIN_STATUS Object is specified for RSVP-TE in [RFC3473].

Two bits are allocated for the administrative control of OAM monitoring. Two bits (IANA to assign) are allocated by this draft: the "OAM Flows Enabled" (M) and "OAM Alarms Enabled" (O) bits. When the "OAM Flows Enabled" bit is set, OAM packets are sent if it is cleared no OAM packets are emitted. When the "OAM Alarms Enabled" bit is set OAM triggered alarms are enabled and associated consequent actions are executed including the notification of the management system. When this bit is cleared, alarms are suppressed and no action is executed and the management system is not notified.

4.4. Handling OAM Configuration Errors

To handle OAM configuration errors a new Error Code (IANA to assign) "OAM Problem" is introduced. To refer to specific problems a set of Error Values is defined.

If a node does not support the establishment of OAM MEP or MIP entities it must use the error value (IANA to assign): "MEP establishment not supported" or "MIP establishment not supported" respectively in the PathErr message.

If a node does not support a specific OAM technology/solution it must

use the error value (IANA to assign): "Unsupported OAM Type" in the PathErr message.

If a different technology specific OAM configuration TLV is included than what was specified in the OAM Type an error must be generated with error value: "OAM Type Mismatch" in the PathErr message.

There is a hierarchy in between the OAM configuration elements. If this hierarchy is broken the error value: "Configuration Error" must be used in the PathErr message.

If a node does not support a specific OAM Function it must use the error value: "Unsupported OAM Function" in the PathErr message.

4.5. Considerations on Point-to-Multipoint OAM Configuration

RSVP-TE extensions for the establishment of point-to-multipoint (P2MP) LSPs are specified in [RFC4875]. A P2MP LSP is comprised of multiple source-to-leaf (S2L) sub-LSPs. These S2L sub-LSPs are set up between the ingress and egress LSRs and are appropriately combined by the branch LSRs using RSVP semantics to result in a P2MP TE LSP. One Path message may signal one or multiple S2L sub-LSPs for a single P2MP LSP. Hence the S2L sub-LSPs belonging to a P2MP LSP can be signaled using one Path message or split across multiple Path messages.

P2MP OAM mechanisms are very specific to the data plane technology, hence in this document we only highlight basic operations for P2MP OAM configuration. We consider only the configuration of the root to leaves OAM flows of P2MP LSPs and as such aspects of any return path are outside the scope of our discussions. We also limit our consideration to cases where all leaves must successfully establish OAM entities in order a P2MP OAM is successfully established. In any case, the discussion set forth below provides only guidelines for P2MP OAM configuration, details SHOULD be specified in technology specific documents.

The root node may select if it uses a single Path message or multiple Path messages to setup the whole P2MP tree. In the case when multiple Path messages are used the root node is responsible also to keep the OAM Configuration information consistent in each of the sent Path messages, i.e., the same information MUST be included in all Path messages used to construct the multicast tree. Each branching node will propagate the Path message downstream on each of the branches, when constructing a Path message the OAM Configuration information MUST be copied unchanged from the received Path message, including the related ADMIN_STATUS bits, LSP Attribute Flags and the OAM Configuration TLV. The latter two also imply that the

LSP_ATTRIBUTES and LSP_REQUIRED_ATTRIBUTES Object MUST be copied for the upstream Path message to the subsequent downstream Path messages.

Leaves MUST create and configure OAM sink functions according to the parameters received in the Path message, for P2MP OAM configuration there is no possibility for parameter negotiation on a per leaf basis. This is due to the fact that the only OAM source function, residing in the root of the tree, can only operate with a single configuration which must be obeyed by all leaves. If a leaf cannot accept the OAM parameters it MUST use the RRO Attributes sub-object [RFC5420] to notify the root of the problem. In particular, if the OAM configuration was successful the leaf would set the "OAM MEP entities desired" flag in the RRO Attributes sub-object in the Resv message, while, if due to any reason, OAM entities could not be established the Resv message should be sent with the "OAM MEP entities desired" bit cleared in the RRO Attributes sub-object. Branching nodes should collect and merge the received RROs according to the procedures described in [RFC4875]. This way, the root when receiving the Resv message (or messages if multiple Path messages were used to setup the tree) will have a clear information on which of the leaves could the OAM sink functions be established. If all leaves established OAM entities successfully, the root can enable the OAM message flow. On the other hand, if at some leaves the establishment was unsuccessful additional actions will be needed before the OAM message flow can be enabled. Such action could be to setup two independent P2MP LSPs. One with OAM Configuration information towards leaves which successfully setup OAM. This can be done by pruning the leaves which failed to setup OAM of the previously signalled P2MP LSP. The other P2MP LSP could be constructed for leaves without OAM entities. What exact procedures are needed are technology specific and SHOULD be described in technology specific documents.

5. IANA Considerations

Two bits ("OAM Alarms Enabled" (O) and "OAM Flows Enabled" (M)) needs to be allocated in the ADMIN_STATUS Object.

Two bits ("OAM MEP entities desired" and "OAM MIP entities desired") needs to be allocated in the LSP Attributes Flags Registry.

This document specifies one new TLV to be carried in the LSP_ATTRIBUTES and LSP_REQUIRED_ATTRIBUTES objects in Path and Resv messages: OAM Configuration TLV.

One new Error Code: "OAM Problem" and a set of new values: "MEP establishment not supported", "MIP establishment not supported", "Unsupported OAM Type", "Configuration Error" and "Unsupported OAM Function" needs to be assigned.

IANA is requested to open a new registry: "RSVP-TE OAM Configuration Registry" that maintains the "OAM Type" code points, an associated sub-TLV space, and the allocations of "OAM Function Flags" within the OAM Configuration TLV.

6. Security Considerations

The signaling of OAM related parameters and the automatic establishment of OAM entities based on RSVP-TE messages adds a new aspect to the security considerations discussed in [RFC3473]. In particular, a network element could be overloaded, if a remote attacker could request liveliness monitoring, with frequent periodic messages, for a high number of LSPs, targeting a single network element. Such an attack can efficiently be prevented when mechanisms for message integrity and node authentication are deployed. Since the OAM configuration extensions rely on the hop-by-hop exchange of existing RSVP-TE messages, procedures specified for RSVP message security in [RFC2747] can be used to mitigate possible attacks.

For a more comprehensive discussion on GMPLS security please see the Security Framework for MPLS and GMPLS Networks [RFC5920]. Cryptography can be used to protect against many attacks described in [RFC5920].

7. Acknowledgements

The authors would like to thank Francesco Fondelli, Adrian Farrel, Loa Andersson, Eric Gray and Dimitri Papadimitriou for their useful comments.

8. References

8.1. Normative References

- [RFC3471] "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC5420] "Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.

8.2. Informative References

- [IEEE-CFM] "IEEE 802.1ag, Draft Standard for Connectivity Fault Management", work in progress.
- [IEEE-PBBTE] "IEEE 802.1Qay Draft Standard for Provider Backbone Bridging Traffic Engineering", work in progress.
- [RFC2747] "RSVP Cryptographic Authentication", RFC 2747, January 2000.
- [RFC3469] "Framework for Multi-Protocol Label Switching (MPLS)-based Recovery", RFC 3469, February 2003.
- [RFC4377] "Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks", RFC 4377, February 2006.
- [RFC4379] "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.
- [RFC4875] "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.
- [RFC5654] "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.
- [RFC5828] "GMPLS Ethernet Label Switching Architecture and Framework", RFC 5828, March 2010.

- [RFC5860] "Requirements for OAM in MPLS Transport Networks",
RFC 5860, May 2010.
- [RFC5920] "Security Framework for MPLS and GMPLS Networks",
RFC 5920, July 2010.
- [RFC5921] "A Framework for MPLS in Transport Networks", RFC 5921,
July 2010.
- [RFC6060] "Generalized Multiprotocol Label Switching (GMPLS) Control
of Ethernet Provider Backbone Traffic Engineering
(PBB-TE)", RFC 6060.

Authors' Addresses

Attila Takacs
Ericsson
Konyves Kalman krt. 11.
Budapest, 1097
Hungary

Email: attila.takacs@ericsson.com

Don Fedyk
Alcatel-Lucent
Groton, MA 01450
USA

Email: donald.fedyk@alcatel-lucent.com

Jia He
Huawei

Email: hejia@huawei.com

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: October 20, 2011

S. Belotti, Ed.
P. Grandi
Alcatel-Lucent
D. Ceccarelli, Ed.
D. Caviglia
Ericsson
F. Zhang
D. Li
Huawei Technologies
April 18, 2011

Information model for G.709 Optical Transport Networks (OTN)
draft-ietf-ccamp-otn-g709-info-model-00

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 October 20, 2011.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
1.1. Terminology	4
2. OSPF-TE requirements overview	4
3. RSVP-TE requirements overview	5
4. G.709 Digital Layer Info Model for Routing and Signaling	5
4.1. Tributary Slot type	8
4.2. Tributary Port Number	9
4.3. Signal type	9
4.4. Bit rate and tolerance	11
4.5. Unreserved Resources	11
4.6. Maximum LSP Bandwidth	11
4.7. Distinction between terminating and switching capability	12
4.8. Priority Support	14
4.9. Multi-stage multiplexing	14
4.10. Generalized Label	14
5. Security Considerations	15
6. IANA Considerations	15
7. Contributors	15
8. Acknowledgements	15
9. References	15
9.1. Normative References	15
9.2. Informative References	16
Authors' Addresses	17

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 either at the same line rate or different line rates (e.g. 40G and 10G). Bundling links at different rates makes the control plane more scalable and permits better networking flexibility.

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 negoziation
- Support for constraint signaling

4. G.709 Digital Layer Info Model for Routing and Signaling

The digital OTN layered structure is comprised of digital path layer networks (ODU) and digital section layer networks (OTU). An OTU section layer supports one ODU path layer as client and provides monitoring capability for the OCh. An ODU path layer may transport a heterogeneous assembly of ODU clients. Some types of ODUs (i.e., ODU1, ODU2, ODU3, ODU4) may assume either a client or server role within the context of a particular networking domain. ITU-T G.872 recommendation provides two tables defining mapping and multiplexing capabilities of OTNs, which are reproduced below.

ODU client	OTU server
ODU 0	-
ODU 1	OTU 1
ODU 2	OTU 2
ODU 2e	-
ODU 3	OTU 3
ODU 4	OTU 4
ODU flex	-

Figure 1: OTN mapping capability

ODU client	ODU server
1,25 Gbps client	ODU 0
-	
2,5 Gbps client	ODU 1
ODU 0	
10 Gbps client	ODU 2
ODU0,ODU1,ODUflex	
10,3125 Gbps client	ODU 2e
-	
40 Gbps client	ODU 3
ODU0,ODU1,ODU2,ODU2e,ODUflex	
100 Gbps client	ODU 4
ODU0,ODU1,ODU2,ODU2e,ODU3,ODUflex	
CBR clients from greater than 2.5 Gbit/s to 100 Gbit/s: or GFP-F mapped packet clients from 1.25 Gbit/s to 100 Gbit/s.	ODUflex
-	

Figure 2: OTN multiplexing capability

How an 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:

- o Case A) one or more wavelength sub-networks connected by optical links or
- o Case B) one or more ODU links (having sub-lambda and/or lambda bandwidth granularity)
- o Case C) a mix of ODU links and wavelength sub-networks.

This document considers the TE information needed for ODU path computation and parameters needed to be signaled for LSP setup.

The following sections list and analyze each type of data that needs to be advertised and signaled in order to support path computation and LSP setup.

4.1. Tributary Slot type

ITU-T recommendations define two types of TS but each link can only support a single type at a given time. The rules to be followed when selecting the TS to be used are:

- If both ends of a link can support both 2.5Gbps TS and 1.25Gbps TS, then the link will work with 1.25Gbps TS.
- If one end can support the 1.25Gbps TS, and another end the 2.5Gbps TS, the link will work with 2.5Gbps TS.

In case the bandwidth accounting is provided in number of TSs, the type of TS is needed to perform correct routing operations. Currently such information is not provided by the routing protocol and not taken into account during LSP signaling.

The tributary slot type information is one of the parameters needed to correctly configure physical interfaces, therefore it has to be signaled via RSVP-TE. This allows the end points of the FA know which TS should be used.

[editor note]: SG15 ITU-T G.798 describes the so called PT=21-to-

PT=20 interworking process that explains how two equipments 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 neighbour. Therefore, in order to let the NE change TS granularity accordingly to the neighbour requirements, the AUTOpayloadtype needs to be configured and the HO ODU source can be either not provisioned (i.e. TS not allocated) or configured following a specific mapping depending of the type of LO ODU carried. In this case the process of auto-negotiation 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.

4.2. 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.

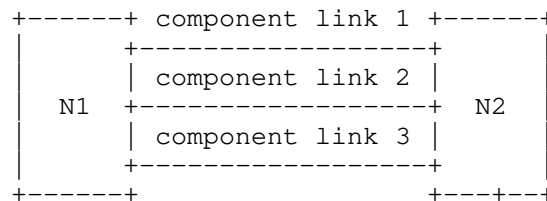


Figure 3: Concurrent path computation

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.

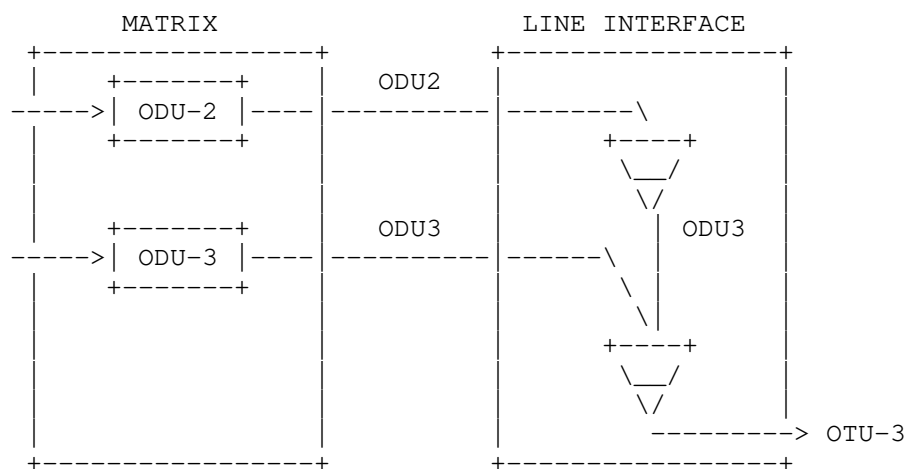


Figure 4: Switching and Terminating capabilities

The figure in the example shows a line interface able to:

- Multiplex an ODU2 coming from the switching matrix into and ODU3 and map it into an OTU3
- Map an ODU3 coming from the switching matrix into an OTU3

In this case the interface bandwidth advertised is ODU2 with switching capability and ODU3 with both switching and terminating capabilities.

This piece of information needs to be advertised together with the related unreserved bandwidth and signal type. As a consequence signaling must have the possibility to setup an LSP allowing the

local selection of resources consistent with the limitations considered during the path computation.

In figures 6 and 7 there are two examples of the need of termination/switching capability differentiation. In both examples all nodes are supposed to support single-stage capability. The figure 6 addresses a scenario in which a failure on link B-C forces node A to calculate another ODU2 LSP path carrying ODU0 service along the nodes B-E-D. Being D a single stage capable node, it is able to extract ODU0 service only from ODU2 interface. Node A has to know that from E to D exists an available OTU2 link from which node D can extract the ODU0 service. This information is required in order to avoid that the OTU3 link is considered in the path computation.

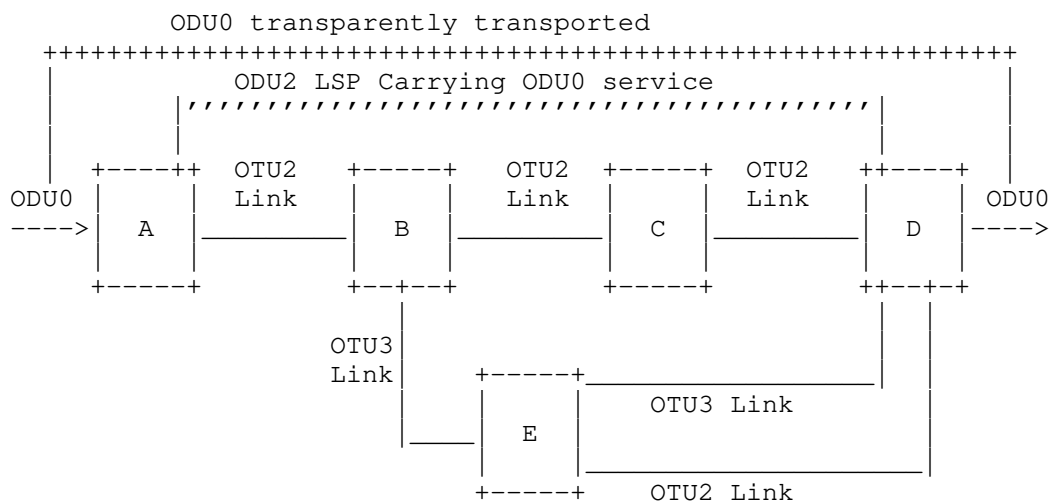


Figure 5: Switching and Terminating capabilities - Example 1

Figure 7 addresses the scenario in which the restoration of the ODU2 LSP (ABCD) is required. The two bundled component links between B and E could be used, but the ODU2 over the OTU2 component link can only be terminated and not switched. This implies that it cannot be used to restore the ODU2 LSP (ABCD). However such ODU2 unreserved bandwidth must be advertised since it can be used for a different ODU2 LSP terminating on E, e.g. (FBE). Node A has to know that the ODU2 capability on the OTU2 link can only be terminated and that the restoration of (ABCD) can only be performed using the ODU2 bandwidth available on the OTU3 link.

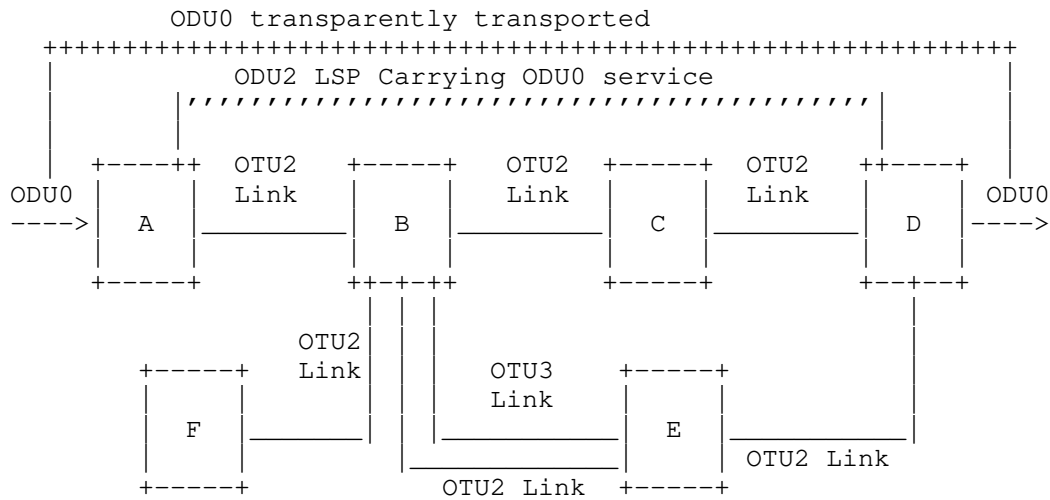


Figure 6: Switching and Terminating capabilities - Example 2

4.8. Priority Support

The IETF foresees that up to eight priorities must be supported and that all of them have to be advertised independently on the number of priorities supported by the implementation. Considering that the advertisement of all the different supported signal types will originate large LSAs, it is advised to advertise only the information related to the really supported priorities.

4.9. Multi-stage multiplexing

With reference to the [OTN-FWK], introduction of multi-stage multiplexing implies the advertisement of cascaded adaptation capabilities together with the matrix access constraints. The structure defined by IETF for the advertisement of adaptation capabilities is ISCD/IACD as in [RFC4202] and [RFC5339]. Modifications to ISCD/IACD, if needed, have to be addressed in the related encoding documents.

4.10. Generalized Label

The ODUk label format defined in [RFC4328] could be updated to support new signal types defined in [G709-V3] but would hardly be further enhanced to support possible new signal types.

Furthermore such label format may have scalability issues due to the

high number of labels needed when signaling large LSPs. For example, when an ODU3 is mapped into an ODU4 with 1.25G tributary slots, it would require the utilization of thirty-one labels ($31 \times 4 \times 8 = 992$ bits) to be allocated while an ODUflex into an ODU4 may need up to eighty labels ($80 \times 4 \times 8 = 2560$ bits).

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

5. Security Considerations

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

[HIER-BIS]

K.Shiomoto, A.Farrel, "Procedure for Dynamically Signaled Hierarchical Label Switched Paths", work in progress draft-ietf-lsp-hierarchy-bis-08, February 2010.

[OTN-OSPF]

D.Ceccarelli, D.Caviglia, F.Zhang, D.Li, Y.Xu, P.Grandi, S.Belotti, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolutive G.709 OTN Networks", work in progress draft-cceccarelli-ccamp-gmpls-ospf-g709-03, August 2010.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.
- [RFC5339] Le Roux, JL. and D. Papadimitriou, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 5339, September 2008.

9.2. Informative References

- [G.709-v1] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), February 2001.
- [G.709-v2] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1), March 2003.
- [G.709-v3] ITU-T, "Rec G.709, version 3", approved by ITU-T on December 2009.

[G.872-am2]

ITU-T, "Amendment 2 of G.872 Architecture of optical transport networks for consent", consented by ITU-T on June 2010.

[OTN-FWK]

F.Zhang, D.Li, H.Li, S.Belotti, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", work in progress draft-ietf-ccamp-gmpls-g709-framework-00, April 2010.

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

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: January 6, 2012

A. Malis, ed.
Verizon Communications
A. Lindem, ed.
Ericsson
July 5, 2011

Updates to ASON Routing for OSPFv2 Protocols (RFC 5787bis)
draft-ietf-ccamp-rfc5787bis-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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2011 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

The ITU-T has defined an architecture and requirements for operating an Automatically Switched Optical Network (ASON).

The Generalized Multiprotocol Label Switching (GMPLS) protocol suite is designed to provide a control plane for a range of network technologies including optical networks such as time division multiplexing (TDM) networks including SONET/SDH and Optical Transport Networks (OTNs), and lambda switching optical networks.

The requirements for GMPLS routing to satisfy the requirements of ASON routing, and an evaluation of existing GMPLS routing protocols are provided in other documents. This document defines extensions to the OSPFv2 Link State Routing Protocol to meet the requirements for routing in an ASON.

Note that this work is scoped to the requirements and evaluation expressed in RFC 4258 and RFC 4652 and the ITU-T Recommendations current when those documents were written. Future extensions or revisions of this work may be necessary if the ITU-T Recommendations are revised or if new requirements are introduced into a revision of RFC 4258.

Table of Contents

1. Introduction	4
1.1. Conventions Used in This Document	5
2. Routing Areas, OSPF Areas, and Protocol Instances	5
3. Terminology and Identification	6
4. Reachability	6
5. Link Attribute	7
5.1. Local Adaptation	7
5.2. Bandwidth Accounting	8
6. Routing Information Scope	8
6.1. Link Advertisement (Local and Remote TE Router ID Sub-TLV)	9
6.2. Reachability Advertisement (Local TE Router ID sub-TLV)	10
7. Routing Information Dissemination	11
7.1 Import/Export Rules	11
7.2 Loop Prevention	11
7.2.1 Inter-RA Export Upward/Downward Sub-TLVs	12
7.2.2 Inter-RA Export Upward/Downward Sub-TLV Processing	13
8. OSPFv2 Scalability	13
9. Security Considerations	14
10. IANA Considerations	14
10.1. Sub-TLVs of the Link TLV	14

10.2. Sub-TLVs of the Node Attribute TLV	15
10.3. Sub-TLVs of the Router Address TLV	15
11. Management Considerations	16
11.1. Routing Area (RA) Isolation	16
11.2 Routing Area (RA) Topology/Configuration Changes	16
12. Comparison to Requirements in RFC 4258	16
13. References	22
13.1. Normative References	22
13.2. Informative References	23
14. Acknowledgements	24
Appendix A. ASON Terminology	25
Appendix B. ASON Routing Terminology	26
Authors' Addresses	27

1. Introduction

The Generalized Multiprotocol Label Switching (GMPLS) [RFC3945] protocol suite is designed to provide a control plane for a range of network technologies including optical networks such as time division multiplexing (TDM) networks including SONET/SDH and Optical Transport Networks (OTNs), and lambda switching optical networks.

The ITU-T defines the architecture of the Automatically Switched Optical Network (ASON) in [G.8080].

[RFC4258] describes the routing requirements for the GMPLS suite of routing protocols to support the capabilities and functionality of ASON control planes identified in [G.7715] and in [G.7715.1].

[RFC4652] evaluates the IETF Link State routing protocols against the requirements identified in [RFC4258]. Section 7.1 of [RFC4652] summarizes the capabilities to be provided by OSPFv2 [RFC2328] in support of ASON routing. This document describes the OSPFv2 specifics for ASON routing.

Multi-layer transport networks are constructed from multiple networks of different technologies operating in a client-server relationship. The ASON routing model includes the definition of routing levels that provide scaling and confidentiality benefits. In multi-level routing, domains called routing areas (RAs) are arranged in a hierarchical relationship. Note that as described in [RFC4652], there is no implied relationship between multi-layer transport networks and multi-level routing. The multi-level routing mechanisms described in this document work for both single-layer and multi-layer networks.

Implementations may support a hierarchical routing topology (multi-level) for multiple transport network layers and/or a hierarchical routing topology for a single transport network layer.

This document describes the processing of the generic (technology-independent) link attributes that are defined in [RFC3630], [RFC4202], and [RFC4203] and that are extended in this document. As described in Section 5.2, technology-specific traffic engineering attributes and their processing may be defined in other documents that complement this document.

Note that this work is scoped to the requirements and evaluation expressed in [RFC4258] and [RFC4652] and the ITU-T Recommendations current when those documents were written. Future extensions of revisions of this work may be necessary if the ITU-T Recommendations are revised or if new requirements are introduced into a revision of

[RFC4258].

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].

The reader is assumed to be familiar with the terminology and requirements developed in [RFC4258] and the evaluation outcomes described in [RFC4652].

General ASON terminology is provided in Appendix A. ASON routing terminology is described in Appendix B.

2. Routing Areas, OSPF Areas, and Protocol Instances

An ASON routing area (RA) represents a partition of the data plane, and its identifier is used within the control plane as the representation of this partition.

RAs are hierarchically contained: a higher-level (parent) RA contains lower-level (child) RAs that in turn MAY also contain RAs, etc. Thus, RAs contain RAs that recursively define successive hierarchical RA levels. Routing information may be exchanged between levels of the RA hierarchy, i.e., Level N+1 and N, where Level N represents the RAs contained by Level N+1. The links connecting RAs may be viewed as external links (inter-RA links), and the links representing connectivity within an RA may be viewed as internal links (intra-RA links). The external links to an RA at one level of the hierarchy may be internal links in the parent RA. Intra-RA links of a child RA MAY be hidden from the parent RA's view. [RFC4258]

An ASON RA can be mapped to an OSPF area, but the hierarchy of ASON RA levels does not map to the hierarchy of OSPF areas. Instead, successive hierarchical levels of RAs MUST be represented by separate instances of the protocol. Thus, inter-level routing information exchange (as described in Section 7) involves the export and import of routing information between protocol instances.

An ASON RA may therefore be identified by the combination of its OSPF instance identifier and its OSPF area identifier. With proper and careful network-wide configuration, this can be achieved using just the OSPF area identifier, and this process is RECOMMENDED in this document. These concepts are discussed in Section 7.

A key ASON requirement is the support of multiple transport planes or layers. Each transport node has associated topology (links and

reachability) which is used for ASON routing.

3. Terminology and Identification

This section describes the mapping of key ASON entities to OSPF entities. Appendix A contains a complete glossary of ASON routing terminology.

There are three categories of identifiers used for ASON routing (G7715.1): transport plane names, control plane identifiers for components, and SCN addresses. This section discusses the mapping between ASON routing identifiers and corresponding identifiers defined for GMPLS routing, and how these support the physical (or logical) separation of transport plane entities and control plane components. GMPLS supports this separation of identifiers and planes.

In the context of OSPF Traffic Engineering (TE), an ASON transport node corresponds to a unique OSPF TE node. An OSPF TE node is uniquely identified by the TE Router Address TLV [RFC3630]. In this document, this TE Router Address is referred to as the TE Router ID, which is in the ASON transport plane name space. The TE Router ID should not be confused with the OSPF Router ID which uniquely identifies an OSPF router within an OSPF routing domain [RFC2328] and is in a name space for control plane components.

Note: The Router Address top-level TLV definition, processing, and usage are unchanged from [RFC3630]. This TLV specifies a stable OSPF TE node IP address, i.e., the IP address is always reachable when there is IP connectivity to the associated OSPF TE node.

ASON defines a Routing Controller (RC) as an entity that handles (abstract) information needed for routing and the routing information exchange with peering RCs by operating on the Routing Database (RDB). ASON defines a Protocol Controller (PC) as an entity that handles protocol-specific message exchanges according to the reference point over which the information is exchanged (e.g., E-NNI, I-NNI), and internal exchanges with the Routing Controller (RC) [RFC4258]. In this document, an OSPF router advertising ASON TE topology information will perform both the functions of the RC and PC. Each OSPF router is uniquely identified by its OSPF Router ID [RFC2328].

4. Reachability

Reachability in ASON refers to the set of endpoints reachable in the transport plane by a node or the reachable endpoints of a level N. Reachable entities are identified in the transport plane name space

(ASON SNPP name space). In order to advertise blocks of reachable address prefixes, a summarization mechanism is introduced that is based on the techniques described in [RFC5786]. For ASON reachability advertisement, blocks of reachable address prefixes are advertised together with the associated data plane node. The data plane node is identified in the control plane by its TE Router ID, as discussed in section 6.

In order to support ASON reachability advertisement, the Node Attribute TLV defined in [RFC5786] is used to advertise the combination of a TE Router ID and its set of associated reachable address prefixes. The Node Attribute TLV can contain the following sub-TLVs:

- TE Router ID sub-TLV: Length: 4; Defined in Section 6.2
- Node IPv4 Local Address sub-TLV: Length: variable; [RFC5786]
- Node IPv6 Local Address sub-TLV: Length: variable; [RFC5786]

A router may support multiple transport nodes as discussed in section 6, and, as a result, may be required to advertise reachability (ASON SNPPs) separately for each transport node. As a consequence, it MUST be possible for the router to originate more than one TE LSA containing the Node Attribute TLV when used for ASON reachability advertisement.

Hence, the Node Attribute TLV [RFC5786] advertisement rules must be relaxed for ASON. A Node Attribute TLV MAY appear in more than one TE LSA originated by the RC when the RC is advertising reachability information for a different transport node identified by the Local TE Router Sub-TLV (refer to section 6.1).

5. Link Attribute

With the exception of local adaptation (described below), the mapping of link attributes and characteristics to OSPF TE Link TLV Sub-TLVs is unchanged [RFC4652]. OSPF TE Link TLV Sub-TLVs are described in [RFC3630] and [RFC4203]. Advertisement of this information SHOULD be supported on a per-layer basis, i.e., one TE LSA per unique switching capability and bandwidth granularity combination.

5.1. Local Adaptation

Local adaptation is defined as a TE link attribute (i.e., sub-TLV) that describes the cross/inter-layer relationships.

The Interface Switching Capability Descriptor (ISCD) TE Attribute [RFC4202] identifies the ability of the TE link to support cross-connection to another link within the same layer. When advertising

link adaptation, it also identifies the ability to use a locally terminated connection that belongs to one layer as a data link for another layer (adaptation capability). However, the information associated with the ability to terminate connections within that layer (referred to as the termination capability) is advertised with the adaptation capability.

For instance, a link between two optical cross-connects will contain at least one ISCD attribute describing the Lambda Switching Capable (LSC) switching capability. Conversely, a link between an optical cross-connect and an IP/MPLS Label Switching Router (LSR) will contain at least two ISCD attributes, one for the description of the LSC termination capability and one for the Packet Switching Capable (PSC) adaptation capability.

In OSPFv2, the Interface Switching Capability Descriptor (ISCD) is a sub-TLV (type 15) of the top-level Link TLV (type 2) [RFC4203]. The adaptation and termination capabilities are advertised using two separate ISCD sub-TLVs within the same top-level Link TLV.

An interface MAY have more than one ISCD sub-TLV, [RFC4202] and [RFC4203]. Hence, the corresponding advertisements should not result in any compatibility issues.

5.2. Bandwidth Accounting

GMPLS routing defines an Interface Switching Capability Descriptor (ISCD) that provides, among other things, the available (maximum/minimum) bandwidth per priority available for Label Switched Path (LSPs). One or more ISCD sub-TLVs can be associated with an interface, [RFC4202] and [RFC4203]. This information, combined with the Unreserved Bandwidth Link TLV sub-TLV [RFC3630], provides the basis for bandwidth accounting.

In the ASON context, additional information may be included when the representation and information in the other advertised fields are not sufficient for a specific technology, e.g., SDH. The definition of technology-specific information elements is beyond the scope of this document. Some technologies will not require additional information beyond what is already defined in [RFC3630], [RFC4202], and [RFC4203].

6. Routing Information Scope

For ASON routing, the control plane component routing adjacency topology (i.e., the associated Protocol Controller (PC) connectivity) and the transport topology are NOT assumed to be congruent [RFC4258]. Hence, a single OSPF router (i.e., the PC) MUST be able to advertise

on behalf of multiple transport layer nodes. The OSPF routers are identified by OSPF Router ID and the transport nodes are identified by TE Router ID.

The Router Address TLV [RFC3630] is used to advertise the TE Router ID associated with the advertising Routing Controller. TE Router IDs for additional transport nodes are advertised through specification of the Local TE Router Identifier in the Local and Remote TE Router TE sub-TLV and the Local TE Router Identifier sub-TLV described in the sections below. These Local TE Router Identifiers are typically used as the local endpoints for TE Label Switched Paths (LSPs) terminating on the associated transport node.

It MAY be feasible for multiple OSPF Routers to advertise TE information for the same transport node. However, this is not considered a required use case and is not discussed further.

6.1. Link Advertisement (Local and Remote TE Router ID Sub-TLV)

An OSPF router advertising on behalf of multiple transport nodes will require additional information to distinguish the link endpoints amongst the subsumed transport nodes. In order to unambiguously specify the transport topology, the local and remote transport nodes MUST be identified by TE router ID.

For this purpose, a new sub-TLV of the OSPFv2 TE LSA top-level Link TLV is introduced that defines the Local and Remote TE Router ID.

The Type field of the Local and Remote TE Router ID sub-TLV is assigned the value 26 (see Section 10). The Length field takes the value 8. The Value field of this sub-TLV contains 4 octets of the Local TE Router Identifier followed by 4 octets of the Remote TE Router Identifier. The value of the Local and Remote TE Router Identifier SHOULD NOT be set to 0.

The format of the Local and Remote TE Router ID sub-TLV is:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Type (26)             |               Length (8)         |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Local TE Router Identifier                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Remote TE Router Identifier                             |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This sub-TLV MUST be included as a sub-TLV of the top-level Link TLV

if the OSPF router is advertising on behalf of one or more transport nodes having TE Router IDs different from the TE Router ID advertised in the Router Address TLV. Therefore, it MUST be included if the OSPF router is advertising on behalf of multiple transport nodes.

Note: The Link ID sub-TLV identifies the other end of the link (i.e., Router ID of the neighbor for point-to-point links) [RFC3630]. When the Local and Remote TE Router ID Sub-TLV is present, it MUST be used to identify local and remote transport node endpoints for the link and the Link-ID sub-TLV MUST be ignored. The Local and Remote ID sub-TLV, if specified, MUST only be specified once.

6.2. Reachability Advertisement (Local TE Router ID sub-TLV)

When an OSPF router is advertising on behalf of multiple transport nodes, the routing protocol MUST be able to associate the advertised reachability information with the correct transport node.

For this purpose, a new sub-TLV of the OSPFv2 TE LSA top-level Node Attribute TLV is introduced. This TLV associates the local prefixes (see above) to a given transport node identified by TE Router ID.

The Type field of the Local TE Router ID sub-TLV is assigned the value 5 (see Section 10). The Length field takes the value 4. The Value field of this sub-TLV contains the Local TE Router Identifier [RFC3630] encoded over 4 octets.

The format of the Local TE Router ID sub-TLV is:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
|      Type (5)                       |      Length (4)                   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
|      Local TE Router Identifier      |                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

This sub-TLV MUST be included as a sub-TLV of the top-level Node Attribute TLV if the OSPF router is advertising on behalf of one or more transport nodes having TE Router IDs different from the TE Router ID advertised in the Router Address TLV. Therefore, it MUST be included if the OSPF router is advertising on behalf of multiple transport nodes.

7. Routing Information Dissemination

An ASON routing area (RA) represents a partition of the data plane, and its identifier is used within the control plane as the representation of this partition. An RA may contain smaller RAs inter-connected by links. ASON RA levels do not map directly to OSPF areas. Rather, hierarchical levels of RAs are represented by separate OSPF protocol instances.

Routing controllers (RCs) supporting multiple RAs disseminate information downward and upward in this ASON hierarchy. The vertical routing information dissemination mechanisms described in this section do not introduce or imply hierarchical OSPF areas. RCs supporting RAs at multiple levels are structured as separate OSPF instances with routing information exchange between levels described by import and export rules between these instances. The functionality described herein does not pertain to OSPF areas or OSPF Area Border Router (ABR) functionality.

7.1 Import/Export Rules

RCs supporting RAs disseminate information upward and downward in the hierarchy by importing/exporting routing information as TE LSAs. TE LSAs are area-scoped opaque LSAs with opaque type 1 [RFC3630]. The information that MAY be exchanged between adjacent levels includes the Router Address, Link, and Node Attribute top-level TLVs.

The imported/exported routing information content MAY be transformed, e.g., filtered or aggregated, as long as the resulting routing information is consistent. In particular, when more than one RC is bound to adjacent levels and both are allowed to import/export routing information, it is expected that these transformations are performed in a consistent manner. Definition of these policy-based mechanisms is outside the scope of this document.

In practice, and in order to avoid scalability and processing overhead, routing information imported/exported downward/upward in the hierarchy is expected to include reachability information (see Section 4) and, upon strict policy control, link topology information.

7.2 Loop Prevention

When more than one RC is bound to an adjacent level of the ASON hierarchy, and is configured to export routing information upward or downward, a specific mechanism is required to avoid looping of routing information. Looping is the re-advertisement of routing information into an RA that had previously advertised that routing

information upward or downward into an upper or lower level RA in the ASON hierarchy. For example, without loop prevention mechanisms, this could happen when the RC advertising routing information downward in the hierarchy is not the same one that advertises routing information upward in the hierarchy.

7.2.1 Inter-RA Export Upward/Downward Sub-TLVs

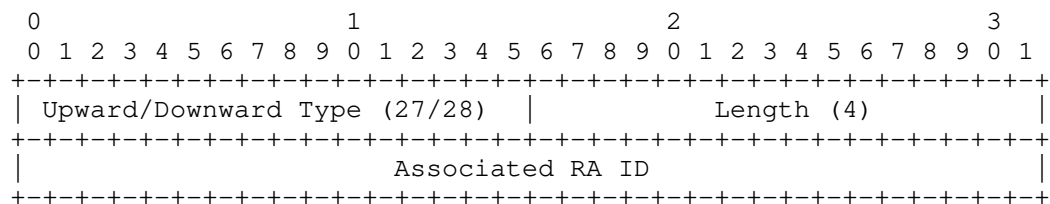
The Inter-RA Export Sub-TLVs can be used to prevent the re-advertisement of OSPF TE routing information into an RA which previously advertised that information. The type value 28 (see Section 10) will indicate that the associated routing information has been exported downward. The type value 27 (see Section 10) will indicate that the associated routing information has been exported upward. While it is not required for routing information exported downward, both Sub-TLVs will include the Routing Area (RA) ID from the which the routing information was exported. This RA is not necessarily the RA originating the routing information but RA from which the information was immediately exported.

These additional Sub-TLVs MAY be included in TE LSAs that include any of the following top-level TLVs:

- Router Address top-level TLV
- Link top-level TLV
- Node Attribute top-level TLV

The Type field of the Inter-RA Export Upward and Inter-RA Export Downward sub-TLVs are respectively assigned the values 27 and 28 (see Section 10). The Length field in these Sub-TLVs takes the value 4. The Value field in these sub-TLVs contains the associated RA ID. The RA ID value must be a unique identifier for the RA within the ASON routing domain.

The format of the Inter-RA Export Upward and Inter-RA Export Downward Sub-TLVs is graphically depicted below:



7.2.2 Inter-RA Export Upward/Downward Sub-TLV Processing

TE LSAs MAY be imported or exported downward or upward in the ASON routing hierarchy. The direction and advertising RA ID are advertised in an Inter-RA Export Upward/Downward Sub-TLV. They MUST be retained and advertised in the receiving RA with the associated routing information.

When exporting routing information upward in the ASON routing hierarchy, any information received from a level above, i.e., tagged with an Inter-RA Export Downward Sub-TLV, MUST NOT be exported upward. Since an RA at level N is contained by a single RA at level N+1, this is the only checking that is necessary and the associated RA ID is used solely for informational purposes.

When exporting routing information downward in the ASON routing hierarchy, any information received from a level below, i.e., tagged with an Inter-RA Export Upward Sub-TLV MUST NOT be exported downward if the target RA ID matches the RA ID associated with the routing information. This additional checking is required for routing information exported downward since a single RA at level N+1 may contain multiple RAs at level N in the ASON routing hierarchy. In order words, routing information MUST NOT be exported downward into the RA from which it was received.

8. OSPFv2 Scalability

The extensions described herein are only applicable to ASON routing domains and it is not expected that the attendant reachability (see Section 4) and link information will ever be mixed with global or local IP routing information. If there were ever a requirement for a given RC to participate in both domains, separate OSPFv2 instances would be utilized. However, in a multi-level ASON hierarchy, the potential volume of information could be quite large and the recommendations in this section SHOULD be followed by RCs implementing this specification.

- Routing information exchange upward/downward in the hierarchy between adjacent RAs SHOULD, by default, be limited to reachability information. In addition, several transformations such as prefix aggregation are RECOMMENDED to reduce the amount of information imported/exported by a given RC when such transformations will not impact consistency.
- Routing information exchange upward/downward in the ASON hierarchy involving TE attributes MUST be under strict policy control. Pacing and min/max thresholds for triggered updates are strongly RECOMMENDED.

- The number of routing levels MUST be maintained under strict policy control.

9. Security Considerations

This document specifies the contents and processing of OSPFv2 TE LSAs [RFC3630] and [RFC4202]. The TE LSA extensions defined in this document are not used for SPF computation, and have no direct effect on IP routing. Additionally, ASON routing domains are delimited by the usual administrative domain boundaries.

Any mechanisms used for securing the exchange of normal OSPF LSAs can be applied equally to all TE LSAs used in the ASON context. Authentication of OSPFv2 LSA exchanges (such as OSPF cryptographic authentication [RFC2328] and [RFC5709]) can be used to secure against passive attacks and provide significant protection against active attacks. [RFC5709] defines a mechanism for authenticating OSPFv2 packets by making use of the HMAC algorithm in conjunction with the SHA family of cryptographic hash functions.

If a stronger authentication were believed to be required, then the use of a full digital signature [RFC2154] would be an approach that should be seriously considered. Use of full digital signatures would enable precise authentication of the OSPF router originating each OSPF link-state advertisement, and thereby provide much stronger integrity protection for the OSPF routing domain.

10. IANA Considerations

This document is classified as Standards Track. It defines new sub-TLVs for inclusion in OSPF TE LSAs. According to the assignment policies for the registries of code points for these sub-TLVs, values must be assigned by IANA [RFC3630].

This draft requests early allocation of IANA code points in accordance with [RFC4020]. [NOTE TO RFC Editor: this paragraph and the RFC 4020 reference can be removed during RFC editing].

The following subsections summarize the required sub-TLVs.

10.1. Sub-TLVs of the Link TLV

This document defines the following sub-TLVs of the Link TLV advertised in the OSPF TE LSA:

- Local and Remote TE Router ID sub-TLV (26)
- Inter-RA Export Upward sub-TLV (27)
- Inter-RA Export Downward sub-TLV (28)

Codepoints for these Sub-TLVs should be allocated from the "Types for sub-TLVs of TE Link TLV (Value 2)" registry standards action range (0 - 32767) [RFC3630].

Note that the same values for the Inter-RA Export Upward sub-TLV and the Inter-RA Export Downward Sub-TLV MUST be used when they appear in the Link TLV, Node Attribute TLV, and Router Address TLV.

10.2. Sub-TLVs of the Node Attribute TLV

This document defines the following sub-TLVs of the Node Attribute TLV advertised in the OSPF TE LSA:

- Local TE Router ID sub-TLV (5)
- Inter-RA Export Upward sub-TLV (27)
- Inter-RA Export Downward sub-TLV (28)

Codepoints for these Sub-TLVs should be assigned from the "Types for sub-TLVs of TE Node Attribute TLV (Value 5)" registry standards action range (0 - 32767) [RFC5786].

Note that the same values for the Inter-RA Export Upward sub-TLV and the Inter-RA Export Downward Sub-TLV MUST be used when they appear in the Link TLV, Node Attribute TLV, and Router Address TLV.

10.3. Sub-TLVs of the Router Address TLV

The Router Address TLV is advertised in the OSPF TE LSA [RFC3630]. Since this TLV currently has no Sub-TLVs defined, a "Types for sub-TLVs of Router Address TLV (Value 1)" registry must be defined.

The registry guidelines for the assignment of types for sub-TLVs of the Router Address TLV are as follows:

- o Types in the range 0-32767 are to be assigned via Standards Action.
- o Types in the range 32768-32777 are for experimental use; these will not be registered with IANA, and MUST NOT be mentioned by RFCs.
- o Types in the range 32778-65535 are not to be assigned at this time. Before any assignments can be made in this range, there MUST be a Standards Track RFC that specifies IANA Considerations that covers the range being assigned.

This document defines the following sub-TLVs for inclusion in the Router Address TLV:

- Inter-RA Export Upward sub-TLV (27)
- Inter-RA Export Downward sub-TLV (28)

Codepoints for these Sub-TLVs should be allocated from the "Types for sub-TLVs of Router Address TLV (Value 1)" registry standards action range (0 - 32767).

Note that the same values for the Inter-RA Export Upward sub-TLV and the Inter-RA Export Downward Sub-TLV MUST be used when they appear in the Link TLV, Node Attribute TLV, and Router Address TLV.

11. Management Considerations

11.1. Routing Area (RA) Isolation

If the RA Identifier is mapped to the OSPF Area ID as recommended in section 2.0, OSPF [RFC2328] implicitly provides isolation. On any intra-RA link, packets will only be accepted if the area-id in the OSPF packet header matches the area ID for the OSPF interface on which the packet was received. Hence, RCs will only establish adjacencies and exchange reachability information (see Section 4.0) with RCs in the same RC. Other mechanisms for RA isolation are beyond the scope of this document.

11.2 Routing Area (RA) Topology/Configuration Changes

The GMPLS Routing for ASON requirements [RFC4258] dictate that the routing protocol MUST support reconfiguration and SHOULD support architectural evolution. OSPF [RFC2328] includes support for the dynamic introduction or removal of ASON reachability information through the flooding and purging of OSPF opaque LSAs [RFC5250]. Also, when an RA is partitioned or an RC fails, stale LSAs SHOULD NOT be used unless the advertising RC is reachable. The configuration of OSPF RAs and the policies governing the redistribution of ASON reachability information between RAs are implementation issues outside of the OSPF routing protocol and beyond the scope of this document.

12. Comparison to Requirements in RFC 4258

The following table shows how this draft complies with the requirements in [RFC4258]. The first column contains a requirements number (1-30) and the relevant section in RFC 4258. The second column describes the requirement, the third column discusses the compliance to that requirement, and the fourth column lists the relevant section in draft, and/or another RFC that already satisfies the requirement.

RFC 4258 Section (Req. Number)	RFC 4258 Requirement	Compliance	Reference
3.0 (1)	The failure of an RC, or the failure of communications between RCs, and the subsequent recovery from the failure condition MUST NOT disrupt call in progress.	Implied by separation of transport and control plane.	Not an attribute of routing protocol.
3.1 (2)	Multiple Hierarchical Level of ASON Routing Areas (RAs) .	Yes	Sections 2 and 3
3.1 (3)	Prior to establishing communications, RCs MUST verify that they are bound to the same parent RA.	Yes when RA maps to OSPF Area ID.	Section 11.1
3.1 (4)	The RC ID MUST be unique within its containing RA.	Yes	RFC 2328 and Section 3.
3.1 (5)	Each RA within a carrier's network SHALL be uniquely identifiable. RA IDs MAY be associated with a transport plane name space, whereas RC IDs are associated with a control plane name space.	Yes - although uniqueness is the operator's responsibility.	Sections 2, 3, and 11.1
3.2 (6)	Hierarchical Routing Information Dissemination	Yes	Section 7
3.2 (7)	Routing Information exchanged between levels N and N+1 via separate instances and import/export.	Yes	Section 7.1

3.2 (8)	Routing Information exchanged between levels N and N+1 via external link (inter-RA links).	No - Not described.	
3.2 (9)	Routing information exchange MUST include reachability information and MAY include, upon policy decision, node and link topology.	Yes	Sections 4, 6, 6.1, 6.2, and 8
3.2 (10)	There SHOULD NOT be any dependencies on the different routing protocols used within an RA or in different RAs.	Yes - separate instances.	Sections 2 and 3
3.2 (11)	The routing protocol SHALL differentiate the routing information originated at a given-level RA from derived routing information (received from external RAs), even when this information is forwarded by another RC at the same level.	Yes	Section 7.2
3.2 (12)	The routing protocol MUST provide a mechanism to prevent information propagated from a Level N+1 RA's RC into the Level N RA's RC from being re-introduced into the Level N+1 RA's RC.	Yes	Section 7.2
3.2 (13)	The routing protocol MUST provide a mechanism to prevent information propagated from a Level N-1 RA's RC into the Level N RA's RC from being re-introduced into the Level N-1 RA's RC.	Yes	Section 7.2

3.2 (14)	Instance of a Level N routing function and an instance of a Level N+1 routing function in the same system.	Yes	Sections 2, 3, and 7
3.2 (15)	The Level N routing function is on a separate system the Level N+1 routing function.	Not described but possible.	N/A
3.3 (16)	The RC MUST support static (i.e., operator assisted) and MAY support automated configuration of the information describing its relationship to its parent and its child within the hierarchical structure (including RA ID and RC ID).	Yes - automation requirement is ambiguous.	Sections 2 and 3. Config is product specific.
3.3 (17)	The RC MUST support static (i.e., operator assisted) and MAY support automated configuration of the information describing its associated adjacencies to other RCs within an RA.	Yes - when OSPF area maps to RA discovery is automatic.	RFC 2328 and Section 11.1
3.3 (18)	The routing protocol SHOULD support all the types of RC adjacencies described in Section 9 of [G.7715]. The latter includes congruent topology (with distributed RC) and hubbed topology (e.g., note that the latter does not automatically imply a designated RC).	Yes	RFC 2328

3.4 (19)	The routing protocol SHOULD be capable of supporting architectural evolution in terms of the number of hierarchical levels of RAs, as well as the aggregation and segmentation of RAs.	Yes	RFC 2328, RFC 5250, and Section 11.2.
3.5.2 (20)	Advertisements MAY contain the following common set of information regardless of whether they are link or node related: <ul style="list-style-type: none"> - RA ID of the RA to which the advertisement is bounded - RC ID of the entity generating the advertisement - Information to uniquely identify advertisements - Information to determine whether an advertisement has been updated - Information to indicate when an advertisement has been derived from a different level RA 	Yes Yes Yes No - Must compare to old Yes	Section 7.2.1 RFC 2328 RFC 2328, RFC 5250 Section 7.2.1
3.5.3 (21)	The Node Attributes Node ID and Reachability must be advertised. It MAY be advertised as a set of associated external (e.g., User Network Interface (UNI)) address/address prefixes or a set of associated SNPP link IDs/SNPP ID prefixes, the selection of which MUST be consistent within the applicable scope.	Yes - Prefixes only for reachability	RFC 5786, Section 4 and 6

3.5.4 (22)	The Link Attributes Local SNPP link ID, Remote SNPP link ID, and layer specific characteristics must be advertised.	Yes	Section 6.1
3.5.4 (23)	Link Signaling Attributes other than Local Adaptation (Signal Type, Link Weight, Resource Class, Local Connection Types, Link Capacity, Link Availability, Diversity Support)	Yes	Section 5, RFC 4652 - Section 5.3.1
3.5.4 (24)	Link Signaling Local Adaptation	Yes	Section 5.1
5 (25)	The routing adjacency topology (i.e., the associated PC connectivity topology) and the transport network topology SHALL NOT be assumed to be congruent.	Yes	Section 2, 3, and 6
5 (26)	The routing topology SHALL support multiple links between nodes and RAs.	Yes	RFC 2328, RFC 3630
5 (27)	The routing protocol SHALL converge such that the distributed RDBs become synchronized after a period of time.	Yes	RFC 2328, RFC 5250
5 (28)	Self-consistent information at the receiving level resulting from any transformation (filter, summarize, etc.) and forwarding of information from one RC to RC(s) at different levels when multiple RCs are bound to a single RA.	Yes - However, this is not a routing protocol function.	Section 7.1

5 (29)	In order to support operator-assisted changes in the containment relationships of RAs, the routing protocol SHALL support evolution in terms of the number of hierarchical levels of RAs. For example: support of non-disruptive operations such as adding and removing RAs at the top/bottom of the hierarchy, adding or removing a hierarchical level of RAs in or from the middle of the hierarchy, as well as aggregation and segmentation of RAs.	Partial - OSPF supports the purging of stale advertisements and origination of new. The non-disruptive behavior is implementation specific.	RFC 2328 and RFC 5250
5 (30)	A collection of links and nodes such as a subnetwork or RA MUST be able to represent itself to the wider network as a single logical entity with only its external links visible to the topology database.	Yes - Within an RA it must be consistent.	Sections 4 and 6

13. References

13.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol

Label Switching (GMPLS)", RFC 4202, October 2005.

- [RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.
- [RFC5786] Aggarwal, R. and K. Kompella, "Advertising a Router's Local Addresses in OSPF TE Extensions", RFC 5786, March 2010.

13.2. Informative References

- [RFC2154] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", RFC 2154, June 1997.
- [RFC4020] Kompella, K. and A. Zinin, "Early IANA Allocation of Standards Track Code Points", BCP 100, RFC 4020, February 2005.
- [RFC4258] Brungard, D., Ed., "Requirements for Generalized Multi-Protocol Label Switching (GMPLS) Routing for the Automatically Switched Optical Network (ASON)", RFC 4258, November 2005.
- [RFC4652] Papadimitriou, D., Ed., Ong, L., Sadler, J., Shew, S., and D. Ward, "Evaluation of Existing Routing Protocols against Automatic Switched Optical Network (ASON) Routing Requirements", RFC 4652, October 2006.
- [RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, October 2009.

For information on the availability of ITU Documents, please see <http://www.itu.int>.

- [G.7715] ITU-T Rec. G.7715/Y.1306, "Architecture and Requirements for the Automatically Switched Optical Network (ASON)", June 2002.
- [G.7715.1] ITU-T Rec. G.7715.1/Y.1706.1, "ASON Routing Architecture and Requirements for Link State Protocols", February 2004.

- [G.805] ITU-T Rec. G.805, "Generic Functional Architecture of Transport Networks)", March 2000.
- [G.8080] ITU-T Rec. G.8080/Y.1304, "Architecture for the Automatically Switched Optical Network (ASON)," June 2006 (and Amendments 1 (March 2008) and 2 (Sept. 2010)).

14. Acknowledgements

The editors would like to thank Dimitri Papadimitriou for editing RFC 5787, from which this document is derived, and Lyndon Ong, Remi Theillaud, Stephen Shew, Jonathan Sadler, Deborah Brungard, and Lou Berger for their useful comments and suggestions.

Appendix A. ASON Terminology

This document makes use of the following terms:

Administrative domain: (See Recommendation [G.805].) For the purposes of [G7715.1], an administrative domain represents the extent of resources that 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.

Control plane: performs the call control and connection control functions. Through signaling, the control plane sets up and releases connections, and may restore a connection in case of a failure.

(Control) Domain: represents a collection of (control) entities that are grouped for a particular purpose. The control plane is subdivided into domains matching administrative domains. Within an administrative domain, further subdivisions of the control plane are recursively applied. A routing control domain is an abstract entity that hides the details of the RC distribution.

External NNI (E-NNI): interfaces located between protocol controllers between control domains.

Internal NNI (I-NNI): interfaces located between protocol controllers within control domains.

Link: (See Recommendation G.805.) A "topological component" that describes a fixed relationship between a "subnetwork" or "access group" and another "subnetwork" or "access group". Links are not limited to being provided by a single server trail.

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, fault, configuration, accounting, and security management.

Management domain: (See Recommendation G.805.) A management domain defines a collection of managed objects that are grouped to meet organizational requirements according to geography, technology, policy, or other structure, and for a number of functional areas such as configuration, security, (FCAPS), for the purpose of providing control in a consistent manner. Management domains can be disjoint, contained, or overlapping. As such, the resources

within an administrative domain can be distributed into several possible overlapping management domains. The same resource can therefore belong to several management domains simultaneously, but a management domain shall not cross the border of an administrative domain.

Subnetwork Point (SNP): The SNP is a control plane abstraction that represents an actual or potential transport plane resource. SNPs (in different subnetwork partitions) may represent the same transport resource. A one-to-one correspondence should not be assumed.

Subnetwork Point Pool (SNPP): A set of SNPs that are grouped together for the purposes of routing.

Termination Connection Point (TCP): A TCP represents the output of a Trail Termination function or the input to a Trail Termination Sink function.

Transport plane: provides bidirectional or unidirectional transfer of user information, from one location to another. It can also provide transfer of some control and network management information. The transport plane is layered; it is equivalent to the Transport Network defined in Recommendation G.805.

User Network Interface (UNI): interfaces are located between protocol controllers between a user and a control domain. Note: There is no routing function associated with a UNI reference point.

Appendix B. ASON Routing Terminology

This document makes use of the following terms:

Routing Area (RA): an RA represents a partition of the data plane, and its identifier is used within the control plane as the representation of this partition. Per [G.8080], an RA is defined by a set of sub-networks, the links that interconnect them, and the interfaces representing the ends of the links exiting that RA. An RA may contain smaller RAs inter-connected by links. The limit of subdivision results in an RA that contains two sub-networks interconnected by a single link.

Routing Database (RDB): a repository for the local topology, network topology, reachability, and other routing information that is updated as part of the routing information exchange and may additionally contain information that is configured. The RDB may contain routing information for more than one routing area (RA).

Routing Components: ASON routing architecture functions. These functions can be classified as protocol independent (Link Resource Manager or LRM, Routing Controller or RC) or protocol specific (Protocol Controller or PC).

Routing Controller (RC): handles (abstract) information needed for routing and the routing information exchange with peering RCs by operating on the RDB. The RC has access to a view of the RDB. The RC is protocol independent.

Note: Since the RDB may contain routing information pertaining to multiple RAs (and possibly to multiple layer networks), the RCs accessing the RDB may share the routing information.

Link Resource Manager (LRM): supplies all the relevant component and TE link information to the RC. It informs the RC about any state changes of the link resources it controls.

Protocol Controller (PC): handles protocol-specific message exchanges according to the reference point over which the information is exchanged (e.g., E-NNI, I-NNI), and internal exchanges with the RC. The PC function is protocol dependent.

Authors' Addresses

Andrew G. Malis
Verizon Communications
117 West St.
Waltham MA 02451 USA

EMail: andrew.g.malis@verizon.com

Acee Lindem
Ericsson
102 Carric Bend Court
Cary, NC 27519

EMail: acee.lindem@ericsson.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2012

A. Takacs
B. Gero
Ericsson
H. Long
Huawei
July 11, 2011

GMPLS RSVP-TE Extensions for Ethernet OAM Configuration
draft-ietf-ccamp-rsvp-te-eth-oam-ext-06

Abstract

The GMPLS controlled Ethernet Label Switching (GELS) work extended GMPLS RSVP-TE to support the establishment of Ethernet LSPs. IEEE Ethernet Connectivity Fault Management (CFM) specifies an adjunct OAM flow to check connectivity in Ethernet networks. CFM can be also used with Ethernet LSPs for fault detection and triggering recovery mechanisms. The ITU-T Y.1731 specification builds on CFM and specifies additional OAM mechanisms, including Performance Monitoring, for Ethernet networks. This document specifies extensions of GMPLS RSVP-TE to support the setup of the associated Ethernet OAM (CFM and Y.1731) entities defining Ethernet technology specific TLV based on [OAM-CONF-FWK].

Requirements Language

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

Status of this Memo

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

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

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

This Internet-Draft will expire on January 7, 2012.

Copyright Notice

Copyright (c) 2011 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. Background	4
2. Overview of Ethernet OAM operation	5
3. GMPLS RSVP-TE Extensions	7
3.1. Operation overview	7
3.2. OAM Configuration TLV	9
3.3. Ethernet OAM Configuration TLV	9
3.3.1. MD Name Sub-TLV	10
3.3.2. Short MA Name Sub-TLV	11
3.3.3. MEP ID Sub-TLV	12
3.3.4. Continuity Check (CC) Sub-TLV	12
3.4. Pro-active Performance Monitoring	13
3.5. Ethernet OAM configuration errors	13
4. IANA Considerations	15
5. Security Considerations	16
6. Acknowledgements	17
7. References	19
7.1. Normative References	19
7.2. Informative References	19
Authors' Addresses	20

1. Background

Provider Backbone Bridging - Traffic Engineering (PBB-TE) [IEEE-PBBTE] decouples the Ethernet data and control planes by explicitly supporting external control/management mechanisms to configure static filtering entries in bridges and create explicitly routed Ethernet connections. In addition PBB-TE defines mechanisms for protection switching of bidirectional Ethernet connections. Ethernet Connectivity Fault Management (CFM) defines an adjunct connectivity monitoring OAM flow to check the liveness of Ethernet networks [IEEE-CFM], including the monitoring of explicitly-routed Ethernet connections.

In IETF the GMPLS controlled Ethernet Label Switching (GELS) work extended the GMPLS control plane to support the establishment of explicitly routed Ethernet connections [RFC5828][RFC6060]. We refer to GMPLS established Ethernet connections as Ethernet LSPs. GELS enables the application of MPLS-TE and GMPLS provisioning and recovery features in Ethernet networks.

2. Overview of Ethernet OAM operation

For the purposes of this document, we only discuss Ethernet OAM [IEEE-CFM] aspects that are relevant for the connectivity monitoring of Ethernet LSPs.

PBB-TE [IEEE-PBBTE] defines point-to-point Ethernet Switched Paths (ESPs) as a provisioned traffic engineered unidirectional connectivity, identified by the 3-tuple [ESP-MAC DA, ESP-MAC SA, ESP-VID] where the ESP-MAC DA is the destination address of the ESP, the ESP-MAC SA is the source address of the ESP, and the ESP-VID is a VLAN identifier allocated for explicitly routed connections. To form a bidirectional PBB-TE connection two co-routed point-to-point ESPs are combined. The combined ESPs must have the same ESP-MAC addresses but may have different ESP-VIDs.

Note that although it would be possible to use GMPLS to setup a single unidirectional ESP, the Ethernet OAM mechanisms are only full functional when bidirectional connections are established with co-routed ESPs. Hence, we focus on bidirectional point-to-point PBB-TE connections only.

At both ends of the bidirectional point-to-point PBB-TE connection one Maintenance Endpoint (MEP) is configured. The MEPs monitoring a PBB-TE connection must be configured with the same Maintenance Domain Level (MD Level) and Maintenance Association Identifier (MAID). Each MEP has a unique identifier, the MEP ID. Besides these identifiers a MEP monitoring a PBB-TE connection must be provisioned with the 3-tuples [ESP-MAC DA, ESP-MAC SA, ESP-VID] of the two ESPs.

In the case of point-to-point VLAN connections, the connection is identified with a single VLAN forwarding traffic in both directions or with two VLANs each forwarding traffic in a single direction. Hence instead of the 3-tuples of the PBB-TE case MEPs must be provisioned with the proper VLAN information, otherwise the same MD Level, MAID, MEP ID configuration is required in this case as well.

MEPs exchange Connectivity Check Messages (CCMs) periodically with fixed intervals. Eight distinct intervals are defined in [IEEE-CFM]:

#	CCM Interval (CCI)	3 bit encoding
0	Reserved	000
1	3 1/3 ms	001
2	10 ms	010
3	100 ms	011
4	1 s	100
5	10 s	101
6	1 min	110
7	10 min	111

Table 1: CCM Interval encoding

If 3 consecutive CCM messages are not received by one of the MEPs it declares a connectivity failure and signals the failure in subsequent CCM messages, by setting the Remote Defect Indicator (RDI) bit, to the remote MEP. If a MEP receives a CCM message with RDI set it immediately declares failure. The detection of a failure may trigger protection switching mechanisms or may be signaled to a management system. However, what happens once a failure is detected is out of the scope of this document.

At each transit node Maintenance Intermediate Points (MIPs) can be established to help failure localization by supporting link trace and loop back functions. MIPs need to be provisioned with a subset of MEP identification parameters described above.

3. GMPLS RSVP-TE Extensions

3.1. Operation overview

To simplify the configuration of connectivity monitoring, when an Ethernet LSP is signaled the associated MEPs should be automatically established. To monitor an Ethernet LSP a set of parameters must be provided to setup a Maintenance Association and related MEPs. Optionally, MIPs may be created at the transit nodes of the Ethernet LSP. The LSP Attributes Flags: "OAM MEP entities desired" and "OAM MIP entities desired", described in [OAM-CONF-FWK] are used to signal that the respective OAM entities must be established. Subsequently, an OAM Configuration TLV is added to the LSP_ATTRIBUTES Object specifying that Ethernet OAM is to be setup for the LSP. The below detailed Ethernet OAM specific information is carried in the new Ethernet OAM Configuration sub-TLV.

- o A unique MAID must be allocated for the PBB-TE connection and both MEPs must be configured with the same information. The MAID consists of an optional Maintenance Domain Name (MD Name) and a mandatory Short Maintenance Association Name (Short MA Name). Various formatting rules for these names have been defined by [IEEE-CFM]. Since this information is also carried in all CCM messages, the combined length of the Names is limited to 44 bytes. How these parameters are determined is out of scope of this document.
- o Each MEP must be provisioned with a MEP ID. The MEP ID uniquely identifies a given MEP within a Maintenance Association. That is, the combination of MAID and MEP ID must uniquely identify a MEP. How the value of the MEP ID is determined is out of scope of this document.
- o The Maintenance Domain Level (MD Level) allows hierarchical separation of monitoring entities. [IEEE-CFM] allows differentiation of 8 levels. How the value of the MD Level is determined is out of scope of this document. Note that most probably for all Ethernet LSPs a single (default) MD Level will be used within a network domain.
- o The desired CCM Interval must be specified by the management system based on service requirements or operator policy. The same CCM Interval must be set in each of the MEPs monitoring a given Ethernet LSP. How the value of the CCM Interval is determined is out of scope of this document.
- o The desired CCM priority to be set by MEPs for the CCM frames can be specified. The same CCM priority must be set in each of the

MEPs monitoring a given Ethernet LSP. How CCM priority is determined is out of scope of this document. Note that the highest priority is used as the default CCM priority.

- o MEPs must be aware of their own and the reachability parameters of the remote MEP. In the case of bidirectional point-to-point PBB-TE connections this requires that the 3-tuples [ESP-MAC A, ESP-MAC B, ESP-VID1] and [ESP-MAC B, ESP-MAC A, ESP-VID2] are configured in each MEP, where the ESP-MAC A is the same as the local MEP's MAC address and ESP-MAC B is the same as remote MEP's MAC address. The GMPLS Ethernet Label for forwarding, as defined in [RFC6060], consists of the ESP-MAC DA and ESP-VID. Hence the necessary reachability parameters for the MEPs can be obtained from Ethernet Labels (i.e., carried in the "downstream" and upstream labels). In the case of point-to-point VLAN connections, MEPs need to be provisioned with the VLAN identifiers, which can be derived similarly from the Ethernet Label.

Assuming the procedures described in [RFC6060] for bidirectional PBB-TE Ethernet LSP establishment the MEP configuration should be as follows. When the RSVP-TE signaling is initiated for the bidirectional Ethernet LSP the local node generates a Path message and:

- o Allocates an Upstream Label from its MAC address (ESP-MAC A) and locally selected VID (ESP-VID1), which will be used to receive traffic;
- o Inserts the OAM Configuration TLV with OAM Type set to Ethernet OAM in the LSP_ATTRIBUTES object;
- o Adds the OAM Function Flags sub-TLV in the OAM Configuration TLV and sets the OAM function flags as needed;
- o Adds an Ethernet OAM Configuration sub-TLV in the OAM Configuration TLV that specifies the CCM Interval and MD Level;
- o Adds an MD Name Sub-TLV (optional) and a Short MA Name Sub-TLV to the Ethernet OAM Configuration TLV, that will unambiguously identify a Maintenance Association for this specific PBB-TE connection. Note that values for these parameters may be derived from the GMPLS LSP identification parameters;
- o Adds a MEP ID Sub-TLV to the Ethernet OAM Configuration TLV. It selects two distinct integer values to identify the local and remote MEPs within the Maintenance Association created for monitoring of the point-to-point PBB-TE connection.

Once the remote node receives the Path message it can use the UPSTREAM_LABEL to extract the reachability information of the initiator. Then it allocates a Label by selecting the MAC address (ESP-MAC B) and VID (ESP-VID2) it would like to use to receive traffic. These parameters determine the reachability information of the local MEP. That is, the 3-tuples [ESP-MAC A, ESP-MAC B, ESP-VID1] and [ESP-MAC B, ESP-MAC A, ESP-VID2] are derived from the Ethernet Labels. In addition the information received in the Ethernet OAM Configuration TLV is used to configure the local MEP.

Once the Resv message successfully arrives to the initiator it can extract the remote side's reachability information from the Label Object whereby this node has also obtained all the information needed to establish its local MEP.

3.2. OAM Configuration TLV

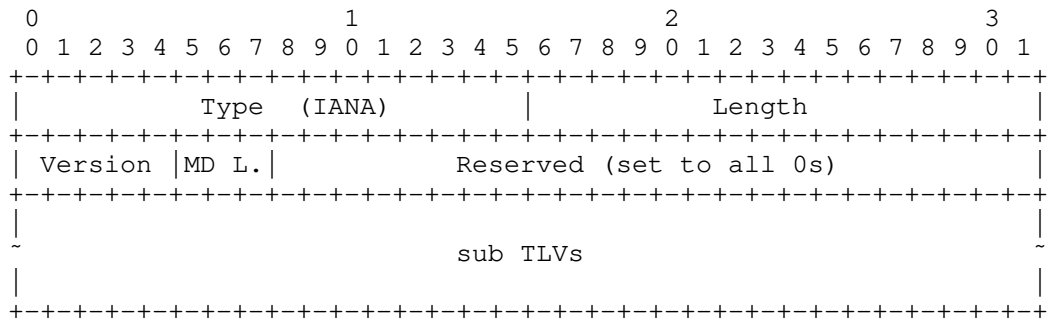
This TLV is specified in [OAM-CONF-FWK] and is used to select which OAM technology/method should be used for the LSP. In this document a new OAM Type: Ethernet OAM is defined.

OAM Type	Description
-----	-----
0	Reserved
1	Ethernet OAM
2-256	Reserved

The receiving node when the Ethernet OAM Type is requested should look for the corresponding technology specific Ethernet OAM Configuration TLV.

3.3. Ethernet OAM Configuration TLV

The Ethernet OAM Configuration TLV (depicted below) is defined for Ethernet OAM specific configuration parameters. The Ethernet OAM Configuration TLV is carried within the OAM Configuration TLV in the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES object in Path messages. This new TLV accommodates generic Ethernet OAM information and carries sub-TLVs.



Type: indicates a new type: the Ethernet OAM Configuration TLV. IANA is requested to assign a value from the "OAM Type sub-TLV" space in the "RSVP-TE OAM Configuration Registry".

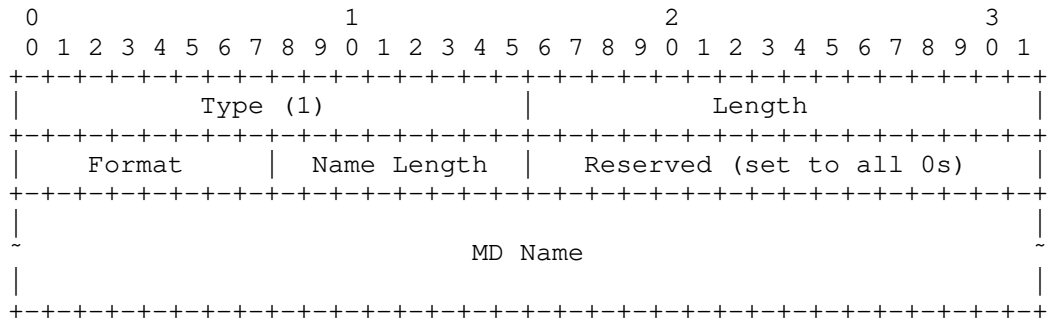
Length: indicates the total length including sub-TLVs.

Version: identifies the CFM protocol version according to [IEEE-CFM]. If a node does not support a specific CFM version an error must be generated: "OAM Problem/Unsupported OAM Version"

MD L. (MD Level): indicates the desired MD Level. The values are according to [IEEE-CFM]. If a node does not support a specific MD Level an error must be generated: "OAM Problem/Unsupported OAM Level".

3.3.1. MD Name Sub-TLV

The optional MD Name sub-TLV is depicted below.



Type: 1, MD Name Sub-TLV.

Length: indicates the total length of the TLV including padding.

Format: according to [IEEE-CFM].

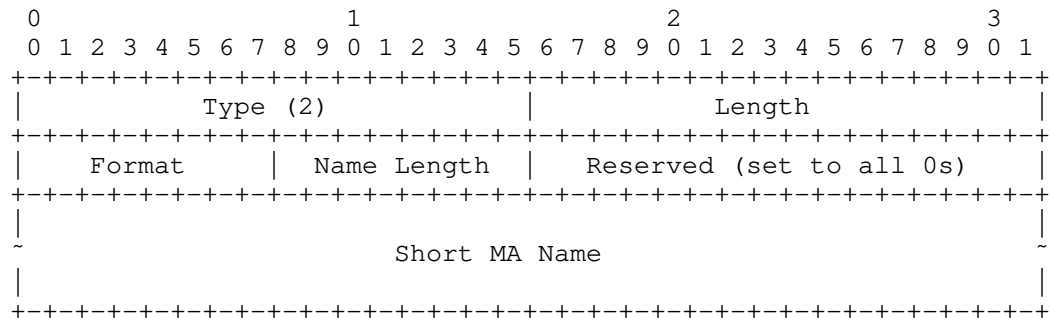
Name Length: the length of the MD Name field in bytes. This is necessary to allow non 4 byte padded MD Name lengths.

MD Name: variable length field, formatted according to the format specified in the Format field.

If an undefined Format is specified an error must be generated: "OAM Problem/Unknown MD Name Format". Also the combined length of MD Name and Short MA Name must be less or equal to 44bytes, if this is violated an error must be generated: "OAM Problem/Name Length Problem". Note that it is allowed to have no MD Name, as such the MD Name sub-TLV is optional. In this case the MA Name must uniquely identify a Maintenance Association.

3.3.2. Short MA Name Sub-TLV

The Short MA Name sub-TLV is depicted below.



Type: 2, Short MA Name Sub-TLV.

Length: indicates the total length of the TLV including padding.

Format: according to [IEEE-CFM].

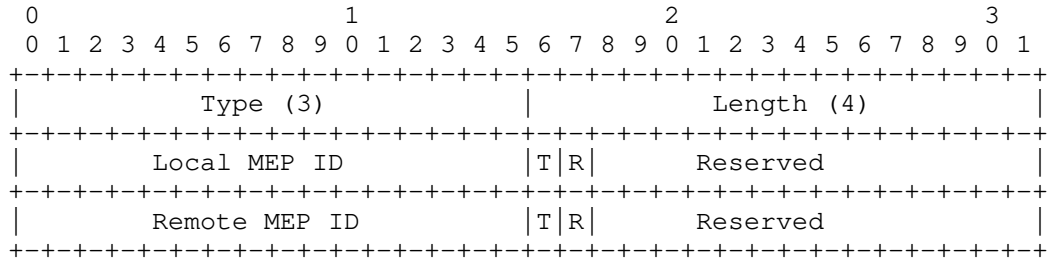
Name Length: the length of the MA Name field in bytes. This is necessary to allow non 4 byte padded MA Name lengths.

Short MA Name: variable length field formatted according to the format specified in the Format field.

If an undefined Format is specified an error must be generated: "OAM Problem/Unknown MA Name Format". Also the combined length of MD Name and Short MA Name must be less or equal to 44bytes, if this is violated an error must be generated: "OAM Problem/Name Length Problem". Note that it is allowed to have no MD Name, in this case the MA Name must uniquely identify a Maintenance Association.

3.3.3. MEP ID Sub-TLV

The MEP ID Sub-TLV is depicted below.



Type: 3, MEP ID Sub-TLV.

Length: indicates the total length of the TLV including padding.

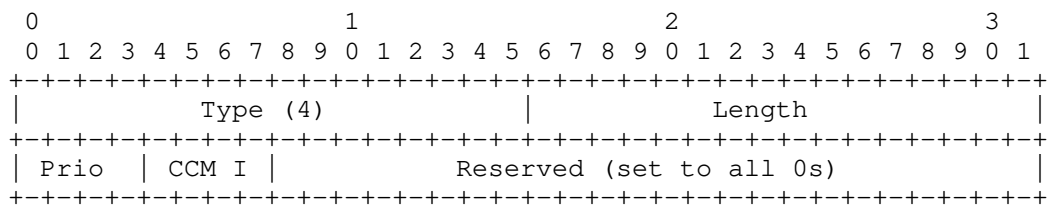
Local MEP ID: a 16 bit integer value in the range 1-8191 of the MEP ID on the initiator side.

Remote MEP ID: a 16 bit integer value in the range 1-8191 of the MEP ID to be set for the MEP established at the receiving side. This value is determined by the initiator node. This is possible, since a new MAID is assigned to each PBB-TE connection, and MEP IDs must be only unique within the scope of the MAID.

Two flags are defined Transmit (T) and Receive (R). When T is set the corresponding MEP must send OAM packets. When R is set the corresponding MEP must expect to receive OAM packets. These flags are used to configure the role of MEPs.

3.3.4. Continuity Check (CC) Sub-TLV

The Continuity Check (CC) sub-TLV is depicted below.



Type: 4, Continuity Check (CC) sub-TLV.

Prio: Indicates the priority to be set for CCM frames. In Ethernet 3 bits carried in VLAN TAGs identify priority information.

CCM I (CCM Interval): CCM Interval, according to the 3 bit encoding [IEEE-CFM] shown in Table 1. If a node does not support the requested CCM Interval an error must be generated: "OAM Problem/Unsupported CC Interval".

3.4. Pro-active Performance Monitoring

Ethernet OAM functions for Performance Monitoring (PM) allow measurements of different performance parameters including Frame Loss Ratio, Frame Delay and Frame Delay variation as defined in the ITU-T Y.1731 recommendation. Only a subset of PM functions are operated in a pro-active fashion to monitor the performance of the connection continuously. Pro-active PM supports Fault Management functions, by providing an indication of decreased service performance and as such may provide triggers to initiate recovery procedures.

While on demand PM functions are always initiated by management commands, for pro-active PM it may be desirable to utilize the control plane for configuration and activation together with Fault Management functions such as Continuity Check.

ITU-T Y.1731 defines dual-ended Loss Measurement as pro-active OAM for performance monitoring and as a PM function applicable to fault management. For dual-ended Loss Measurement each MEP piggy-backs transmitted and received frame counters on CC messages; to support and synchronize bidirectional Loss Measurements at the MEPs. Dual-ended Loss Measurement is invoked by setting the Performance Monitoring/Loss OAM Function Flag in the OAM Function Flags Sub-TLV [OAM-CONF-FWK]. Besides configuring the Continuity Check functionality, no additional configuration is required for this type of Loss Measurement.

3.5. Ethernet OAM configuration errors

In addition to error values specified in [OAM-CONF-FWK] this document defines the following values for the "OAM Problem" Error Code.

- o If a node does not support a specific CFM version an error must be generated: "OAM Problem/Unsupported OAM Version".
- o If a node does not support a specific MD Level an error must be generated: "OAM Problem/Unsupported OAM Level".
- o If an undefined MD name format is specified an error must be generated: "OAM Problem/Unknown MD Name Format".

- o If an undefined MA name format is specified an error must be generated: "OAM Problem/Unknown MA Name Format".
- o If the combined length of MD Name and Short MA Name must be less or equal to 44bytes, if this is violated an error must be generated: "OAM Problem/Name Length Problem".
- o If a node does not support the requested CCM Interval an error must be generated: "OAM Problem/Unsupported CC Interval".

4. IANA Considerations

This document specifies the Ethernet OAM Configuration sub-TLV to be carried in the OAM Configuration TLV in LSP_ATTRIBUTES and LSP_REQUIRED_ATTRIBUTES objects in Path messages.

IANA is requested to allocate the value 1 for Ethernet OAM from the OAM Type space in the "RSVP-TE OAM Configuration Registry" and allocate type 1 for the Ethernet OAM Configuration sub-TLV from the OAM Type sub-TLV space in the "RSVP-TE OAM Configuration Registry".

The following values need to be assigned under the Error Code: "OAM Problem": "Unsupported OAM Version", "Unsupported OAM Level", "Unknown MD Name Format", "Unknown MA Name Format", "Name Length Problem", "Unsupported CC Interval".

5. Security Considerations

This document does not introduce any additional security issue to those discussed in [OAM-CONF-FWK].

6. Acknowledgements

The authors would like to thank Francesco Fondelli, Adrian Farrel, Loa Andersson, Eric Gray and Dimitri Papadimitriou for their useful comments.

Contributors

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

Dinesh Mohan

7. References

7.1. Normative References

[OAM-CONF-FWK]

"OAM Configuration Framework for GMPLS RSVP-TE", Internet Draft, work in progress.

[RFC5828] "GMPLS Ethernet Label Switching Architecture and Framework", RFC 5828, March 2010.

[RFC6060] "Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)", RFC 6060.

7.2. Informative References

[IEEE-CFM]

"IEEE 802.1ag, Draft Standard for Connectivity Fault Management", work in progress.

[IEEE-PBBTE]

"IEEE 802.1Qay Draft Standard for Provider Backbone Bridging Traffic Engineering", work in progress.

Authors' Addresses

Attila Takacs
Ericsson
Laborc u. 1.
Budapest, 1037
Hungary

Email: attila.takacs@ericsson.com

Balazs Gero
Ericsson
Laborc u. 1.
Budapest, 1037
Hungary

Email: balazs.gero@ericsson.com

Hao Long
Huawei

Email: lonho@huawei.com

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2012

E. Bellagamba, Ed.
L. Andersson, Ed.
Ericsson
P. Skoldstrom, Ed.
Acreo AB
D. Ward
Juniper
A. Takacs
Ericsson
July 11, 2011

Configuration of Pro-Active Operations, Administration, and Maintenance
(OAM) Functions for MPLS-based Transport Networks using RSVP-TE
draft-ietf-ccamp-rsvp-te-mpls-tp-oam-ext-06

Abstract

This specification describes the configuration of pro-active MPLS-TP Operations, Administration, and Maintenance (OAM) Functions for a given LSP using a set of TLVs that are carried by the RSVP-TE protocol.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunication Union Telecommunication Standardization Sector (ITU-T) effort to include an MPLS Transport Profile within the IETF MPLS and PWE3 architectures to support the capabilities and functionalities of a packet transport network.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 12, 2012.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
1.1. Contributing Authors	4
1.2. Requirements Language	4
2. Overview of MPLS OAM for Transport Applications	4
3. Theory of Operations	5
3.1. MPLS OAM Configuration Operation Overview	5
3.1.1. Configuration of BFD sessions	5
3.1.2. Configuration of Performance Monitoring	6
3.1.3. Configuration of Measurements and FMS	6
3.2. OAM Configuration TLV	6
3.3. BFD Configuration sub-TLV	9
3.3.1. Local Discriminator sub-TLV	10
3.3.2. Negotiation Timer Parameters sub-TLV	11
3.3.3. BFD Authentication sub-TLV	12
3.4. Performance Monitoring sub-TLV	12
3.4.1. MPLS OAM PM Loss sub-TLV	13
3.4.2. MPLS OAM PM Delay sub-TLV	15
3.5. MPLS OAM FMS sub-TLV	16
4. IANA Considerations	17
5. BFD OAM configuration errors	17
6. Acknowledgements	17
7. Security Considerations	17
8. References	18
8.1. Normative References	18
8.2. Informative References	19
Authors' Addresses	20

1. Introduction

This document describes the configuration of pro-active MPLS-TP Operations, Administration, and Maintenance (OAM) Functions for a given LSP using TLVs carried by RSVP-TE [RFC3209]. In particular it specifies the mechanisms necessary to establish MPLS-TP OAM entities for monitoring and performing measurements on an LSP, as well as defining information elements and procedures to configure pro-active MPLS OAM functions. Initialization and control of on-demand MPLS OAM functions are expected to be carried out by directly accessing network nodes via a management interface; hence configuration and control of on-demand OAM functions are out-of-scope for this document.

The Transport Profile of MPLS must, by definition [RFC5654], be capable of operating without a control plane. Therefore there are three options for configuring MPLS-TP OAM, without a control plane by either using an NMS or LSP Ping, or with a control plane using GMPLS (specifically RSVP-TE) .

Pro-active MPLS OAM is performed by three different protocols, Bidirectional Forwarding Detection (BFD) [RFC5880] for Continuity Check/Connectivity Verification, the delay measurement protocol (DM) [MPLS-PM] for delay and delay variation (jitter) measurements, and the loss measurement protocol (LM) [MPLS-PM] for packet loss and throughput measurements. Additionally there is a number of Fault Management Signals that can be configured.

BFD is a protocol that provides low-overhead, fast detection of failures in the path between two forwarding engines, including the interfaces, data link(s), and to the extent possible the forwarding engines themselves. BFD can be used to track the liveness and detect data plane failures of MPLS-TP point-to-point and might also be extended to support point-to-multipoint connections.

The delay and loss measurements protocols [MPLS-PM] use a simple query/response model for performing bidirectional measurements that allows the originating node to measure packet loss and delay in both directions. By timestamping and/or writing current packet counters to the measurement packets at four times (Tx and Rx in both directions) current delays and packet losses can be calculated. By performing successive delay measurements the delay variation (jitter) can be calculated. Current throughput can be calculated from the packet loss measurements by dividing the number of packets sent/received with the time it took to perform the measurement, given by the timestamp in LM header. Combined with a packet generator the throughput measurement can be used to measure the maximum capacity of a particular LSP.

MPLS Transport Profile (MPLS-TP) describes a profile of MPLS that enables operational models typical in transport networks, while providing additional OAM, survivability and other maintenance functions not currently supported by MPLS. [RFC5860] defines the requirements for the OAM functionality of MPLS-TP.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunication Union Telecommunication Standardization Sector (ITU-T) effort to include an MPLS Transport Profile within the IETF MPLS and PWE3 architectures to support the capabilities and functionalities of a packet transport network.

1.1. Contributing Authors

This document is the result of a large team of authors and contributors. The following is a list of the co-authors:

John Drake

Benoit Tremblay

1.2. Requirements Language

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

2. Overview of MPLS OAM for Transport Applications

[MPLS-TP-OAM-FWK] describes how MPLS OAM mechanisms are operated to meet transport requirements outlined in [RFC5860].

[BFD-CCCV] specifies two BFD operation modes: 1) "CC mode", which uses periodic BFD message exchanges with symmetric timer settings, supporting Continuity Check, 2) "CV/CC mode" which sends unique maintenance entity identifiers in the periodic BFD messages supporting Connectivity Verification as well as Continuity Check.

[MPLS-PM] specifies mechanisms for performance monitoring of LSPs, in particular it specifies loss and delay measurement OAM functions.

[MPLS-FMS] specifies fault management signals with which a server LSP can notify client LSPs about various fault conditions to suppress alarms or to be used as triggers for actions in the client LSPs. The following signals are defined: Alarm Indication Signal (AIS), Link Down Indication (LDI) and Locked Report (LKR). To indicate client faults associated with the attachment circuits Client Signal Failure

Indication (CSF) can be used. CSF is described in [MPLS-TP-OAM-FWK] and in the context of this document is for further study.

[MPLS-TP-OAM-FWK] describes the mapping of fault conditions to consequent actions. Some of these mappings may be configured by the operator, depending on the application of the LSP. The following defects are identified: Loss Of Continuity (LOC), Misconnectivity, MEP Misconfiguration and Period Misconfiguration. Out of these defect conditions, the following consequent actions may be configurable: 1) whether or not the LOC defect should result in blocking the outgoing data traffic; 2) whether or not the "Period Misconfiguration defect" should result in a signal fail condition.

3. Theory of Operations

3.1. MPLS OAM Configuration Operation Overview

RSVP-TE, or alternatively LSP Ping [LSP-PING CONF], can be used to simply enable the different OAM functions, by setting the corresponding flags in the "OAM Functions TLV". Additionally one may include sub-TLVs for the different OAM functions in order to specify different parameters in detail.

3.1.1. Configuration of BFD sessions

For this specification, BFD MUST be run in either one of the two modes:

- Asynchronous mode, where both sides should be in active mode.
- Unidirectional mode

In the simplest scenario LSP Ping, or alternatively RSVP-TE [RSVP-TE CONF], is used only to bootstrap a BFD session for an LSP, without any timer negotiation.

Timer negotiation can be performed either in subsequent BFD control messages (in this case the operation is similar to LSP Ping based bootstrapping described in [RFC5884]) or directly in the LSP ping configuration messages.

When BFD Control packets are transported in the G-ACh they are not protected by any end-to-end checksum, only lower-layers are providing error detection/correction. A single bit error, e.g. a flipped bit in the BFD State field could cause the receiving end to wrongly conclude that the link is down and in turn trigger protection switching. To prevent this from happening the "BFD Configuration

sub-TLV" has an Integrity flag that when set enables BFD Authentication using Keyed SHA1 with an empty key (all 0s) [RFC5880]. This would make every BFD Control packet carry an SHA1 hash of itself that can be used to detect errors.

If BFD Authentication using a pre-shared key / password is desired (i.e. authentication and not only error detection) the "BFD Authentication sub-TLV" MUST be included in the "BFD Configuration sub-TLV". The "BFD Authentication sub-TLV" is used to specify which authentication method that should be used and which pre-shared key / password that should be used for this particular session. How the key exchange is performed is out of scope of this document.

3.1.2. Configuration of Performance Monitoring

It is possible to configure Performance Monitoring functionalities such as Loss, Delay and Throughput as described in [MPLS-PM].

When configuring Performance monitoring functionalities it can be chosen either the default configuration (by only setting the respective flags in the "OAM functions TLV") or a customized configuration (by including the respective Loss and/or Delay sub-TLVs).

3.1.3. Configuration of Measurements and FMS

Additional OAM functions may be configured by setting the appropriate flags in the "OAM Functions TLV", these include Performance Measurements (packet loss, throughput, delay, and delay variation) and Fault Management Signal handling.

By setting the PM Loss flag in the "OAM Functions TLV" and including the "MPLS OAM PM Loss sub-TLV" one can configure the measurement interval and loss threshold values for triggering protection.

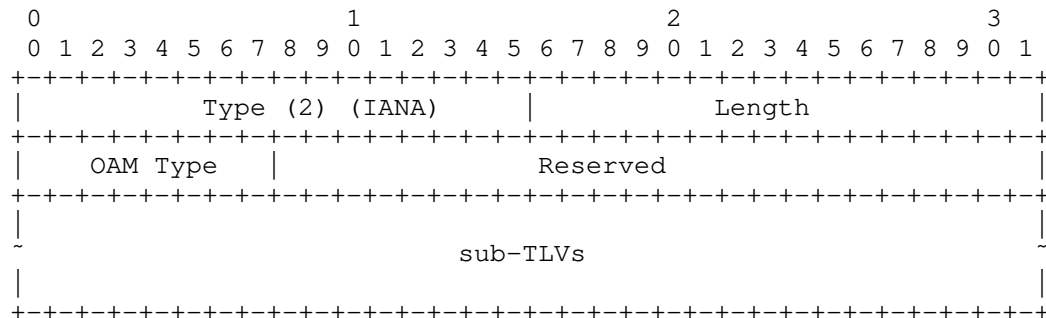
Delay measurements are configured by setting PM Delay flag in the "OAM Functions TLV" and including the "MPLS OAM PM Loss sub-TLV" one can configure the measurement interval and the delay threshold values for triggering protection.

To configure Fault Monitoring Signals and their refresh time the FMS flag in the "OAM Functions TLV" MUST be set and the "MPLS OAM FMS sub-TLV" included.

3.2. OAM Configuration TLV

The "OAM Configuration TLV" is depicted in the following figure. It specifies the OAM functions that are to be used for the LSP and it is

defined in [OAM-CONF-FWK]. The "OAM Configuration TLV" is carried in the LSP_ATTRIBUTES object in Path and Resv messages.



Type: indicates the "OAM Configuration TLV" (2) (IANA to assign).

OAM Type: one octet that specifies the technology specific OAM Type. If the requested OAM Type is not supported, an error must be generated: "OAM Problem/Unsupported OAM Type".

This document defines a new OAM Type: "MPLS OAM" (suggested value 2, IANA to assign) from the "RSVP-TE OAM Configuration Registry". The "MPLS OAM" type is set to request the establishment of OAM functions for MPLS-TP LSPs. The specific OAM functions are specified in the "Function Flags" sub-TLV as depicted in [OAM-CONF-FWK].

The receiving edge LSR when the MPLS-TP OAM Type is requested should check which OAM Function Flags are set in the "Function Flags TLV" (also defined in [OAM-CONF-FWK]) and look for the corresponding technology specific configuration TLVs.

Additional corresponding sub-TLVs are as follows:

- "BFD Configuration sub-TLV", which MUST be included if the CC and/or the CV OAM Function flag is set. This sub-TLV MUST carry a "BFD Local Discriminator sub-TLV" and a "Timer Negotiation Parameters sub-TLV" if the N flag is cleared. If the I flag is set, the "BFD Authentication sub-TLV" may be included.
- "MPLS OAM PM Loss sub-TLV" within the "Performance Monitoring sub-TLV", which MAY be included if the PM/Loss OAM Function flag is set. If the "MPLS OAM PM Loss sub-TLV" is not included, default configuration values are used. Such sub-TLV MAY also be included in case the Throughput function flag is set and there is

the need to specify measurement interval different from the default ones. In fact the throughput measurement make use of the same tool as the loss measurement, hence the same TLV is used.

- "MPLS OAM PM Delay sub-TLV" within the "Performance Monitoring sub-TLV", which MAY be included if the PM/Delay OAM Function flag is set. If the "MPLS OAM PM Delay sub-TLV" is not included, default configuration values are used.

- "MPLS OAM FMS sub-TLV", which MAY be included if the FMS OAM Function flag is set. If the "MPLS OAM FMS sub-TLV" is not included, default configuration values are used.

Moreover, if the CV or CC flag is set, the CC flag MUST be set at the same time. The format of an MPLS-TP CV/CC message is shown in [BFD-CCCV] and it requires, together with the BFD control packet information, the "Unique MEP-ID of source of BFD packet". [MPLS-TP-IDENTIF] defines the composition of such identifier as:

```
<"Unique MEP-ID of source of BFD packet"> ::=  
<src_node_id><src_tunnel_num><lsp_num>
```

GMPLS signaling [RFC3473] uses a 5-tuple to uniquely identify an LSP within an operator's network. This tuple is composed of a Tunnel Endpoint Address, Tunnel_ID, Extended Tunnel ID, and Tunnel Sender Address and (GMPLS) LSP_ID.

Hence, the following mapping is used without the need of redefining a new TLV for MPLS-TP proactive CV purpose.

- Tunnel ID = src_tunnel_num
- Tunnel Sender Address = src_node_id
- LSP ID = LSP_Num

"Tunnel ID" and "Tunnel Sender Address" are included in the "SESSION" object [RFC3209], which is mandatory in both Path and Resv messages.

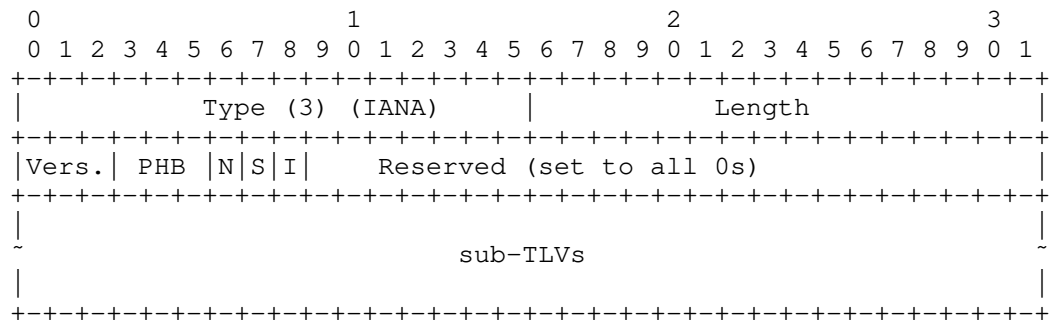
"LSP ID" will be the same on both directions and it is included in the "SENDER_TEMPLATE" object [RFC3209] which is mandatory in Path messages.

[Author's note: the same "Unique MEP-ID of source" will be likely required for Performance monitoring purposes. This need to be agreed with [MPLS-PM] authors.]

3.3. BFD Configuration sub-TLV

The "BFD Configuration sub-TLV" (depicted below) is defined for BFD OAM specific configuration parameters. The "BFD Configuration sub-TLV" is carried as a sub-TLV of the "OAM Configuration TLV".

This TLV accommodates generic BFD OAM information and carries sub-TLVs.



Type: indicates a new type, the "BFD Configuration sub-TLV" (IANA to define).

Length: indicates the total length including sub-TLVs.

Version: identifies the BFD protocol version. If a node does not support a specific BFD version an error must be generated: "OAM Problem/Unsupported OAM Version".

PHB: Identifies the Per-Hop Behavior (PHB) to be used for periodic continuity monitoring messages.

BFD Negotiation (N): If set timer negotiation/re-negotiation via BFD Control Messages is enabled, when cleared it is disabled.

Symmetric session (S): If set the BFD session MUST use symmetric timing values.

Integrity (I): If set BFD Authentication MUST be enabled. If the "BFD Configuration sub-TLV" does not include a "BFD Authentication sub-TLV" the authentication MUST use Keyed SHA1 with an empty pre-shared key (all 0s).

Encapsulation Capability (G): if set, it shows the capability of encapsulating BFD messages into G-Ach channel. If both the G bit and U bit are set, configuration gives precedence to the G bit.

Encapsulation Capability (U): if set, it shows the capability of encapsulating BFD messages into UDP packets. If both the G bit and U bit are set, configuration gives precedence to the G bit.

Bidirectional (B): if set, it configures BFD in the Bidirectional mode. If it is not set it configures BFD in unidirectional mode. In the second case, the source node does not expect any Discriminator values back from the destination node.

The "BFD Configuration sub-TLV" MUST include the following sub-TLVs in the Path message:

- "Local Discriminator sub-TLV";
- "Negotiation Timer Parameters sub-TLV" if the N flag is cleared.

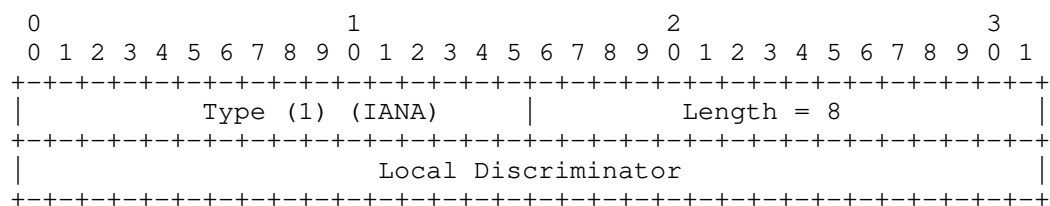
The "BFD Configuration sub-TLV" MUST include the following sub-TLVs in the Resv message:

- "Local Discriminator sub-TLV";
- "Negotiation Timer Parameters sub-TLV" if:
 - the N and S flags are cleared
 - the N flag is cleared and the S flag is set and a timing interval larger than the one received needs to be used

Reserved: Reserved for future specification and set to 0.

3.3.1. Local Discriminator sub-TLV

The "Local Discriminator sub-TLV" is carried as a sub-TLV of the "BFD Configuration sub-TLV" and is depicted below.



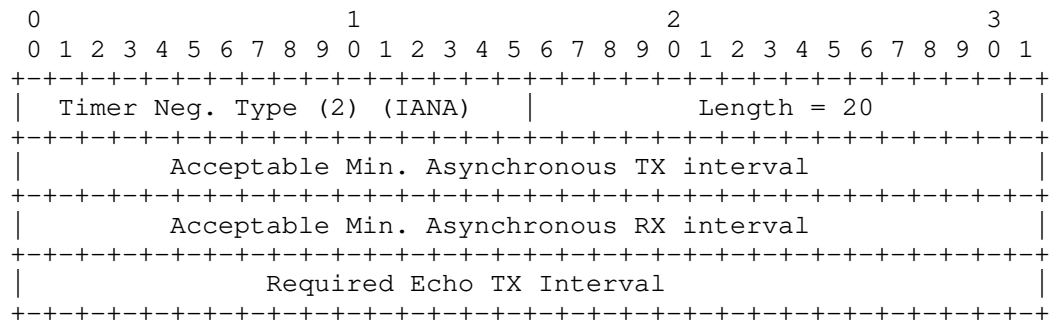
Type: indicates a new type, the Local Discriminator sub-TLV (1) (IANA to define).

Length: indicates the TLV total length in octets.

Local Discriminator: A unique, nonzero discriminator value generated by the transmitting system and referring to itself, used to demultiplex multiple BFD sessions between the same pair of systems.

3.3.2. Negotiation Timer Parameters sub-TLV

The "Negotiation Timer Parameters sub-TLV" is carried as a sub-TLV of the "BFD Configuration sub-TLV" and is depicted below.



Type: indicates a new type, the "Negotiation Timer Parameters sub-TLV" (IANA to define).

Length: indicates the TLV total length in octets. (20)

Acceptable Min. Asynchronous TX interval: in case of S (symmetric) flag set in the "BFD Configuration sub-TLV", it expresses the desired time interval (in microseconds) at which the ingress LER intends to both transmit and receive BFD periodic control packets. If the receiving edge LSR can not support such value, it can reply with an interval greater than the one proposed.

In case of S (symmetric) flag cleared in the "BFD Configuration sub-TLV", this field expresses the desired time interval (in microseconds) at which a edge LSR intends to transmit BFD periodic control packets in its transmitting direction.

Acceptable Min. Asynchronous RX interval: in case of S (symmetric) flag set in the "BFD Configuration sub-TLV", this field MUST be equal to "Acceptable Min. Asynchronous TX interval" and has no additional meaning respect to the one described for "Acceptable Min. Asynchronous TX interval".

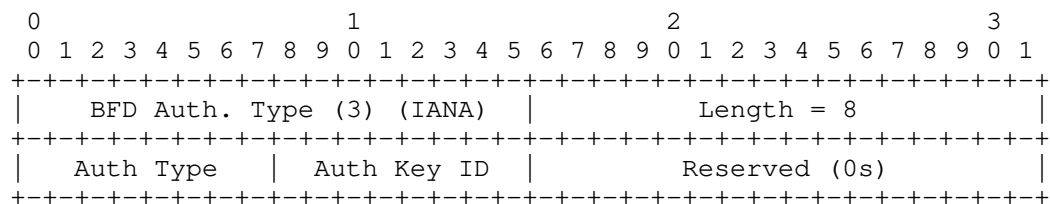
In case of S (symmetric) flag cleared in the "BFD Configuration sub-TLV", it expresses the minimum time interval (in microseconds) at which edge LSRs can receive BFD periodic control packets. In case

this value is greater than the "Acceptable Min. Asynchronous TX interval" received from the other edge LSR, such edge LSR MUST adopt the interval expressed in this "Acceptable Min. Asynchronous RX interval".

Required Echo TX Interval: the minimum interval (in microseconds) between received BFD Echo packets that this system is capable of supporting, less any jitter applied by the sender as described in [RFC5880] sect. 6.8.9. This value is also an indication for the receiving system of the minimum interval between transmitted BFD Echo packets. If this value is zero, the transmitting system does not support the receipt of BFD Echo packets. If the receiving system can not support this value an error MUST be generated "Unsupported BFD TX rate interval".

3.3.3. BFD Authentication sub-TLV

The "BFD Authentication sub-TLV" is carried as a sub-TLV of the "BFD Configuration sub-TLV" and is depicted below.



Type: indicates a new type, the "BFD Authentication sub-TLV" (IANA to define).

Length: indicates the TLV total length in octets. (8)

Auth Type: indicates which type of authentication to use. The same values as are defined in section 4.1 of [RFC5880] are used.

Auth Key ID: indicates which authentication key or password (depending on Auth Type) should be used. How the key exchange is performed is out of scope of this document.

Reserved: Reserved for future specification and set to 0.

3.4. Performance Monitoring sub-TLV

If the "OAM functions TLV" has either the L (Loss), D (Delay) or T (Throughput) flag set, the "Performance Monitoring sub-TLV" MUST be present.

In case the vlues needs to be different than the default ones the "Performance Monitoring sub-TLV", "MPLS OAM PM Loss sub-TLV" MAY include the following sub-TLVs:

- "MPLS OAM PM Loss sub-TLV" if the L flag is set in the "OAM functions TLV";
- "MPLS OAM PM Delay sub-TLV" if the D flag is set in the "OAM functions TLV";

The "Performance Monitoring sub-TLV" depicted below is carried as a sub-TLV of the "OAM Functions TLV".

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   Perf Monitoring Type (IANA)   |           Length           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|D|L|J|Y|K|C|           Reserved (set to all 0s)           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     ~                         |
|                                     sub-TLVs                  |
|                                     ~                         |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

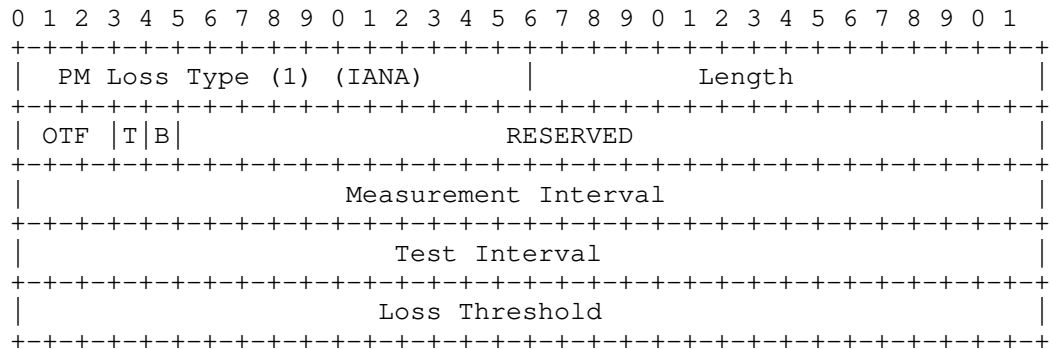
```

Configuration Flags, for the specific function description please refer to [MPLS-PM]:

- D: Delay inferred/direct (0=INFERRED, 1=DIRECT)
- L: Loss inferred/direct (0=INFERRED, 1=DIRECT)
- J: Delay variation/jitter (1=ACTIVE, 0=NOT ACTIVE)
- Y: Dyadic (1=ACTIVE, 0=NOT ACTIVE)
- K: Loopback (1=ACTIVE, 0=NOT ACTIVE)
- C: Combined (1=ACTIVE, 0=NOT ACTIVE)

3.4.1. MPLS OAM PM Loss sub-TLV

The "MPLS OAM PM Loss sub-TLV" depicted below is carried as a sub-TLV of the "Performance Monitoring sub-TLV".



Type: indicates a new type, the "MPLS OAM PM Loss sub-TLV" (IANA to define, suggested value 1).

Length: indicates the length of the parameters in octets (12).

OTF: Origin Timestamp Format of the Origin Timestamp field described in [MPLS-PM]. By default it is set to IEEE 1588 version 1.

Configuration Flags, please refer to [MPLS-PM] for further details:

- T: Traffic-class-specific measurement indicator. Set to 1 when the measurement operation is scoped to packets of a particular traffic class (DSCP value), and 0 otherwise. When set to 1, the DS field of the message indicates the measured traffic class. By default it is set to 1.
- B: Octet (byte) count. When set to 1, indicates that the Counter 1-4 fields represent octet counts. When set to 0, indicates that the Counter 1-4 fields represent packet counts. By default it is set to 0.

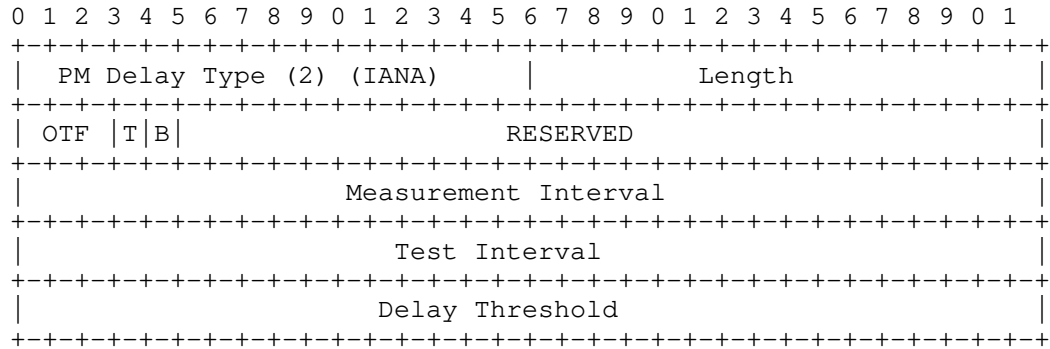
Measurement Interval: the time interval (in microseconds) at which Loss Measurement query messages MUST be sent on both directions. If the edge LSR receiving the Path message can not support such value, it can reply back with a higher interval. By default it is set to (TBD).

Test Interval: test messages interval as described in [MPLS-PM]. By default it is set to (TBD).

Loss Threshold: the threshold value of lost packets over which protections MUST be triggered. By default it is set to (TBD).

3.4.2. MPLS OAM PM Delay sub-TLV

The "MPLS OAM PM Delay sub-TLV" depicted below is carried as a sub-TLV of the "OAM Functions TLV".



Type: indicates a new type, the "MPLS OAM PM Loss sub-TLV" (IANA to define, suggested value 1).

Length: indicates the length of the parameters in octets (12).

OTF: Origin Timestamp Format of the Origin Timestamp field described in [MPLS-PM]. By default it is set to IEEE 1588 version 1.

Configuration Flags, please refer to [MPLS-PM] for further details:

- T: Traffic-class-specific measurement indicator. Set to 1 when the measurement operation is scoped to packets of a particular traffic class (DSCP value), and 0 otherwise. When set to 1, the DS field of the message indicates the measured traffic class. By default it is set to 1.
- B: Octet (byte) count. When set to 1, indicates that the Counter 1-4 fields represent octet counts. When set to 0, indicates that the Counter 1-4 fields represent packet counts. By default it is set to 0.

Measurement Interval: the time interval (in microseconds) at which Delay Measurement query messages MUST be sent on both directions. If the edge LSR receiving the Path message can not support such value, it can reply back with a higher interval. By default it is set to (TBD).

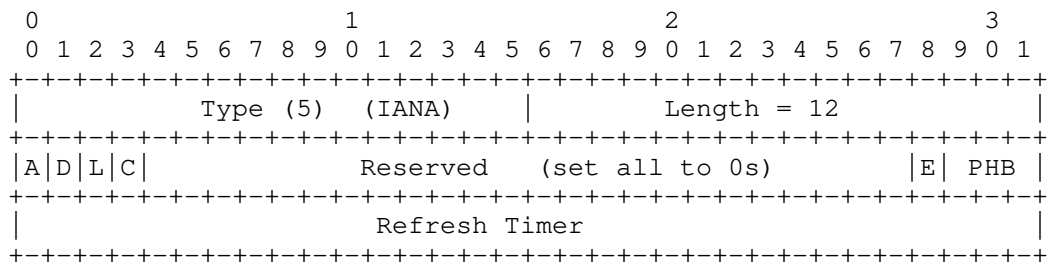
Test Interval: test messages interval as described in [MPLS-PM]. By

default it is set to (TBD).

Delay Threshold: the threshold value of measured delay (in microseconds) over which protections MUST be triggered. By default it is set to (TBD).

3.5. MPLS OAM FMS sub-TLV

The "MPLS OAM FMS sub-TLV" depicted below is carried as a sub-TLV of the "OAM Configuration sub-TLV".



Type: indicates a new type, the "MPLS OAM FMS sub-TLV" (IANA to define).

Length: indicates the TLV total length in octets.

Signal Flags: are used to enable the following signals:

- A: Alarm Indication Signal (AIS) as described in [MPLS-FMS]
- D: Link Down Indication (LDI) as described in [MPLS-FMS]
- L: Locked Report (LKR) as described in [MPLS-FMS]
- C: Client Signal Failure (CSF) as described in [MPLS-CSF]
- Remaining bits: Reserved for future specification and set to 0.

Configuration Flags:

- E: used to enable/disable explicitly clearing faults
- PHB: identifies the per-hop behavior of packets with fault management information

Refresh Timer: indicates the refresh timer (in microseconds) of fault indication messages. If the edge LSR receiving the Path message can

not support such value, it can reply back with a higher interval.

4. IANA Considerations

This document specifies the following new TLV types:

- "BFD Configuration" type: 2;
- "MPLS OAM PM Loss" type: 3;
- "MPLS OAM PM Delay" type: 4;
- "MPLS OAM FMS" type: 5.

sub-TLV types to be carried in the "BFD Configuration sub-TLV":

- "Local Discriminator" sub-TLV type: 1;
- "Negotiation Timer Parameters" sub-TLV type: 2.
- "BFD Authentication" sub-TLV type: 3.

5. BFD OAM configuration errors

In addition to error values specified in [OAM-CONF-FWK] and [ETH-OAM] this document defines the following values for the "OAM Problem" Error Code:

- "MPLS OAM Unsupported Functionality";
- "OAM Problem/Unsupported TX rate interval".

6. Acknowledgements

The authors would like to thank David Allan, Lou Berger, Annamaria Fulignoli, Eric Gray, Andras Kern, David Jocha and David Sinicrope for their useful comments.

7. Security Considerations

The signaling of OAM related parameters and the automatic establishment of OAM entities introduces additional security considerations to those discussed in [RFC3473]. In particular, a network element could be overloaded if an attacker were to request

high frequency liveliness monitoring of a large number of LSPs, targeting a single network element.

Security aspects will be covered in more detailed in subsequent versions of this document.

8. References

8.1. Normative References

- [MPLS-FMS] Swallow, G., Fulignoli, A., Vigoureux, M., Boutros, S., and D. Ward, "MPLS Fault Management OAM", 2009, <draft-ietf-mpls-tp-fault>.
- [MPLS-PM] Bryant, S. and D. Frost, "Packet Loss and Delay Measurement for the MPLS Transport Profile", 2010, <draft-ietf-mpls-loss-delay>.
- [MPLS-PM-Profile] Bryant, S. and D. Frost, "A Packet Loss and Delay Measurement Profile for MPLS-based Transport Networks", 2010, <draft-ietf-mpls-tp-loss-delay-profile>.
- [MPLS-TP-IDENTIF] Bocci, M., Swallow, G., and E. Gray, "MPLS-TP Identifiers", 2010, <draft-ietf-mpls-tp-identifiers>.
- [OAM-CONF-FWK] Takacs, A., Fedyk, D., and J. van He, "OAM Configuration Framework for GMPLS RSVP-TE", 2009, <draft-ietf-ccamp-oam-configuration-fwk>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", RFC 5586, June 2009.
- [RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.
- [RFC5860] Vigoureux, M., Ward, D., and M. Betts, "Requirements for Operations, Administration, and Maintenance (OAM) in MPLS Transport Networks", RFC 5860, May 2010.
- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.
- [RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, June 2010.

8.2. Informative References

- [BFD-CCCV] Allan, D., Swallow, G., and J. Drake, "Proactive Connectivity Verification, Continuity Check and Remote Defect indication for MPLS Transport Profile", 2010, <draft-ietf-mpls-tp-bfd-cc-cv-rdi>.
- [BFD-Ping] Bahadur, N., Aggarwal, R., Ward, D., Nadeau, T., Sprecher, N., and Y. Weingarten, "LSP Ping and BFD encapsulation over ACH", 2010, <draft-ietf-mpls-tp-lsp-ping-bfd-procedures-02>.
- [ETH-OAM] Takacs, A., Gero, B., Fedyk, D., Mohan, D., and D. Long, "GMPLS RSVP-TE Extensions for Ethernet OAM", 2009, <draft-ietf-ccamp-rsvp-te-eth-oam-ext>.
- [LSP Ping] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", 2006, <RFC 3479>.
- [LSP-PING CONF] Bellagamba, E., Andersson, L., Ward, D., and P. Skoldstrom, "Configuration of pro-active MPLS-TP Operations, Administration, and Maintenance (OAM) Functions Using LSP Ping", 2010, <draft-ietf-mpls-lsp-ping-mpls-tp-oam-conf>.

[MPLS-TP OAM Analysis]

Sprecher, N., Weingarten, Y., and E. Bellagamba, "MPLS-TP OAM Analysis", 2011, <draft-ietf-mpls-tp-oam-analysis>.

[MPLS-TP-OAM-FWK]

Bocci, M. and D. Allan, "Operations, Administration and Maintenance Framework for MPLS-based Transport Networks", 2010, <draft-ietf-mpls-tp-oam-framework>.

[RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.

[RFC5921] Bocci, M., Bryant, S., Frost, D., Levrau, L., and L. Berger, "A Framework for MPLS in Transport Networks", RFC 5921, July 2010.

Authors' Addresses

Elisa Bellagamba (editor)
Ericsson
Torshamnsgatan 48
Kista, 164 40
Sweden

Email: elisa.bellagamba@ericsson.com

Loa Andersson (editor)
Ericsson
Torshamnsgatan 48
Kista, 164 40
Sweden

Phone:
Email: loa.andersson@ericsson.com

Pontus Skoldstrom (editor)
Acreo AB
Electrum 236
Kista, 164 40
Sweden

Phone: +46 8 6327731
Email: pontus.skoldstrom@acreo.se

Dave Ward
Juniper

Phone:
Email: dward@juniper.net

Attila Takacs
Ericsson
1. Laborc u.
Budapest,
HUNGARY

Phone:
Email: attila.takacs@ericsson.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 29, 2011

A. Kern
A. Takacs
Ericsson
April 27, 2011

GMPLS RSVP-TE Extensions for SONET/SDH and OTN OAM Configuration
draft-ietf-ccamp-rsvp-te-sdh-otn-oam-ext-02

Abstract

GMPLS has been extended to support connection establishment in both SONET/SDH [RFC4606] and OTN [RFC4328] networks. However support for the configuration of the supervision functions is not specified. Both SONET/SDH and OTN implement supervision functions to qualify the transported signals. This document defines extensions to RSVP-TE for SONET/SDH and OTN OAM configuration based on the OAM Configuration Framework defined in [GMPLS-OAM-FWK].

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 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 October 2011

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 BSD License.

Table of Contents

1. Introduction	4
2. Overview of SONET/SDH and OTN OAM related functions	5
2.1. Continuity supervision	5
2.2. Connectivity supervision	5
2.2.1. SONET/SDH	5
2.2.2. OTN	5
2.3. Signal quality supervision	5
2.3.1. SONET/SDH	6
2.3.2. OTN	6
3. RSVP-TE signaling extensions	7
3.1. Operation overview	7
3.1.1. Continuity Check supervision	7
3.1.2. Connectivity Monitoring supervision	7
3.1.2.1. SDH/SONET	7
3.1.2.2. OTN	8
3.1.3. Signal quality supervision	9
3.2. Signaling support of Virtual Concatenation Groups (VCG)	9
3.3. OAM types and functions	10
3.4. SONET/SDH OAM Configuration sub-TLV	10
3.5. OTN OAM Configuration sub-TLV	11
3.6. TTI Configuration Sub-TLV	11
3.6.1. SDH TTI Configuration Sub-TLV	11
3.6.2. OTN TTI Configuration Sub-TLV	12
3.7. Degraded signal thresholds Sub-TLV	13
4. Error handling	15
5. IANA Considerations	16
6. Security Considerations	17
7. Acknowledgements	18
8. References	19
8.1. Normative References	19
8.2. Informative References	19
Authors' Addresses	21

1. Introduction

Both SONET/SDH and OTN implement supervision functions to qualify the transported signals. Supervision functions include continuity, connectivity, signal quality, alignment and payload supervision. The ITU-T G.806 [G.806] recommendation defines the generic framework of the supervision functions, which are then further specified for SONET/SDH and OTN in technology specific documents.

GMPLS has been extended to support connection establishment in both SONET/SDH [RFC4606] and OTN [RFC4328] networks. These documents however do not support the configuration of the respective supervision functions.

[GMPLS-OAM-FWK] defines a technology-agnostic framework for GMPLS to support the establishment and configuration of the pro-active OAM functions of signalled connections. The properties of the OAM functions are exchanged during connection establishment and may be modified during the life of the connection. The technology specific parameters to be exchanged are to be described in accompanying documents. This document defines the extensions for SONET/SDH and OTN OAM configuration for end-to-end monitoring.

2. Overview of SONET/SDH and OTN OAM related functions

SONET/SDH [G.707] and OTN [G.709] provide a variety of supervision functions. Here we only consider continuity, connectivity and signal quality supervision functions, as these are the candidates for GMPLS based configuration.

2.1. Continuity supervision

Continuity supervision provides methods monitoring the health of a connection (trail).

2.2. Connectivity supervision

The connectivity supervision function provides a method to detect misconnections. The detection procedure is based on emitting a Trace Trail Identifier (TTI) known by both endpoints. The TTI is included by the source node as an overhead signal for each connection. The receiver node then compares the received TTI with the expected value and decides if a miss-connection occurred.

2.2.1. SONET/SDH

In case of SONET/SDH, connectivity supervision is implemented in the Regeneration Section (RS) and in the lower and higher order path layers (LOVC and HOVC). In all layers the TTI encodes only the Access Point Identifier (API) of the source node. In the various layers the lengths of these TTIs are different. In RS the TTI (encoded in J0 octet) is either 1 or 16 octets long. In higher order paths the TTI (encoded in J1), is either 16 or 64 octet long. In lower order paths the TTI is transmitted in the J2 byte and is 16 octet long.

2.2.2. OTN

In case of OTN, connectivity supervision is supported by the OTUk and ODUk digital hierarchy layers. In both layers, the length of the TTI is 64 octets, but only the first 32 octets are considered for connectivity supervision. This first part is further divided into a Source Access Point Identifier (SAPI) and a Destination Access Point Identifier (DAPI). Connectivity supervision may consider either the SAPI or DAPI only or both. The structure of the SAPI and DAPI is specified in [G.709].

2.3. Signal quality supervision

The quality of the transmitted signal is monitored as a ratio of bad frames. If the number of such frames reaches a threshold a defect

state is declared. To detect the correctness of the frames an Error Detection Code (EDC), such as Bit Interleaved Parity (BIP), is used. The distribution of the errors is assumed to follow either Poisson or a bursty distribution. For Poisson distribution an EDC violation ratio is defined as the threshold; while for the bursty model the threshold is defined as a number of consecutive 1-second time intervals in which the EDC violation exceeds a predefined ratio. In case of Poisson error distribution two defect state levels are defined: the Excessive Error and Degraded Signal defect. In the case of the bursty model, only the Degraded Signal defect level is considered.

2.3.1. SONET/SDH

SONET/SDH supports both Excessive Error and Degraded Signal defect levels and supports both Poisson and bursty error distribution models. These signal quality parameters are configured for the Multiplexing Section (MS) and the LOVC and HOVC path layers.

2.3.2. OTN

For OTN, in the digital transport layers (OTUk and ODUk) only the bursty error distribution model errors with the Degraded Signal defect level is supported. Two parameters are defined: Ratio of the bad frames in a one second interval (0% to 100% or 0 to number of frames per 1-second interval) and Number of consecutive intervals (between 2 and 10). Signal quality supervision in the optical transport layers is not specified by [G.798], it is indicated to be for further study.

3. RSVP-TE signaling extensions

3.1. Operation overview

RFC 4606 and RFC 4328 define the RSVP-TE extensions necessary to manage SDH/SONET and OTN optical and digital hierarchy connections. The monitoring functions associated to these connections may be configured together with configuring the connection itself.

The LSP Attribute Flag "OAM MEP entities desired" [GMPLS-OAM-FWK] is used to signal that the monitoring functions at the endpoints must be established. The "OAM MIP entities desired" flag must be set to 0 and must be ignored.

To configure OAM parameters the OAM Configuration TLV can be included in the LSP_ATTRIBUTES object. The TLV identifies which OAM technology ("OAM Type" field) to be used as well as which OAM functions are to be enabled (OAM Function Flags sub-TLV). For SONET/SDH and OTN the "Continuity Check" and "Connectivity Verification" flags control the Continuity and Connectivity supervision functions, while the "Performance Monitoring/Loss" flag enables the Signal Quality supervision function. Since delay monitoring is not used for SONET/SDH or OTN the "Performance Monitoring/Delay" flag must be cleared.

For additional details the appropriate technology specific sub-TLV can be carried in the OAM Configuration TLV.

3.1.1. Continuity Check supervision

In case of both discussed technologies, setting up continuity supervision function for a connection does not need further configuration besides enabling it. Therefore, by setting the "Connectivity Monitoring" Flag of OAM Function implicitly enables the continuity supervision function as well.

3.1.2. Connectivity Monitoring supervision

3.1.2.1. SDH/SONET

[G.707] defines three bytes (signals) for connectivity supervision purposes: the J0 byte in RS layer, the J1 and J2 bytes in HOVC and LOVC layers. These bytes encode 1 octet, 16 octet or 64 octet long unstructured octet streams. The source node emits this stream and the destination node matches it with an expected one. When the destination detects mismatch, defect state will be declared.

Since these streams encode an identifier defined for the source node,

different stream will to be emitted in the upstream and downstream directions for bidirectional connections. During the configuration the egress node has to be configured with the TTI value to be expected in the downstream direction and the TTI value to be emitted in the upstream direction. Therefore the SONET/SDH OAM Configuration TLV carries two Connectivity Supervision TLVs.

3.1.2.2. OTN

[G.709] defines a 64 octet long TTI, where the first 32 octets have a generic structure: a zero octet, a 15 octet long SAPI, a second zero octet and finally the 15 octet long DAPI.

For a unidirectional connection a single Connection Supervision TLV encodes elements of the TTI to be emitted. This TLV also specifies which parts of the TTI are compared to the expected values (only SAPI, only DAPI, both SAPI and DAPI).

In case of a bidirectional connection an endpoint can use a common API value for SAPI (for transmitted signal) and DAPI (for received signal). (See Figure 1.) The TTI values used in downstream and upstream directions are derived from the two API values: the downstream TTI will have the form of [0, API_a, 0, API_z] while the upstream TTI will use the form of [0, API_z, 0 API_a].

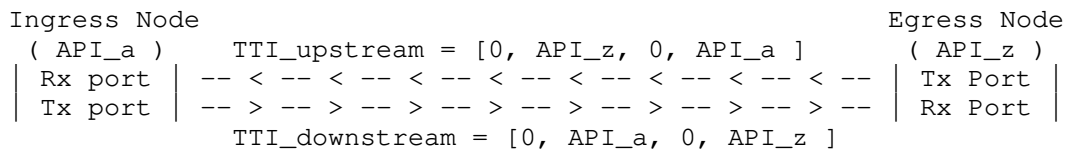


Figure 1: TTI construction when a single API identifies the receiver and transmitter interfaces

Then, a single Connectivity Supervision TLV is defined. The SAPI field carries the API of the ingress node (API_a) that initiates the signaling, while the DAPI carries the API of the egress node (API_z).

On the other hand, it is possible that the endpoints use different values as SAPI and DAPI to identify the transmitter and receiver ports of a bidirectional connection (See Figure 2). In this case the TTIs to be used in the two directions are independent, thus, they must be explicitly configured. Therefore, two Connectivity Supervision TLVs are added to the OTN OAM Configuration TLV. Each TLV encodes whether it defines the downstream or the upstream TTI.

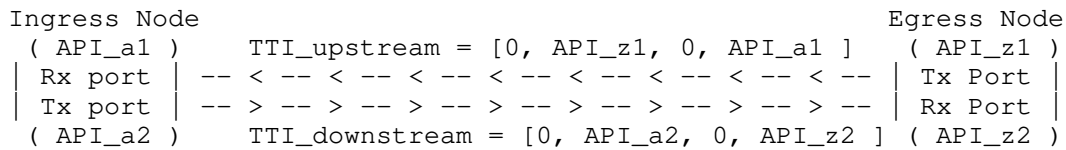


Figure 2: TTI construction when dedicated APIs identify the receiver and transmitter interfaces

3.1.3. Signal quality supervision

Signal quality supervision function is implemented in MS, HOVC, LOVC layers of SDH/SONET. All three layers support exceeded error level with Poisson error distribution model and degraded signal defect level with both, Poisson and bursty error distribution model. Dedicated Signal quality supervision TLVs encode each level, therefore when the "Performance Monitoring/Loss" flag is set; several such TLVs can be added to the SONET/SDH OAM Configuration TLV. If a configuration TLV for a particular level is missing the default parameters for that level is to be applied.

The OTN supports only Degraded Signal defect with bursty error model in OTUk and ODUk layers. Thus, the only parameters to be encoded are: the threshold for bad frames in a 1-second interval and the number of consecutive 1-second intervals with excessive bad frames. Furthermore, as only one level is allowed a single Signal quality supervision TLV is added to the OTN OAM Configuration TLV.

3.2. Signaling support of Virtual Concatenation Groups (VCG)

A key capability of both, SONET/SDH and OTN is the support of virtual concatenation. This inverse multiplexing method uses multiplicity of parallel basic signals. The supervision function parameters of these basic signals can be different.

[GMPLS-VCAT-LCAS] summarises GMPLS signaling capabilities to support virtual concatenation and proposes extensions to that. A Virtual Concatenated Group (VCG) is constructed from several individual data plane signals. The co-routed signals of a VCG could be provisioned together using a single RSVP-TE session (co-signaled). As different OAM configuration may be applied to each of these individual signals, the OAM configuration extension is applied as follows.

We assume that the same OAM type and the same set of OAM functions apply to every individual signal of the VCG. A single OAM Configuration TLV is carried in the LSP_ATTRIBUTES Object, while multiple instances of technology specific OAM configuration sub-TLVs are added: one instance per individual signal. The order of these

TLVs refers to the logical order of the basic signals (as they are listed in the Label Object).

[GMPLS-VCAT-LCAS] allows extension/pruning of a VCG. To achieve it the traffic descriptor, which encodes how the VCG is structured, in the RSVP-TE session is updated. If the VCG is updated the contents of the OAM Configuration TLV needs to be updated accordingly.

3.3. OAM types and functions

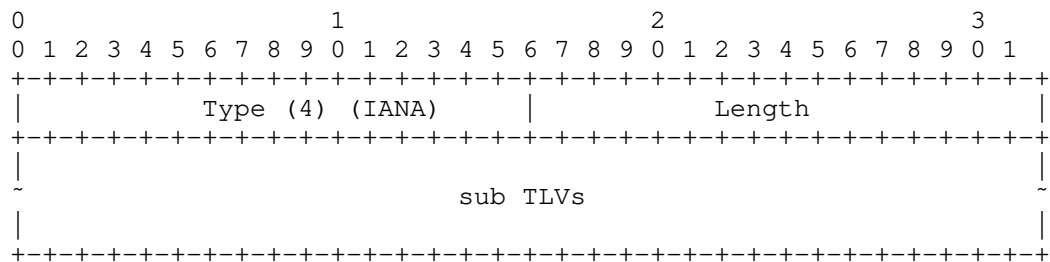
This document defines two new code points for the "OAM Type" field of the OAM Configuration TLV, defined in [GMPLS-OAM-FWK]: SONET/SDH OAM and OTN Digital Hierarchy OAM.

OAM Type	Description
3	SONET/SDH OAM
4	OTN Digital Hierarchy OAM

The "OAM Function Flags sub-TLV", defined in [GMPLS-OAM-FWK]. SONET/SDH and OTN supervision functions are defined in this document for the following flags: "Continuity Check", "Connectivity Verification" and "Performance Monitoring/Loss". As delay measurement is not supported, requesting that function SHOULD generate an error with code/value "OAM Problem/Unsupported OAM Function".

3.4. SONET/SDH OAM Configuration sub-TLV

SONET/SDH OAM Configuration sub-TLV is defined to encode the parameters of continuity, connectivity and signal quality supervision functions for SONET/SDH networks.

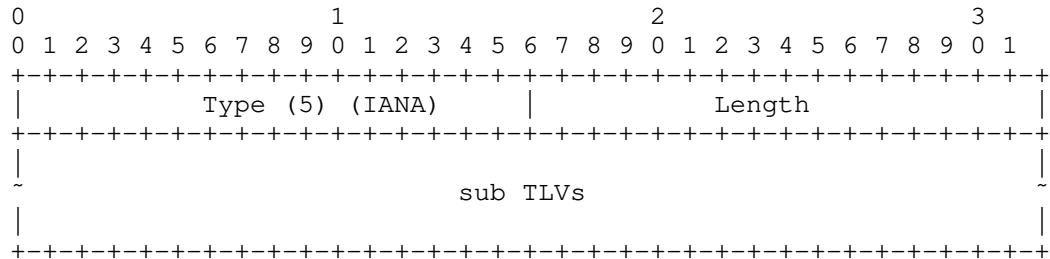


Type: indicates a new type: the SONET/SDH OAM Configuration TLV (IANA to define).

Length: indicates the total length including sub-TLVs

3.5. OTN OAM Configuration sub-TLV

OTN OAM Configuration TLV is defined to encode the parameters of continuity, connectivity and signal quality supervision functions for OTN.



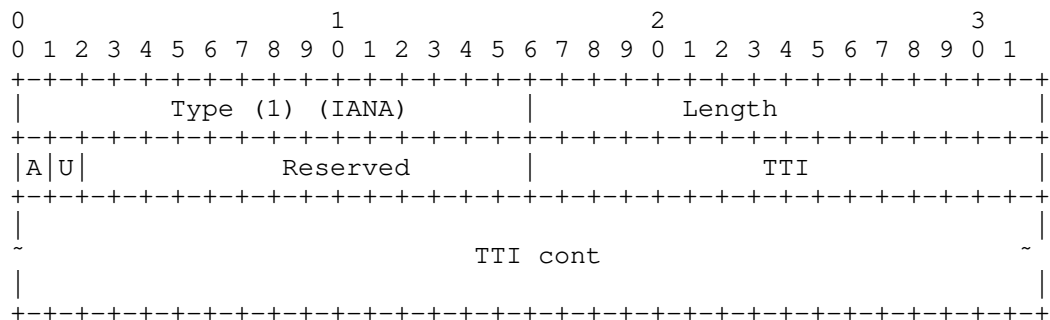
Type: indicates a new type: the OTN OAM Configuration TLV (IANA to define).

Length: indicates the total length including sub-TLVs

3.6. TTI Configuration Sub-TLV

3.6.1. SDH TTI Configuration Sub-TLV

This sub-TLV is carried in the SONET/SDH OAM Configuration sub-TLV, if the Connectivity Verification OAM Function Flag is set. In every supporting layers the TTI identifies the source interface (SAPI); however, the length of this identifier varies layer-by-layer (See Section 2.2.1). Therefore, a generic TLV is defined supporting various TTI lengths.



Flag "A", when set enables the AIS insertion on detecting TTI mismatch.

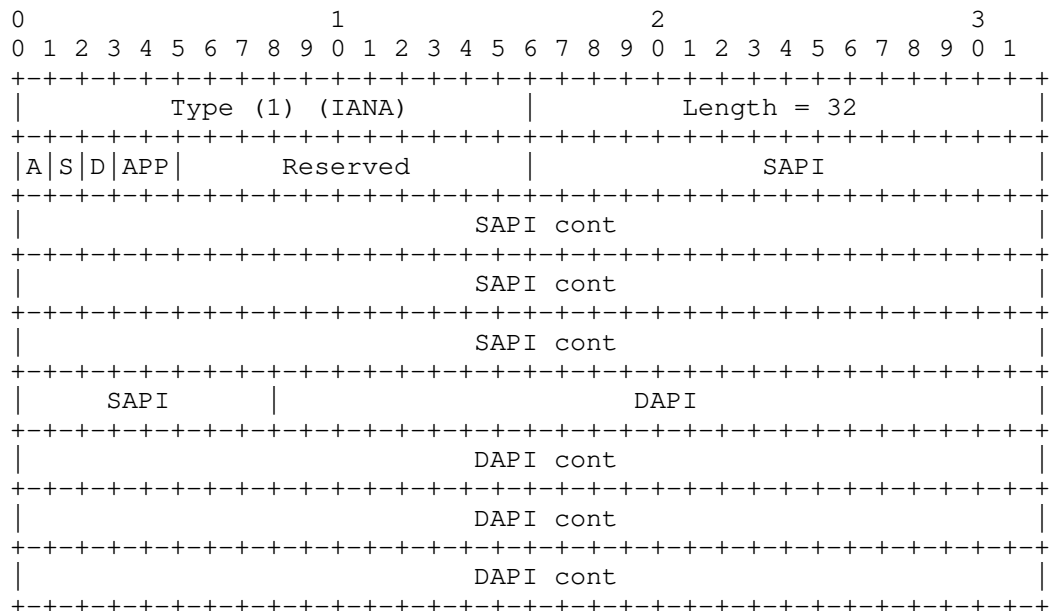
Flat "U" encodes if the TTI refers to the downstream TTI (U=0) or the upstream one (U=1).

The TTI field carries the TTI to be transmitted by the source node and to be expected by the sink. The TLV is padded to 4-octets.

If the specified length and format of the TTI carried in this TLV is not supported by the referred SONET/SDH layer, error must be generated: "OAM Problem/TTI Length Mismatch".

3.6.2. OTN TTI Configuration Sub-TLV

This sub-TLV is carried in the OTN OAM Configuration sub-TLV, if the Connectivity Verification OAM Function Flag is set.



Three control flags are defined. Flag "A" indicates that AIS insertion on detecting TTI mismatch (failing the connectivity verification) is required (A=1) or not (A=0). The next two flags define which parts of the received TTI are compared to the expected one. If flag "S" is set the TTI octets 1 to 15 are matched to the expected SAPI value. If the flag "D" is set the TTI octets 17 to 31 are matched to the expected DAPI value. If both "S" and "D" are set both parts of TTI are compared to SAPI and DAPI values. Setting both "S" and "D" bits to 0 is invalid, and if encountered error must be generated: "OAM Problem/Invalid CC/CV configuration".

The next two bits "APP" encode the applicability of the TTI configuration and the following code points are defined:

0 - Single TTI configuration: the TTI configuration is done according only to this TLV and no further TTI configuration TLVs are expected. This code point is used for unidirectional connections and for bidirectional connections with common APIs (See Figure 1)

1 - Downstream TTI for double TTI configuration: the current TLV instruct the configuration of the TTI to be used in downstream direction (See Figure 2).

2 - Upstream TTI for double TTI configuration: the current TLV instruct the configuration of the TTI to be used in upstream direction (See Figure 2).

3 - Invalid.

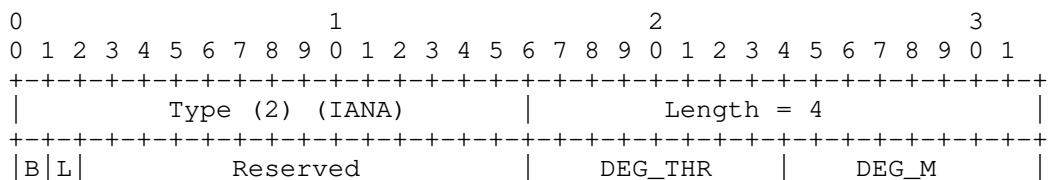
If the APP is set to 1 and the next or the previous sub-TLV is not an OTN TTI Configuration TLV with APP code point 2, then an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Missing Upstream TTI configuration".

If the APP is set to 2 and the next or the previous sub-TLV is not an OTN TTI Configuration TLV with APP code point 1, then an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Missing Downstream TTI configuration".

If the APP is set to either 1 or 2 and the unidirectional LSP is signaled (no UPSTREAM_LABEL is added to the message) or the APP is set to 3, an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Invalid applicability code"

3.7. Degraded signal thresholds Sub-TLV

The Degraded signal thresholds Sub-TLV instructs the configuration of the signal quality supervision function. This sub-TLV is applicable in both SONET/SDH and OTN cases. This sub-TLV can be carried in both the SONET/SDH OAM Configuration sub-TLV or OTN OAM Configuration sub-TLV, if the PerformanceMonitoring/Loss OAM Function Flag is set.



Two flags are defined to encode the signal quality measurement. The bit "B" encodes if distribution of errors is either Poisson (B=0) or Bursty (B=1). In case of Poisson distribution of errors two levels of defects are defined and encoded with bit "L": excessive error (L=0) and degraded signal (L=1). Since in case of Bursty distribution of errors only degraded signal defect is to be detected, therefore, in this latter case (B=1) the "L" bit must be set. Otherwise error must be generated: "OAM Problem/Invalid Performance Monitoring/Loss configuration".

In the second case (B=1) it encodes ratio of the bad frames in a 1-second period and can be set between 0 and 100, interpreted as ratios in percentage.

The field "DEG_M" defines monitoring time-frame in 1 second periods assuming bursty distribution of errors. The valid values are 2 to 10 periods.

4. Error handling

In addition to error values specified in [GMPLS-OAM-FWK] this document defines the following values for the "OAM Problem" Error Code.

- o If Performance Measurement/Delay flag is set in the OAM Functions Flag sub-TLV, an error must be generated "OAM Problem/Unsupported OAM Function".
- o In case of SONET/SDH OAM when the length or format of the TTI to be configured is not supported by the referred SONET/SDH layer, an error must be generated: "OAM Problem/TTI Length Mismatch".
- o If both "S" and "D" bits in OTN TTI Configuration TLV are set to 0, error must be generated: "OAM Problem/Invalid CC/CV configuration".
- o If the APP is set to 1 and the next or the previous sub-TLV is not an OTN TTI Configuration TLV with APP code point 2, then an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Missing Upstream TTI configuration".
- o If the APP is set to 2 and the next or the previous sub-TLV is not an OTN TTI Configuration TLV with APP code point 1, then an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Missing Downstream TTI configuration".
- o If the APP is set to either 1 or 2 and the unidirectional LSP is signaled (no UPSTREAM_LABEL is added to the message) or the APP is set to 3, an error must be generated "OAM Problem/Invalid OTN TTI Configuration/Invalid applicability code".
- o If flag "B" in Degraded signal thresholds Sub-TLV is set to 1 and flag "L" in the same sub-TLV is set to 0 error must be generated "OAM Problem/Invalid Performance Monitoring/Loss configuration".

5. IANA Considerations

This document specifies two new sub-TLVs to be carried in the OAM Configuration TLV in the LSP_ATTRIBUTES or LSP_REQUIRED_ATTRIBUTES Objects in Path and Resv messages. The document assigns values 3 and 4 from the "OAM Type" field of the OAM Configuration TLV.

The following error values need to be assigned under "OAM Problem" error code: "OAM Problem/Unsupported OAM Function", "OAM Problem/TTI Length Mismatch", "OAM Problem/Invalid CC/CV configuration", "OAM Problem/Invalid OTN TTI Configuration/Missing Upstream TTI configuration", "OAM Problem/Invalid OTN TTI Configuration/Missing Downstream TTI configuration", "OAM Problem/Invalid OTN TTI Configuration/Invalid applicability code", "OAM Problem/Invalid Performance Monitoring/Loss configuration".

6. Security Considerations

Security aspects are addressed in the OAM configuration framework document [GMPLS-OAM-FWK]

7. Acknowledgements

The authors would like to thank Francesco Fondelli for his useful comments.

8. References

8.1. Normative References

[GMPLS-OAM-FWK]

Takacs, A., Fedyk, D., and H. Jia, "OAM Configuration Framework and Requirements for GMPLS RSVP-TE", draft-ietf-ccamp-oam-configuration-fwk-01 (work in progress), March 2009.

8.2. Informative References

[G.707] International Telecommunications Union, "Network node interface for the synchronous digital hierarchy (SDH)", ITU-T Recommendation G.707, January 2007.

[G.709] International Telecommunications Union, "Interfaces for the Optical Transport Network (OTN)", ITU-T Recommendation G.709, March 2003.

[G.798] International Telecommunications Union, "Characteristics of optical transport network hierarchy equipment functional blocks", ITU-T Recommendation G.798, December 2006.

[G.806] International Telecommunications Union, "Characteristics of transport equipment - Description methodology and generic functionality", ITU-T Recommendation G.806, January 2009.

[GMPLS-VCAT-LCAS]

Bernstein, G., Rabbat, R., and H. Helvoort, "Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS)", draft-ietf-ccamp-gmpls-vcat-lcas-07 (work in progress), December 2008.

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

[RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.

[RFC4606] Mannie, E. and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous

Digital Hierarchy (SDH) Control", RFC 4606, August 2006.

Authors' Addresses

Andras Kern
Ericsson
Laborc u. 1.
Budapest, 1037
Hungary

Email: andras.kern@ericsson.com

Attila Takacs
Ericsson
Laborc u. 1.
Budapest, 1037
Hungary

Email: attila.takacs@ericsson.com

Network Working Group
Internet Draft
Intended status: Informational
Expires: September 2011

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

March 14, 2011

Routing and Wavelength Assignment Information Model for Wavelength
Switched Optical Networks

draft-ietf-ccamp-rwa-info-11.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 14, 2011.

Copyright Notice

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

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

Abstract

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

Table of Contents

1. Introduction.....	3
1.1. Revision History.....	4
1.1.1. Changes from 01.....	4
1.1.2. Changes from 02.....	4
1.1.3. Changes from 03.....	4
1.1.4. Changes from 04.....	5
1.1.5. Changes from 05.....	5
1.1.6. Changes from 06.....	5
1.1.7. Changes from 07.....	5
1.1.8. Changes from 08.....	5
1.1.9. Changes from 09.....	5
1.1.10. Changes from 10.....	6
2. Terminology.....	6
3. Routing and Wavelength Assignment Information Model.....	6
3.1. Dynamic and Relatively Static Information.....	7
4. Node Information (General).....	7
4.1. Connectivity Matrix.....	8
4.2. Shared Risk Node Group.....	8
5. Node Information (WSON specific).....	9
5.1. Resource Accessibility/Availability.....	10

5.2. Resource Signal Constraints and Processing Capabilities..	14
5.3. Compatibility and Capability Details.....	15
5.3.1. Shared Ingress or Egress Indication.....	15
5.3.2. Modulation Type List.....	15
5.3.3. FEC Type List.....	15
5.3.4. Bit Rate Range List.....	15
5.3.5. Acceptable Client Signal List.....	16
5.3.6. Processing Capability List.....	16
6. Link Information (General).....	16
6.1. Administrative Group.....	17
6.2. Interface Switching Capability Descriptor.....	17
6.3. Link Protection Type (for this link).....	17
6.4. Shared Risk Link Group Information.....	17
6.5. Traffic Engineering Metric.....	17
6.6. Port Label (Wavelength) Restrictions.....	17
6.6.1. Port-Wavelength Exclusivity Example.....	19
7. Dynamic Components of the Information Model.....	20
7.1. Dynamic Link Information (General).....	21
7.2. Dynamic Node Information (WSON Specific).....	21
8. Security Considerations.....	21
9. IANA Considerations.....	22
10. Acknowledgments.....	22
11. References.....	23
11.1. Normative References.....	23
11.2. Informative References.....	24
12. Contributors.....	25
Author's Addresses.....	25
Intellectual Property Statement.....	26
Disclaimer of Validity.....	27

1. Introduction

The purpose of the following information model for WSONs is to facilitate constrained optical path 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 optical path 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 [WSON-Frame]. 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.

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

WDM: Wavelength Division Multiplexing.

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

3. Routing and Wavelength Assignment Information Model

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

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

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

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

3.1. Dynamic and Relatively Static Information

All the RWA information of concern in a WSON network is subject to change over time. Equipment can be upgraded; links may be placed in or out of service and the like. However, from the point of view of RWA computations there is a difference between information that can change with each successive connection establishment in the network and that information that is relatively static on the time scales of connection establishment. A key example of the former is link wavelength usage since this can change with connection setup/teardown and this information is a key input to the RWA process. Examples of relatively static information are the potential port connectivity of a WDM ROADM, and the channel spacing on a WDM link.

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

4. Node Information (General)

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

Formally,

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

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

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

4.1. Connectivity Matrix

The connectivity matrix (ConnectivityMatrix) represents either the potential connectivity matrix for asymmetric switches (e.g. ROADMs and such) or fixed connectivity for an asymmetric device such as a multiplexer. Note that this matrix does not represent any particular internal blocking behavior but indicates which ingress 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 ingress ports and N the number of egress 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 ingress port i can connect to egress 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 [WSON-Frame] 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 [WSON-Frame] 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 ingress wavelength on a particular ingress port to a desired egress wavelength on a particular egress port.
3. Limitations on the types of signals that can be converted and the conversions that can be performed.

For modeling purposes and encoding efficiency identical processing resources such as regenerators or wavelength converters with identical limitations, and processing and accessibility properties are grouped into "blocks". Such blocks can consist of a single resource, though grouping resources into blocks leads to more efficient encodings. The resource pool model is composed of one or more resource blocks where the accessibility to and from any resource within a block is the same.

This leads to the following formal high level model:

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

Where

```
<ResourcePool> ::= <ResourceBlockInfo>...  
[<ResourceBlockAccessibility>...] [<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 [WSON-Frame] 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 ingress or egress 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 ingress to each resource block since each ingress represents a single wavelength signal, while in Figure 2 shows a single multiplexed WDM ingress or egress, e.g., a fiber, to/from each set of block.

This model assumes N ingress ports (fibers), P resource blocks containing one or more identical resources (e.g. wavelength converters), and M egress ports (fibers). Since not all ingress ports can necessarily reach each resource block, the model starts with a resource pool ingress matrix $RI(i,p) = \{0,1\}$ whether ingress 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 ingress 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 ingress port constraint is modeled via a static wavelength set mechanism and the case of shared access to a set of blocks is modeled via a dynamic wavelength set mechanism.

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

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

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

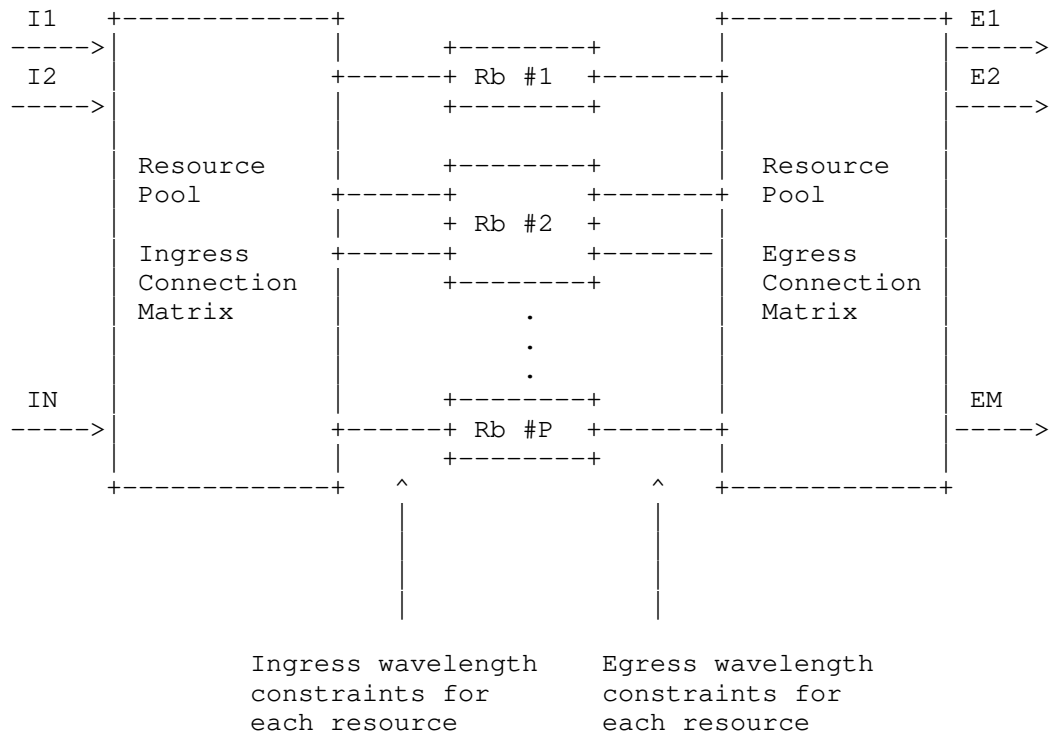


Figure 1 Schematic diagram of resource pool model.

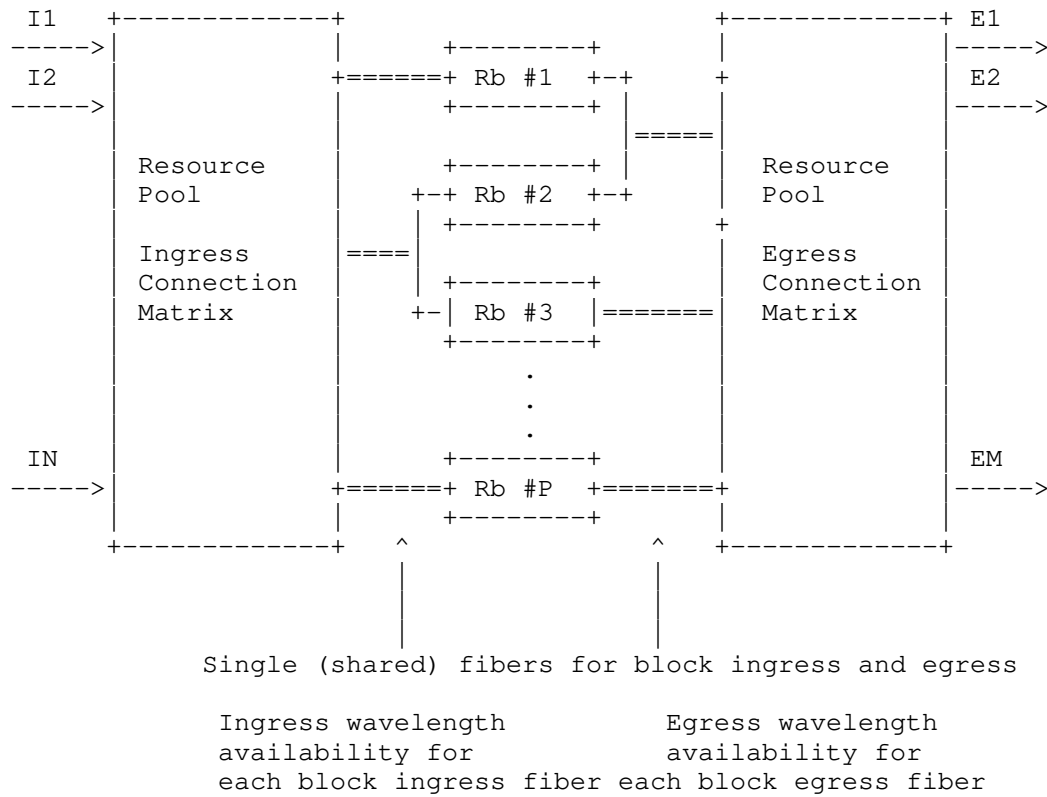


Figure 2 Schematic diagram of resource pool model with shared block accessibility.

Formally the model can be specified as:

```
<ResourceBlockAccessibility> ::= <PoolIngressMatrix>
<PoolEgressMatrix>
```

```
<ResourceWaveConstraints> ::= <IngressWaveConstraints>
<EgressWaveConstraints>
```

```
<RBPoolState>
::=(<ResourceBlockID><NumResourcesInUse><InAvailableWavelengths><OutA
vailableWavelengths>)...

```

Note that except for <RBPoolState> all the other components of <ResourcePool> are relatively static. Also the <InAvailableWavelengths> and <OutAvailableWavelengths> are only used in the cases of shared ingress or egress 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 <EgressWaveConstraints> previously discussed. As discussed in [WSON-Frame] 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> ::= <SharedIngress> <ModulationTypeList>
<FETypeList> <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 egress constraint indication.

```
<OutputConstraints> ::= <SharedEgress> <ModulationTypeList>
<FETypeList>
```

5.3. Compatibility and Capability Details

5.3.1. Shared Ingress or Egress Indication

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

5.3.2. Modulation Type List

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

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

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

5.3.3. FEC Type List

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

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

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

5.3.4. Bit Rate Range List

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

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

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

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

5.3.5. Acceptable Client Signal List

The list is simply:

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

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

5.3.6. Processing Capability List

The ProcessingCapabilities were defined in Section 5.2 as follows:

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

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

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

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

6. Link Information (General)

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

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

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

6.1. Administrative Group

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

6.2. Interface Switching Capability Descriptor

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

6.3. Link Protection Type (for this link)

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

6.4. Shared Risk Link Group Information

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

6.5. Traffic Engineering Metric

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

6.6. Port Label (Wavelength) Restrictions

Port label (wavelength) restrictions (PortLabelRestriction) model the label (wavelength) restrictions that the link and various optical devices such as OXCs, ROADMs, and waveband multiplexers may impose on

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

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

Where

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

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

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

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

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

WAVEBAND1: Waveband device with a tunable center frequency and passband. This constraint is characterized by the MaxWaveBandWidth parameters which indicates the maximum width of the waveband in terms of channels. Note that an additional wavelength set can be used to

indicate the overall tuning range. Specific center frequency tuning information can be obtained from dynamic channel in use information. It is assumed that both center frequency and bandwidth (Q) tuning can be done without causing faults in existing signals.

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

LabelSet is a conceptual set of labels (wavelengths).

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

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

PortSet is a conceptual set of ports.

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

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

6.6.1. Port-Wavelength Exclusivity Example

Although there can be many different ROADM or switch architectures that can lead to the constraint where a lambda (label) maybe used at most once on a set of ports Figure 3 shows a ROADM architecture based on components known as a Wavelength Selective Switch (WSS)[OFC08]. This ROADM is composed of splitters, combiners, and WSSes. This ROADM has 11 egress ports, which are numbered in the diagram. Egress ports 1-8 are known as drop ports and are intended to support a single wavelength. Drop ports 1-4 egress from WSS #2, which is fed from WSS #1 via a single fiber. Due to this internal structure a constraint is placed on the egress ports 1-4 that a lambda can be only used once over the group of ports (assuming uni-cast and not multi-cast operation). Similarly the egress ports 5-8 have a similar constraint due to the internal structure.

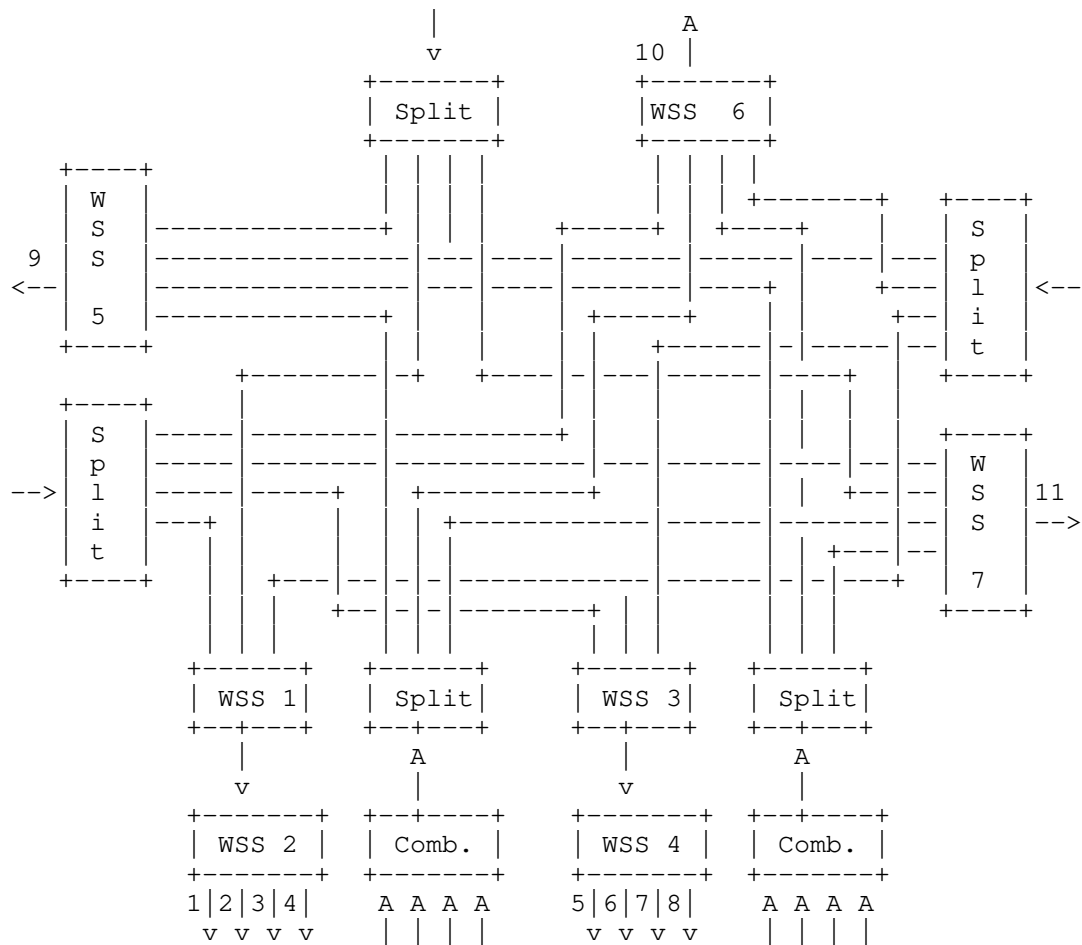


Figure 3 A ROADM composed from splitter, combiners, and WSSs.

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.

7.2. Dynamic Node Information (WSON Specific)

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

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

8. Security Considerations

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

9. IANA Considerations

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

10. Acknowledgments

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

11. References

11.1. Normative References

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

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

11.2. Informative References

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

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

Author's Addresses

Greg M. Bernstein (ed.)

Grotto Networking

Fremont California, USA

Phone: (510) 573-2237

Email: gregb@grotto-networking.com

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

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

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

Phone: +86-755-28973237
Email: danli@huawei.com

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

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

Intellectual Property Statement

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

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

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

Disclaimer of Validity

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

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: September 2011

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

March 14, 2011

Routing and Wavelength Assignment Information Encoding for
Wavelength Switched Optical Networks

draft-ietf-ccamp-rwa-wson-encode-11.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 14, 2011.

Copyright Notice

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

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

Abstract

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

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

Conventions used in this document

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

Table of Contents

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

1.1.4. Changes from 03 draft.....	5
1.1.5. Changes from 04 draft.....	5
1.1.6. Changes from 05 draft.....	5
1.1.7. Changes from 06 draft.....	5
1.1.8. Changes from 07 draft.....	5
1.1.9. Changes from 08 draft.....	6
1.1.10. Changes from 09 draft.....	6
1.1.11. Changes from 10 draft.....	6
2. Terminology.....	6
3. Resource Pool Accessibility/Availability.....	7
3.1. Resource Pool Accessibility Sub-TLV.....	8
3.2. Resource Block Wavelength Constraints Sub-TLV.....	10
3.3. Resource Pool State Sub-TLV.....	10
3.4. Block Shared Access Wavelength Availability sub-TLV.....	12
4. Resource Properties Encoding.....	13
4.1. Resource Block Information Sub-TLV.....	13
4.2. Input Modulation Format List Sub-Sub-TLV.....	14
4.2.1. Modulation Format Field.....	15
4.3. Input FEC Type List Sub-Sub-TLV.....	17
4.3.1. FEC Type Field.....	17
4.4. Input Bit Range List Sub-Sub-TLV.....	19
4.4.1. Bit Range Field.....	19
4.5. Input Client Signal List Sub-Sub-TLV.....	20
4.6. Processing Capability List Sub-Sub-TLV.....	21
4.6.1. Processing Capabilities Field.....	21
4.7. Output Modulation Format List Sub-Sub-TLV.....	23
4.8. Output FEC Type List Sub-Sub-TLV.....	23
5. Security Considerations.....	23
6. IANA Considerations.....	23
7. Acknowledgments.....	23
APPENDIX A: Encoding Examples.....	23
A.1. Wavelength Converter Accessibility Sub-TLV.....	23
A.2. Wavelength Conversion Range Sub-TLV.....	23
A.3. An OEO Switch with DWDM Optics.....	23
8. References.....	23
8.1. Normative References.....	23
8.2. Informative References.....	23
9. Contributors.....	23
Authors' Addresses.....	23
Intellectual Property Statement.....	23
Disclaimer of Validity.....	23

1. Introduction

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

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

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

Note that encodings of information needed by the routing and label assignment process applicable to general networks beyond WSON are addressed in a separate document [Gen-Encode].

1.1. Revision History

1.1.1. Changes from 00 draft

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

Clarification of TBD on connection matrix type and possibly numbering.

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

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

1.1.2. Changes from 01 draft

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

1.1.3. Changes from 02 draft

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

Added E bit to the Wavelength Converter Set Field.

Added bidirectional connectivity matrix example. Added simple link set example. Edited examples for consistency.

1.1.4. Changes from 03 draft

Removed encodings for general concepts to [Gen-Encode].

Added in WSON signal compatibility and processing capability information encoding.

1.1.5. Changes from 04 draft

Added encodings to deal with access to resource blocks via shared fiber.

1.1.6. Changes from 05 draft

Revised the encoding for the "shared access" indicators to only use one bit each for ingress and egress.

1.1.7. Changes from 06 draft

Removed section on "WSON Encoding Usage Recommendations"

1.1.8. Changes from 07 draft

Section 3: Enhanced text to clarify relationship between pools, blocks and resources. Section 3.1, 3.2: Change title to clarify Pool-Block relationship. Section 3.3: clarify block-resource state.

Section 4: Deleted reference to previously removed RBNF element. Fixed TLV figures and descriptions for consistent sub-sub-TLV nomenclature.

1.1.9. Changes from 08 draft

Fixed ordering of fields in second half of sub-TLV example in Appendix A.1.

Clarifying edits in section 3 on pools, blocks, and resources.

1.1.10. Changes from 09 draft

Fixed the "Block Shared Access Wavelength Availability sub-TLV" of section 3.4 to use an "RB set field" rather than a single RB ID. Removed all 1st person idioms.

1.1.11. Changes from 10 draft

Removed remaining 1st person idioms. Updated IANA section.

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

WDM: Wavelength Division Multiplexing.

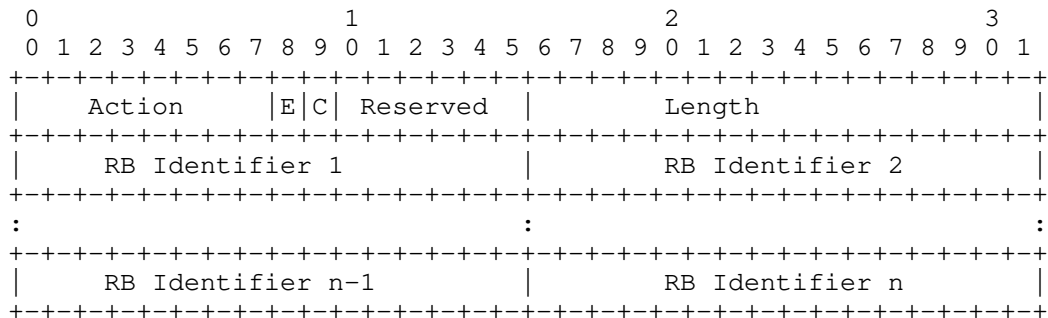
Wavelength Switched Optical Network (WSO): A WDM based optical network in which switching is performed selectively based on the center wavelength of an optical signal.

3. Resource Pool Accessibility/Availability

This section defines the sub-TLVs for dealing with accessibility and availability of resource blocks within a pool of resources. These include the ResourceBlockAccessibility, ResourceWaveConstraints, and RBPooolState sub-TLVs. All these sub-TLVs are concerned with sets of resources. As described in [WSOON-Info] a resource pool is composed of blocks of resources with similar properties and accessibility characteristics.

In a WSON node that includes resource blocks (RB) denoting subsets of these blocks allows one to efficiently describe common properties the blocks and to describe the structure, if non-trivial, of the resource pool. The RB Set field is defined in a similar manner to the label set concept of [RFC3471].

The information carried in a RB set field is defined by:



Action: 8 bits

0 - Inclusive List

Indicates that the TLV contains one or more RB elements that are included in the list.

2 - Inclusive Range

Indicates that the TLV contains a range of RBs. The object/TLV contains two WC elements. The first element indicates the start of the range. The second element indicates the end of the range. A value of zero indicates that there is no bound on the corresponding portion of the range.

E (Even bit): Set to 0 denotes an odd number of RB identifiers in the list (last entry zero pad); Set to 1 denotes an even number of RB identifiers in the list (no zero padding).

C (Connectivity bit): Set to 0 to denote fixed (possibly multi-cast) connectivity; Set to 1 to denote potential (switched) connectivity. Used in resource pool accessibility sub-TLV. Ignored elsewhere.

Reserved: 6 bits

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

Length: 16 bits

The total length of this field in bytes.

RB Identifier:

The RB identifier represents the ID of the resource block which is a 16 bit integer.

3.1. Resource Pool Accessibility Sub-TLV

This sub-TLV describes the structure of the resource pool in relation to the switching device. In particular it indicates the ability of an ingress port to reach a resource block and of a resource block to reach a particular egress port. This is the PoolIngressMatrix and PoolEgressMatrix of [WSO-Info].

The resource pool accessibility sub-TLV is defined by:

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Connectivity |                               Reserved |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Ingress Link Set Field A #1 |
:                                                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               RB Set Field A #1           |
:                                                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
| Additional Link set and RB set pairs as needed to       |
: specify PoolIngressMatrix                                :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Egress Link Set Field B #1   |
:                                                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               RB Set B Field #1 (for egress connectivity)
:                                                           :
+-----+-----+-----+-----+-----+-----+-----+-----+
| Additional Link Set and RB set pairs as needed to       |
: specify PoolEgressMatrix                                 :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where

Connectivity indicates how the ingress/egress ports connect to the resource blocks.

0 -- the device is fixed (e.g., a connected port must go through the resource block)

1 -- the device is switched (e.g., a port can be configured to go through a resource but isn't required)

The Link Set Field is defined in [Gen-Encode].

Note that the direction parameter within the Link Set Field is used to indicate whether the link set is an ingress or egress link set, and the bidirectional value for this parameter is not permitted in this sub-TLV.

See Appendix A.1 for an illustration of this encoding.

3.2. Resource Block Wavelength Constraints Sub-TLV

Resources, such as wavelength converters, etc., may have a limited input or output wavelength ranges. Additionally, due to the structure of the optical system not all wavelengths can necessarily reach or leave all the resources. These properties are described by using one or more resource wavelength restrictions sub-TLVs as defined below:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     RB Set Field                                     |
|                                                                                     |
:                                                                                     :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Input Wavelength Set Field                       |
|                                                                                     |
:                                                                                     :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Output Wavelength Set Field                     |
|                                                                                     |
:                                                                                     :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

RB Set Field:

A set of resource blocks (RBs) which have the same wavelength restrictions.

Input Wavelength Set Field:

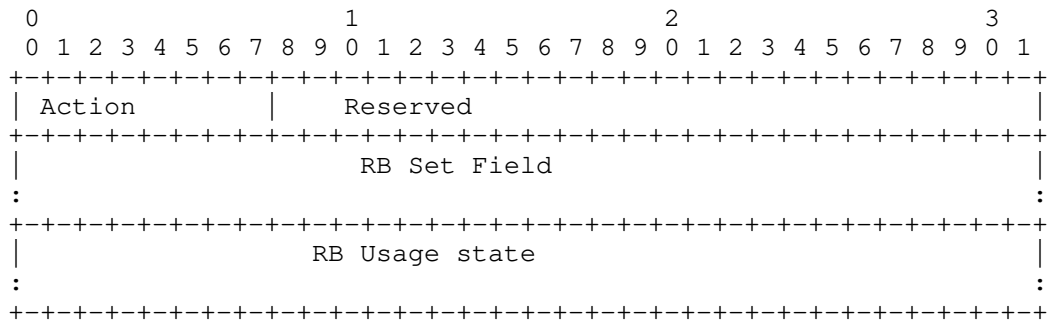
Indicates the wavelength input restrictions of the RBs in the corresponding RB set.

Output Wavelength Set Field:

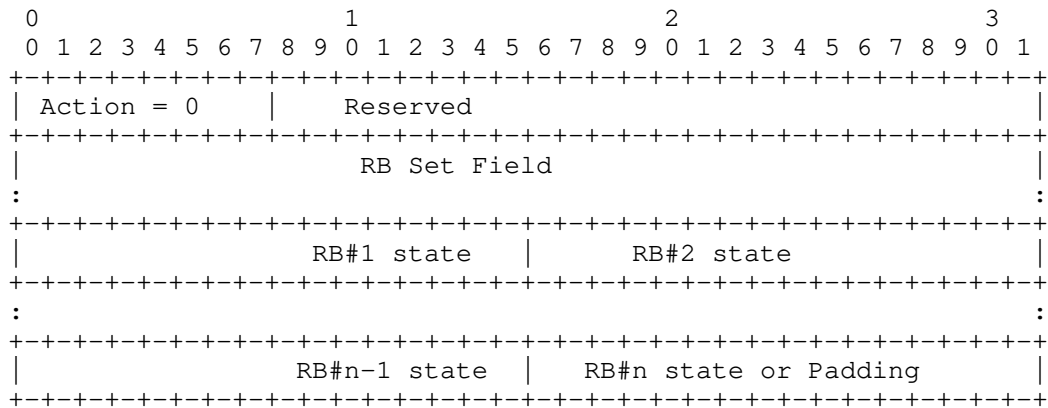
Indicates the wavelength output restrictions of RBs in the corresponding RB set.

3.3. Resource Pool State Sub-TLV

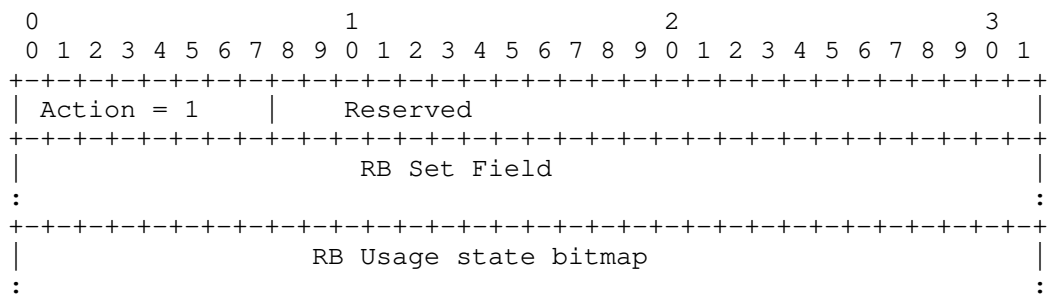
The state of the pool is given by the number of resources available in each block. The usage state of resources within a block is encoded as either a list of 16 bit integer values or a bit map indicating whether a single resource is available or in use. The bit map encoding is appropriate when resource blocks consist of a single resource. This information can be relatively dynamic, i.e., can change when a connection is established or torn down.



Where Action = 0 denotes a list of 16 bit integers and Action = 1 denotes a bit map. In both cases the elements of the RB Set field are in a one-to-one correspondence with the values in the usage RB usage state area.



Whether the last 16 bits is a wavelength converter (RB) state or padding is determined by the number of elements in the RB set field.



```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               .....                               | Padding bits |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

RB Usage state: Variable Length but must be a multiple of 4 bytes.

Each bit indicates the usage status of one RB with 0 indicating the RB is available and 1 indicating the RB is in used. The sequence of the bit map is ordered according to the RB Set field with this sub-TLV.

Padding bits: Variable Length

3.4. Block Shared Access Wavelength Availability sub-TLV

Resources blocks may be accessed via a shared fiber. If this is the case then wavelength availability on these shared fibers is needed to understand resource availability.

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| I | E |                               Reserved                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               RB Set Field                               |
:                                                                           :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Ingress Available Wavelength Set Field       |
:                               (Optional)                                   :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Egress Available Wavelength Set Field         |
:                               (Optional)                                   :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

I bit:

Indicates whether the ingress available wavelength set field is included (1) or not (0).

E bit:

Indicates whether the egress available wavelength set field is included (1) or not (0).

RB Set Field:

A Resource Block set in which all the members share the same ingress or egress fiber or both.

Ingress Available Wavelength Set Field:

Indicates the wavelengths currently available (not being used) on the ingress fiber to this resource block.

Egress Available Wavelength Set Field:

Indicates the wavelengths currently available (not being used) on the egress fiber from this resource block.

4. Resource Properties Encoding

Within a WSON network element (NE) there may be resources with signal compatibility constraints. Such resources typically come in "blocks" which contain a group of identical and indistinguishable individual resources. These resource blocks may consist of regenerators, wavelength converters, etc... Such resource blocks may also constitute the network element as a whole as in the case of an electro optical switch. This section primarily focuses on the signal compatibility and processing properties of such a resource block, the accessibility aspects of a resource in a shared pool, except for the shared access indicators, were encoded in the previous section.

The fundamental properties of a resource block, such as a regenerator or wavelength converter, are:

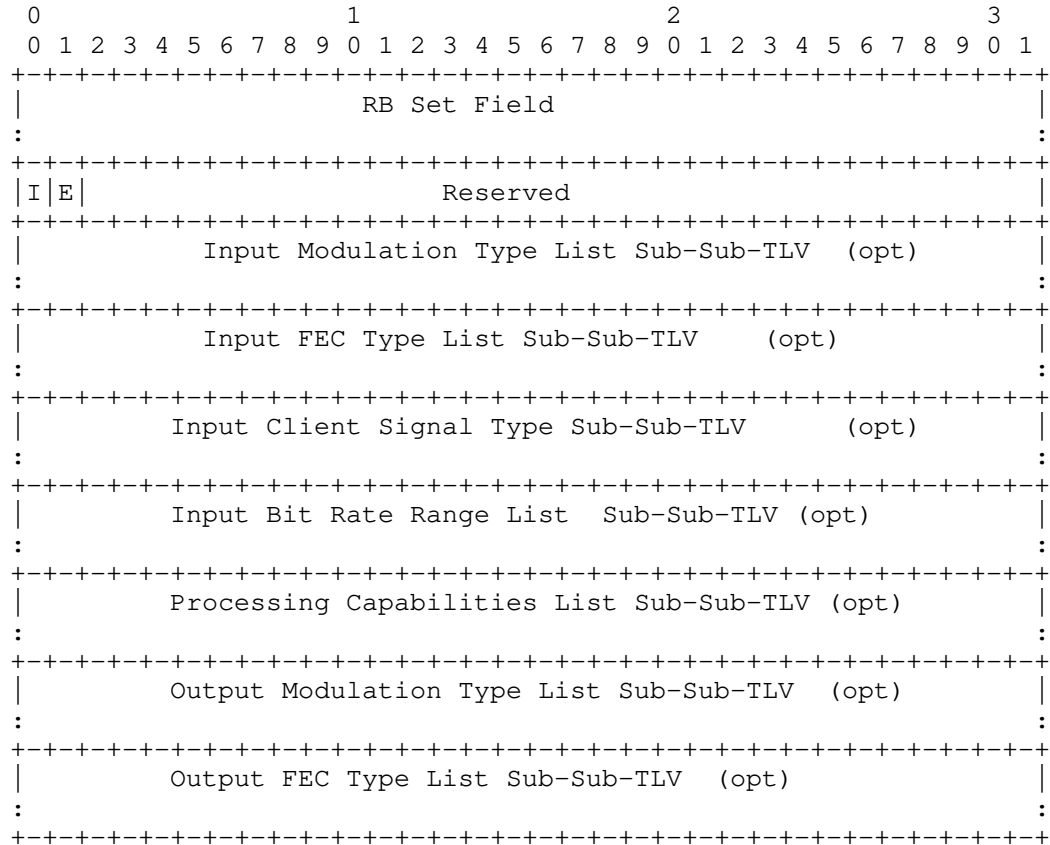
- (a) Input constraints (shared ingress, modulation, FEC, bit rate, GPID)
- (b) Processing capabilities (number of resources in a block, regeneration, performance monitoring, vendor specific)
- (c) Output Constraints (shared egress, modulation, FEC)

4.1. Resource Block Information Sub-TLV

Resource Block descriptor sub-TLVs are used to convey relatively static information about individual resource blocks including the

resource block properties of section 3. and the number of resources in a block.

This sub-TLV has the following format:



Where I and E, the shared ingress/egress indicator, is set to 1 if the resource blocks identified in the RB set field utilized a shared fiber for ingress/egress access and set to 0 otherwise.

4.2. Input Modulation Format List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable input modulation formats.

Type := Input Modulation Format List

Value := A list of Modulation Format Fields

4.2.1. Modulation Format Field

Two different types of modulation format fields are defined: a standard modulation field and a vendor specific modulation field. Both start with the same 32 bit header shown below.

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|S|I|          Modulation ID          |          Length          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where S bit set to 1 indicates a standardized modulation format and S bit set to 0 indicates a vendor specific modulation format. The length is the length in bytes of the entire modulation type field.

Where I bit set to 1 indicates it is an input modulation constraint and I bit set to 0 indicates it is an output modulation constraint.

Note that if an output modulation is not specified then it is implied that it is the same as the input modulation. In such case, no modulation conversion is performed.

The format for the standardized type for the input modulation is given by:

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|1|          Modulation ID          |          Length          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Possible additional modulation parameters depending upon |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:   the modulation ID                                     :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Modulation ID (S bit = 1); Input modulation (I bit = 1)

Takes on the following currently defined values:

0 Reserved

- 1 optical tributary signal class NRZ 1.25G
- 2 optical tributary signal class NRZ 2.5G
- 3 optical tributary signal class NRZ 10G
- 4 optical tributary signal class NRZ 40G
- 5 optical tributary signal class RZ 40G

Note that future modulation types may require additional parameters in their characterization.

The format for vendor specific modulation field (for input constraint) is given by:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0|1| Vendor Modulation ID | Length |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Enterprise Number |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
: Any vendor specific additional modulation parameters :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Vendor Modulation ID

This is a vendor assigned identifier for the modulation type.

Enterprise Number

A unique identifier of an organization encoded as a 32-bit integer. Enterprise Numbers are assigned by IANA and managed through an IANA registry [RFC2578].

Vendor Specific Additional parameters

There can be potentially additional parameters characterizing the vendor specific modulation.

4.3. Input FEC Type List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable FEC types.

Type := Input FEC Type field List

Value := A list of FEC type Fields

4.3.1. FEC Type Field

The FEC type Field may consist of two different formats of fields: a standard FEC field or a vendor specific FEC field. Both start with the same 32 bit header shown below.

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
+-----+			
S I	FEC ID	Length	
+-----+			
Possible additional FEC parameters depending upon			
+-----+			
: the FEC ID		:	
+-----+			

Where S bit set to 1 indicates a standardized FEC format and S bit set to 0 indicates a vendor specific FEC format. The length is the length in bytes of the entire FEC type field.

Where I bit set to 1 indicates it is an input FEC constraint and I bit set to 0 indicates it is an output FEC constraint.

Note that if an output FEC is not specified then it is implied that it is the same as the input FEC. In such case, no FEC conversion is performed.

The length is the length in bytes of the entire FEC type field.

The format for input standard FEC field is given by:

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|1|          FEC ID          |          Length          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Possible additional FEC parameters depending upon          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:   the FEC ID   :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

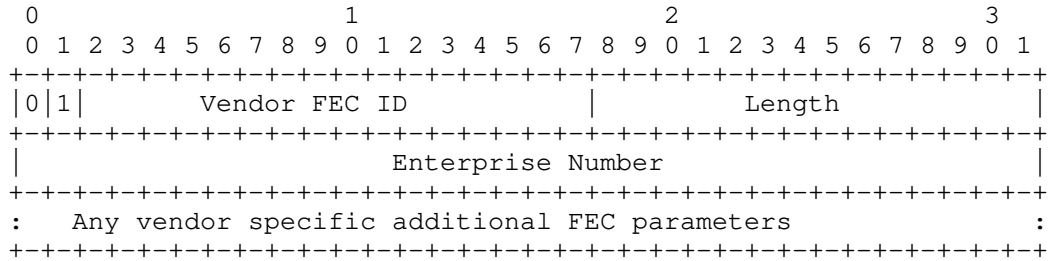
```

Takes on the following currently defined values for the standard FEC ID:

- | | |
|----|---|
| 0 | Reserved |
| 1 | G.709 RS FEC |
| 2 | G.709V compliant Ultra FEC |
| 3 | G.975.1 Concatenated FEC
(RS(255,239)/CSOC(n0/k0=7/6, J=8)) |
| 4 | G.975.1 Concatenated FEC (BCH(3860,3824)/BCH(2040,1930)) |
| 5 | G.975.1 Concatenated FEC (RS(1023,1007)/BCH(2407,1952)) |
| 6 | G.975.1 Concatenated FEC (RS(1901,1855)/Extended Hamming
Product Code (512,502)X(510,500)) |
| 7 | G.975.1 LDPC Code |
| 8 | G.975.1 Concatenated FEC (Two orthogonally concatenated
BCH codes) |
| 9 | G.975.1 RS(2720,2550) |
| 10 | G.975.1 Concatenated FEC (Two interleaved extended BCH
(1020,988) codes) |

Where RS stands for Reed-Solomon and BCH for Bose-Chaudhuri-Hocquengham.

The format for input vendor-specific FEC field is given by:



Vendor FEC ID

This is a vendor assigned identifier for the FEC type.

Enterprise Number

A unique identifier of an organization encoded as a 32-bit integer. Enterprise Numbers are assigned by IANA and managed through an IANA registry [RFC2578].

Vendor Specific Additional FEC parameters

There can be potentially additional parameters characterizing the vendor specific FEC.

4.4. Input Bit Range List Sub-Sub-TLV

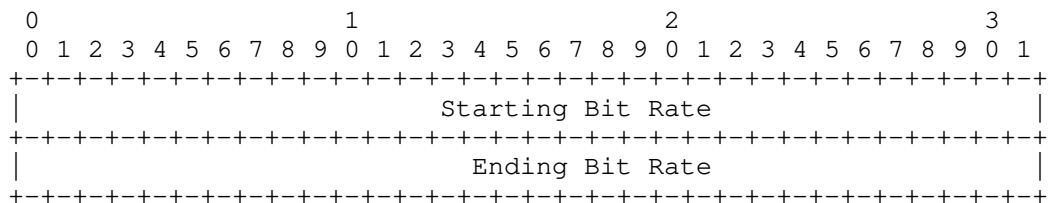
This sub-sub-TLV contains a list of acceptable input bit rate ranges.

Type := Input Bit Range List

Value := A list of Bit Range Fields

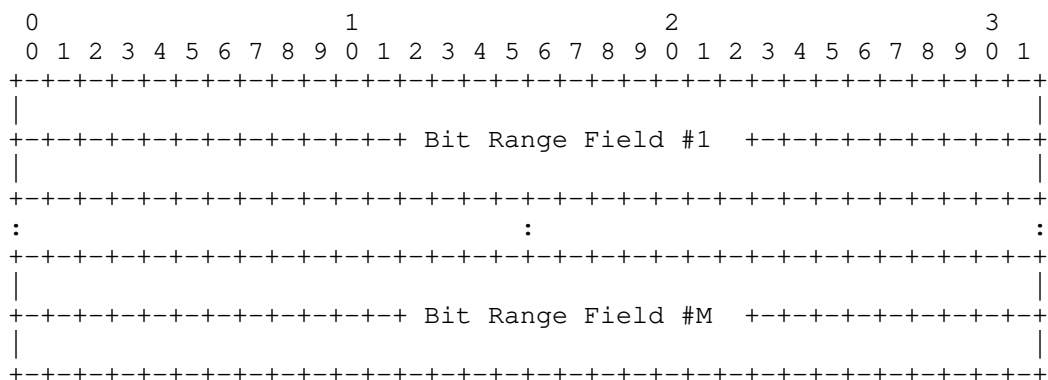
4.4.1. Bit Range Field

The bit rate range list sub-TLV makes use of the following bit rate range field:



The starting and ending bit rates are given as 32 bit IEEE floating point numbers in bits per second. Note that the starting bit rate is less than or equal to the ending bit rate.

The bit rate range list sub-TLV is then given by:



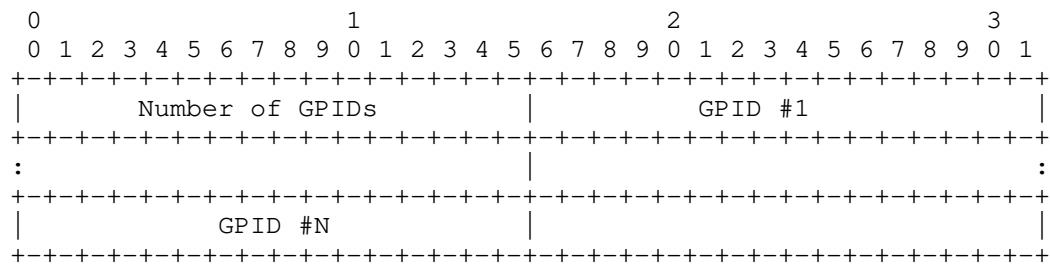
4.5. Input Client Signal List Sub-Sub-TLV

This sub-sub-TLV contains a list of acceptable input client signal types.

```
Type := Input Client Signal List
```

Value := A list of GPIDs

The acceptable client signal list sub-TLV is a list of Generalized Protocol Identifiers (GPIDs). GPIDs are assigned by IANA and many are defined in [RFC3471] and [RFC4328].



4.6. Processing Capability List Sub-Sub-TLV

Value := A list of Processing Capabilities Fields

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

4.6.1. Processing Capabilities Field

The processing capability field is then given by:

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Processing Cap ID                               |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Possible additional capability parameters depending upon                       |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
:   the processing ID                                                           :
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

When the processing Cap ID is "number of resources" the format is simply:

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Processing Cap ID                               |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Number of resources per block                   |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

When the processing Cap ID is "regeneration capability", the following additional capability parameters are provided in the sub-TLV:

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|  T  |  C  |                               Reserved                               |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

Where T bit indicates the type of regenerator:

T=0: Reserved

T=1: 1R Regenerator

T=2: 2R Regenerator

T=3: 3R Regenerator

Where C bit indicates the capability of regenerator:

C=0: Reserved

C=1: Fixed Regeneration Point

C=2: Selective Regeneration Point

Note that when the capability of regenerator is indicated to be Selective Regeneration Pools, regeneration pool properties such as ingress and egress restrictions and availability need to be specified. This encoding is to be determined in the later revision.

4.7. Output Modulation Format List Sub-Sub-TLV

This sub-sub-TLV contains a list of available output modulation formats.

Type := Output Modulation Format List

Value := A list of Modulation Format Fields

4.8. Output FEC Type List Sub-Sub-TLV

This sub-sub-TLV contains a list of output FEC types.

Type := Output FEC Type field List

Value := A list of FEC type Fields

5. 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.

6. IANA Considerations

This document provides general protocol independent information encodings. There is no IANA allocation request for the TLVs defined in this document. IANA allocation requests will be addressed in protocol specific documents based on the encodings defined here.

7. Acknowledgments

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

APPENDIX A: Encoding Examples

A.1. Wavelength Converter Accessibility Sub-TLV

Example:

Figure 1 shows a wavelength converter pool architecture know as "shared per fiber". In this case the ingress and egress pool matrices are simply:

$$WI = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline 1 & 1 \\ \hline \end{array}, \quad WE = \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & 1 \\ \hline \end{array}$$

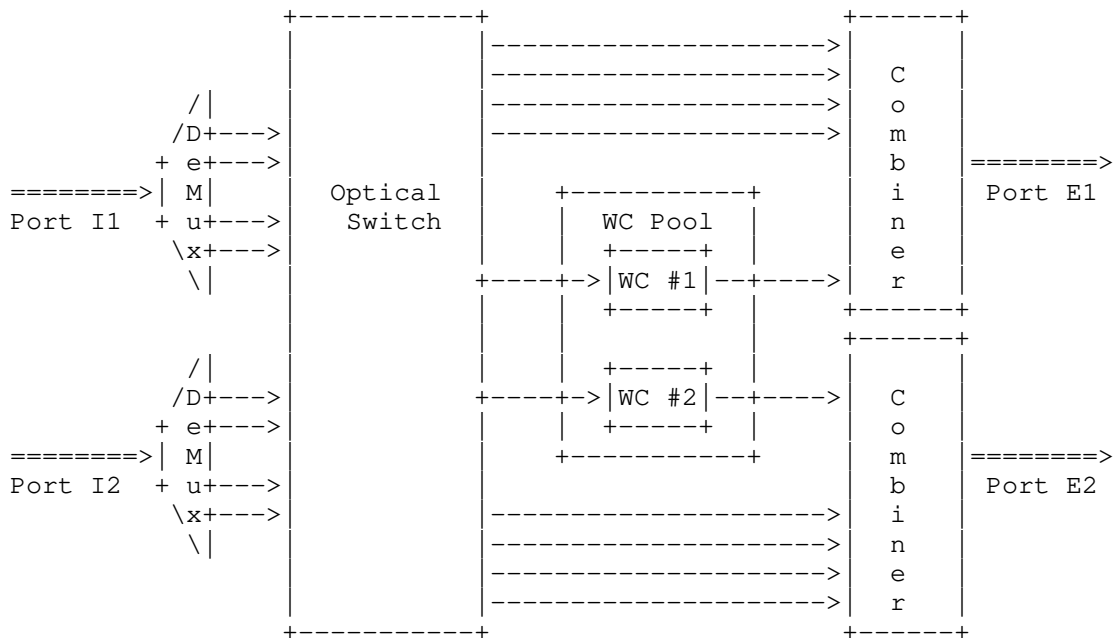


Figure 1 An optical switch featuring a shared per fiber wavelength converter pool architecture.

This wavelength converter pool can be encoded as follows:

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Connectivity=1 |                               Reserved |
|               | Note: I1,I2 can connect to either WC1 or WC2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 0 1 | 0 0 0 0 0 0 |               Length = 12 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Link Local Identifier = #1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Link Local Identifier = #2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 1 | Reserved |               Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | RB ID = #1 |               RB ID = #2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Note: WC1 can only connect to E1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 1 0 | 0 0 0 0 0 0 |               Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Link Local Identifier = #1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 0 | Reserved |               Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | RB ID = #1 |               zero padding |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Note: WC2 can only connect to E2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 1 0 | 0 0 0 0 0 0 |               Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | Link Local Identifier = #2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0      | 0 |               Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|               | RB ID = #2 |               zero padding |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

A.2. Wavelength Conversion Range Sub-TLV

Example:

This example, based on figure 1, shows how to represent the wavelength conversion range of wavelength converters. Suppose the

wavelength range of input and output of WC1 and WC2 are {L1, L2, L3, L4}:

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      Note: WC Set
+-----+-----+-----+-----+-----+-----+-----+-----+
| Action=0 | 1 | Reserved | Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| WC ID = #1 | WC ID = #2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
      Note: wavelength input range
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2 | Num Wavelengths = 4 | Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Grid | C.S. | Reserved | n for lowest frequency = 1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
      Note: wavelength output range
+-----+-----+-----+-----+-----+-----+-----+-----+
| 2 | Num Wavelengths = 4 | Length = 8 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Grid | C.S. | Reserved | n for lowest frequency = 1 |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

A.3. An OEO Switch with DWDM Optics

Figure 2 shows an electronic switch fabric surrounded by DWDM optics. In this example the electronic fabric can handle either G.709 or SDH signals only (2.5 or 10 Gbps). To describe this node, the following information is needed:

```
<Node_Info> ::= <Node_ID>[Other GMPLS sub-
TLVs][<ConnectivityMatrix>...] [<ResourcePool>][<RBPoolState>]
```

In this case there is complete port to port connectivity so the <ConnectivityMatrix> is not required. In addition since there are sufficient ports to handle all wavelength signals the <RBPoolState> element is not needed.

Hence the attention will be focused on the <ResourcePool> sub-TLV:

```
<ResourcePool> ::=
<ResourceBlockInfo>[<ResourceBlockAccessibility>...][<ResourceWaveCon-
straints>...]
```

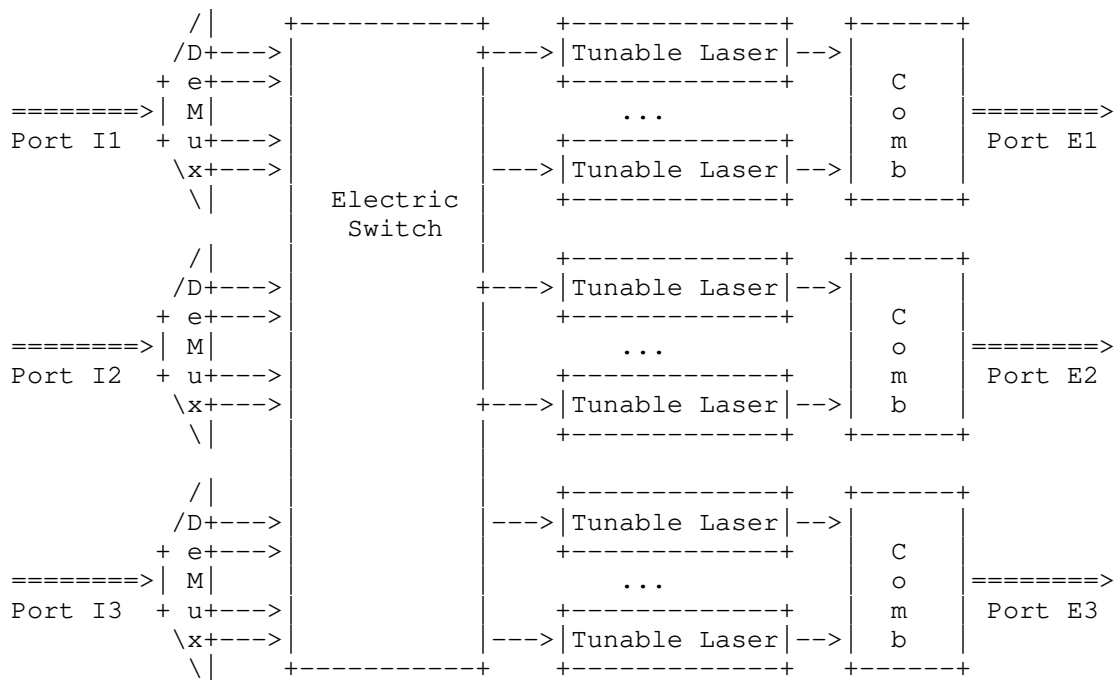


Figure 2 An optical switch built around an electronic switching fabric.

The resource block information will tell us about the processing constraints of the receivers, transmitters and the electronic switch. The resource availability information, although very simple, tells us that all signals must traverse the electronic fabric (fixed connectivity). The resource wavelength constraints are not needed since there are no special wavelength constraints for the resources that would not appear as port/wavelength constraints.

<ResourceBlockInfo>:


```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Connectivity=1 | Reserved      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Ingress Link Set Field A #1      |
:                                     (All ingress links connect to resource) :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     RB Set Field A #1                  |
:                                     (trivial set only one resource block) :
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Egress Link Set Field B #1         |
:                                     (All egress links connect to resource) :
+-----+-----+-----+-----+-----+-----+-----+-----+
```

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC4328] Papadimitriou, D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [G.694.1] ITU-T Recommendation G.694.1, "Spectral grids for WDM applications: DWDM frequency grid", June, 2002.

8.2. Informative References

- [G.694.1] ITU-T Recommendation G.694.1, Spectral grids for WDM applications: DWDM frequency grid, June 2002.
- [G.694.2] ITU-T Recommendation G.694.2, Spectral grids for WDM applications: CWDM wavelength grid, December 2003.
- [Gen-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", work in progress: draft-ietf-ccamp-general-constraint-encode.
- [Otani] T. Otani, H. Guo, K. Miyazaki, D. Caviglia, "Generalized Labels for G.694 Lambda-Switching Capable Label Switching Routers", work in progress: draft-ietf-ccamp-gmpls-g-694-lambda-labels.
- [WSO-Frame] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-wavelength-switched-framework.

[WSO-Info] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-info.

9. 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.

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 2, 2012

R. Kunze, Ed.
Deutsche Telekom AG
July 1, 2011

A framework for Management and Control of optical interfaces supporting
G.698.2
draft-kunze-g-698-2-management-control-framework-00

Abstract

This document provides a framework that describes a solution space for the control and management of optical interfaces according to the Black Link approach as specified by ITU-T [ITU.G698.2] and further revisions. In particular, it examines topological elements and related network management measures.

Optical Routing and Wavelength assignment based on WSON is out of scope. This document concentrates on the management of optical interfaces. The application of a dynamic control plane, e.g. for auto-discovery or for the distribution of interface parameters, is complementary. Anyway, this work is not in conflict with WSON but leverages and supports related work already done for management plane and control plane.

The framework document will not address the client mapping into G.709. This document only addresses the lower layers. Furthermore, support for Fast Fault Detection, to e.g. trigger Protection Switching is provided by the WDM interface capability of the client interface (e.g. ITU-T G.709) is out of scope for 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.

Status of This Memo

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

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

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

material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 2, 2012.

Copyright Notice

Copyright (c) 2011 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	5
2. Terminology and Definitions	5
3. DWDM Black Link Management Solution Space	6
3.1. Description of Client Network Layer – WDM connection	7
3.1.1. Traditional WDM deployments	7
3.1.2. Black Link Deployments	8
4. Black Link Operation scenarios	10
4.1. Bringing into service	10
4.2. Configuration Management	10
4.3. In service (performance management)	10
4.4. Fault Clearance	10
5. Black Link Control and Management Solutions	10
5.1. BL Separate Operation and Management Approaches	11
5.1.1. Direct connection to the management system	12
5.1.2. Indirect connection to the WDM management system	14
5.2. Control Plane Considerations	15
5.2.1. Black Link deployment with common control plane	15
5.2.2. Black Link deployment with an separate control plane	16
6. Requirements for BL and FW deployments	16
6.1. Interoperability Aspects	16
7. Acknowledgements	17
8. IANA Considerations	18
9. Security Considerations	18
10. Contributors	18
11. References	18
11.1. Normative References	18
11.2. Informative References	19

1. Introduction

The usage of the Black Link approach in carrier long haul and aggregation networks adds a further option for operators to facilitate their networks. The integration of optical coloured interfaces into routers and other types of clients could lead to a lot of benefits regarding an efficient and optimized data transport for higher layer services.

Carriers deploy their networks today as a combination of transport and packet infrastructure. This ensures high available and flexible data transport. Both network technologies are managed usually 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 coloured interface moves into the client (e.g. router), it is necessary to establish a management connection between the client providing the coloured interface and the corresponding EMS (Element Management System) of the transport network to ensure that the coloured interface parameters can be managed in the same way as traditional deployments allow this.

The objective of this document is to provide a framework that describes the solution space for the control and management of WDM Black Links as specified by ITU-T [ITU.G698.2] and further revisions. In particular, it examines topological elements and related network management measures.

Optical Routing and Wavelength assignment based on WSON is out of scope. This document concentrates on the management of optical interfaces. The application of a dynamic control plane, e.g. for auto-discovery or distribute interface parameters, is complementary. Anyway, this work is not in conflict with WSON but leverages and supports related work already done for management plane and control plane.

Furthermore, support for Fast Fault Detection, to e.g. trigger Protection Switching is provided by the WDM interface capability of the client interface (e.g. ITU-T G.709) is out of scope for 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 Control and Management Plane 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 colored interfaces.

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. Terminology and Definitions

Black Link: The Black Link [ITU.G698.2] allows supporting an optical transmitter/receiver pair of one or different vendors to inject a DWDM channel and run it over an optical network composed of amplifiers, filters, add-drop multiplexers from a different vendor. Therefore the standard defines the ingress and egress parameters at the interface Ss and Rs.

Coloured Interface: The term coloured interface defines the single channel optical interface that is used to bridge long distances and is directly connected with a DWDM system. Coloured interfaces operate on a fix wavelength or within a wavelength band (tunability). Coloured interface is a more generic term and it is a superset of the Black Link.

Friendly Wavelength: A friendly wavelength is a wavelength that is generated or originated by an optical interface that is not part of the WDM system but completely managed and known by the WDM system.

Alien Wavelength: Alien Wavelength: An alien wavelength is a wavelength that is generated or originated by an optical interface that is not part of the WDM system and not managed and known by the WDM system.

Forward error correction (FEC): FEC is an important way of improving the performance of high-capacity long haul 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).

Intra-domain Interface (IaDI): The intra-domain interface (line site of the optical system) is a physical interface within an optical administrative or vendor domain and is implemented as:

- a. standardized single channel interface specified according to G.698.2 (standardized optical interface AND OTUk according G.709) or
- b. proprietary single channel interface proprietary optical interface OR functionally specified OTUkV according G.709, i.e. proprietary FEC.

Inter-Domain Interface(IrDI): The inter-domain interface is a physical interface that represents the boundary between two administrative domains as well as the boundary between client and optical domain.

Management Plane: Management Plane: The management plane supports FCAPS (Fault, Configuration, Accounting, Performance and Security Management) capabilities for carrier networks.

Control Plane: The control plane supports signalling, path computation, routing, path setup and restoration.

Client Network Layer: The client network layer is the layer above (on top) the WDM layer, from the perspective of the WDM layer.

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. DWDM Black Link Management Solution Space

Basically the management of optical interfaces using a Black Link deals with aspects needed for setup, tear down and maintenance of wavelengths and all related optical parameters, which are demanded by a client network layer (the layer above WDM). The following types of WDM networks are considered for a management of optical interfaces using a black link:

- a. Passive WDM
- b. Legacy point to point WDM systems
- c. Legacy WDM systems with OADMs
- d. Transparent optical networks supporting specific IPoDWDM functions, interfaces or protocols

Table 1 provides a list of tasks, which are related to BL management, It is indicated which domain (optical or client) is responsible for a task. The relevance of a task for each type of WDM network is also indicated.

Task	Domain	a	b	c	d
determination of centre frequency	client	R	R	R	R
configuration of centre frequency at colored IF	optical	NR	NR	R	R
path computation of wavelength	optical	NR	NR	R	R
routing of wavelength	optical	NR	NR	R	R
wavelength setup across optical network	client	?	?	R	R
detection of wavelength fault	optical	R	R	R	R
fault isolation, identification of root failure	optical	NR	R	R	R
repair actions within optical network	optical	R	R	R	R
protection switching of wavelength	optical	NR	NR	R	R
restoration of wavelength	optical	NR	NR	R	R

Table 1: List of tasks related to BL management

Furthermore the following deployment cases will be considered:

- a. Exclusive Black Link deployment
- b. Black Link deployment in combination with grey client network interfaces

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

3.1. Description of Client Network Layer - WDM connection

3.1.1. Traditional WDM deployments

The ordinary connection of a client layer network towards a WDM system is based today on client interfaces (grey) bridging short or intermediate distances between client and WDM system. The Optical Signal incoming into the WDM system must be converted (OEO conversion) to corresponding WDM wavelength grid and the power level that is applicable for the WDM transmission path. This conversion is done by a component termed as transponder (see Figure 1).

After that OEO conversion the signal complies with the parameters that are specified for a certain WDM link.

Figure 1 shows the traditional Client - WDM interconnection using transponders for wavelength conversion. IrDI and IaDI as defined in Section 2 specifying the different demarcation areas related to external and internal connections

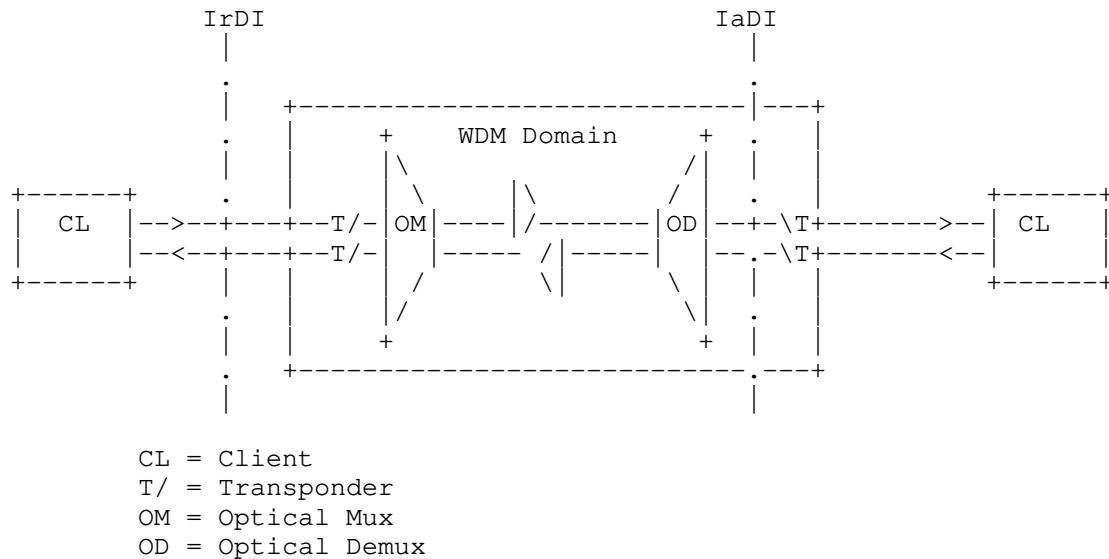


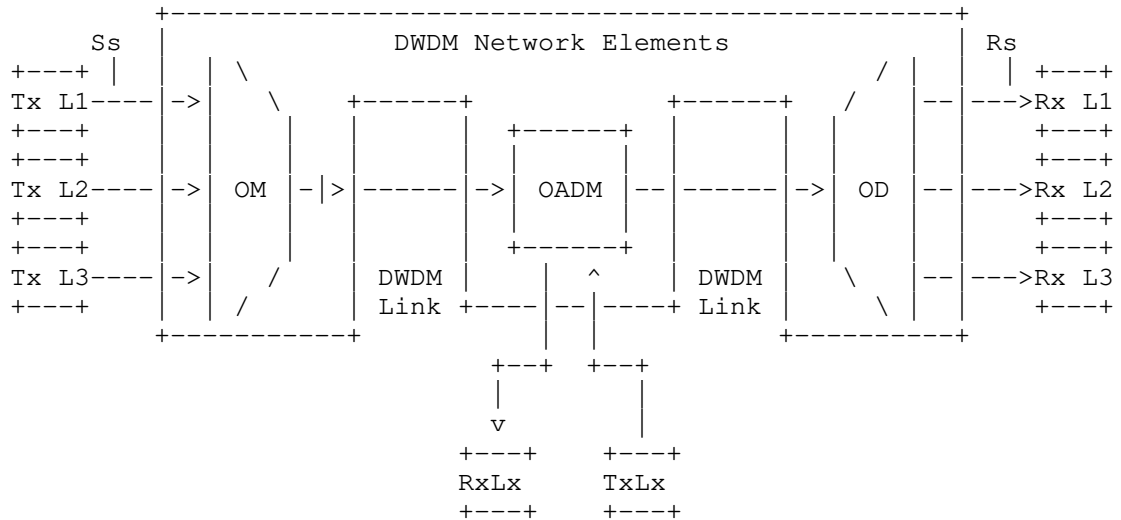
Figure 1: Inter and Intra-Domain Interface Identification

The management and control of WDM and client layer is done by different control and management solutions. Different operational units are responsible for client and WDM layer.

3.1.2. Black Link Deployments

In case of a black link deployment Figure 2 the DWDM transceiver is located directly at the client and the grey interfaces will be saved. In that case a solution must be found to manage that coloured interface in the same way as in the traditional case. This requirement must be fulfilled especially in the cases where legacy equipment and Black Link Wavelength interfaces will be used in parallel or together and the operational situation is unchanged.

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 WDM 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 2: Linear Black Link

Independent from the WDM networks that are considered the usage of colored interfaces must perform as well in mixed setups with both legacy and colored interface equipment using the BL.

4. Black Link Operation scenarios

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

tbd.

4.2. Configuration Management

tbd.

4.3. In service (performance management)

tbd.

4.4. Fault Clearance

tbd.

5. Black Link Control and Management Solutions

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. Each WDM vendor provides a management system that also administrates the wavelengths.

This old operational approach was predicted on a high amount/rate of connection oriented traffic in carrier networks. This behaviour has been changed completely. Today IP is the dominating traffic in the network and from the operating perspective it is more beneficial to use a common management and operation approach. Due to a long history of operational separation it must be possible to manage and operate Black Link deployments with the traditional approach too.

Therefore from the operational point of view in a pure Black Link or in a mixed setup with legacy equipment (transponders) there are two approaches to manage and operate the network.

1. Separate operation and management of client and Transport network
 - a. Direct link to the management system (e.g. EMS, OSS)
 - b. Indirect link to the management system; using a protocol

between the peer node and the directly connected WDM system node to exchange management information

2. Common operation and management of IP and 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 option should be favoured if the underlying WDM transport network is mainly used to interconnect IP nodes and the service creation and restoration will be done on higher layers (e.g. IP/MPLS). Then it is more beneficial have a higher level of integration and a common management will be more efficient.

5.1. BL Separate Operation and Management Approaches

5.1.1.1. Direct connection to the management system

As depicted in Figure 3 one possibility to manage the optical interface within the client is a direct connection to the management system of the optical domain. This ensures manageability as usual.

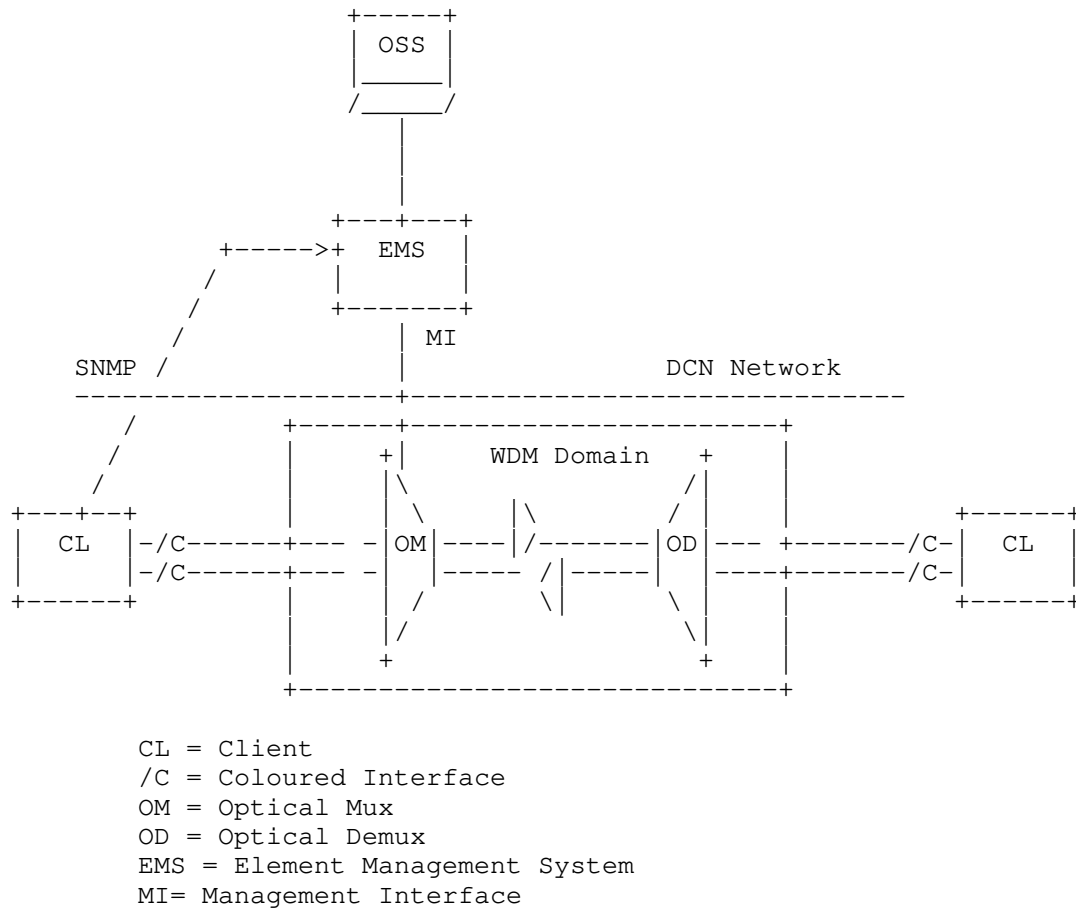


Figure 3: Connecting BL on Transport Management

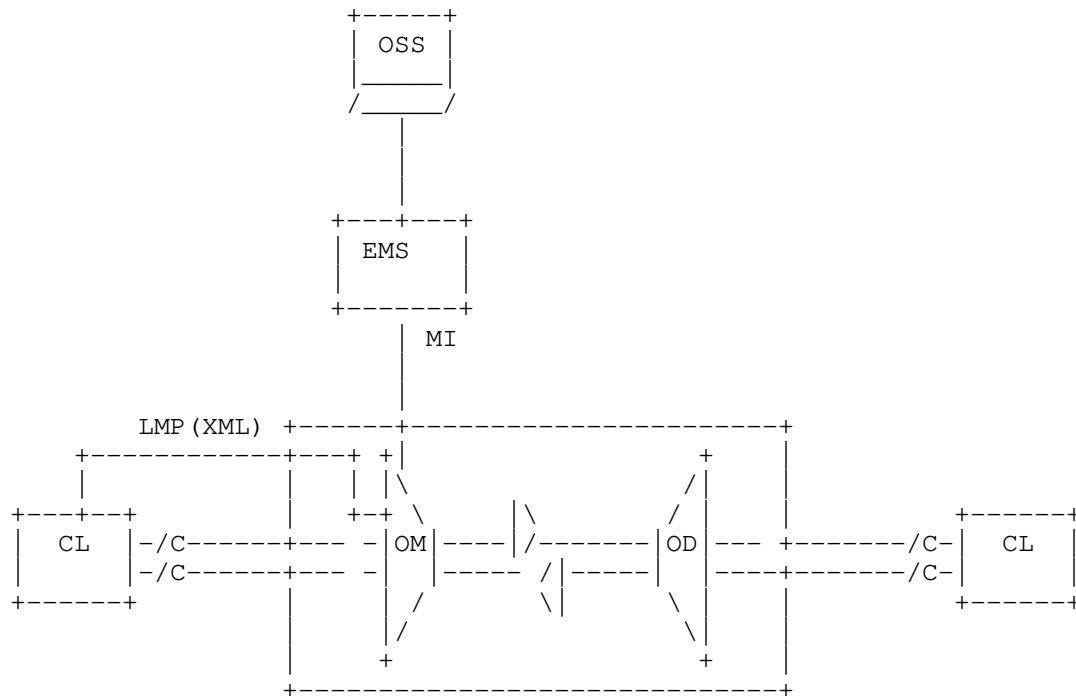
The exchange of management information between client and management system assumes that some form of a direct link exists between the client node and the WDM management system (e.g. EMS). This may be an Ethernet Link or a DCN connection.

It must be ensured that the optical interface can be managed in a standardized way to enable interoperable solutions between different optical interface vendors and vendors of the optical network management software. RFC 3591 [RFC3591] defines manage objects for the optical interface type but does not cover the scenarios described 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 client and management system of the WDM domain.

Note that a software update of the interface components of the client does not lead obligatory to an update of the software of the EMS and vice versa.

5.1.2. Indirect connection to the WDM management system

The alternative as shown in Figure 4 can be used in cases where a more automated relationship between transport node and router is aspired. In that case a combination of rudimentary control plane features and manual management will be used. It is a first step into a more control plane oriented operation model.



CL = Client
 /C = Coloured 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 client 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 support the transmission of configuration data (information). This functionality has to be added to the existing extensions of the protocol. The use 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 a DCN. It is proposed to use an out of band signalling over a separate link or DCN to ensure a high availability.

5.2. Control Plane Considerations

Basically it is not mandatory necessary to run a control plane in Black Link scenarios at least not in simple black link case where clients will be connected point to point using a simple WDM infrastructure (multiplexer and amplifier). As a first step it is possible to configure the entire link using the standard management system and a direct connection of the router or client to the EMS of the transport network. Configuration information will be exchanged using SNMP (see sections Section 5.1.1).

Looking at the control plane the following two scenarios may be considered:

- a. A common control plane for transport and client network; this implies a single operation unit responsible for both client and transport network management.
- b. A separate control plane for client and optical network without any interaction

As mentioned in chapter Section 5.1.2 some control plane features like LMP in an enhanced version could be used.

In such simple scenario it is imaginable to use only LMP to exchange information between the nodes of the optical domain. LMP must be run between the both end-points of the link and between the edge node and the first optical network node.

5.2.1. Black Link deployment with common control plane

tbd.

5.2.2. Black Link deployment with an separate control plane

tbd.

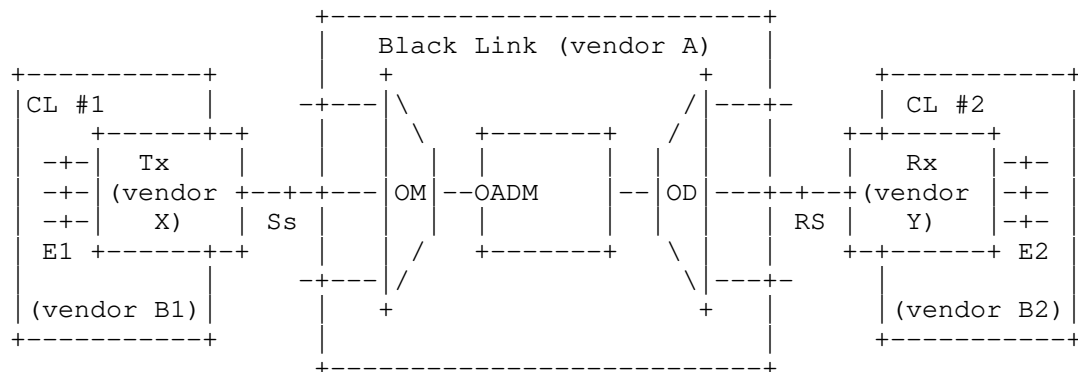
6. Requirements for BL and FW deployments

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. clients from different vendors and a WDM transport system using short-reach, grey interfaces. Applying the Black Link (BL) concept, clients (e.g. routers) now become 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 three different vendors, namely vendor A (black link designer), and vendors X and Y for the Tx/Rx. 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. The electrical interfaces of the Tx/Rx are indicated by E1/E2. The client equipment CL #1 containing the Tx is provided by a vendor B1 who is not necessarily the same vendor as the vendor X of the Tx. Multi-source agreements (MSA) for pluggable modules (e.g. for SFPs, XFPs, etc.) specify form factors and electrical interfaces thus enabling in principle to use a pluggable transmitter from any vendor who complies with the MSA. Similarly, the client equipment at the receiver side can be provided by a vendor B2 different from the vendor Y of the pluggable receiver. This modularity feature, i.e. the capability to use in principle different vendors for the client equipment, modular (pluggable) transmitters/receivers, and the black link is a key requirement addressing interoperability issues of the black link approach



CL = Client
 /C = Coloured Interface
 OM = Optical Mux
 OD = Optical Demux
 EMS= Element Management System
 MI= Management Interface

Figure 5: Interoperability aspects

In practice, a network operator may not use five different vendors when implementing black link systems. A simplified use case could be to choose the same vendor B for the client equipment on both sides (i.e. vendor B1 = vendor B2 = vendor B) and to choose the same vendor X for the Tx and Rx (i.e. vendor X = vendor Y) thus enabling to use universal pluggable modules for the optical transmitters and receivers.

An even more simplified use case could be to choose the same vendor B for all client equipment and Tx/Rx (i.e. B = B1 = X = B2 = Y) thus having only two vendors for the whole set-up, namely vendor A and vendor B, but to give up the possibility to use universal pluggable modules.

Other vendor combinations could also be realized (e.g. vendor X = vendor Y = vendor A).

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

This document has no requirement for a change to the security models within GMPLS, associated protocols and management interfaces. As well as the LMP security models could be operated unchanged.

10. Contributors

Arnold Mattheus
Deutsche Telekom
Darmstadt
Germany
email arnold.Mattheus@telekom.de

Manuel Paul
Deutsche Telekom
Berlin
Germany
email Manuel.Paul@telekom.de

Josef Roese
Deutsche Telekom
Darmstadt
Germany
email j.roese@telekom.de

Frank Luennemann
Deutsche Telekom
Muenster
Germany
email Frank.Luennemann@telekom.de

11. References

11.1. Normative References

- [ITU.G.872] International Telecommunications Union,
 "Architecture of optical transport networks", ITU-
 T Recommendation G.872, November 2001.
- [ITU.G698.2] International Telecommunications Union, "Amplified

multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.

- [ITU.G709] International Telecommunications Union, "Interface for the Optical Transport Network (OTN)", ITU-T Recommendation G.709, March 2003.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3591] Lam, H-K., Stewart, M., and A. Huynh, "Definitions of Managed Objects for the Optical Interface Type", RFC 3591, September 2003.
- [RFC4209] Fredette, A. and J. Lang, "Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems", RFC 4209, October 2005.

11.2. Informative References

- [Black-Link-MIB] Internet Engineering Task Force, "A SNMP MIB to manage the optical parameters characteristic of a DWDM Black-Link", draft-galimbe-kunze-black-link-mib-00 draft-galimbe-kunze-black-link-mib-00, July 2011.

Author's Address

Ruediger Kunze (editor)
Deutsche Telekom AG
Berlin, 10589
DE

Phone: +49 30 3497 3152
EMail: ruediger.kunze@telekom.de

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 5, 2012

Y. Li
F. Zhang
ZTE
R. Casellas
CTTC
July 4, 2011

Flexible Grid Label Format in Wavelength Switched Optical Network
draft-li-ccamp-flexible-grid-label-00

Abstract

Flexible grid is regarded as an efficient way to improve the network capacity utilization. Mixed bit rate transmission systems can allocate their channel with different spectral bandwidths so that they can be optimized for the bandwidth requirements of the particular bit rate and modulation scheme of the individual channels. To support the flexible grid technique, this document extends the wavelength label to accommodate this new specification. It is demonstrated that the extended label format is compatible to the rigid one and can be used in the routing and signaling procedure in the Wavelength Switched Optical Network (WSO).

Status of this Memo

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

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

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

This Internet-Draft will expire on January 5, 2012.

Copyright Notice

Copyright (c) 2011 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	3
3. Label format	4
3.1. Label values	4
3.2. Flexible Label	5
4. Flexible label applications	7
4.1. Application for Routing	7
4.2. Applications for Signaling	8
4.3. Applications for PCE	8
5. Acknowledgements	9
6. IANA Considerations	9
7. Security Considerations	9
8. References	9
8.1. Normative references	9
8.2. Informative References	9
Authors' Addresses	10

1. Introduction

Dense Wavelength Division Multiplexing (DWDM) optical network is widely deployed by telecom operators to carry their data service. With the continuing exponential growth of internet traffic, more efficient utilization of optical network bandwidth for extremely high data rates is required. Although multi-level modulation formats and advanced photonics techniques have enabled 100 G/s transmission within a 50 GHz DWDM fixed grid (or channel spacing), much higher speed traffic, such as 400 Gbit/s and 1 Tbit/s signals are not expected to adapt such a narrow channel. So a wider fixed grid like 100 GHz spacing is required to enable these new transmission formats without inter-channel crosstalk. However, the total available spectrum resource of the specific band is limited (about 4.4 THz in C band). If a wider grid is chosen, the fewer wavelengths can be allocated to carry the data. Not to mention that some low bitrate signals will occupy too much spectral bandwidth so that the total utilization efficiency of the spectrum resource is relatively low.

The recent revision of ITU-T Recommendation [G.694.1] has decided to introduce the flexible grid DWDM technique which provide a new tool that operators can implement to provide a higher degree of network optimization than fixed grid systems. Flexible grid network is composed of arbitrarily assigned spectral slices. That means in such networks the adjacent channel spacing and assigned spectral bandwidth per wavelength are variable. Mixed bitrate transmission systems can allocate their channel with different spectral bandwidths so that they can be optimized for the bandwidth requirements of the particular bit rate and modulation scheme of the individual channels. This technique is regarded to be a promising way to improve the network utilization efficiency and fundamentally reduce the cost of the IP core network.

Based on the DWDM technique, Wavelength Switched Optical Network (WSO) uses the control plane to dynamically provide Label Switched Paths (LSPs) for the requested end to end connections. The label switching is performed selectively on wavelength label representing the center wavelength/frequency of the optical signal. To support the flexible grid technique, this document extends the wavelength label defined in [RFC6205] to accommodate the new specification. It is proved that the extended label format is compatible to the rigid one and can be used in the routing and signaling procedure in WSON and generic GMPLS network.

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. Label format

3.1. Label values

The wavelength label format is defined in [RFC6205] and the corresponding wavelength or frequency value is referred to ITU-T Recommendations [G.694.1] and [G.694.2] for DWDM and CWDM grid respectively. The ITU-T fixed grid is based on nominal center frequency/wavelength.

For DWDM system, the nominal center frequency is calculated as:

Frequency (THz)=193.1 THz+n* channel spacing

In the context of rigid grid, the channel spacing of DWDM can support 12.5 GHz, 25 GHz, 50 GHz, or 100 GHz. However, once chosen, the adjacent channel spacing of the wavelengths is fixed. As mentioned in the section 1, 50 GHz channel spacing is most commonly used.

The recent revision of [G.694.1] has defined suggested values for the flexible DWDM grid. The concept of "frequency slot" is introduced to describe the frequency range allocated to a channel. A frequency slot is defined by its nominal central frequency and its required slot width values.

For the flexible DWDM grid, the allowed frequency slots have a nominal central frequency (in THz) defined by:

Frequency (THz)=193.1 THz + n * 0.00625

and a slot width (the same meaning as the spectral bandwidth) defined by:

12.5 GHz * m

where m is a positive integer.

The nominal center frequency representations of the fixed grid and flexible grid types are similar except that the latter has a more precise channel spacing granularity (6.25 GHz). Meanwhile the adjacent channel spacing (the spacing of the adjacent nominal center frequency) is implied to be (n1-n2) * 6.25 GHz, where n1 and n2 represent the n number defined above for the nominal center frequency of the adjacent frequency slots respectively (n is an integer

including positive, negative integer and 0). The slot width assigned to a frequency slot is arbitrary times of the slot width granularity. It was agreed on flexible grids with a granularity of 6.25 GHz for the central frequency and slot width of a multiple of 12.5 GHz. The slot width granularity is twice the channel spacing granularity, so that by carefully choosing n and m, the spectral resources can be allocated without leaving any gaps between slots. Therefore, in contrast to the rigid label, the new flexible label should have a capability to indicating the slot width allocation.

Note that in this document, the concepts "slot width" and "frequency slot" are similar to "spectral bandwidth" and "wavelength channel" respectively.

3.2. Flexible Label

To accommodate the new feature mentioned above, the wavelength label supporting flexible grid is illustrated as follows :

```

      0               1               2               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|Grid | C.S. | Identifier | n |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Additional slot width parameters |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Additional slot width parameters:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| m | Reserved |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Grid:

One new Grid type called "Flexible DWDM" is defined.

Grid	Value
Reserved	0
ITU-T DWDM	1
ITU-T CWDM	2

Flexible DWDM	3
Future use	4-7

C.S.:

For Grid=1 and 2, C.S. is referred to DWDM and CWDM channel spacing [RFC6205], which indicates that the adjacent channel spacing is constant. In this situation, the spectral bandwidth value allocated to every single channel is equal to value of the channel spacing.

For Grid=3, C.S. is referred to channel spacing granularity, accordingly the slot width granularity is twice of the C.S.. Minimum channel spacing granularity of 6.25 GHz with a slot width granularity of 12.5 GHz is supported.

C.S. (GHz)	Value
Reserved	0
100	1
50	2
25	3
12.5	4
6.25	5
Future use	6-15

Identifier:

The identifier field in the flexible label format is left unmodified compared with [RFC6205]. It is defined to distinguish which transmitter is used to carry the lambda. This identifier only has a local significance that should be indicated in the signaling message for LSP establishment. For routing information flooding, this field is meaningless and should be ignored on receipt.

n:

This field is used to compute the nominal center frequency/wavelength

of the channel mentioned above. Together with the channel spacing granularity (C.S.), the spacing of the adjacent channel is $(n1-n2) * 6.25$ GHz in flexible grid network (see definition of $n1$ and $n2$ in section 3.1).

Additional slot width parameters.

The slot width parameters field is mandatory only when Grid is set to 3 for flexible grid condition. These 5 bits field are used to represent how many slot width granularity the label has occupied. As the granularity is defined to be twice of the channel spacing granularity, so the slot width is calculated to be $m * 2 * C.S..$

4. Flexible label applications

This section illustrated the routing, signaling, PCE application of the extended flexible grid label.

4.1. Application for Routing

Flexible grid is regarded as an enabler for another kind of networks, requiring network elements, or nodes, that go past beyond the functional requirements of OXCs or ROADMs, in the sense that they do switching based on a frequency range. This means that a new swithing type called e.g. "Spectrum Selective Switching" in Interface Switching Capability Descriptor (ISCD) SHOULD be defined. However this is beyond the scope of this document and will be studied in the routing draft.

In addition to the topology information, wavelength constraints information like Port Label Restrictions, Shared Backup Labels, Resource Pool Wavelength Constrains, Resource Block Available Wavelengths detailed in [I-D.ietf-ccamp-rwa-info] should be flooded in the network through routing protocol like OSPF-TE. All the information is described by the label set object. The general label set is described in [RFC3471] and specific wavelength label set in [I-D.ietf-ccamp-general-constraint-encode] . There are 5 ways to represent the wavelength label set

1. Inclusive list
2. Exclusive list
3. Inclusive range
4. Exclusive range
5. Bitmap set

For flexible grid optical network, the label set should be more actually to represent the spectral resources constraints. For type 1

and 2, flexible label with different slot width is acceptable to put into the list. For type 3 and 4, start label and end label with minimal slot width (while it is not mandatory) is RECOMMENDED. For type 5, the base label/frequency slot is REQUIRED to have a minimum slot width ($m=1$). As there MAY exist some situations that the unused bandwidth between two occupied bandwidth is odd times of the channel spacing granularity (not integral times of the slot with granularity), two bits are needed to represent a single slot. 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 wavelength constraints flooding, any combination of the 5 types of label sets is feasible.

4.2. Applications for Signaling

In flexibel grid network, flexible label representing frequency "slots" or "ranges" rather than individual wavelengths is requested to establish the LSP. The extensions to the Genralized Label Request object and TSPEC object are needed, this will be studied in the future.

To establish a label switched path, an available wavelength label satisfying the wavelength continuity constraints is reserved with signaling protocol like RSVP-TE. For the flexible grid DWDM network, this procedure should be modified to assign available spectral resources. In other words, the label is not only assigning the nominal center frequency of wavelength but also the slot width for the LSP. The slot width is definitely clarified through the field *m* in the label. Nevertheless in the procedure, wavelength continuity constraint is unchanged.

4.3. Applications for PCE

[RFC6163] describes a Path Computation Element (PCE) can be used to performing routing and wavelength assignment in WSON. [RFC5440] details the path computation element communication protocol messages for this purpose. According to the modulation format, FEC type, client bitrates[I-D.ietf-ccamp-rwa-info][I-D.ietf-ccamp-rwa-wson-encode], and physical impairment, the required frequency slot indicated by flexible label should be calculated out by the PCE to carry the client signal.

5. Acknowledgements

6. IANA Considerations

A future revision of this document will present requests to IANA for codepoint allocation.

7. Security Considerations

8. References

8.1. Normative references

- [G.694.1] International Telecommunications Union, "Spectral grids for WDM applications: DWDM frequency grid", Recommendation G.694.1, June 2002 .
- [G.694.2] International Telecommunications Union, "Spectral grids for WDM applications: CWDM wavelength grid", Recommendation G.694.2, December 2003 .
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSOs)", RFC 6163, April 2011.
- [RFC6205] Otani, T. and D. Li, "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, March 2011.

8.2. Informative References

- [I-D.ietf-ccamp-general-constraint-encode]
Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "General

Network Element Constraint Encoding for GMPLS Controlled Networks", draft-ietf-ccamp-general-constraint-encode-05 (work in progress), May 2011.

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

Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info-11 (work in progress), March 2011.

[I-D.ietf-ccamp-rwa-wson-encode]

Bernstein, G., Lee, Y., Li, D., Imajuku, W., and J. Han, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-wson-encode-11 (work in progress), March 2011.

Authors' Addresses

Yao Li
ZTE
P.R.China

Phone: +86 025 52871109
Email: li.yao3@zte.com.cn

Zhang Fei
ZTE
P.R.China

Phone: +86 025 52871109
Email: zhang.fei3@zte.com.cn

Ramon Casellas
CTTC
Spain

Phone: +34 936452916
Email: ramon.casellas@cttc.es

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 5, 2012

G. Martinelli, Ed.
G. Galimberti
Cisco
July 4, 2011

WSON Optical Interface Class
draft-martinelli-wson-interface-class-00

Abstract

Current work on wavelength switched optical network includes several considerations regarding the interface signal compatibility. In particular ingress and egress optical interfaces will require a check on several optical parameters to assess if the signal generated by the ingress interface can be compatible with the receiving interface. Current solution available encode all parameters in WSON protocol extensions while in this draft will propose an alternative method to keep into account the signal compatibility issue at protocol level.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 5, 2012.

Copyright Notice

Copyright (c) 2011 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 and Procedures	4
3.2. Encoding	5
4. Optical Interface Class Semantic	6
5. Acknowledgements	6
6. IANA Considerations	6
7. Security Considerations	7
8. References	7
8.1. Normative References	7
8.2. Informative References	8
Appendix A. Encoding example	8
Authors' Addresses	8

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) are under development as well as new modulation formats and it is not clear if and when there will be a dominating technology. In a current realistic scenario DWDM optical networks manage different bit-rates as well as different modulation formats over the same link. So in general different signal characteristics will coexist at the same time.

To a further extent, the WSON activity will consider the case where the control plane has optical impairments awareness as detailed in [I-D.ietf-ccamp-wson-impairments]. The Control Plane function related to impairment awareness might require some additional interface parameters to assess the optical feasibility path. In such a case is likely further protocol extensions might be required just to add some parameters.

Scope of this draft is to propose an Optical Interface Class identifier as a solution for the WSON signal compatibility problem. To some extent the idea is have protocol extensions independent from optical technology evolution by keeping the semantic of optical characteristics separated from protocol scope. The final goal is a simplified but general representation rather than encoding saving.

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 list of parameters is considered:

- o Modulation Format. Only NRZ currently defined.
- o FEC, according to G.709 and G.975.
- o Bit Rate.

Note that this list of parameters is defined by ITU and might be subject to change due to internal physics.

The above encoding is going to be used within several WSON specific protocol extensions.

- o OSPF [I-D.ietf-ccamp-wson-signal-compatibility-ospf] since the path computation function need to consider optical interface parameters during the RWA process.
- o RSVP [I-D.ietf-ccamp-wson-signaling] since during the signaling phase there is the need to know optical ingress and egress interface properties (and eventually interfaces at regeneration point).
- o In addition, PCEP extension might need 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 and Procedures

The Optical Interface Class will be a unique number that identify all information related to optical characteristic's of such interface.

In term of RWA process the only operation required to assess the endpoint compatibility (interfaces or regeneration points) is to check if the two LSP endpoint have the same Class value. The procedure of signal compatibility assessment become just a numbers comparison: if two Optical Interface Class are equals the signal compatibility constrain is satisfied.

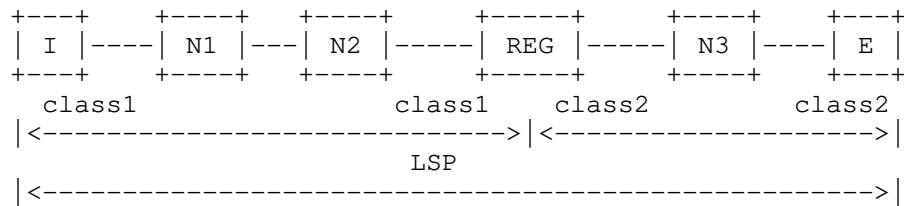


Figure 1

In case the RWA process will result in a path that need a wavelength conversion each interface involved in the wavelength conversion must satisfy the Optical Interface Class constrain. As represented in Figure 1, two different Optical Interface Classes are required for the given LSPs.

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 parameter values are standardized (e.g. new modulation formats become standard) the wson protocols and RWA don't need any extensions.

3.2. Encoding

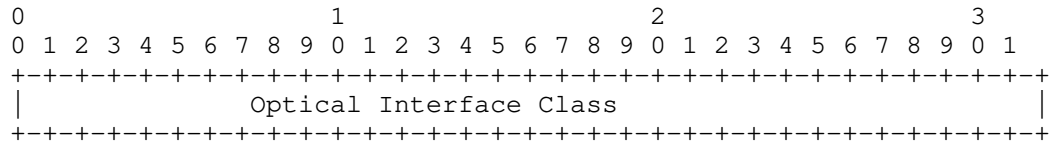


Figure 2: Optical Interface Class

The Optical Interface Class will be used in proper TLVs for different WSON protocol extensions .

In case an optical interface or a regeneration point will support multiple capabilities, a list of Interface Classes can be used as defined in [I-D.ietf-ccamp-rwa-wson-encode].

[Editor Note: the number could be 32 or 64 bits, size to be defined.]

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 elements. In this way the same class value will have the same meaning on every 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 a number represented by the Optical Interface Class. 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

This memo includes no request to IANA.

All drafts are required to have an IANA considerations section (see the update of RFC 2434 [I-D.narten-iana-considerations-rfc2434bis] for a guide). If the draft does not require IANA to do anything, the

section contains an explicit statement that this is the case (as above). If there are no requirements for IANA, the section will be removed during conversion into an RFC by the RFC Editor.

7. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

8. References

8.1. Normative References

- [I-D.ietf-ccamp-rwa-wson-encode]
Bernstein, G., Lee, Y., Li, D., Imajuku, W., and J. Han,
"Routing and Wavelength Assignment Information Encoding
for Wavelength Switched Optical Networks",
draft-ietf-ccamp-rwa-wson-encode-11 (work in progress),
March 2011.
- [I-D.ietf-ccamp-wson-signal-compatibility-ospf]
Lee, Y. and G. Bernstein, "OSPF Enhancement for Signal and
Network Element Compatibility for Wavelength Switched
Optical Networks",
draft-ietf-ccamp-wson-signal-compatibility-ospf-04 (work
in progress), March 2011.
- [I-D.ietf-ccamp-wson-signaling]
Bernstein, G., "Signaling Extensions for Wavelength
Switched Optical Networks",
draft-ietf-ccamp-wson-signaling-01 (work in progress),
March 2011.
- [ITU-G.698.1]
International Telecommunications Union, "Multichannel DWDM
applications with single-channel optical interfaces", ITU-
T Recommendation G.698.1, December 2006.
- [ITU-G.698.2]
International Telecommunications Union, "Amplified
multichannel DWDM applications with single channel optical
interfaces", ITU-T Recommendation G.698.2, July 2007.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

- [I-D.ietf-ccamp-wson-impairments]
Lee, Y., Bernstein, G., Li, D., Martinelli, G., Chen, M., Han, J., Galimberti, G., Tanzi, A., Bianchi, D., Kattan, M., Schroetter, D., Ceccarelli, D., Bellagamba, E., and D. Caviglia, "A Framework for the Control of Wavelength Switched Optical Networks (WSON) with Impairments", draft-ietf-ccamp-wson-impairments-07 (work in progress), April 2011.
- [I-D.lee-pce-wson-rwa-ext]
Lee, Y., Casellas, R., Margaria, C., and O. Dios, "PCEP Extension for WSON Routing and Wavelength Assignment", draft-lee-pce-wson-rwa-ext-01 (work in progress), March 2011.
- [I-D.narten-iana-considerations-rfc2434bis]
Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", draft-narten-iana-considerations-rfc2434bis-09 (work in progress), March 2008.
- [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.
- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.
- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

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 support for new optical technologies.

Authors' Addresses

Giovanni Martinelli (editor)
Cisco
via Philips 12
Monza 20900
IT

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

Gabriele M Galimberti
Cisco
Via Philips,12
20052 - Monza
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2012

P. Peloso, Ed.
Alcatel-Lucent
G. Martinelli
Cisco
J. Meuric
France Telecom
C. Margaria
Nokia Siemens Networks
July 11, 2011

OSPF-TE Extensions for WSON-specific Network Element Constraints
draft-peloso-ccamp-wson-ospf-oeo-04

Abstract

The original content of this internet draft was to propose some extensions to OSPF encoding in the context of Wavelength Switched Optical Networks, especially for internal constraints of optical network elements. General description can be found in the framework document.

This update of the document still aims at specifying the detailed structure of OSPF LSAs for WSONs. Nevertheless, the proposed LSA layout slightly differs from the current content of the information model and encodings drafts. As a result, the following sections highlight the differences between both approaches and summarize why the authors think these CCAMP's drafts would benefit from an update according to the proposed description.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 12, 2012.

Copyright Notice

Copyright (c) 2011 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. Information Model	4
2.1. Summary of Information Model Changes	5
2.2. Node Information (WSON specific)	6
2.2.1. Label Restrictions	6
2.2.2. Resource Pools, Resource Blocks and Resource Description Containers	7
2.2.3. Resource Pool Accessibility	11
2.2.4. Resource Pool ID	12
2.2.5. Resource Block State	12
2.2.6. Resource Description	13
2.2.7. Resource Pool Wavelength Constraints	14
2.2.8. Shared Access Available Wavelengths	15
3. Encoding	15
3.1. Node related generic encodings	15
3.2. Node related WSON specific encodings	16
3.2.1. Label Restrictions	16
3.2.2. Id Set Field	16
3.2.3. Resource Pool Accessibility	17
3.2.4. Resource Block State	18
3.2.5. Resource Description	18
3.2.6. Resource Pool Wavelength Constraints	21
3.2.7. Shared Access Available Wavelengths	22
3.2.8. Resource Pool	22
3.2.9. Resource Description Container	23
3.3. Link related encodings	23
4. OSPF-TE Extensions	24
4.1. Introduction	24
4.2. Link top level TLV	25
4.3. Node Attribute top level TLV	26
4.4. Resource Pool top level TLV	26
4.5. Resource Description Container top level TLV	26
4.5.1. Resource Description sub-TLV	27
5. Acknowledgements	27
6. Contributors	27
7. IANA Considerations	27
8. Security Considerations	28
9. References	28
9.1. Normative References	28
9.2. Informative References	29
Appendix A. Solution(s) Evaluation	29
A.1. RBNFs Comparison	30
A.2. Depiction of the considered cases for evaluation	32
A.3. Comparing evaluation of the solutions	34
Authors' Addresses	35

1. Introduction

The original content of this internet draft was to propose some extensions to OSPF encoding in the context of Wavelength Switched Optical Networks, especially for internal constraints of optical network elements. General description can be found in the framework document [RFC6163].

This update of the document still aims at specifying the detailed structure of OSPF LSAs for WSONs. Nevertheless, the proposed LSA layout slightly differs from the current content of the information model [I-D.ietf-ccamp-rwa-info] and encodings [I-D.ietf-ccamp-rwa-wson-encode] drafts. As a result, the following sections highlight the differences between both approaches and summarize why the authors think these CCAMP's drafts would benefit from an update according to the proposed description.

More specifically, the sections below follow the scope of current documents, namely information model, encodings and OSPF-TE extensions. Building the latter allowed to identify some improvements which are described in the two former parts. In both, the line has been drawn between the optical information that can be specified by using generic protocol extensions and the one requiring some WSON-specific objects, as agreed by the working group.

1.1. Requirements Language

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

2. Information Model

This section 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 optical path 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.

The reduced Backus-Naur form (RBNF) syntax of [RFC5511] is used to aid in defining the RWA information model.

The text in the following reports every WSON information model modification compared to [I-D.ietf-ccamp-rwa-info]. Whenever a RBNF term is used without explicit definition we assume the same format

and semantic of the original information model.

An initial sub section here below reports a summary of changes introduced by this document.

2.1. Summary of Information Model Changes

In this document, most of the concepts and definitions from [I-D.ietf-ccamp-rwa-info] remain the same. For instance, a "Resource Block" is still "a group of devices with same features and same connectivity constraints".

Compared to the aforementioned document, the following main changes should be noticed:

1. The Resource Pool entity is introduced into the model, allowing the definition of several resource entities per node, which can be advertised independantly. A "Resource Pool" is defined as a group of resource blocks with same connectivity constraints. Several Resource Pools can be defined to associate them with different properties. The goal is to decrease the size of OSPF advertisement upon LSP changes (setup or tear down).
2. The connectivity matrix, defining the node capabilities on interconnection of external links, is used in order to describe connectivity constraints between node-external links and the resource pools. Two advantages can be stressed. First, it gathers all the static information into a node LSA, which OSPF-TE is not required to advertise upon LSP updates. Then it limits the number of connectivity representations introduced by [draft-ietf-rwa-info] (which proposes similar TLVs in different LSAs).
3. The scope of Resource Block Information is reduced, and focuses only on resource/device description. The described device are then efficiently instantiated by referring to these defined types. This allows to separate the physical resource characteristics from the way they are arranged in the node, thus having the description completely independent from the node design.

As a result, this method allows to share resource description for all the identical blocks of a node, thus decreasing the total size of information. Furthermore, as this information is very static and common to several resource blocks, it can be advertised and refreshed independently to any other information.

2.2. Node Information (WSON specific)

As presented 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.

2.2.1. Label Restrictions

This section is a preamble presenting the Label Restriction entity, which is referred many times later in this document.

Wavelength constraint are used in different part of the information model, either as static constraints (in the resource pool as RPWvlConstraints, and the resource block IngressWaveConstraint and EgressWaveConstraint) or representing dynamic properties of a given element (SharedAccessWvls in resource pool). In the GMPLS context Wavelengths are physical instance of Labels.

The wavelength constraints used in this document, although having different semantic, refer to the same notion of list of wavelength. Those constraints apply in addition to either the incoming part of a device (or group of device), the outgoing part or both if the constraint is the same, which is for instance not unusual for static wavelength constraint.

To support this concept, this section defines a field:

`LABEL_RESTRICTIONS`

that carry a label set information and for which direction this label restriction is valid. The directions considered is upstream, downstream or both. The label set information is the one defined in [I-D.ietf-ccamp-rwa-info] as AvailableLabel.

This encoding is reused in different TLV or sub-TLV for different semantic but do not require to define a TLV per direction.

DELTA:

- Define a generic information for label restrictions
- Reuse generic label set and provide a compact representation

2.2.2. Resource Pools, Resource Blocks and Resource Description Containers

As presented in [RFC6163], a WSON node may include regenerators or wavelength converters arranged in shared pools, and can include OEO based WDM switches as well. There are plenty approaches used in the design of WDM switches containing regenerator or converters. 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 OEO devices to convert from a given ingress wavelength on a particular ingress port to a desired egress wavelength on a particular egress port, which are summarized under the accessibility constraints.
3. Limitations on the types of signals that can be converted and the conversions that can be performed, namely the processing capabilities.

For modeling purposes and encoding efficiency regenerators or wavelength converters with identical limitations and/or processing and accessibility constraints are grouped into "blocks". Such blocks can consist of a single resource, though grouping resources into blocks leads to more efficient encodings. Then, these resource blocks are gathered once more into resource pool, for which the blocks share the same accessibility constraints. OEO devices sharing accessibility constraints are likely to being multiplexed on a given piece of equipment (like an Optical Amplifier, a splitter, a Wavelength Selective Switch port, a length of fiber...).

Definitions:

- Resource Block: A group of resources sharing both the same processing properties and the same accessibility constraints. Each Resource Block can contain a different number of resources, but all the resources constituting the block are identical devices.
- Resource Pool: A group of resources sharing the same accessibility constraints, hence a Resource Pool becomes a group of Resource Blocks sharing the same accessibility constraints. Each Resource Pool can contain a different number of blocks, each

of different size, as long as all the devices in the pool are subject to the same accessibility constraints regarding the way these are linked to ingress and egress links of the WSON node containing the pool.

The following picture represents the model of WSON nodes with the help of Resource Blocks and Resource Pools entities.

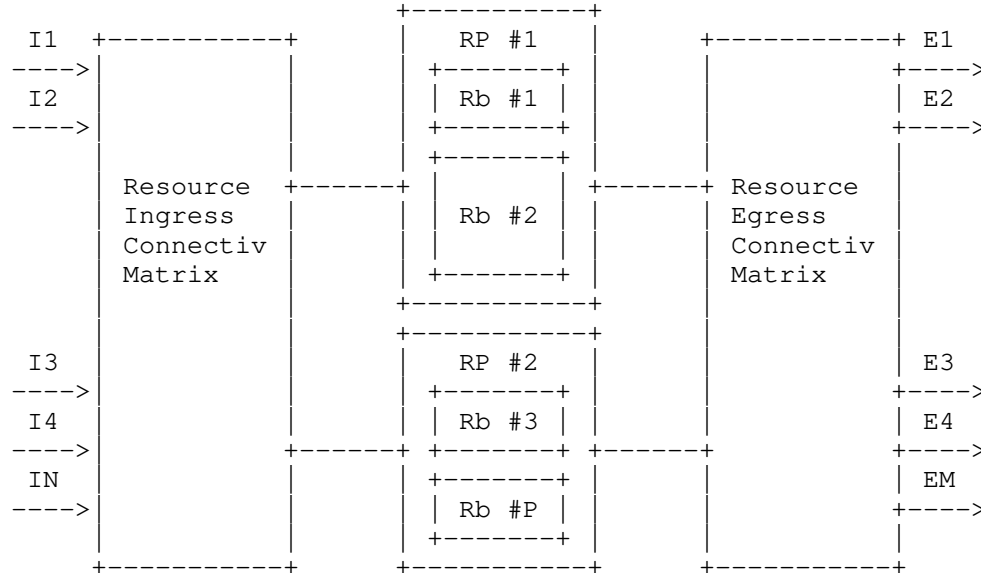


Figure 1

This figure shows a Resource Ingress Connectivity Matrix and another one of the egress, the model from [I-D.ietf-ccamp-rwa-info] gathers both these connectivity matrix inside a Resource Pool Accessibility item, which would lead to the following definition of a Resource Pool.

The following picture represents an abstracted model of the preceding node, that corresponds to the information model chosen in this document.

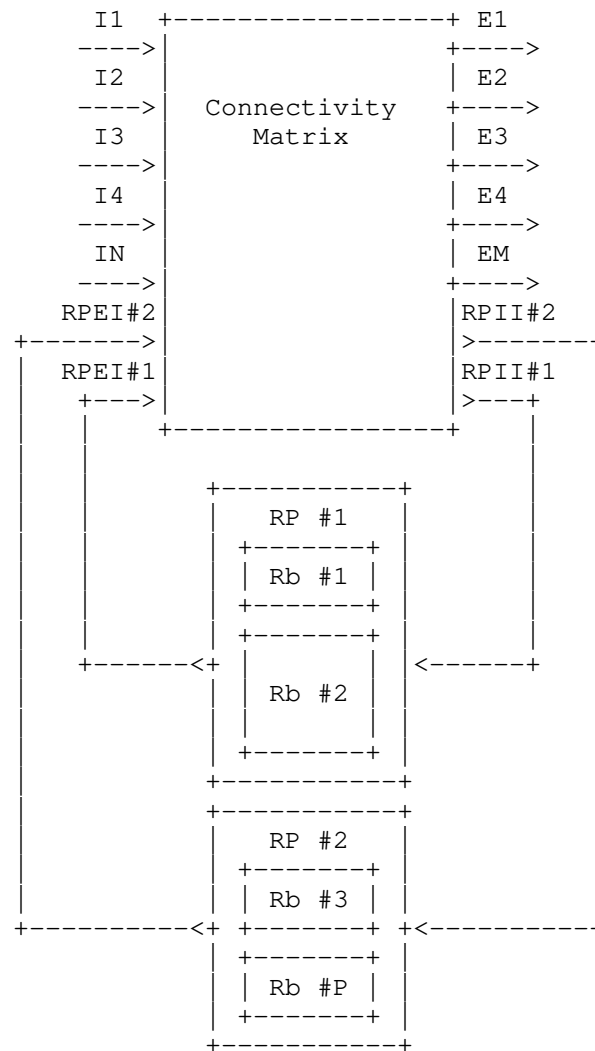


Figure 2

Since by definitions Resource Pools indentify wavelength accessibility to regeneration resources, Section 2.2.3 details how to deal with accessibility constrains. This lead to the following definition of a Resource Pool.

```
<ResourcePool> ::= <ResourcePoolID> ([<SharedAccessWvls>]...)
                    ([<ResourceBlockState>]...)
                    ([<ResourcePoolWvlConstraints>]...)
```

- ResourcePoolID a unique (within node scope) number used to identify the pool,
- SharedAccessWvls represents the dynamic spectral availability coming from the usage of wavelengths by activated resources inside the pool,
- ResourceBlockStates are used to provide the dynamic availability of resources inside the pool.
- ResourcePoolWvlConstraints may be used to define the structural (static) spectral constraints of accessibility of the pool.

A WSON node having some OEO resource might have from 1 to P resource pools. The ResourcePool is created as an entity that will fit in a dedicated TLV (as sub-TLV) so the case of multiple Resource Pools will be handled by fitting one or more Resource Pool entity in each advertisement. The unique identifier ResourcePoolID allows to distinguish among all available pools.

As this document means to have one Resource Pool entity per physical pool of resources inside the node, inside a given node there is no reason for its pools not to share type of resources, hence their modeled representations refer to identical Resource Descriptions entities. In order to avoid unnecessary information flooding, this document gathers all these Resource Descriptions inside a dedicated entity, that is named Resource Description Container.

<ResourceDescriptionContainer> ::= <ResourceDescription>...

The Resource Description Container is a list of Resource Descriptions which, in turn, defines the features (i.e. physical characteristics) of each type of resources held inside the pools (of a given node).

DELTA:

- Introduced definition of Resource Pool.
- Introduced definition of Resource Pool ID.
- Introduced definition of Resource Description Container.
- Changed accordingly Figure 1 and 2 from [I-D.ietf-ccamp-rwa-info].

- Changed the RBNF from [I-D.ietf-ccamp-rwa-info].
- Changed the Resource Block Info into Resource Description (small semantic change, due to minor internal changes).
- Adapted some pieces of models which were related to Resource Block, to the Resource Pool level, namely: RPWvlConstraints

2.2.3. Resource Pool Accessibility

Every device inside a Resource Pool shares the same accessibility constraints, hence the accessibility is a property related to the pool. In order to depict the accessibility of a given pool, two pieces of information need to be described:

- Which ingress links of the node can be connected to the entry of the Resource Pool,
- Which egress links of the node can be connected to the exit of the Resource Pool.

Following remarks can be made concerning these accessibility information:

- These information share the same nature as the one of the Connectivity Matrix,
- These information are relatively static, changing only when the switching fabric of the node is changing (either failure or upgrade),

Hence, the accessibility information of every Resource Pool are embedded together inside the node own's Connectivity Matrix. The solution used to do that consists in using both Local Link Identifiers and Resource Pool Identifiers inside the Link Sets of the Connectivity Matrix. To keep unchanged the definition of the Link Set, 32 bits unnumbered IDs for the Resource Pool are needed (see Section 2.2.4). Thanks to this in the context of a node, the Connectivity Matrix is then providing associations between:

- On one side a set composed of a mix of: (1) ingress link(s) and (2) exit(s) of resource pool(s),
- On the other side a set composed of a mix of: (1) egress link(s) and (2) entry(ies) of resource pool(s).

Then the RBNF for the Connectivity Matrix becomes,

```
<ConnectivityMatrix> ::= <MatrixID> <ConnectType>
    (<IngressSetOfMixedLink&Pool> <EgressSetOfMixedLink&PoolSet>)...
```

The Resource Pool Accessibility information are optional, if not defined, Resource Pool is meant to have no accessibility constraints: from every node ingress port it's possible to reach the pool and the pool egress can reach every egress port of the node.

 DELTA:

This section could be compared to the Resource Block Accessibility constraint, and this is a major change that is proposed here.

2.2.4. Resource Pool ID

In order to encode directly resource pools accessibility, inside the node's connectivity matrix, each Resource Pool needs to be identified alike an internal link with one ID on each side (ingress and egress), and then requires a Resource Pool ID. For each Resource Pool, WSON node assigns one identifier to each side of the pool. This identifier is a non-zero 32-bit number that is unique within the scope of the WSON node that assigns it, hence the Resource Pool ID is composed by a couple of unique numbers.

Consider a (resource) pool inside WSON node A. WSON node A chooses two distincts identifiers for the pool (one for the ingress side and one for the egress side). Considering these identifiers being unique inside the scope of the WSON node A, implies that: no other (resource) pool inside WSON node A may be assigned the value corresponding to any of these two identifiers, neither any (unnumbered) link between WSON node A and any other node may be assigned a link local identifier (from the WSON node A perspective) value corresponding to any of these two identifiers.

Support for resource pools in routing includes carrying information about the identifiers of these pools. Specifically, when an LSR advertises a resource pool, the advertisement carries both the ingress and the egress identifiers of the link.

```
<RPoolID> ::= <RESOURCE_INGRESS_ID> <RESOURCE_EGRESS_ID>
```

2.2.5. Resource Block State

The Resource Block State keep track of the current usage of a resource block within a resource pool.

The state indicate for the resource the number of available resources

and optionally the total number (or maximum number) of resources. decoupling ResourceDescription from the ResourceBlock configuration and allowing a better aggregation of the ResourceDescription. The state available in info model is the following:

Resource Block State definition

```
<ResourceBlockState> ::= <ResourceBlockID> [<CountMaxResources>]
    <CountAvailableResources>
```

DELTA:

This definition of the Resource Block State allow to separate the total number of resources from the resource description (differing in this from [I-D.ietf-ccamp-rwa-info]). This enable a sharing of the resource description between all the pools, while the other solution requires that each pool holds the same number of devices to share the same ResourceBlockDescription (see Section 2.2.6).

2.2.6. Resource Description

The resource block information contains the pieces of information needed to fully identify the resource block static and dynamic information. The static information consist of the characteristics that do not depend on the LSPs using the resource block. In particular the wavelength constraints are the one of the OEO and are independent of the LSPs. the static information is described by a ResourceDescription, which can be valid for several resource blocks, then referenced by their ResourceBlockID.

The ResourceBlockID identifies a resource block, it is a node wide stable and unique identifier (inside the node context). The ResourceBlockID is defined in the ResourceBlockState TLV held in the Resource Pool TLV and used in the Resource Description TLV.

```
<ResourceDescription> := <ResourceBlockID>... <InputConstraints>
    <ProcessingCapabilities> <OutputConstraints>
```

with,

```
<InputConstraints> ::= [<IngressWaveConstraint>] [<modulation-list>]
    [<fec-list>] [<rate-range-list>] [<client-signal-list>]
```

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

```
<OutputConstraints> ::= [<EgressWaveConstraint>] [<modulation-list>]  
                        [<fec-list>]
```

IngressWaveConstraint and EgressWaveConstraint are described in Section 2.2.7. The modulation-list and fec-list represent the list of modulation formats and FEC encoding available within the resource block. This information MAY be present in the advertisement, the absence of this information means that potentially all Modulation and FEC are accepted and possible cranchback may occur.

DELTA:

- Split between static (can be in a separate LSA or in the resource pool) and dynamic information.
- The maximum number of resource is in the state to allow better summarization of the resourceDescription
- The static information is describing the properties, the ResourceDescription is more explicit than resourceInfo in this context
- Changed the RBNF from [I-D.ietf-ccamp-rwa-info], make use of generic label restriction for the wavelength restrictions.

2.2.7. Resource Pool Wavelength Constraints

This field defines any constraint at wavelength level within a resource pool, and is meaningful only when a subset of wavelengths could be configurable within the Pool. This information is static since it depends on specific physical resources within the pools and changes only if there is a node reconfiguration (OEO pools added or removed from an optical node, change in the mux or demuxing devices). As there is an ingress side and an egress side of a pool, this item needs to modelize the wavelength usage on each side.

This field takes the format of a Label_Restrictions Section 2.2.1. At most two instances of this item can be needed: one for each sides (incoming / outgoing) of the pool.

The field is optional, when this field is not present it means there are no specific wavelength constraints imposed by pool. As an example this field is equivalent to the Maximum Bandwidth field defined within [RFC3630]. As the Maximum Bandwith represents the true link capacity, the RESOURCE_POOL_WAVELENGTH_CONSTRAINTS represent the set of wavelengths that can possibly be configured on

the pool.

Note that the usable set of wavelengths could be limited by other constraints: e.g. currently in-use wavelength (see Section 2.2.8) or due to OEO device constraint on compliant wavelengths (see Wavelength Constraints in Section 2.2.6).

DELTA:

Only wavelength constrain. While physical constraints are grouped in another set.

2.2.8. Shared Access Available Wavelengths

The SHARED_ACCESS_AVAILABLE_WAVELENGTHS represents wavelength usage in a Resource Pool hence it is related with the Resource Pool dynamic state.

If a wavelength is in use within a pool, the same wavelength cannot be reused in the same pool however the pool will be available for a different wavelength depending on free resource blocks (Resource Pool definition as in Section 2.2.2). As there is an ingress side and egress side of a pool, this item needs to modelize the wavelength usage on each side. Hence, this representation automatically considers the case of wavelength conversion happening inside the pool.

This field takes the format of a Label_Restrictions Section 2.2.1. At most two instances of this item can be needed: one for each sides (incoming / outgoing) of the pool.

N.B.: Hence, SHARED_ACCESS_AVAILABLE_WAVELENGTHS has the same format as RESOURCE_POOL_WAVELENGTH_CONSTRAINTS defined in Section 2.2.7.

DELTA:

Only wavelength constraint. While physical constraints are grouped in another set.

3. Encoding

3.1. Node related generic encodings

In this section we propose modification to [I-D.ietf-ccamp-general-constraint-encode].

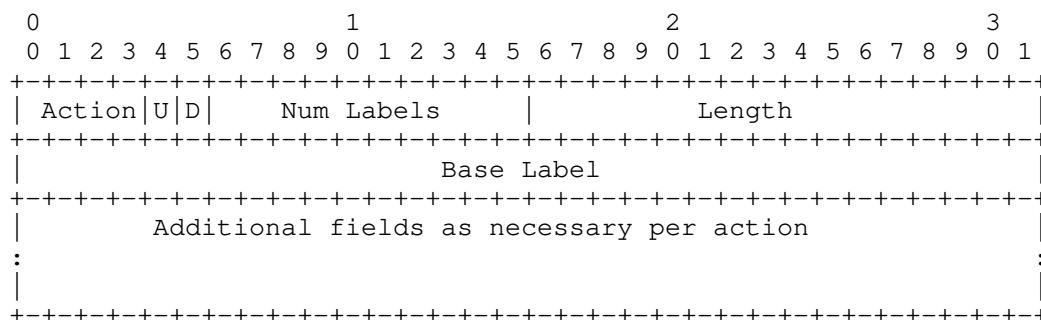
3.2. Node related WSON specific encodings

This section refer to [I-D.ietf-ccamp-rwa-wson-encode]

3.2.1. Label Restrictions

Relatively to section 2.2 of

[I-D.ietf-ccamp-general-constraint-encode] the LABEL_SET field is here slightly modified in order to define a Label Restrictions field.



Although it make sense only using the actions 0-Inclusive List, 2-Inclusive Range or 4-Bitmap. The U bit indicate a label set restriction valid at the upstream direction/incoming side of a resource pool/resource block. The D bit indicate a label set restriction valid at the downstream/outgoing side of a resource pool/resource block. At least one of U or D bit MUST be set, both U and D bit MAY be set.

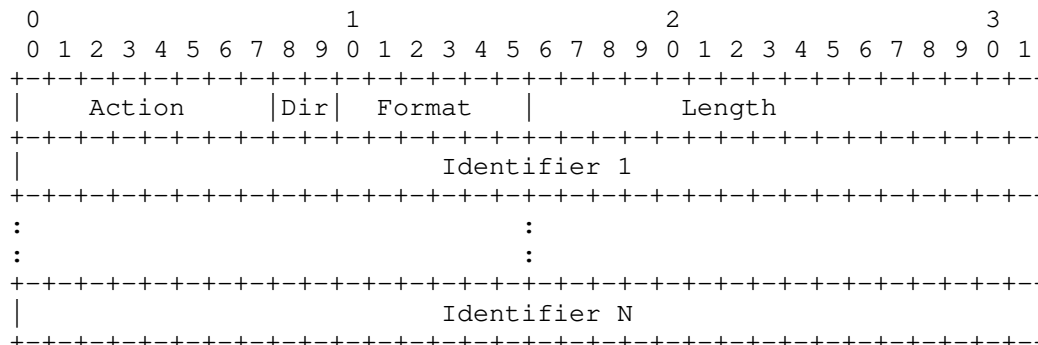
DELTA:

The Num Labels field become 10 bits and this leave room for 1024 labels represented by this encoding. This encoding will be reused in specific TLVs, in case more than 1024 labels are needed multiple fields within TLVs can be used.

3.2.2. Id Set Field

With the introduction of resource description describing properties for a group of resource block we need to efficiently represent a set of IDs. To do so we introduce an IDSet field which has the same encoding as the LinkSet field defined in [I-D.ietf-ccamp-general-constraint-encode] but with a more generic description.

ID Set Field



The Action, Dir have the same encoding as in [I-D.ietf-ccamp-general-constraint-encode]. The Format field indicates the format and length of the Identifier:

- 0 -- 32 Bit unnumbered identifier
- 1 -- IPv4 identifier
- 2 -- IPv6 identifier

This field is used later to define a set of resource blocks (e.g. to list the resource blocks sharing the same resource description).

3.2.3. Resource Pool Accessibility

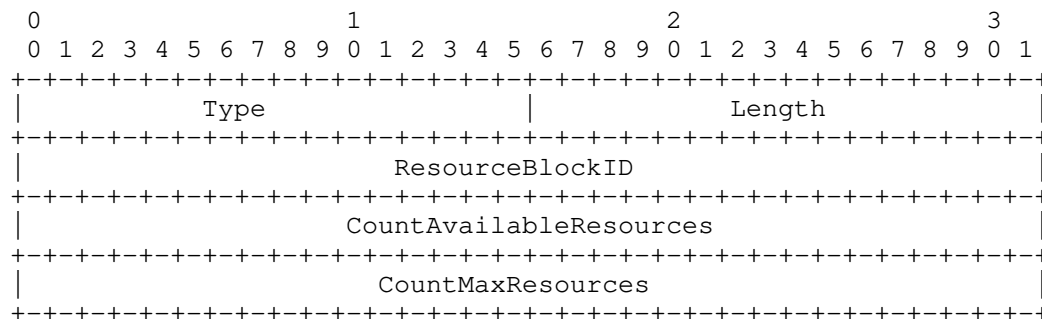
The Resource Pool Accessibility needs no encoding of its own. As explained in the Section 2.2.3 this piece of information is merged inside the Connectivity Matrix object which is actually not impacted by this solution.

Nota: The Link Sets held inside the Connectivity Matrix are composed of LINK_LOCAL_IDENTIFIERS (32 bits identifiers), and the solution to describe the Resource Pool Accessibility consists in using either RESOURCE_INGRESS_ID or RESOURCE_EGRESS_ID (also 32 bits identifiers) which are by definition different from the LINK_LOCAL_IDENTIFIERS (see Section 2.2.4).

DELTA: A major change here as the content of this field are moved inside Connectivity Matrix.

3.2.4. Resource Block State

This TLV indicate the state of a resource block as defined in Section 2.2.5. It defines the ResourceBlockId, and provides the number of free resources and maximum in this resource block. The ResourceBlockID field is a 32 bit node-wide identifier,



The information of the maximum number of resource is optional, this is encoded with a value of 0 in the CountMaxResource field, or with a Length value set to 8 instead of 12.

----- DELTA:

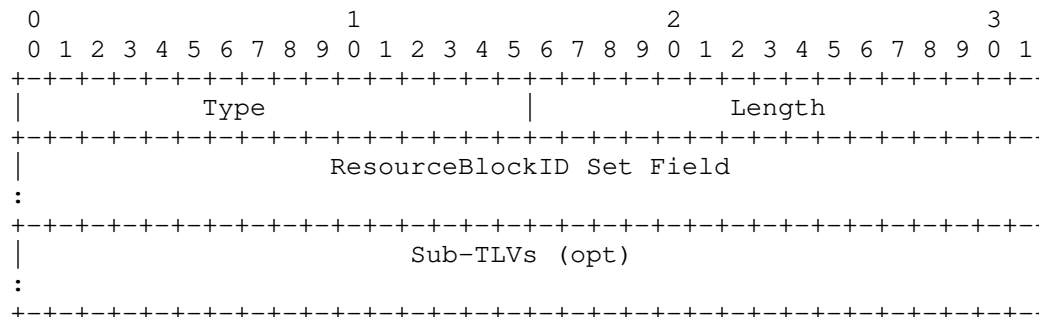
This is an adaptation of the resource pool status that fits the new definition of resource description.

3.2.5. Resource Description

Resource Description sub-TLVs represent the information described in Section 2.2.6.

The resource description TLV encoding follow the definition from Section 2.2.6 with a list of sub-sub TLV.

Resource Description TLV



The ResourceBlockID Set Field is encoded using the IDSet field encoding using the ResourceBlockID as identifier with format 0.

The Sub-Sub TLVs are defined as follow, the order does not matter. Each of the Sub-Sub-TLV defined in this document MAY be repeated more than once, on receipt all Sub-Sub-TLV MUST be taken into account. The resulting information is the union of all the element of the Sub-Sub-TLVs (all Sub-Sub-TLVs of this document describe lists). For example an implementation may choose to indicate that in total 4 label can be used as 4 Label constraint Sub-Sub-tlv, each of them with 1 label.

Info model	Type	Encoding
IngressWaveConstraint	Label Constraints	Label restriction, see Section 3.2.1.
Input modulation-list	Modulation List	A list of Modulation Format Fields, described in [I-D.ietf-ccamp-rwa-wson-encode] section 4.2.1.
Input fec-list	FEC List	A list of FEC type, described in [I-D.ietf-ccamp-rwa-wson-encode] section 4.3.1.
Input rate-range-list	Rate Range	A list of rate range field, described in [I-D.ietf-ccamp-rwa-wson-encode] section 4.4.1.
Input client-signal-list	Client Signal List	A list of GPids, described in [I-D.ietf-ccamp-rwa-wson-encode] section 4.5.

ProcessingCapabilities	Processing Capabilities	A list of Processing Capabilities Fields, except processing cap "Number of Resources", described in [I-D.ietf-ccamp-rwa-wson-encode] section 4.6.1.
EgressWaveConstraint	Label Constraints	Label restriction, see Section 3.2.1.
Output modulation-list	Modulation List	see Input modulation-list
Output fec-list	FEC List	see Input fec-list

Resource description Sub-Sub-TLVs and relation to info model

The Label Constraints Sub-Sub-TLV is used for IngressWaveConstraint and EgressWaveConstraint as the Label Restriction field carries the U and D bit to allow to distinguish a label restriction valid for incoming, outgoing or both.

The Modulation List Sub-Sub-TLV is similarly used for the input and output modulation list. The Sub-Sub-TLV contains a list of Modulation format field, which indicate if they are valid for the input (I bit set to 1) or for the output (I bit cleared). The list of Modulation format field MUST contain at least one ingress FEC modulation format. If no Egress modulation format is present in the list it is implied that no modulation format conversion is impossible, the egress modulation list is the same as the ingress modulation list and modulation format is not performed.

The FEC list Sub-Sub-TLV is also representing both Input and Output FEC list. The Sub-Sub-TLV is defined as a list of FEC Fields, conceptually being Sub-Sub-Sub-TLVs indicating via the I bit if they are valid for ingress or egress. At least one ingress FEC MUST be present in the list, if no egress modulation format is present in the list it is implied that the egress FEC list is the same as the ingress FEC list. In such case FEC format conversion MAY be performed.

The Processing Capabilities Sub-Sub-TLV is the same as in [I-D.ietf-ccamp-rwa-wson-encode] section 4.6.1. except for the maximum number of resource which is represented in the ResourceBlockState. The FEC and Modulation format conversion capabilities are expressed via the Modulation and FEC list by not including any egress modulation/fec in the respective lists.

Bit-Rate Range and Client Signal lists are unchanged from [I-D.ietf-ccamp-rwa-wson-encode]

 DELTA:

- use a common TLV for the label restriction
- use a common TLV for the FEC list
- use a common TLV for the Modulation format list
- re-use indirectly (via ID Set) the general encoding LinkSet for RBlockId set
- More explicit statement on FEC and Modulation format conversion capabilities

3.2.6. Resource Pool Wavelength Constraints

This TLV is used to describe static wavelength constraint, it follows the encoding of Label_Restrictions field Section 3.2.1

RESOURCE_POOL_WAVELENGTH_CONSTRAINTS TLV

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type										Length																													
Label_Restrictions field(s)																																							
:																																							

The Label_Restrictions field might be repeated several times depending on the U and D bit flags. In case multiple fields with the same U and D bits set to 1, the final resulting constrain will be the intersection of all Label_Restrictions. If multiple TLVs are present the resulting constraint is the intersection of all the TLV.

Example below:

- No RESOURCE_POOL_WAVELENGTH_CONSTRAINTS TLV meaning that these type of constraints are not described.
- A TLV present with one Label_Restrictions field with both the U or D bits MUST be set to 1. Which means the same constrains apply to both sides of the pool.

- A TLV present with three Label_Restrictions field presents, one field with U=1 so applicable upstream. The two other fields with D=1 so applicable downstream

DELTA: Small delta, just using the add-on bits to provide a direction/side semantic.

3.2.7. Shared Access Available Wavelengths

This TLV is used to describe dynamic wavelength availability, it follows the encoding of Label_Restrictions field. Section 3.2.1

SHARED_ACCESS_AVAILABLE_WAVELENGTH TLV

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Type                 |               Length                 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Label_Restrictions field(s)                               |
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The same rules and usage defined in Section 3.2.6 apply here.

3.2.8. Resource Pool

The RESOURCE_POOL TLV contains the preceding TLVs.

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Type                 |               Length                 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               RESOURCE_INGRESS_ID                                         |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               RESOURCE_EGRESS_ID                                          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|               Sub-TLVs as needed (Opt)                                   |
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

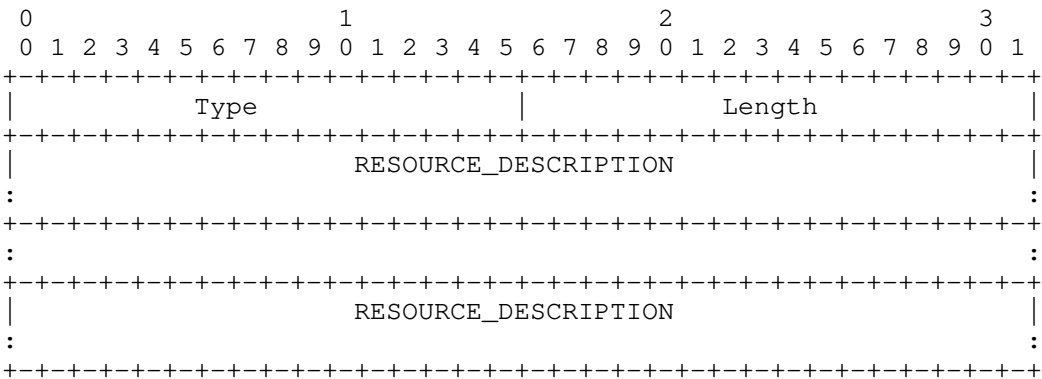
List of possible Sub-TLVs:

Name	Static/Dynamic
Resource Block State	Dynamic
Shared Access Available Wavelength	Dynamic
Resource Pool Wavelength Constraints	Static

DELTA:
Similar to Resource Pool inside [I-D.ietf-ccamp-rwa-wson-encode] with a different internal layout that allows for multiple instances.

3.2.9. Resource Description Container

The RESOURCE_DESCRIPTION_CONTAINER is a list of RESOURCE_DESCRIPTION. This one MAY be used to extract the static content of the previous TLV, in order to hold all this content inside this purposely defined static TLV. Then each one can be in separatly flooded entities (e.g. in separated LSAs see Section 4.1.



DELTA:
New item.

3.3. Link related encodings

This section does not differ from the equivalent in [I-D.ietf-ccamp-general-constraint-encode]

4. OSPF-TE Extensions

This section handles OSPF-TE extensions.

It starts with introducing the top view of the extensions provided by this draft. Then a sub-section dedicated for each top level TLV details the extensions relevant for this top level TLV.

4.1. Introduction

This introduction provides the layout of the preceding information model (Section 2) and encodings (Section 3) into top-level TLVs of opaque LSAs.

[RFC3630] introduces Link top level TLV (type 2). This document extends its content with the encodings depicted in Section 3.3. These extensions offer the capability to advertise restrictions on the list of available labels.

N.B.: This capability is specifically useful when these labels have a network wide semantic like suggested in a WSON context.

[RFC5786] introduces Node Attribute top level TLV (type 5). This document extends its content with the encodings depicted in Section 3.1. These extensions offer the capability to advertise restrictions on the switching capabilities of the node.
N.B.: This TLV is unique for a given node and contains static information only, hence no more than one LSA per node is expected to host such a TLV.

This document introduces a new top level TLV named RESOURCE_POOL (type value to be defined), which encodings are depicted in Section 3.2. RESOURCE_POOL TLV offers the capability to advertise one or multiple pools of OEO devices held in a given node. This object can carry resource descriptions, the available resources inside the pool(s) and the availability of wavelengths to reach the pool (refer to pool definition inside Section 2.2.2).
N.B.: A LSA can contain more than one RESOURCE_POOL top level TLV (allowing one LSA to advertise the description of all the pools of the originating node). Alternatively, a node can originate more than one LSA containing each RESOURCE_POOL top level TLVs (allowing each LSA to advertise an individual pool). In that case all the RESOURCE_POOL originated by the same node MUST have different RESOURCE_POOL_ID. As most of the information contained inside a RESOURCE_POOL are dynamic, an implementer may well choose to define one LSA per pool of resources in order to reduce the quantity of information flooded upon change in resource usage.

This document introduces another new top level TLV named

RESOURCE_DESCRIPTION_CONTAINER (type value to be defined), which encoding is depicted in Section 3.2.9.

RESOURCE_DESCRIPTION_CONTAINER TLV contains a list of RESOURCE_DESCRIPTION valid in the scope of the originating node. A given node MUST NOT originate more than one LSA containing RESOURCE_DESCRIPTION_CONTAINER TLV. An LSA containing a RESOURCE_DESCRIPTION_CONTAINER TLV MUST NOT contain any additional top level TLV.

N.B.: This TLV is designed to be unique in the scope of the originating node and to gather all the resource descriptions relevant in this scope.

Summarizing Table

Top-TLV	Type	Name	Instances	Static/Dynamic
	2	Link	1 per fiber	Mix
	5	Node Attribute	1 per Node	Static
	TBD	Resource Pool	1 per Pool	Dynamic
	TBD	Resource Desc Cont	1 per Node	Static

DELTA:

- Renamed the Node Optical Property tlv into Resource Pool TLV
- Allow multiple instance of Resource Pool TLV
- Introduced an optional new TLV named Resource Description

4.2. Link top level TLV

This section refer to
[I-D.ietf-ccamp-gmpls-general-constraints-ospf-te].

The following new sub-TLVs are added to the Link top level TLV (type 2).

Sub-TLV Type	Length	Name
TBD	variable	Port Label Restrictions
TBD	variable	Available Wavelengths
TBD	variable	Shared Backup Wavelengths

In Link TLV, all the sub-TLV listed above are optional.

4.3. Node Attribute top level TLV

This section refer to
[I-D.ietf-ccamp-gmpls-general-constraints-ospf-te].

The following new sub-TLVs are added to the Node Attribute top level TLV (type 5).

Sub-TLV Type	Length	Name
TBD	variable	Connectivity Matrix
TBD	variable	Port Label Restrictions
TBD	variable	Shared Risk Node Group

In Node Attribute, all the sub-TLV listed above are optional. None of them contain sub-TLV.

4.4. Resource Pool top level TLV

This section refer to [I-D.ietf-ccamp-wson-signal-compatibility-ospf]

The following sub-TLVs are created for the Resource Pool top level TLV.

Sub-TLV Type	Length	Name
TBD	variable	Resource Block State
TBD	variable	Shared Access Available Wavelength
TBD	variable	Resource Pool Wavelength Constraints

In Resource Pool, all the sub-TLV listed above are optional.

4.5. Resource Description Container top level TLV

This section refer to [I-D.ietf-ccamp-wson-signal-compatibility-ospf]

The following sub-TLVs are created for the Resource Description Container top level TLV.

Sub-TLV Type	Length	Name
TBD	variable	Resource Description

4.5.1. Resource Description sub-TLV

The following sub-TLVs are created for the Resource Pool top level TLV.

Sub-TLV Type	Length	Name
TBD	variable	Modulation List
TBD	variable	FEC List
TBD	variable	Rate Range List
TBD	variable	Client Signal List
TBD	variable	Processing Capabilities
TBD	variable	Label Constraints

In Resource Description, all the sub-TLV listed above are optional.

5. Acknowledgements

This template was derived from an initial version written by Pekka Savola and contributed by him to the xml2rfc project.

This document shares common material with the documents quoted, which seems fair as the target of this version is to highlight differences.

The editors wish to thank Ramon Casellas for his constructive comments.

6. Contributors

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

Email: daniele.ceccarelli@ericsson.com

7. IANA Considerations

This memo requires many requests to IANA, which will be completed in a latter version.

8. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

9. References

9.1. Normative References

- [I-D.ietf-ccamp-general-constraint-encode]
Bernstein, G., "General Network Element Constraint Encoding for GMPLS Controlled Networks",
draft-ietf-ccamp-general-constraint-encode-04 (work in progress), December 2010.
- [I-D.ietf-ccamp-gmpls-general-constraints-ospf-te]
Zhang, F., Lee, Y., Han, J., Bernstein, G., Xu, Y., Zhang, G., Li, D., Chen, M., and Y. Ye, "OSPF-TE Extensions for General Network Element Constraints",
draft-ietf-ccamp-gmpls-general-constraints-ospf-te-00 (work in progress), March 2011.
- [I-D.ietf-ccamp-rwa-wson-encode]
Bernstein, G., Lee, Y., Li, D., Imajuku, W., and J. Han, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks",
draft-ietf-ccamp-rwa-wson-encode-11 (work in progress), March 2011.
- [I-D.ietf-ccamp-wson-signal-compatibility-ospf]
Lee, Y. and G. Bernstein, "OSPF Enhancement for Signal and Network Element Compatibility for Wavelength Switched Optical Networks",
draft-ietf-ccamp-wson-signal-compatibility-ospf-04 (work in progress), March 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.
- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering

(TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax Used to Form Encoding Rules in Various Routing Protocol Specifications", RFC 5511, April 2009.
- [RFC5786] Aggarwal, R. and K. Kompella, "Advertising a Router's Local Addresses in OSPF Traffic Engineering (TE) Extensions", RFC 5786, March 2010.

9.2. Informative References

- [I-D.ietf-ccamp-rwa-info] Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-info-11 (work in progress), March 2011.
- [I-D.narten-iana-considerations-rfc2434bis] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", draft-narten-iana-considerations-rfc2434bis-09 (work in progress), March 2008.
- [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

Appendix A. Solution(s) Evaluation

Within this section we try evaluate the amount of information that needs to be exchanged through routing advertisements. For this evaluation we consider some minimum optical node reference design to make a OEO extension future proof.

This sections starts with summarizing the LSAs needed to depict a node with both the solution depicted by this document and the solution depicted by [I-D.ietf-ccamp-rwa-info]. Afterwards, the hypothesis concerning the node features that will serve as a basis for the solution evaluation will be presented, before the actual results of the solutions evaluations (both the one of this document

and the one of [I-D.ietf-ccamp-rwa-info]).

A.1. RBNFs Comparison

In this section we try compare the how TLVs are composed according two this draft proposal versus existing WSON solutions. The goal here is to provide the all reference for and easy understanding where two solutions are different. Numbers will be provided in the next section.

The evaluation will be done on the Resource Pool top-level TLV since the Node address and Link TLV are considered equivalent.

WSON Drafts. According to [I-D.ietf-ccamp-wson-signal-compatibility-ospf] in section 2 defines the Optical Node Property TLV which collect the WSON specific information. This TLV is composed of the following:

```
<ResourcePool> ::= [<ResourceBlockInformation>]...  
    [<ResourceBlockAccessibility>]... [<ResourceBlockWvlConstraint>]...  
    [<ResourceBlockPoolState>...] [<SharedAccessWvls>...]
```

- a) Resource Block Information. Defined as : ([<ResourceSet>] <InputConstraints> <ProcessingCapabilities> <OutputConstraints>). A resource block information defines here the number of devices inside the block.
- b) Resource Block Accessibility. Defined as (<PoolIngressMatrix> <PoolEgressMatrix>) which is expanded in tuples like (<INGRESS_LINK_SET><ResourceSet>)* and (<EGRESS_LINK_SET><ResourceSet>)*. Note that INGRESS/EGRESS_LINK_SET is a name defined here for the link set field used in the [I-D.ietf-ccamp-rwa-info] document.
- c) Resource Block Wavelength Constraints. Defined as <IngressWaveConstraints><EgressWaveConstraints>. This is expanded in <ResourceSet>INPUT_WAVELENGTH_SET OUTPUT_WAVELENGTH_SET, for the static constraints of resource blocks.
- d) Shared Access Wavelengths. Defined as <IngressWaveConstraints><EgressWaveConstraints>. This is expanded in <ResourceSet>INPUT_WAVELENGTH_SET OUTPUT_WAVELENGTH_SET, for the shared fibers between blocks.

- e) Resource Block Pool State. <ResourceSet> <USAGE_STATE_BITMAP>

In current proposal there are two types of TLV.

First the Resource Pool TLV (with an instance per pool) is composed of the following information:

```
<ResourcePool> ::= <ResourcePoolID> [<ResourceDescription>]...
    [<ResourcePoolWvlConstraints>]... [<SharedAccessWvls>]...
    [<ResourceBlockState>]...
```

- a) Resource Description. Which is defined as: (<RBlockID>...) <InputConstraints> <ProcessingCapabilities> <OutputConstraints>. This is equivalent to the item a) above without the number of devices inside the resource block, which allow this definition to be usable by any block. The number of available resource of a given type inside the pool being specified by the Resource Block State below. When a Resource Description Container TLV is defined by a Node, the Resource Pool TLV of this same node SHOULD NOT contain any Resource Description sub-TLV.
- b) Resource Block State. Where RBlockState is defined as <RBlockID> [<NumResources>] <NumberOfAvailableResources>. This field efficiently report how many of a given resource type is available inside the pool or not.
- c) Shared Access Available Wavelength. This is composed of a Label Restriction field and SHOULD used to depict the dynamic constraints of the pool.
- d) Resource Pool Wavelength Constraints. This is composed of a Label Restriction field and MAY be used to depict the static constraints of the pool.

Second the Resource Descriptor Container TLV (with a single instance per node) is used to gather all the Resource Descriptions of a given node, as these are static information composed of the following information:

```
<ResourceDescriptionContainer> ::= <ResourceDescription>...
```

- a) Resource Description. Which is defined as: (<RBlockID>...) <InputConstraints> <ProcessingCapabilities> <OutputConstraints>. This is equivalent to the item a) above.

A.2. Depiction of the considered cases for evaluation

For the sake of the comparison we have considered the following parameters and values characterizing the optical node design:

- o Node Degree Connectivity: 4, 8 and 16.
- o WDM capacity: 100 wavelengths.
- o Switching capacity. Defines the total node switching capability and is calculated as Node Degree Connectivity x 100 wavelengths.
- o Regeneration Capability. We assume a value of 5% of the total switching capacity.
- o Add/Drop Capability. We assume a typical value of 25% of the switching capacity. So in the average up to 30 wavelengths per incoming fiber can be added/dropped within the optical node.
- o Resource pool setup and capabilities. A physical resource pool contains a mix of Add/Drop and Regeneration capabilities. This has the effect of increasing the number of resource pool advertized. Resource pool can be fully flexible (connected to any port), partial (only to some port) or Fixed (can only be connected to one direction). This parameter influences the complexity of the connectivity matrix.
- o Number of Regenerator types. For a given node the number of OEO capabilities is limited, it is typically decided by the type of electrical equipment and optical modules (emitting laser and optical receiver).
- o Blocking Ratio. The Spatial/Spectral blocking ratio indicates how much port-based/wavelength based blocking a node is experiencing.

For example considering the typical design it results in the following static layout:

- o 3 OEO pools each having 3 Resource Block inside.
- o Connectivity Matrix: (8+30+30) 64x64 if considering one connectivity matrix. Ingress=64x3, Egress=3x64 (considering the OEO access with a multiple-wavelength link).

The following types of nodes and node designs were considered in this evaluation:

Node Types and designs

Node Type	Nodal Degree	Pool Type	Blocking
Small(S), Flexible	4	Partial	None
Small(S), Fixed(port)	4	Fixed	Port
Small(S), Fixed(label)	4	Partial	Lambda
Middle(M), Flexible	8	Flexible	None
Large(L), Flexible	16	Flexible	None

For the small nodes, 5 different type of regenerators are considered, for the Middle and Large ones 10 different type of regenerators are considered. Based on those designs we derived the following important figures:

- o Number of resourcePool : depends on the pool type and connectivity, which depend on the port blocking and number of Add/Drop and Regenerator capacity.
- o Number of resourceBlock. There is two numbers to be considered here : the number of resourceBlock for a given resource pool (this document) and total number of resourceBlock ([I-D.ietf-ccamp-rwa-info]). In this document the number of resource block within a resource pool is, worst case, the number of possible regenerator types, whereas in [I-D.ietf-ccamp-rwa-info] the number of resource block depends on the number of OEO types and on the connectivity.
- o Number of connectivity matrix/number of pairs/link per pairs. The number of sub-matrix increase depending on the port blocking ratio, the number of pair in one connectivity matrix depends on the wavelength restrictions. Those two criteria do not depend on which information model is considered. The number of link per set is increased by the number of resource pool in this draft.

Those numbers for each node are shown in the following table:

Details of information elements per node

Node Type	# Pools	Resource Blocks	Matrix/Pair/Links
S, Flexible	6	5 (30)	1/1/10 (1/1/1)
S, Fixed(port)	12	5 (60)	4/4/4 (4/4/1)
S, Fixed(label)	6	5 (30)	4/1/10 (4/1/1)
M, Flexible	3	10 (30)	1/1/11 (1/1/1)
L, Flexible	5	10 (50)	1/1/21 (1/1/1)

Nota: Values for [I-D.ietf-ccamp-rwa-wson-encode] are between brackets.

For further reading easiness the above table could be further expanded as the following one:

Details of information elements per node

Node Type	#Pools	#Device Type	#Blocks	#ResProp TLV	Matrix/Pair/Links
S, Flexible	6	5 (30)	30	5 (25)	1/1/10 (1/1/1)
S, Fixed(port)	12	5 (60)	60	5 (45)	4/4/4 (4/4/1)
S, Fixed(label)	6	5 (30)	30	5 (25)	4/1/10 (4/1/1)
M, Flexible	3	10 (30)	30	10 (35)	1/1/11 (1/1/1)
L, Flexible	5	10 (50)	50	10 (40)	1/1/21 (1/1/1)

Nota: Values for [I-D.ietf-ccamp-rwa-wson-encode] are between brackets.

A.3. Comparing evaluation of the solutions

Based on those key information model elements both the tables "LSA size" indicate the size of the LSAs in this document and in [I-D.ietf-ccamp-rwa-wson-encode]. Number of flooded LSAs of a given type are indicated between brackets (when bigger than 1).

Solution of this document - Average size (and number) of LSAs per node type (unit: bytes)

Node Type	Node Attr LSA	Resource Pool LSA	Resource Desc LSA
S, Flexible	117	120 (6)	524
S, Fixed(port)	692	120 (12)	644
S, Fixed(label)	620	120 (6)	524
M, Flexible	127	120 (3)	904
L, Flexible	209	120 (5)	984

Solution of [I-D.ietf-ccamp-rwa-wson-encode] - Average size (and number) of LSAs per node type (unit: bytes)

Node Type	Node Attr LSA	Optical Node LSA
S, Flexible	49	2801
S, Fixed(port)	340	2980
S, Fixed(label)	132	4118
M, Flexible	52	2980
L, Flexible	54	2809

The Resource Description Container LSA contains several resource description TLVs. This LSA is smaller than the corresponding in [I-D.ietf-ccamp-rwa-wson-encode] mainly because the resource description do not depend on the port/lambda connectivity and number of device per block, thus allowing a better sharing of the information depicting the oeo capabilities.

The following summarizing table indicates the size of the sum of all LSA and the average size per update. In this document all the dynamic part is in the resource pool, allowing a more efficient updating behavior. The evaluation for [I-D.ietf-ccamp-rwa-wson-encode] are best case/worst case; the best case being an update of the RBState TLV and SharedAccessPool TLV only, which requires a multi-instance implementation of OSPF.

Summarizing Table (unit:bytes)

Node Type	Total LSA size	Total number of LSA	Avg size of an update
S, Flexible	1361 (2850)	8 (2)	120 (616/2801)
S, Fixed(port)	2776 (5411)	14 (2)	120 (1192/2980)
S, Fixed(label)	1864 (2941)	8 (2)	120 (616/4118)
M, Flexible	1391 (3032)	5 (2)	120 (448/2980)
L, Flexible	1793 (4172)	7 (2)	120 (720/2809)

Nota: Values for [I-D.ietf-ccamp-rwa-wson-encode] are between brackets

The node design considered are typical case, a worst case can be a node with high nodal degree, with lots of port and wavelength constraints. With considering a nodal degree of 8, resulting in 28 resource pool and 140 resource blocks, the total size is 9816 (11820) with 30 (2) LSAs.

Authors' Addresses

Pierre Peloso (editor)
Alcatel-Lucent
R.te de Villejust
Nozay, 91620
France

Phone: +33 130 702 662
Email: pierre.peloso@alcatel-lucent.com

Giovanni Martinelli
Cisco
Monza, 20900
Italy

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

Julien Meuric
France Telecom
2, av. Pierre Marzin
Lannion, 22307
France

Phone: +33 296 052 828
Email: julien.meuric@orange-ftgroup.com

Cyril Margaria
Nokia Siemens Networks
St-Martin str. 76
Munchen, 81541
Germany

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

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: January 4, 2012

D. Shimazaki
R. Hayashi
K. Shiimoto
NTT Corporation
July 3, 2011

Requirement and protocol for WSON and non-WSON interoperability
draft-shimazaki-ccamp-wson-interoperability-00.txt

Abstract

GMPLS protocol enabled network operator to setup optical path network rapidly and dynamically. Recently, WSON [8] is standardized to achieve WDM core network. However, it is difficult that all network equipment supports WSON protocol. Therefore, interoperability between WSON and non-WSON nodes is needed to construct path network.

This document describes requirement for interoperability between WSON node and non-WSON nodes and functions that routing and signaling protocol should support.

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on December, 2011.

Copyright Notice

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

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

Conventions Used In This Document

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

Table of Contents

1.	Introduction	4
2.	Requirements	4
2.1.	Requirements for non-WSON nodes	4
2.2.	Requirements for WSON nodes	5
3.	Protocols	5
3.1.	OSPF-TE	5
3.2.	RSVP-TE	6
3.3.	PCE	6
4.	Path setup scenario	6
5.	Security considerations	7
6.	IANA Considerations	7
7.	Acknowledgments	7
8.	References	7
	Authors' Addresses	8

1. Introduction

This document proposes interoperability mechanism between non-WSON nodes and WSON nodes.

GMPLS [1] defines routing and signaling protocol extension to control multilayer network. [2] describes GMPLS extended OSPF. [2] added four sub-TLVs to Link TLV that is defined in OSPF-TE [3]. [4] describes GMPLS extended RSVP-TE. [4] added new object to RSVP-TE [5].

[6] defines PCE basic architecture. It defines path computation entity separated from network element for traffic engineering.

WSON network consists of WDM link, tunable transmitter receiver, ROADM, wavelength converter, and electro-optical network elements. [8] defines control framework of these components with GMPLS and PCE protocol. WSON protocol extension is extension to GMPLS protocols [2], [5]. For example, WSON extended OSPF draft mentions that wavelength Availability information is added to Link TLV of GMPLS extended OSPF-TE, as well as GMPLS extended OSPF added four sub-TLVs to Link TLV of OSPF-TE [3].

The Wavelength Switched Optical Network (WSON), referring to Wavelength Division Multiplexing (WDM) based optical network in which switching is performed selectively based on the wavelength of an optical signal, is a promising optical network in that it provides broadband and energy-saving transmission. Recently, WSON-supported ROADMs and OXCs appear and interoperating experiments have been demonstrated. On the other hand, there are few commercially-available routers working in WSON.

This document describes requirement for interoperability between WSON and non-WSON nodes and functions that routing and signaling protocol should support. Under the proposed operation, non-WSON nodes (ex. routers sending/receiving electrical signal) do not send GMPLS protocol messages related to WSON, while necessary message objects are exchanged and relayed among WSON nodes (ex. ROADMs and OXCs).

2. Requirements

2.1. Requirements for non-WSON nodes

Non-WSON nodes need to have no impact on interoperating WSON nodes. Detail is described in below.

When non-WSON node receives routing protocol information that includes WSON extended information, the node should ignore the WSON

extended information and understand the conventional GMPLS information and add this information to traffic engineering database (TED). It must transfer routing information to neighbor node.

Non-WSON node should combine two TEDs, WSON and non-WSON and make one TED. Non-WSON node can calculate the path route under the whole network topology including WSON network. In WSON network, both route and wavelength should be determined. However, non-WSON node does not have wavelength information in the TED, so it calculates only route without considering available wavelength information.

Non-WSON node can send signaling message to next hop node and setup path strictly or loosely.

When non-WSON node receive path request message of signaling protocol with WSON information in explicit route object, such as lambda label, the node should cancel the signaling and send path error message to source node. In the other hand, it receives path request message with WSON information in record route object, the node can ignore the incomprehensible information and continue path setup procedure.

2.2. Requirements for WSON nodes

WSON nodes need to handle not only WSON-extended information but also no WSON information. WSON nodes at the border of WSON need to compute available wavelength along the assigned path by non-WSON nodes, or compute both a route and available wavelength at the same time if non-WSON nodes doesn't assign a route in WSON strictly or there are no available wavelength along the assigned route by non-WSON nodes.

WSON nodes at the border of WSON need to add WSON-extended objects to a signalling protocol messages after receiving them from non-WSON nodes. WSON nodes at the border of WSON need to take WSON-extended objects from a signalling protocol before relaying them to non-WSON nodes. Note that this function is optional if non-WSON nodes neglect WSON-extended information in a signalling protocol received.

3. Protocols

3.1. OSPF-TE

Non-WSON nodes ignore available lambda information. When LSA include both lambda and other information, for example adjacency or SC information in Link-TLV, it is desirable that non-WSON nodes ignore only lambda information. ROADMs/OXCs advertise and share available lambda information.

3.2. RSVP-TE

Non-WSON nodes send RSVP-TE PATH message including just route information. WSON nodes add lambda information to be used in WSON to RSVP-TE PATH message. Additionally, it is desirable that OXC at the egress border delete lambda information from PATH message.

3.3. PCE

There are several computation models in terms of "who computes what". non-WSON nodes calculate just path route. In WSON area, WSON nodes or PCE calculate wavelength selection or RWA problem.

What	Who
Path route	non-WSON nodes
Wavelength or RWA	Nodes at the border of WSON or PCE

Table 1: Function deployment model

4. Path setup scenario

Under the proposed path set-up control, a source non-WSON nodes compute a path route with constraint shortest path fast (CSPF). A border node of WSON computes an available wavelength and, if necessary, a route. Then, the border node adds wavelength information, which is used in WSON, to signaling message. The other border node of WSON takes the wavelength information from the signaling message and sends it to a node outside WSON.

One of recommended GMPLS RSVP-TE signaling scenarios is described as follows based on Fig.1, in which non-WSON nodes don't support WSON and are outside WSON, while ROADMs are in WSON. Firstly, a source non-WSON node R1 sends RSVP-TE signaling by assigning only R2 loosely as a path route and switching type as Label Switching Capable (LSC). Here, R1 is assumed to understand a topology in LSC region including all of the non-WSON nodes and ROADMs with OSPF-TE. When a ROADM O1 receives the signaling message from R1, it sends path computation request to PCE with the route information which R1 calculates to select the wavelength in WSON. When a PCE receives the path computation request, its routing and wavelength assignment (RWA) algorithm computes one of available wavelengths along the route O1-O2 which R1 calculates. If there is no available wavelength, the PCE's RWA algorithm computes both a route and one of available wavelengths. After calculation, PCE replies the route and wavelength to O1. When

O1 receives route wavelength reply message from PCE, then sends the signaling which assigns explicit route object (ERO) and wavelength label to the next node.

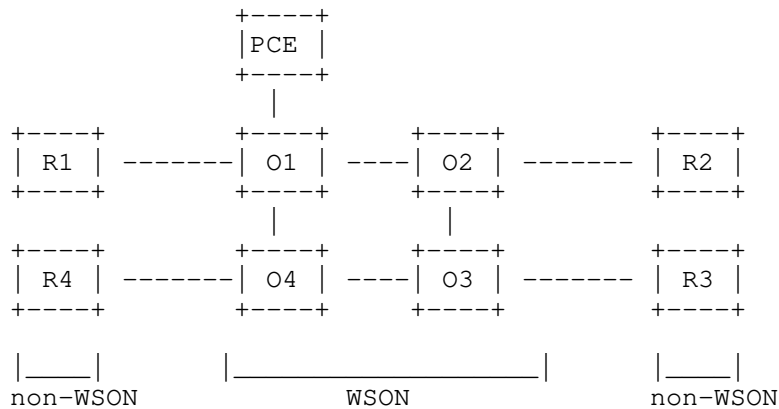


Figure 1 WSON including ROADMs and non-WSN including routers

5. Security considerations

This document does not require changes to the security models within GMPLS and associated protocols. That is, the OSPF-TE, RSVP-TE, and PCEP security models could be operated unchanged.

6. IANA Considerations

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

7. Acknowledgments

Anyone who provide comments and helpful inputs.

8. References

- [1] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [2] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203,

October 2005.

- [3] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [4] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [5] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [6] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [7] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [8] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.
- [9] Otani, T. and D. Li, "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, March 2011.
- [10] Lee, Y., Casellas, R., Margaria, C., and O. Dios, "PCEP Extension for WSON Routing and Wavelength Assignment", draft-lee-pce-wson-rwa-ext-01 (work in progress), March 2011.
- [11] Zhang, F., Lee, Y., Han, J., Bernstein, G., Xu, Y., Zhang, G., Li, D., Chen, M., and Y. Ye, "OSPF Extensions in Support of Routing and Wavelength Assignment (RWA) in Wavelength Switched Optical Networks (WSONs)", draft-zhang-ccamp-rwa-wson-routing-ospf-03 (work in progress), March 2010.
- [12] Bernstein, G., Lee, Y., Li, D., and W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", draft-ietf-ccamp-general-constraint-encode-05 (work in progress), May 2011.
- [13] Katz, D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 9999, September 2003.

Authors' Addresses

Shimazaki Daisaku
NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585
Japan

Phone: +81 422 59 7443
Email: shimazaki.daisaku@lab.ntt.co.jp

Hayashi Rie
NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585
Japan

Phone: +81 422 59 3180
Email: hayashi.rie@lab.ntt.co.jp

Shiomoto Kohei
NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585
Japan

Phone: +81 422 59 4402
Email: shiomoto.kohei@lab.ntt.co.jp

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 3, 2012

W. Sun
SJTU
T. Nadeau
Lucidvision
M. Morrow
Cisco Systems
G. Zhang
CATR
W. Hu
SJTU
July 2, 2011

Label Switched Path (LSP) Provisioning Performance Management
Information Base for Generalized MPLS (GMPLS) / MPLS-TE networks
draft-sun-ccamp-gmpls-perf-mib-00.txt

Abstract

This memo defines Management Information Bases (MIBs) for performances of provisioning Label Switched Paths (LSPs) in Generalized MPLS or MPLS-TE networks.

When Generalized MPLS/MPLS-TE is used to provision LSPs, it is useful to record the performance of the provisioning process, such as the delay in creating and deleting the LSPs. The managed information may be retrieved by the Management System and visualized on the GUI, so that the performance of dynamic provisioning may be monitored in a timely manner.

This work is a continuation of the work in [RFC5814] and [I-D.ietf-ccamp-dpm], where the provisioning performance values are obtained through active measurements.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 3, 2012.

Copyright Notice

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

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

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

Table of Contents

1. Introduction	4
2. Conventions Used in This Document	5
3. The Internet-Standard Management Framework	6
4. Brief Description of LSP performance MIB Objects	7
4.1. gmplsPerfMaxEntries	7
4.2. gmplsPerfTunnelConfigured	7
4.3. gmplsPerfTable	7
5. GMPLS Performance MIB Module	8
6. References	15
6.1. Normative References	15
6.2. Informative References	15
Authors' Addresses	16

1. Introduction

When Label Switched Paths (LSPs) are provisioned dynamically within an operational network, it is helpful to monitor and record the related performance information, such as the experienced provisioning delay and error events. Such information may help operators to ensure correct operation of dynamic LSP provisioning in their network, or possibly identify performance degradation in the control plane.

This memo defines a set of objects that can reveal the performance of an operational network in terms of dynamic LSP provisioning. It is intended to complement the performance objects, such as the number of packets received and sent, per LSP tunnel, in [RFC3812] and [RFC4802].

Unlike the work in [RFC5814] and [I-D.ietf-ccamp-dpm], where the performance values are obtained through active measurements, this document focuses on the performance values in operational environments. The actual value of the performance in this document is recorded only when an LSP is provisioned, and is thus collected passively. Hence such information reflects only the performance at specific and discrete times. However, when properly used, they can be helpful in identifying performance degradation, or even malfunctioning, in the network control plane.

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. 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 [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].

4. Brief Description of LSP performance MIB Objects

4.1. gmplsPerfMaxEntries

Defines the maximum number of rows stored in the gmplsPerfTable. An implementation MUST start assigning gmplsPerfEntryIndex values at 1 and wrap after exceeding the maximum possible value, as defined by the limit of this object.

4.2. gmplsPerfTunnelConfigured

Defines the The total number of tunnels configured.

4.3. gmplsPerfTable

The performance of past LSP provisioning process is stored in this table. To handle possible provisioning failures, start and complete timestamp of a provisioning operation is recorded. For example, for LSP creation process, the timestamps of creation initiation and completion are recorded seperatly. It is up to the users to determine the actual performance value, or identify a possible creation/deletion failure. The maximum number of entries stored in this table is determined by the value of gmplsPerfMaxEntries.

5. GMPLS Performance MIB Module

```
GMPLS-PROV-PERF-STD-MIB DEFINITIONS ::= BEGIN

IMPORTS
    gmplsTeStdMIB
        FROM GMPLS-TE-STD-MIB
    mplsStdMIB,
    MplsTunnelIndex,
    MplsExtendedTunnelId
        FROM MPLS-TC-STD-MIB                -- RFC 3811
    TimeStamp
        FROM SNMPv2-TC
    MODULE-IDENTITY, OBJECT-TYPE,
    Gauge32, Unsigned32
        FROM SNMPv2-SMI
    OBJECT-GROUP
        FROM SNMPv2-CONF;

gmplsPerfMIB MODULE-IDENTITY
    LAST-UPDATED "201104180654Z"      -- Apr 18, 2011 6:54:00 AM
    ORGANIZATION "IETF Common Control and Measurement Plane Working
Group"
    CONTACT-INFO
        "Weiqiang Sun
        Shanghai Jiao Tong University (SJTU)
        Email: sunwq@mit.edu

        Thomas D. Nadeau
        Email: thomas.nadeau@huawei.com"
    DESCRIPTION
        "Copyright (C) The Internet Society (2011).  This version of
        this MIB module is part of RFC XXX; see the RFC itself for
        full legal notices.

        This MIB module defines managed object definitions
        for dynamic LSP provisioning."
    REVISION "201104180654Z"          -- Apr 18, 2011 6:54:00 AM
    DESCRIPTION
        "Initial version."
    -- 1.3.6.1.2.1.10.166.13.1
    ::= { gmplsTeStdMIB 1 }

gmplsPerfTunnelConfigured OBJECT-TYPE
    SYNTAX      Unsigned32
    MAX-ACCESS  read-only
    STATUS      current
```

DESCRIPTION

"The total number of tunnels configured."

-- 1.3.6.1.2.1.10.166.13.1.3

::= { gmplsPerfMIB 3 }

gmplsPerfMaxEntries OBJECT-TYPE

SYNTAX Gauge32

MAX-ACCESS read-only

STATUS current

DESCRIPTION

"An implementation MUST start assigning gmplsPerfEntryIndex values at 1 and wrap after exceeding the maximum possible value, as defined by the limit of this object.

A value of 0 for this object disables creation of gmplsPerfEntry."

-- 1.3.6.1.2.1.10.166.13.1.2

::= { gmplsPerfMIB 2 }

--Performance Table

gmplsPerfTable OBJECT-TYPE

SYNTAX SEQUENCE OF GmplsPerfEntry

MAX-ACCESS not-accessible

STATUS current

DESCRIPTION

"Defines a table for storing the results of LSP provisioning operations. It allows the provisioning performance be retrieved later for monitoring or diagnostic purposes. The recorded performance information is intended to complement the existing performance statistics in the MPLS-TE-STD-MIB and GMPLS-TE-STD-MIB.

Note that the creation and tear-down operation performances are stored in one table, ie., gmplsPerfTable. When an LSP tunnel creation operation is initiated, an entry MUST be added in this table and Tunnel ID as well as the time of initiation MUST be recorded. Upon completion of the creation process, ie., a positive signaling feedback is received by the ingress LSR, this complete time object in this entry MUST be updated.

When an LSP tunnel deletion process is initiated, the corresponding entry with the same tunnel ID MUST be located and updated with time of the deletion initiation time. When the deletion operation is complete, the entry MUST again

be updated with the completion time.

Under circumstances that the creation or deletion operation may fail, an entry may be partially updated. Eg., when a creation operation timeouts without a positive signaling feedback, the creation completion time may never be updated. When a tear-down operation is caused by nodes other than the Ingress LSR, the tear-down start time may not be known to the ingress LSR. In such cases, the user of the MIB MUST be aware of such events and treat the performance information accordingly.

The number of entries in this table is limited by the value of the corresponding gmplsPerfMaxEntries object. An implementation MUST start assigning gmplsPerfEntryIndex at 1 and wrap after exceeding the maximum possible value, as defined by the limit of gmplsPerfMaxEntries. An implementation of this MIB will remove the oldest entry in the gmplsPerfTable to allow the addition of a new entry once the number of rows in the gmplsPerfTable reaches the value specified by gmplsPerfMaxEntries."

```
-- 1.3.6.1.2.1.10.166.13.1.1
::= { gmplsPerfMIB 1 }
```

gmplsPerfEntry OBJECT-TYPE

SYNTAX GmplsPerfEntry

MAX-ACCESS not-accessible

STATUS current

DESCRIPTION

"Defines an entry in the gmplsPerfTable. An entry can be created when an LSP tunnel is signaled. An implementation of this MIB MAY choose to disable the creation of performance entry, when an LSP is provisioned through SNMP."

INDEX {

gmplsPerfEntryIndex,
gmplsPerfTunnelID }

```
-- 1.3.6.1.2.1.10.166.13.1.1.1
```

```
::= { gmplsPerfTable 1 }
```

GmplsPerfEntry ::= SEQUENCE {

gmplsPerfEntryIndex	Gauge32,
gmplsPerfTunnelID	MplsTunnelIndex,
gmplsPerfCurrentStatus	INTEGER,
gmplsPerfSrcID	MplsExtendedTunnelId,

```
gmplsPerfDstID          MplsExtendedTunnelId,
gmplsPerfCreateStartTime TimeStamp,
gmplsPerfCreateCompleteTime TimeStamp,
gmplsPerfDeleteStartTime TimeStamp,
gmplsPerfDeleteCompleteTime TimeStamp }

gmplsPerfEntryIndex OBJECT-TYPE
    SYNTAX  Gauge32
    MAX-ACCESS not-accessible
    STATUS  current
    DESCRIPTION
        "The index of the performance entry. The number of entries
        in this table is limited by the value of the corresponding
        gmplsPerfMaxEntries object. An implementation MUST start
        assigning gmplsPerfEntryIndex at 1 and wrap after exceeding
        the maximum possible value, as defined by the limit of
        gmplsPerfMaxEntries. An implementation of this MIB will
        remove the oldest entry in the gmplsPerfTable to allow the
        addition of an new entry once the number of rows in the
        gmplsPerfTable reaches the value specified by
        gmplsPerfMaxEntries."
    -- 1.3.6.1.2.1.10.166.13.1.1.1.1
    ::= { gmplsPerfEntry 1 }

gmplsPerfTunnelID OBJECT-TYPE
    SYNTAX  MplsTunnelIndex
    MAX-ACCESS read-create
    STATUS  current
    DESCRIPTION
        "The ID of the tunnel being provisioned."
    REFERENCE
        "RFC 3812"
    -- 1.3.6.1.2.1.10.166.13.1.1.1.2
    ::= { gmplsPerfEntry 2 }

gmplsPerfCurrentStatus OBJECT-TYPE
    SYNTAX  INTEGER {
        CreationInProgress(0),
        Up(1),
        DeletionInProgress(2),
        Deleted(3) }
    MAX-ACCESS read-create
    STATUS  current
    DESCRIPTION
        "This object defines the current status of the LSP tunnel."
```

CreationInProgress

The corresponding LSP tunnel is being created, but the creation operation has not finished yet.

Up

The corresponding LSP tunnel has been created successfully.

DeletionInProgress

The corresponding LSP tunnel is being deleted, but the deletion process has not finished yet.

Deleted

The corresponding LSP tunnel has been deleted."

-- 1.3.6.1.2.1.10.166.13.1.1.1.3

::= { gmplsPerfEntry 3 }

gmplsPerfSrcID OBJECT-TYPE

SYNTAX MplsExtendedTunnelId

MAX-ACCESS read-create

STATUS current

DESCRIPTION

"The address of the ingress LSR ID."

-- 1.3.6.1.2.1.10.166.13.1.1.1.5

::= { gmplsPerfEntry 5 }

gmplsPerfDstID OBJECT-TYPE

SYNTAX MplsExtendedTunnelId

MAX-ACCESS read-create

STATUS current

DESCRIPTION

"The address of the egress LSR ID."

-- 1.3.6.1.2.1.10.166.13.1.1.1.6

::= { gmplsPerfEntry 6 }

gmplsPerfCreateStartTime OBJECT-TYPE

SYNTAX TimeStamp

MAX-ACCESS read-create

STATUS current

DESCRIPTION

"The time when the tunnel setup operation is initiated."

-- 1.3.6.1.2.1.10.166.13.1.1.1.7

::= { gmplsPerfEntry 7 }


```
gmplsPerfCreateCompleteTime OBJECT-TYPE
    SYNTAX      TimeStamp
    MAX-ACCESS  read-create
    STATUS      current
    DESCRIPTION
        "The time when the LSP tunnel create operation
         is complete."
    -- 1.3.6.1.2.1.10.166.13.1.1.1.8
    ::= { gmplsPerfEntry 8 }

gmplsPerfDeleteStartTime OBJECT-TYPE
    SYNTAX      TimeStamp
    MAX-ACCESS  read-create
    STATUS      current
    DESCRIPTION
        "The time when the LSP Tunnel tear-down operation
         is initiated."
    -- 1.3.6.1.2.1.10.166.13.1.1.1.9
    ::= { gmplsPerfEntry 9 }

gmplsPerfDeleteCompleteTime OBJECT-TYPE
    SYNTAX      TimeStamp
    MAX-ACCESS  read-create
    STATUS      current
    DESCRIPTION
        "The time when an LSP tear-down operation
         is complete."
    -- 1.3.6.1.2.1.10.166.13.1.1.1.10
    ::= { gmplsPerfEntry 10 }

gmplsPerfGroups OBJECT IDENTIFIER
    -- 1.3.6.1.2.1.10.166.13.1.4
    ::= { gmplsPerfMIB 4 }

gmplsDeletionGroup OBJECT-GROUP
    OBJECTS {
        gmplsPerfTunnelID,
        gmplsPerfCurrentStatus,
        gmplsPerfSrcID,
        gmplsPerfDstID,
        gmplsPerfDeleteStartTime,
        gmplsPerfDeleteCompleteTime }
    STATUS      current
    DESCRIPTION
        "The group of object that constitute the LSP tunnel
```

```
        deletion performance."
-- 1.3.6.1.2.1.10.166.13.1.4.1
::= { gmplsPerfGroups 1 }

gmplsCreationGroup OBJECT-GROUP
  OBJECTS {
    gmplsPerfTunnelID,
    gmplsPerfCurrentStatus,
    gmplsPerfSrcID,
    gmplsPerfDstID,
    gmplsPerfCreateStartTime,
    gmplsPerfCreateCompleteTime }
  STATUS current
  DESCRIPTION
    "The group of object that constitute the LSP tunnel
    creation performance."
-- 1.3.6.1.2.1.10.166.13.1.4.2
::= { gmplsPerfGroups 2 }

gmplsPerfBasicGroup OBJECT-GROUP
  OBJECTS {
    gmplsPerfEntryIndex,
    gmplsPerfTunnelID,
    gmplsPerfMaxEntries,
    gmplsPerfCurrentStatus,
    gmplsPerfCreateStartTime,
    gmplsPerfCreateCompleteTime,
    gmplsPerfDeleteStartTime,
    gmplsPerfDeleteCompleteTime,
    gmplsPerfDstID,
    gmplsPerfSrcID,
    gmplsPerfTunnelConfigured,
    gmplsPerfErrThreshold }
  STATUS current
  DESCRIPTION
    "Basic objects."
-- 1.3.6.1.2.1.10.166.13.1.4.3
::= { gmplsPerfGroups 3 }

END
```

6. References

6.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, April 1999.
- [RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, April 1999.
- [RFC2580] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Conformance Statements for SMIv2", STD 58, RFC 2580, April 1999.
- [RFC3410] Case, J., Mundy, R., Partain, D., and B. Stewart, "Introduction and Applicability Statements for Internet-Standard Management Framework", RFC 3410, December 2002.
- [RFC3812] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812, June 2004.
- [RFC4802] Nadeau, T. and A. Farrel, "Generalized Multiprotocol Label Switching (GMPLS) Traffic Engineering Management Information Base", RFC 4802, February 2007.

6.2. Informative References

- [I-D.ietf-ccamp-dpm] Sun, W. and G. Zhang, "Label Switched Path (LSP) Data Path Delay Metrics in Generalized MPLS/ MPLS-TE Networks", draft-ietf-ccamp-dpm-03 (work in progress), May 2011.
- [RFC5814] Sun, W. and G. Zhang, "Label Switched Path (LSP) Dynamic Provisioning Performance Metrics in Generalized MPLS Networks", RFC 5814, March 2010.

Authors' Addresses

Weiqiang Sun
Shanghai Jiao Tong University
800 Dongchuan Road
Shanghai 200240
China

Phone: +86 21 3420 5359
Email: sunwq@mit.edu

Thomas D. Nadeau
Lucidvision

Email: tnadeau@lucidvision.com

Monique Morrow
Cisco Systems
Richistrasse 7
CH-8304 Zurich-Wallisellen
Switzerland

Phone: +41 44 878 9412
Email: mmorrow@cisco.com

Guoying Zhang
China Academy of Telecommunication Research, MII.
No.52 Hua Yuan Bei Lu, Haidian District
Beijing 100083
China

Phone: +86-1062300106
Email: zhangguoying@mail.ritt.com.cn

Weisheng Hu
Shanghai Jiao Tong University
800 Dongchuan Road
Shanghai 200240
China

Phone: +86 21 3420 5419
Email: wshu@sjtu.edu.cn

Network Working Group
Internet Draft
Category: Standards Track

Fatai Zhang
Huawei
Guoying Zhang
CATR
Sergio Belotti
Alcatel-Lucent
D. Ceccarelli
Ericsson
Khuzema Pithewan
Infinera
July 8, 2011

Expires: January 8, 2012

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

draft-zhang-ccamp-gmpls-evolving-g709-08.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on January 8, 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
3.1. Requirements for supporting services over hierarchical OTN network	5
4. Extensions for Traffic Parameters for the Evolving G.709	8
4.1. Usage of ODUFlex(CBR) Traffic Parameter	9
4.2. Example of ODUFlex(CBR) Traffic Parameter	10
5. Generalized Label	11
5.1. New definition of Single-stage ODUk Generalized Label ...	11
5.1.1. Examples	14
5.1.2. Label Distribution Procedure	16
5.1.2.1. Notification on Label Error	17
5.1.3. Supporting Virtual Concatenation and Multiplication.	17
5.1.4. Supporting Multiplexing Hierarchy	18
5.1.5. Supporting One-hop Multiplexing Hierarchy via Single Session	19
5.1.5.1. Multiplexing Hierarchy and Solution Alternatives	19
5.1.5.2. Multi Stage Label Format	19
5.1.5.3. Label format for NVC or Multiplier > 1	20

5.1.5.4. Usage of Multi-stage Label in Multi Stage Muxing	21
5.2. New definition of Multi-stage ODUk Generalized Label	22
5.2.1. Multi-stage Label	23
5.2.2. Label format for NVC or Multiplier > 1	24
5.2.3. Usage of Multi-stage Label	24
5.2.4. Label Distribution Rules	26
5.2.5. Examples	27
5.3. Control Plane Backward Compatibility Considerations	29
6. Security Considerations	30
7. IANA Considerations	30
8. References	31
8.1. Normative References	31
8.2. Informative References	32
9. Authors' Addresses	33
Acknowledgment	35

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] extends MPLS to include Layer-2 Switching (L2SC), Time-Division Multiplex (e.g., SONET/SDH, PDH, and ODU), Wavelength (OCh, Lambdas) Switching, and Spatial Switching (e.g., incoming port or fiber to outgoing port or fiber). [RFC3471] presents a functional description of the extensions to Multi-Protocol Label Switching (MPLS) signaling required to support Generalized MPLS. RSVP-TE-specific formats and mechanisms and technology specific details are defined in [RFC3473].

With the evolution and deployment of G.709 technology, it is necessary that appropriate enhanced control technology support be provided for G.709. [RFC4328] describes the control technology details that are specific to foundation G.709 Optical Transport Networks (OTN), as specified in the ITU-T Recommendation G.709 [G709-V1], for ODUk deployments without multiplexing.

In addition to increasing need to support ODUk multiplexing, the evolution of OTN has introduced additional containers and new flexibility. For example, ODU0, ODU2e, ODU4 containers and ODUFlex are developed in [G709-V3].

In addition, the following issues require consideration:

- Support for hitless adjustment of ODUFlex, which is to be specified in ITU-T G.hao.
- Support for Tributary Port Number. The Tributary Port Number has to be negotiated on each link for flexible assignment of

tributary ports to tributary slots in case of LO-ODU over HO-ODU (e.g., ODU2 into ODU3).

Therefore, it is clear that [RFC4328] has to be updated or superseded in order to support ODUk multiplexing, as well as other ODU enhancements introduced by evolution of OTN standards.

This document updates [RFC4328] extending the G.709 ODUk traffic parameters and also presents a new OTN label format which is very flexible and scalable.

2. Terminology

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

3. GMPLS Extensions for the Evolving G.709 - Overview

New features for the evolving OTN, for example, new ODU0, ODU2e, ODU4 and ODUflex containers are specified in [G709-V3]. The corresponding new signal types are summarized below:

- Optical Channel Transport Unit (OTUk):
 - . OTU4
- Optical Channel Data Unit (ODUk):
 - . ODU0
 - . ODU2e
 - . ODU4
 - . ODUflex

A new Tributary Slot (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 will be used for 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 ($j = 0, 1, 2, 2e, 3, \text{flex}$) into an ODUk ($k > j$), as described in Section 3.1.2 of [OTN-frwk].

Virtual Concatenation (VCAT) of OPUk (OPUk-Xv, $k = 1/2/3$, $X = 1 \dots 256$) are also supported by [OTN-V3]. Note that VCAT of OPU0 / OPU2e / OPU4 / OPUflex are 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 OTN label format which is very flexible and scalable. Additionally, procedures about Tributary Port Number assignment through control plane are also provided in this document.

3.1. Requirements for supporting services over hierarchical OTN network

[Editor's Note] The section 3.1 about requirements will be moved to the framework document after discussion.

- 1.[R1] Support signaling mechanism to instantiate ODU_j service layer on an ODU_k link via single stage muxing.

An ODU_j LSP could involve zero ($j=k$) or one stage ($j<k$) multiplexing on a given ODU_k link. Here both Control-plane and Data-plane entities are created for the ODU_j service layer. ODU_k link could be a point-to-point OTU_k link or an H-LSP. This is the most foundational and important requirement the control plane should support.

- 2.[R2] Support signaling mechanism to instantiate ODU_j LSP involving one or more intermediate ODU layers (either pre-existing or not) which cross multiple ODU_k links.

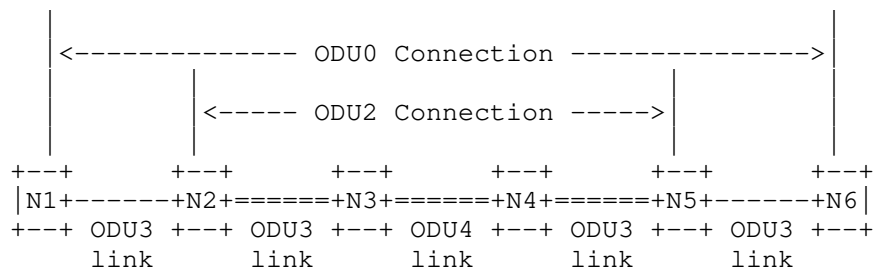


Figure 1 - Requirement 2

Figure 1 shows an example where the ODU0 LSP is multiplexed into an intermediate ODU2, which crosses three ODU links between N2 and N5.

There are two typical scenarios requesting two or more stage multiplexing crossing multiple ODUk links:

- Tunnel scenario: Assume that N3 and N4 in figure 1 are legacy nodes which don't support ODU0 or ODUFlex cross-connection. In order to create ODU0 or ODUFlex service between N1 and N6, an intermediate ODU2 connection can be created between N2 and N5. Then, the ODU0 or ODUFlex can be multiplexed into this ODU2 connection. In this case, N3 and N4 only need to perform ODU2 cross-connection and are not aware of ODU0 or ODUFlex service inside.
- Carrier-in-carrier scenario: Assume that N2, N3, N4 and N5 in figure 1 belong to carrier A, while N1 and N6 belong to carrier B. Carrier B may lease an ODU2 pipe between N2 and N5, which is pre-provisioned by carrier A, to carry LO ODU services between N1 and N6.

More specifically, this requirement can be further divided into two items:

[R2.1] Support signaling mechanism to trigger the creation of one or more intermediate ODU layers over multiple ODUk links based on the ODUj LSP creation request.

[R2.2] Support signaling mechanism to instantiate ODUj service layer on multiple ODUk links where one or more intermediate ODU layers may be pre-existing.

- 3.[R3] Support signaling mechanism to instantiate ODUj LSP involving one or more intermediate ODU layers (either pre-existing or not) on one hop ODUk link.

More specifically, this requirement can be further divided into two items:

[R3.1] Support signaling mechanism to instantiate one or more intermediate layers on one hop ODUk link in order to support the ODUj service layer.

An ODUj LSP could involve two or more stage multiplexing on a given ODUk link. These intermediate layers may be implicitly created as a part of ODUj service LSP creation. In this case, both control plane and data plane entities will be created for the ODUj service layer. However, intermediate ODU layer(s) (implicitly created) will have data plane representation only.

[R3.2] Support signaling mechanism to instantiate ODUj service layer on an ODUk link where one or more intermediate ODU layers may be pre-existing.

An ODUj LSP could involve two or more stage multiplexing on a given ODUk link. These intermediate layers may be pre-existing as a result of another LSP creation on the same ODU hierarchy or explicitly configured through management interface.

- 4.[R4] Support controllable and manageable capability for the intermediate ODU layers which cross one or more hops of ODUk links and which is used for carrying ODUj services.

Once the intermediate ODU layers are created by control plane (may be triggered by the ODUj service or by management plane), they should be under the control of control plane or management plane. The following typical scenarios should be considered:

- The control/management plane should have the capability to reroute the intermediate ODU layers to recover all the contained ODUj layer services to improve the recovery performance after network failure occurs in the intermediate ODU layers.
- The control/management plane should have the capability to delete an empty intermediate ODU connection (i.e., without any ODUj service inside it) to release the bandwidth resource of ODUk link. For example, the management plane may request the control plane to delete an empty intermediate ODU2 in an ODU4 link so that the ODU4 link has enough bandwidth resource to carry a new ODU3 service.

- 5.[R5] Support signaling mechanism where ODUj service LSP creation may involve varying mux hierarchies on each hop.

An end-to-end ODUj service LSP creation may involve zero or more stage ODU multiplexing on every hop in the path. Basically, the scenarios discussed in R1 to R3 could be associated with any of the hops involved.

- 6.[R6] Support signaling mechanism for egress control of OTN interfaces.

An egress interface of an ODUj LSP could involve single or multiple stage multiplexing. Egress Label sub-object defined in [RFC-4003] must be used to signal hierarchical multiplexing information pertaining to the egress interface of the LSP.

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

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

0	1	2	3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1			
Signal Type	Tolerance	NMC	
NVC	Multiplier (MT)		
Bit_Rate			

The Signal Type should be extended to cover the new Signal Type introduced by the evolving OTN. The new Signal Type is extended as follows:

Value	Type
0	Not significant
1	ODU1 (i.e., 2.5 Gbps)
2	ODU2 (i.e., 10 Gbps)
3	ODU3 (i.e., 40 Gbps)
4	ODU4 (i.e., 100 Gbps)
5	Reserved (for future use)
6	OCh at 2.5 Gbps
7	OCh at 10 Gbps
8	OCh at 40 Gbps
9	OCh at 100 Gbps
10	ODU0 (i.e., 1.25 Gbps)
11	ODU2e (i.e., 10Gbps for FC1200 and GE LAN)
12~19	Reserved (for future use)
20	ODUflex(CBR) (i.e., 1.25*N Gbps)
21	ODUflex(GFP-F), resizable (i.e., 1.25*N Gbps)
22	ODUflex(GFP-F), non resizable (i.e., 1.25*N Gbps)
23~255	Reserved (for future use)

In case of ODUflex(CBR), the Bit_Rate and Tolerance fields are used together to represent the actual bandwidth of ODUflex, where:

- The Bit_Rate field indicates the nominal bit rate of ODUflex(CBR) encoded as a 32-bit IEEE single-precision floating-point number (referring to [RFC4506] and [IEEE]).

- The Tolerance field indicates the bit rate tolerance (part per million, ppm) of the ODUflex(CBR) encoded as an unsigned integer, which is bounded in 0~100ppm.

For example, for an ODUflex(CBR) service with Bit_Rate = 2.5Gbps and Tolerance = 100ppm, the actual bandwidth of the ODUflex is:

$$2.5\text{Gbps} * (1 - 100\text{ppm}) \sim 2.5\text{Gbps} * (1 + 100\text{ppm})$$

In case of other ODUk signal types, the Bit_Rate and Tolerance fields are not necessary and MUST be filled with 0.

The usage of the NMC, NVC and Multiplier (MT) fields are the same as [RFC4328].

4.1. Usage of ODUflex(CBR) Traffic Parameter

In case of ODUflex(CBR), the information of Bit_Rate and Tolerance in the ODUflex traffic parameter is used to determine the total number of tributary slots N in the HO ODUk link to be reserved. Here:

$$N = \text{Ceiling of}$$

$$\frac{\text{ODUflex(CBR) nominal bit rate} * (1 + \text{ODUflex(CBR) bit rate tolerance})}{\text{ODUk.ts nominal bit rate} * (1 - \text{HO OPUk bit rate tolerance})}$$

Therefore, a node receiving a Path message containing ODUflex(CBR) traffic parameter can allocate precise number of tributary slots and set up the cross-connection for the ODUflex service.

Table 1 below shows the actual bandwidth of the tributary slot of ODUk (in Gbps), referring to [G709-V3].

Table 1 - Actual TS bandwidth of ODUk

ODUk	Minimum	Nominal	Maximum
ODU2	1.249 384 632	1.249 409 620	1.249 434 608
ODU3	1.254 678 635	1.254 703 729	1.254 728 823
ODU4	1.301 683 217	1.301 709 251	1.301 735 285

Note that:

$$\text{Minimum bandwidth of ODUTk.ts} = \text{ODUTk.ts nominal bit rate} * (1 - \text{HO OPUk bit rate tolerance})$$

Maximum bandwidth of ODTUk.ts =
 ODTUk.ts nominal bit rate * (1 + HO OPUk bit rate tolerance)

Where: HO OPUk bit rate tolerance = 20ppm

For different ODUk, the bandwidths of the tributary slot 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. This is why the traffic parameter should bring the actual bandwidth information other than the NMC field.

4.2. Example of ODUFlex(CBR) Traffic Parameter

This section gives an example to illustrate the usage of ODUFlex(CBR) traffic parameter.

As shown in Figure 2, 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 parameter indicates that Signal Type is 33 (ODUFlex(CBR)), Bit_Rate is 2.5Gbps and Tolerance is 100ppm.

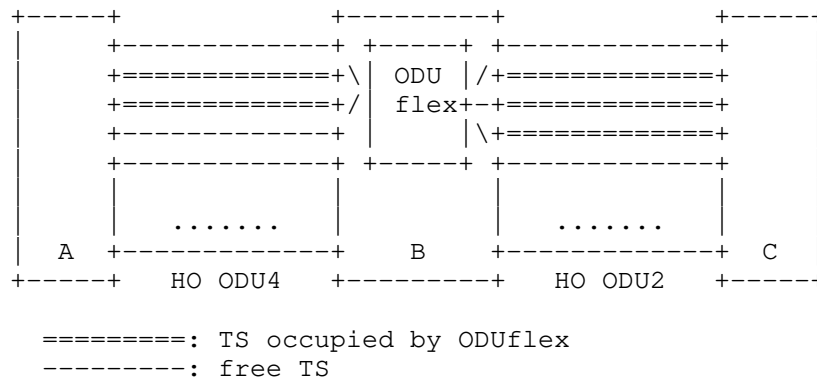


Figure 2 - Example of ODUFlex(CBR) Traffic Parameter

- On the HO ODU4 link between node A and B:

The maximum bandwidth of the ODUflex equals $2.5\text{Gbps} * (1 + 100\text{ppm})$, and the minimum bandwidth of the tributary slot of ODU4 equals $1.301\ 683\ 217\text{Gbps}$, so the total number of tributary slots N1 to be reserved on this link is:

$$N1 = \text{ceiling} (2.5\text{Gbps} * (1 + 100\text{ppm}) / 1.301\ 683\ 217) = 2$$

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

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

$$N2 = \text{ceiling} (2.5\text{Gbps} * (1 + 100\text{ppm}) / 1.249\ 384\ 632) = 3$$

5. Generalized Label

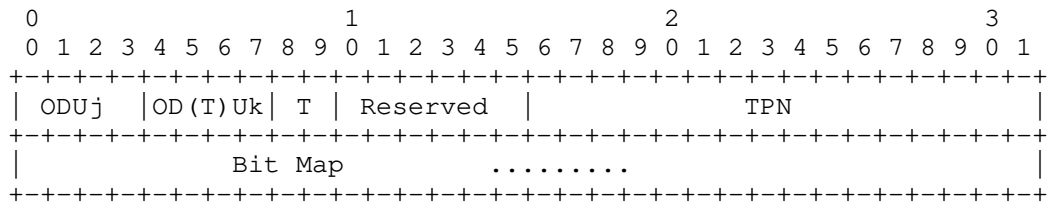
[RFC3471] has defined the Generalized Label which extends the traditional label by allowing the representation of not only labels which travel in-band with associated data packets, but also labels which identify time-slots, wavelengths, or space division multiplexed positions. The format of the corresponding 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.

[RFC 6107] defines using hierarchical LSP for MLN. The H-LSPs can be setup manually or dynamically (induced FAs) for multi-stage multiplexing scenarios. Service creation in hierarchical OTN network can be achieved in following 2 ways.

5.1. New definition of Single-stage ODUk Generalized Label

In order to be compatible with new types of ODU signal and new types of tributary slot, the following new ODUk label format is defined:



ODUj and OD(T)Uk (4 bits respectively): indicate that LO ODUj is multiplexed into HO ODUk (k>j), or LO ODUj is mapped into OTUk (j=k).

ODUj field	Signal type
-----	-----
0	LO ODU0
1	LO ODU1
2	LO ODU2
3	LO ODU3
4	LO ODU4
5	LO ODU2e
6	LO ODUflex
7-15	Reserved (for future use)

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

T (2 bits): indicates the type of tributary slot of HO ODUk when LO ODUj is multiplexed into the HO ODUk (j<k). Currently, two types of tributary slot are defined in [G709-V3], the 1.25Gbps tributary slot and the 2.5Gbps tributary slot.

T field	TS type
-----	-----
0	1.25Gbps TS granularity
1	2.5Gbps TS granularity
2-3	Reserved (for future use)

In case of LO ODU_j mapped into OTU_k (j=k), this field is not necessary and should be ignored.

TPN (16 bits): indicates the Tributary Port Number (TPN) for the assigned Tributary Slot(s).

- In case of LO ODU_j multiplexed into HO ODU1/ODU2/ODU3, only the lower 6 bits of TPN field is significant and the other bits of TPN MUST be set to 0.
- In case of LO ODU_j multiplexed into HO ODU4, only the lower 7 bits of TPN field is significant and the other bits of TPN MUST be set to 0.
- In case of ODU_j mapped into OTU_k (j=k), the TPN is not needed and this field MUST be set to 0.

As per [G709-V3], The TPN is used to allow for correct demultiplexing in the data plane. When an LO ODU_j is multiplexed into HO ODU_k occupying one or more TSs, a new TPN value is configured at the two end of the HO ODU_k link and is put into the related MSI byte(s) in the OPU_k overhead at the (traffic) ingress end of the link, so that the other end of the link can learn which TS(s) is/are used by the LO ODU_j in the data plane.

According to [G709-V3], the rules of TPN assignment should be as the following tables:

Table 2 - TPN Assignment Rules (2.5Gbps TS granularity)			
HO ODU _k	LO ODU _j	TPN	TPN Assignment Rules
ODU2	ODU1	1~4	Fixed, = TS# occupied by ODU1
ODU3	ODU1	1~16	Fixed, = TS# occupied by ODU1
	ODU2	1~4	Flexible, != other existing LO ODU2s' TPNs

Table 3 - TPN Assignment Rules (1.25Gbps TS granularity)

HO ODUk	LO ODUj	TPN	TPN Assignment Rules
ODU1	ODU0	1~2	Fixed, = TS# occupied by ODU0
ODU2	ODU1	1~4	Flexible, != other existing LO ODU1s' TPNs
	ODU0 & ODUflex	1~8	Flexible, != other existing LO ODU0s and ODUflexes' TPNs
ODU3	ODU1	1~16	Flexible, != other existing LO ODU1s' TPNs
	ODU2	1~4	Flexible, != other existing LO ODU2s' TPNs
	ODU0 & ODU2e & ODUflex	1~32	Flexible, != other existing LO ODU0s and ODU2es and ODUflexes' TPNs
ODU4	Any ODU	1~80	Flexible, != ANY other existing LO ODUs' TPNs

Note that in the case of "Flexible", the value of TPN is not relevant to the TS number as per [G709-V3].

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.

The size of the bit map equals to the total number of the tributary slots of HO ODUk, which is deduced by the ODU(T)k and T fields.

In case of an ODUk mapped into OTUk, it's no need to indicate which tributary slots will be used, so the size of Bit Map is 0.

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.

5.1.1. 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 ODU1 (ODU2 or ODU3 or ODU4) is directly mapped into the corresponding OTU1 (OTU2 or OTU3 or OTU4). The following example label indicates an ODU1 mapped into OTU1.

```

      0              1              2              3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 0 0 1 | 0 0 0 1 | 0 0 | Reserved |               All 0s |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

(2) ODUj into ODUk multiplexing:

In such conditions, this label indicates that an ODUj is multiplexed into several tributary slots of OPUk and then mapped into OTUk. Some instances are shown as follow:

- ODU0 into ODU2 Multiplexing:

```

      0              1              2              3
      0 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 0 0 0 | 0 0 1 0 | 0 0 | Reserved |               TPN = 2 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 1 0 0 0 0 0 0 |               Padded Bits (0) |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This above label indicates an ODU0 multiplexed into the second tributary slot of ODU2, wherein the type of the tributary slot is 1.25Gbps, and the TPN value is 2.

- ODU1 into ODU2 Multiplexing with 1.25Gbps TS granularity:

```

      0              1              2              3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 0 0 1 | 0 0 1 0 | 0 0 | Reserved |               TPN = 1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 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 the type of the tributary slot is 1.25Gbps, and the TPN value is 1.

- ODU2 into ODU3 Multiplexing with 2.5Gbps TS granularity:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0 0 1 0|0 0 1 1|0 1| Reserved |               TPN = 1               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0 1 1 0 1 0 1 0 0 0 0 0 0 0 0 0| Padded Bits (0) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

This above label indicates an ODU2 multiplexed into the 2nd, 3rd, 5th and 7th tributary slot of ODU3, wherein the type of the tributary slot is 2.5Gbps, and the TPN value is 1.

5.1.2. Label Distribution Procedure

This document does not change the existing label distribution procedures [RFC4328] for GMPLS except that the new ODUk label should 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 should 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 should 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 should also assign a valid TPN, which does not collided 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.

When receiving an ODU label from its downstream neighbor node, the node should learn which ODU signal type is multiplexed or mapped into which ODU signal type by analyzing the ODUj and the ODUk fields.

In case of ODUj to ODUk multiplexing, the node should firstly determine the size of the Bit Map field according to the signal type

and the tributary slot type of ODUk, and then obtain which tributary slots in ODUk are reserved by its downstream neighbor node according to the position of the bits that are set to 1 in the Bit Map field, so that the node can multiplex the ODUj into the reserved tributary slots of ODUk after the LSP is established. The node should also get 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 is 0 and no additional procedure is needed.

Note that the procedures of other label related objects (e.g., Upstream Label, Label Set) are similar as 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.

5.1.2.1. 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:

- The ODUj field does not match with the Traffic Parameters;
- The OD(T)Uk field does not match with the type of the selected link;
- The selected link only supports 2.5Gbps TS granularity while the T field in the label indicates the 1.25Gbps TS granularity;
- The label includes an invalid TPN value that breaks the TPN assignment rules;
- Not enough bits of Bit Map, or Bit Map with non-zero padding bits;
- The reserved resources (i.e., the number of "1" in the Bit Map field) do not match with the Traffic Parameters.

5.1.3. Supporting Virtual Concatenation and Multiplication

As per [VCAT], 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 [VCAT].

5.1.4. Supporting Multiplexing Hierarchy

As described in [OTN-FRWK], one ODU_j connection can be nested into another ODU_k ($j < k$) 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 ODU_k but not ODU_j switching.

For example, in Figure 3, 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 inner ODU0.

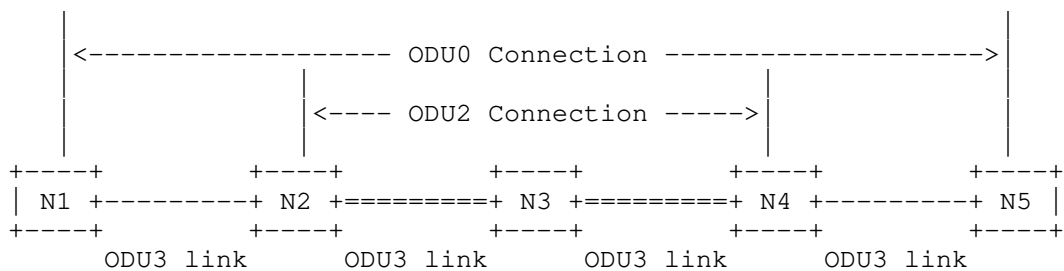


Figure 3 - Example of multiplexing hierarchy

The control plane signaling should support the provisioning of hierarchical multiplexing. Two methods are provided below (taking Figure 3 as example):

- The outer ODU2 connection is created in advance based on network planning, which is treated as a Forwarding Adjacency (FA). Then the inner ODU0 can be created using the resource of the ODU2 FA. In this case, the outer ODU2 and inner ODU0 connections are created separately, and the normal ODU connection creation procedure described in this document can be used.
- 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 an FA after it is created. And then the signaling procedure for the ODU0 connection can be continued using the resource of the ODU2 FA.

5.1.5. Supporting One-hop Multiplexing Hierarchy via Single Session

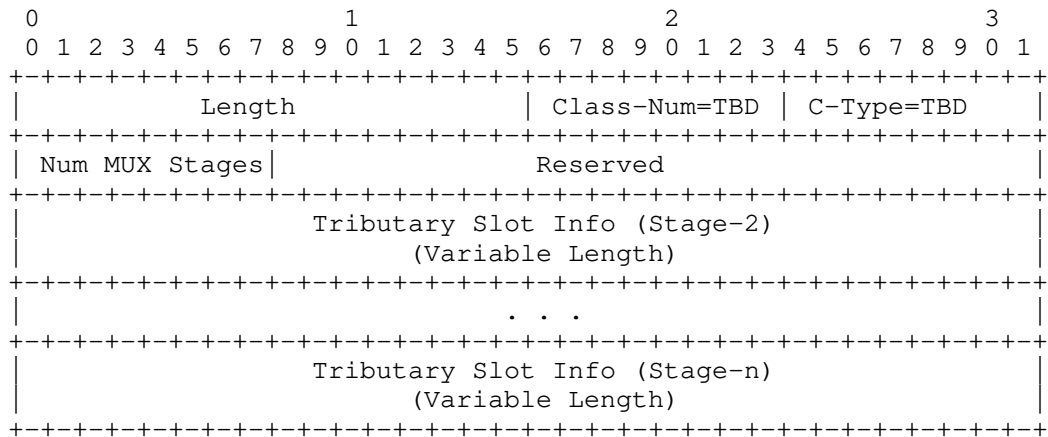
5.1.5.1. Multiplexing Hierarchy and Solution Alternatives

In order to support instantiating ODU_j LSP involving one or more intermediate ODU layers on an ODU_k link (i.e., the scenario described in Requirement 3 of Section 3.1), there are two approaches to achieve the objective. The existing approach is the hierarchical LSP (H-LSP) approach described in Section 5.5, and another one is to use the multi-stage label approach.

For the multi-stage label approach, the whole multiplexing structure on the ODU_k link (i.e., ODU_j service multiplexed into one or more intermediate ODU layers and then multiplexed into ODU_k link) is included in the signaling message which is used for creating the ODU_j service. After receiving the message, both ends of the ODU_k link will construct the multi-stage multiplexing in the data plane. In this way, creation of intermediate ODU layers is treated as part of creation of the ODU_j service, without any intermediate ODU FA on the ODU_k link. Note that the ODU_k link can either be mapped to an OTU_k link directly, or be a multi-hop FA created in advance crossing multiple OTU links (using H-LSP mechanism).

5.1.5.2. Multi Stage Label Format

In this document, a new optional object named MULTI-STAGE LABEL Object is introduced to indicate how the intermediate ODU layers are multiplexed into ODU_k link in the one-hop multi-stage multiplexing scenario. The format of this object is shown below (The Class-Num and the C-Type of this new object are TBD):



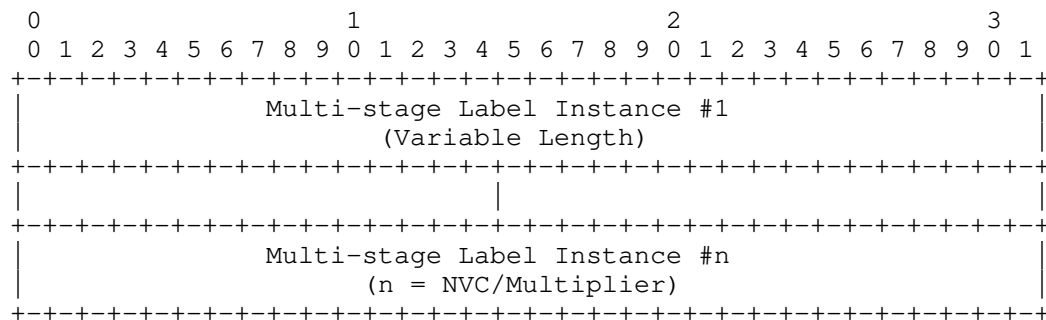
Num MUX Stages: This field indicates the number of multiplexing stages specified by the label.

Tributary Slot Info: This field has the same format as the ODUk label format described in Section 5.1. In the case of n-step multiplexing (e.g., ODUj into ODUi1 into ODUi2 ... into ODUi(n-1) into ODUk multiplexing), The Tributary Slot Info (Stage-2) indicates how ODUi1 is multiplexed into ODUi2; the Tributary Slot Info (Stage-3) indicates how ODUi2 is multiplexed into ODUi3 ... and the Tributary Slot Info (Stage-n) indicates how ODUi(n-1) is multiplexed into the ODUk link. Note that how ODUj is multiplexed into ODUi1 is indicated by the generalized label and is not included in this object.

Note that the MULTI-STAGE LABEL Object is not necessary and must not be included in the signaling message in case the signaling message is used for creating only one ODU layer connection via single stage muxing. One example is to instantiate ODUj service on an ODUk link via single stage muxing. Another example is to use H-LSP mechanism to instantiate ODUj service involving one or more intermediate ODU FAs, where multiple RSVP sessions will be created separately, each of which is used to create one ODU-FA layer connection. In such cases, the generalized label is used without the multi-stage label, as described in Section 5.

5.1.5.3. Label format for NVC or Multiplier > 1

For NVC or Multiplier field value > 1, the multi-stage label format defined in Section 6.2 needs to be repeated NVC/multiplier times.



5.1.5.4. Usage of Multi-stage Label in Multi Stage Muxing

When an ODUj LSP is requested where one or more intermediate ODU layers are involved on an ODUk link, the multi-stage label together with the generalized label can be used to indicate the multi-stage multiplexing structure. The generalized label, as described in Section 5, is used to indicate how the ODUj service is multiplexed into the first intermediate ODU layer and the multi-stage label is used to indicate how the intermediate ODU layers are multiplexed into the ODUk link.

Take Figure 4 as an example. Assume on an OTU3 Link, a restrictive MUX hierarchy is supported on the associated interfaces. In order to switch ODU1 on this Link, ODU3 and ODU2 need to be terminated on the same span as the OTU3 link.

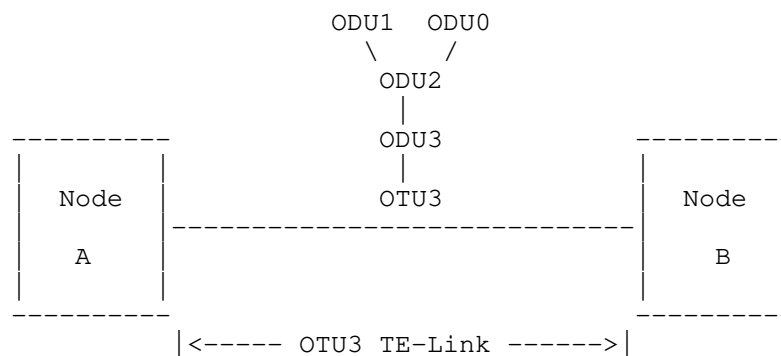


Figure 4 - Multi-stage Label on OTUk Link

In this example, the generalized label is used to indicate how the ODU1 service is multiplexed into the intermediate ODU2, the procedures are the same as described in Section 5. An example generalized label is shown below, assuming that the ODU1 is multiplexed into the 2nd and the 4th tributary slot of ODU2, wherein the type of the tributary slot is 1.25Gbps, and the TPN value is 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
+-----+-----+-----+-----+
| 0 0 0 1 | 0 0 1 0 | 0 0 | Reserved | TPN = 1 |
+-----+-----+-----+-----+
| 0 1 0 1 0 0 0 0 | Padded Bits (0) |
+-----+-----+-----+-----+

```

At the same time, the MULTII-STAGE LABEL Object is also included in the signaling message, which is used to indicate how the intermediate ODU2 is multiplexed into the ODU3. An example multi-stage label is shown below, assuming that the ODU2 is multiplexed into the 2nd, 3rd, 5th and 7th tributary slot of ODU3, wherein the type of the tributary slot is 2.5Gbps, and the TPN value is 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
+-----+-----+-----+-----+
| MUX-Stages=2 | Reserved |
+-----+-----+-----+-----+
| 0 0 1 0 | 0 0 1 1 | 0 1 | Reserved | TPN = 1 |
+-----+-----+-----+-----+
| 0 1 1 0 1 0 1 0 0 0 0 0 0 0 0 0 | Padded Bits (0) |
+-----+-----+-----+-----+

```

5.2. New definition of Multi-stage ODUk Generalized Label

Multi-stage label is a composite label, which can carry timeslot information for one or more ODU layers.

ODUk-----ODUj-----ODUh

TS/TPN for stage-1 TS/TPN for stage-2

Figure 5 - Multi-stage Label

In an OTN network, path of an LSP could be going through links that support restrictive hierarchy. Multi-stage Label is needed when

Service ODU layer requires termination of more than one HO-ODUs on a given OTU/ODU Link.

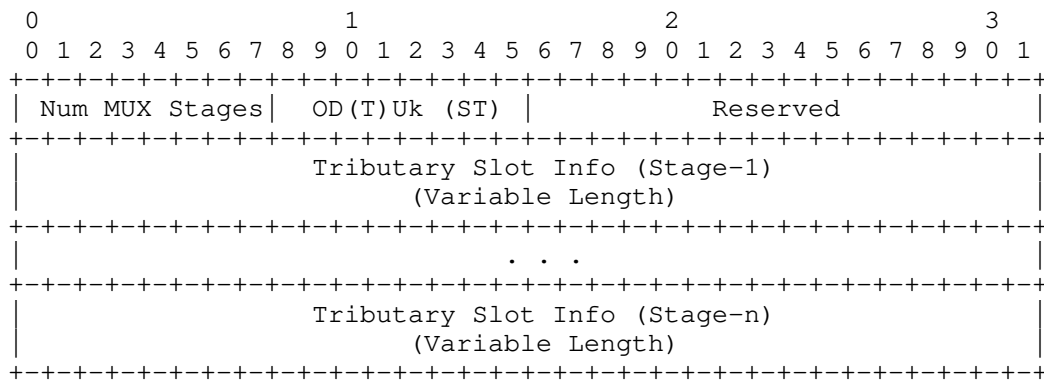
Multi stage label allows implicit creation of intermediate ODU layers for supporting the instantiation of service ODU layer on a given hop, thus eliminating the need for one hop H-LSPs pertaining to intermediate ODU layers.

If higher order ODU layers spans more than one hop due to switching restrictions, H-LSP needs to be used in tandem with multi-stage Label to facilitate end to end service creation.

5.2.1. Multi-stage Label

A multi-stage label includes TS and TPN information for all the stages of a multi-stage multiplexing hierarchy.

The format of a multi-stage label is explained below.



Num MUX Stages:

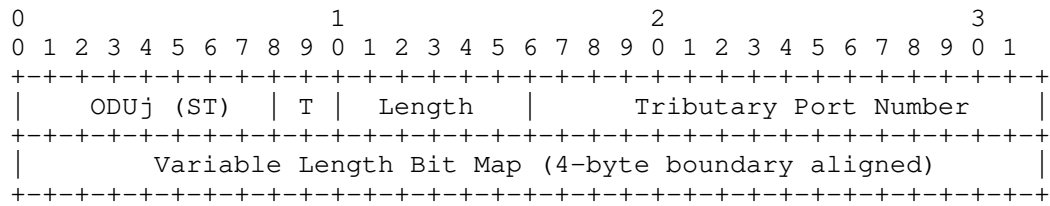
This field indicates the number of multiplexing stages specified by the label.

OD(T)Uk:

This field encodes the signal type of HO OD(T)Uk container.

Tributary Slot Info:

Tributary Slot Information for a single stage is encoded as follows.



ODUj:

This field indicates the signal type of a LO-ODU being multiplexed into its immediate HO-ODU.

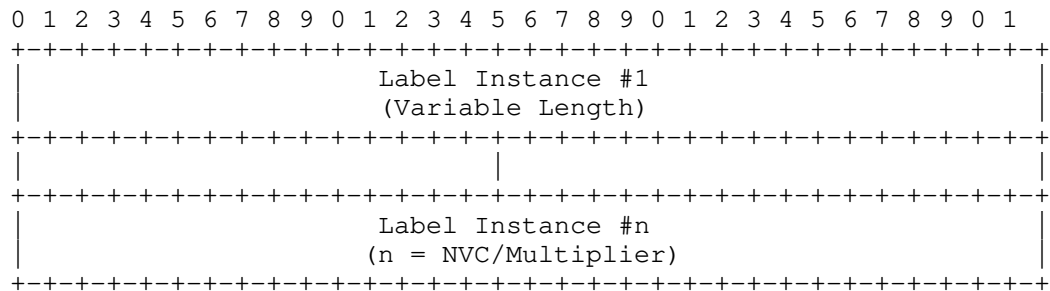
Length:

This field indicates the number of valid Bits in the Bit Map excluding the filler bits.

T & Tributary Port Number & Bit Map: See section 5.1.

5.2.2. Label format for NVC or Multiplier > 1

For NVC or Multiplier field value > 1, the label format defined in section 5 needs to be repeated NVC/multiplier times.



5.2.3. Usage of Multi-stage Label

Multi-stage Label is needed when switching of an ODU Layer requires termination of more than one HO-ODUs on a given OTU/ODU Link. This eliminates the need for creating H-LSPs whose span matches its parent TE-Link.

Example-1:

Assume on an OTU3 Link, a restrictive MUX hierarchy (as shown in figure 6) is supported on the associated interfaces. In order to switch ODU1 on this Link, ODU3 and ODU2 need to be terminated on the same span as the OTU3 link. If multi-stage Label is not supported, H-LSP need to be created for ODU3 and ODU2 layers (or just ODU2 layer at the minimum) in order to support ODU1 LSP. Creation of ODU3 and ODU2 H-LSP on top of OTU3 Link on the same span is not really required as bandwidth management for all ODU layers can still be managed on the OTU3 Link itself.

Multi-stage Label helps in implicit creation of ODU3 and ODU2 layers as part of ODU1 LSP setup and thus eliminates the need for the creation of the H-LSP on every hop.

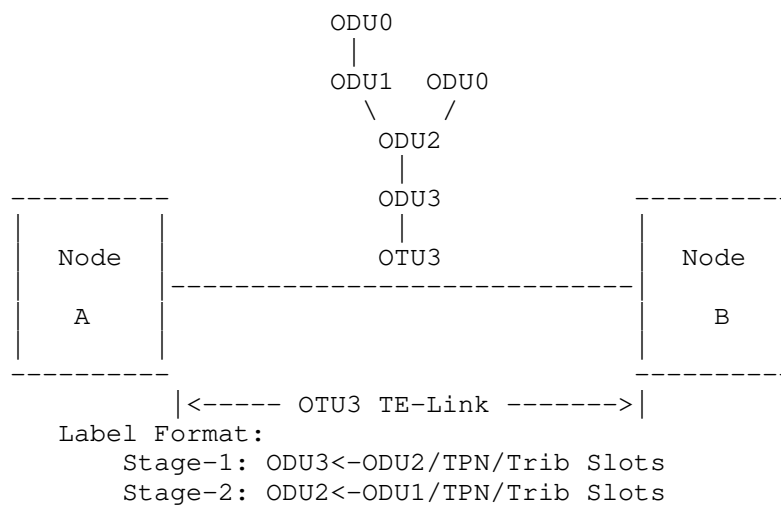


Figure 6 - Multi-stage Label on OTUk Link

Example-2:

Assume on an ODU3 H-LSP (B-C-D), signaling of ODU1 LSP requires termination of ODU2. Multi-stage Label helps in implicit creation of ODU2 layer as part of ODU1 LSP setup (A-B-D-E).

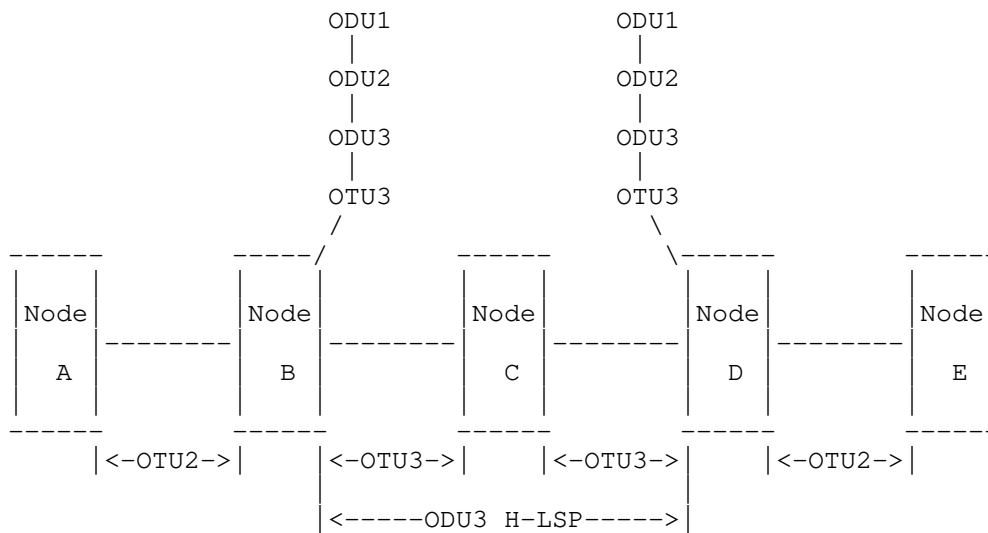


Figure 7 - Multi-stage Label on ODUk Link

Note: Multi-stage Label is NOT intended to facilitate the creation of H-LSP or Hierarchical LSP. It is basically used to eliminate the need for H-LSP in some obvious scenarios.

5.2.4. Label Distribution Rules

This document does not change the existing label distribution procedures defined in [RFC4328] except that the new ODU label should be processed as follows.

A. Sending Side

When Generalized Label Request is received on given node for setting up an ODU LSP from its upstream neighbor, it reserves the bandwidth required for the ODU Layer being switched and also the terminating HO-ODUs layers involved. It sends upstream label and suggested label (if applicable) to the downstream node and downstream label via PATH Message and downstream label to the upstream node via RESV Message.

Note that Label can also be explicitly specified by source node.

The encoding of Generalized Label is as follows:

Case-1: ODUk mapping into OTUk

Number of MUX stages = 0

Tributary Slot information is not included.

Case-2: ODUj mux into ODUk
 Number of MUX Stages = 1.
 Stage-1: Length = <number of TSs on ODUk>.
 TPN = <specified as per Section 5>
 TS BitMap = <TSs reserved for ODUj are set to 1>

Case-3 ODUh mux into ODUj into ODUk
 Number of MUX Stages = 2.
 Stage-1: Length = <number of TSs on ODUk>.
 TPN = <specified as per Section 5>
 TS BitMap = <TSs reserved for ODUj are set to 1>
 Stage-2: Length = <number of TSs on ODUj>.
 TPN = <specified as per Section 5>
 TS BitMap = <TSs reserved for ODUh are set to 1>

B. Receiving Side

The decoding of the Generalized Label is as follows:

Case-1: ODUk mapping into OTUk
 For ODUk to OTUk mapping, the Tributary Slot Information is not expected.

Case-2: ODUj mux into ODUk
 For ODUj to ODUk multiplexing, one MUX stage Label is expected.
 The node extracts the Bit Map field in Tributary Slot Info using the Length field. The position of Bit in the Bitmap interpreted as the Tributary Slot Number. The value stored in the bit indicates if it is reserved for the ODUj.

Case-3: ODUh mux into ODUj into ODUk
 For ODUh mux into ODUj into ODUk, two MUX stage Label is expected.
 Each stage is further decoded as explained in case-2 above.

5.2.5. Examples

Example-1: ODUj LSP over OTUk Links

Consider the network topology shown in the Figure 8 below:

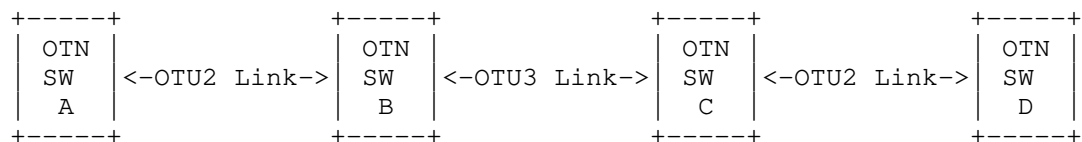


Figure 8 - OTN Signaling Example

Assumptions:

(1) ODU2 links between OTN-Switches A & B and C & D support 1.25Gbps TS Granularity.

(2) ODU3 link between OTN-Switches B & C supports TS Granularity of 2.5Gbps only. Hence, ODU0 switching on this link is possible only through ODU3-ODU2-ODU0 or ODU3-ODU1-ODU0 multiplexing hierarchies.

G.709 Traffic Parameters and Generalized Label for ODU0 LSP from node A to D is captured below:

A. G.709 Traffic Parameters

Signal Type = ODU0
 NMC/Tolerance = 0 // NMC is not used.
 NVC = 0
 Multiplier (MT) = 1
 Bit_Rate = 0

B. Generalized Label Format:

	A to B	B to C	C to D
# of Stages	1	2	1
Stage-1	ODU2<--ODU0 TSG = 1.25G #TSs = 8 TPN = <1..8> BMap = 4bytes	ODU3<--ODU2 TSG = 2.5G #TSs = 16 TPN = <1..4> BMap = 4bytes	ODU2<--ODU0 TSG = 1.25G #TSs = 8 TPN = <1..8> BMap = 4bytes
Stage-2	N/A	ODU2<--ODU0 TSG = 1.25G #TSs = 8 TPN = <1..8> BMap = 4bytes	N/A

Example 2: ODUj LSP over ODUk H-LSP

Refer to Figure 7. The G.709 Traffic Parameters and Generalized Label for ODU1 LSP from Node A to E are captured below:

A. G.709 Traffic Parameters:

Signal Type = ODU1
 NMC/Tolerance = 0 // NMC is not used.

NVC = 0
Multiplier (MT) = 1
Bit_Rate = 0

B. Generalized Label Format:

	A to B	B to D	D to E
# of Stages	1	2	1
Stage-1	ODU2<--ODU1 TSG = 1.25G #TSs = 8 TPN = <1..4> BMap = 4bytes	ODU3<--ODU2 TSG = 2.5G #TSs = 16 TPN = <1..4> BMap = 4bytes	ODU2<--ODU1 TSG = 1.25G #TSs = 8 TPN = <1..4> BMap = 4bytes
Stage-2	N/A	ODU2<--ODU1 TSG = 1.25G #TSs = 8 TPN = <1..4> BMap = 4bytes	N/A

5.3. Control Plane Backward Compatibility Considerations

Since the [RFC4328] has been deployed in the network for the nodes that support [G709-V1] (herein we call them "legacy nodes"), backward compatibility SHOULD be taken into consideration when the new nodes (i.e., nodes that support [G709-V3]) and the legacy nodes are interworking.

For backward compatibility consideration, the new node SHOULD have the ability to generate and parse legacy labels.

- o For the legacy node, it always generates and sends legacy label to its upstream node, no matter the upstream node is new or legacy, as described in [RFC4328].
- o For the new node, it will generate and send legacy label if its upstream node is a legacy one, and generate and send new label if its upstream node is a new one.

One backwards compatibility example is shown in Figure 9:

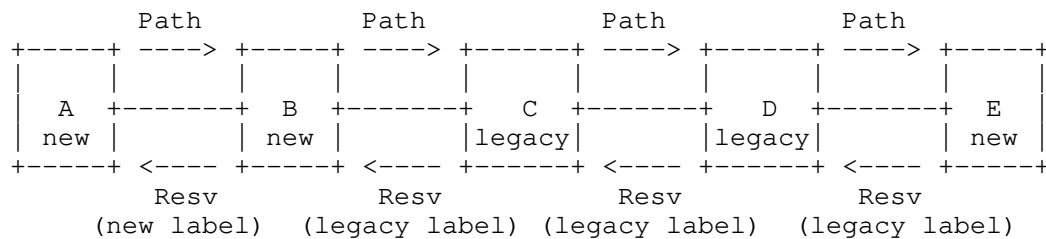


Figure 9 - Backwards compatibility example

As described above, for backward compatibility considerations, it is necessary for a new node to know whether the neighbor node is new or legacy.

One optional method is manual configuration. But it is recommended to use LMP to discover the capability of the neighbor node automatically, as described in [OTN-LMP].

When performing the HO ODU link capability negotiation:

- o If the neighbor node only support the 2.5Gbps TS and only support ODU1/ODU2/ODU3, the neighbor node should be treated as a legacy node.
- o If the neighbor node can support the 1.25Gbps TS, or can support other LO ODU types defined in [G709-V3]), the neighbor node should be treated as new node.
- o If the neighbor node returns a LinkSummaryNack message including an ERROR_CODE indicating nonsupport of HO ODU link capability negotiation, the neighbor node should be treated as a legacy node.

6. 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.

7. IANA Considerations

- G.709 SENDER_TSPEC and FLOWSPEC objects:

The traffic parameters, which are carried in the G.709 SENDER_TSPEC and FLOWSPEC objects, do not require any new object class and type based on [RFC4328]:

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

- o G.709 FLOWSPEC Object: Class = 9, C-Type = 5 [RFC4328]

- Generalized Label Object:

The new defined ODU label (session 5) 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.

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, Jan 2006.
- [RFC3209] D. Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC3209, December 2001.
- [RFC3471] Berger, L., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] L. Berger, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.
- [VCAT] G. Bernstein et al, "Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS)", draft-ietf-ccamp-gmpls-vcat-lcas-13.txt, May 4, 2011.

- [RFC4206] K. Kompella, Y. Rekhter, Ed., " Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC6107] K. Shiimoto, A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC6107, February 2011.
- [RFC6001] Dimitri Papadimitriou et al, "Generalized Multi-Protocol Label Switching (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC6001, February 21, 2010.
- [OTN-frwk] Fatai Zhang et al, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", draft-ietf-ccamp-gmpls-g709-framework-04.txt, March 11, 2011.
- [OTN-info] S. Belotti et al, "Information model for G.709 Optical Transport Networks (OTN)", draft-ietf-ccamp-otn-g709-info-model-00.txt, April 18, 2011.
- [OTN-LMP] Fatai Zhang, Ed., "Link Management Protocol (LMP) extensions for G.709 Optical Transport Networks", draft-zhang-ccamp-gmpls-g.709-lmp-discovery-04.txt, April 6, 2011.
- [G709-V3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709/Y.1331, December 2009.

8.2. Informative References

- [G709-V1] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation (and Amendment 1), February 2001 (November 2001).
- [G709-V2] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation, March 2003.
- [G798-V2] ITU-T, "Characteristics of optical transport network hierarchy equipment functional blocks", G.798, December 2006.
- [G798-V3] ITU-T, "Characteristics of optical transport network hierarchy equipment functional blocks", G.798v3, consented June 2010.
- [RFC4506] M. Eisler, Ed., "XDR: External Data Representation Standard", RFC 4506, May 2006.

[IEEE] "IEEE Standard for Binary Floating-Point Arithmetic",
ANSI/IEEE Standard 754-1985, Institute of Electrical and
Electronics Engineers, August 1985.

[GMPLS-SEC] Fang, L., Ed., "Security Framework for MPLS and GMPLS
Networks", Work in Progress, October 2009.

9. Authors' Addresses

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Guoying Zhang
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.China
Phone: +86-10-68094272
Email: zhangguoying@mail.ritt.com.cn

Sergio Belotti
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6863033
Email: sergio.belotti@alcatel-lucent.it

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: daniele.ceccarelli@ericsson.com

Khuzema Pithewan
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089, USA
Email: kpithewan@infinera.com

Yi Lin
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972914
Email: yi.lin@huawei.com

Yunbin Xu
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.China
Phone: +86-10-68094134
Email: xuyunbin@mail.ritt.com.cn

Pietro Grandi
Alcatel-Lucent
Optics CTO
Via Trento 30 20059 Vimercate (Milano) Italy
+39 039 6864930
Email: pietro_vittorio.grandi@alcatel-lucent.it

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy
Email: diego.caviglia@ericsson.com

Mohit Misra
Infinera Corporation

169, Java Drive
Sunnyvale, CA-94089, USA
Email: mmisra@infinera.com

Rajan Rao
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089, USA
Email: rrao@infinera.com

Ashok Kunjidhapatham
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089, USA
Email: akunjidhapatham@infinera.com

Biao Lu
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089, USA
Email: blu@infinera.com

Lyndon Ong
Ciena
PO Box 308, Cupertino, CA 95015, USA
EMail: lyong@ciena.com

Igor Bryskin
Adva Optical
EMail: IBryskin@advaoptical.com

Acknowledgment

The authors would like to thank Jonathan Sadler and John E Drake 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.

Network Working Group
Internet-Draft
Intended status: Standards Track

Fatai Zhang
Dan Li
Huawei
O. Gonzalez de Dios
Telefonica Investigacion y Desarrollo
C. Margaria. C
Nokia Siemens Networks
July 8, 2011

Expires: January 8, 2012

RSVP-TE Extensions for Configuration SRLG of an FA
draft-zhang-ccamp-srlg-fa-configuration-03.txt

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on January 8, 2012.

Abstract

This memo provides extensions for the Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) for the support of the automatic discovery of SRLG of an LSP.

Conventions used in this document

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

Table of Contents

1. Introduction	2
2. RSVP-TE Requirements.....	4
2.1. SRLG Collection Indication.....	4
2.2. SRLG Collection.....	4
2.3. SRLG Update	4
3. RSVP-TE Extensions	4
3.1. SRLG Collection Flag.....	4
3.2. SRLG sub-object.....	5
4. Signaling Procedures.....	6
4.1. SRLG Collection.....	6
4.2. SRLG Update	6
5. Manageability Considerations.....	7
5.1. Policy Configuration.....	7
5.2. Coherent SRLG IDs.....	8
6. IANA Considerations	8
6.1. RSVP Attribute Bit Flags.....	8
6.2. ROUTE_RECORD Object.....	8
7. Security Considerations.....	9
8. References	9
9. Authors' Addresses	11

1. Introduction

As described in [RFC4206], H-LSP (Hierarchical LSP) can be used for carrying one or more other LSPs. [RFC6107] further mentions the implementation of H-LSP. In packet networks, e.g. MPLS networks, H-LSP mechanism can be implemented by MPLS label stack. In non-packet networks where the label is implicit, label stacks are not possible, and H-LSPs rely on the ability to nest switching technologies. Thus, for example, a lambda switch capable (LSC) LSP can carry a time division multiplexing (TDM) LSP, but cannot carry another LSC LSP.

S-LSP (LSP Stitching), which is defined in [RFC5150], is an LSP that represents a segment of another LSP, i.e., the S-LSP is viewed as one hop by another LSP. As described in [RFC6107], in the data plane the LSPs are stitched so that there is no label stacking or nesting.

Thus, an S-LSP must be of the same switching technology as the end-to-end LSP that it facilitates.

Therefore, H-LSP mechanism can be used in both multi-domain and multi-layer scenarios and S-LSP mechanism can only be used in multi-domain scenario.

Both of the H-LSP and S-LSP can be advertised as a TE link in a GMPLS routing instance for path computation purpose. As described in [RFC6107], if the LSP (H-LSP or S-LSP) is advertised in the same instance of the control plane that advertises the TE links from which the LSP is constructed, the LSP is called an FA.

In multi-domain or multi-layer context, the path information of an LSP may not be provided to the ingress node for confidential reasons and the ingress node may not run the same routing instance with the intermediate nodes traversed by the path. In such scenarios, the ingress node can not get the SRLG information of the path information which the LSP traverse.

Even if the ingress node has the same routing instance with the intermediate nodes traversed by the path, the path information of the H-LSP or S-LSP may not be provided to the ingress node. Hence the ingress node may also not know the SRLG of the path the LSP traverses.

In the case that the ingress node does not get the SRLG of the path the LSP traverses (i.e. H-LSP or S-LSP), there are disadvantages as follows:

- o SRLG-disjoint path, for instance in case of end-to-end path protection, cannot be calculated
- o Intermediate nodes of a pre-planned shared restoration LSP cannot correctly decide on the SRLG-disjointness between two PPRO (PRIMARY_PATH_ROUTE Object)
- o In case that an LSP is advertised as a TE-Link, the ingress node cannot provide the correct SRLG for the TE-Link automatically

In case that an LSP is advertised as a TE-Link, the SRLG information of the TE link needs to be configured manually or automatically. However, for manual configuration, there are some disadvantages (e.g., require configuration coordination and additional management; manual errors may be introduced) mentioned in Section 1.3.4 of [RFC6107].

In addition, Section 1.2 of [RFC6107] describes it is desirable to have a kind of automatic mechanism to advertise the FA (i.e., to signal an LSP and automatically coordinate its use and advertisement in any of the ways with minimum involvement from an operator).

Thus, in order to provide the SRLG information to the TE link automatically when an LSP (H-LSP or S-LSP) is advertised as a TE link, allow disjoint path calculation at ingress and allow correct pre-planned shared LSP to correctly share resource, this document provides an automatic mechanism to collect the SRLG used by a LSP automatically.

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 needs to be capable of updating the newly SRLG information to the endpoints.

3. RSVP-TE Extensions

3.1. SRLG Collection Flag

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

SRLG Collection flag (to be assigned by IANA, recommended bit zero)

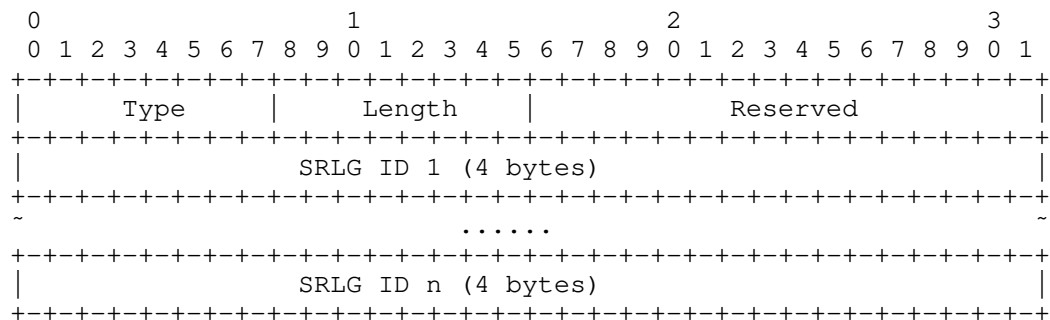
The SRLG Collection flag is meaningful on a Path message. If the SRLG Collection flag is set to 1, it means that the SRLG information

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

A new SRLG sub-object is defined for RRO (ROUTE_RECORD Object) to record the SRLG information of the LSP. Its format is modeled on the RRO sub-objects defined in [RFC3209].



Type

The type of the sub-object, to be assigned by IANA, which is recommended 34.

Length

The Length contains the total length of the sub-object in bytes, including the Type and Length fields. The Length depends on the number of SRLG IDs.

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 (RSVP sessions) using that link should be aware of the changes. Note that,

as stated in RFC 3209, the RRO collects up-to-date detailed path information hop-by-hop about RSVP sessions, providing valuable information to the sender or receiver. Thus, in a similar way, the RRO should also collect the up-to-date SRLG information. The procedure for the update is described below.

When the SRLG of a link is changed, the endpoints of the link need to check the information of all the LSPs that traverse the link in order to find out the LSPs which have requested the SRLG recording. The new SRLG information needs to be updated if the SRLG Collection flag is set on a Path message.

When an endpoint of the link finds out an LSP has requested SRLG recording, it should send a NOTIFY message to the head node of the LSP informing that SRLG information needs to be recollected.

Then, the head node must send a path message with the SRLG recording bit set that should be processed by the intermediate nodes. As described in [RFC2961] section 4.5, changing the message_id of the path message would force the intermediate nodes to fully process the message. Then, the intermediate nodes will all update the SRLG information in the SRLG sub-object. Then the tail node receives the new Path message, fully processes the message and gets the new SRLG information of the LSP from it.

After the tail node of the LSP gets the new SRLG information, it should update the SRLG information in the corresponding Resv message which will be sent to the upstream node as a trigger message. The new Resv message should be fully processed and forwarded in the upstream direction until it arrives at the head node. Then the head node receives the new Resv message, fully processes the message and gets the new SRLG information of the LSP from it.

5. Manageability Considerations

5.1. Policy Configuration

In a border node of inter-domain or inter-layer network, the following SRLG processing policy should be capable of being configured:

- o whether the SRLG IDs of the domain or specific layer network can be exposed to the nodes outside the domain or layer network.

If the SRLG IDs should not be exposed to the nodes outside of the domain or specific layer network by policy, the border node should 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. IANA Considerations

6.1. 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

6.2. 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 [RFC3209] portions of this registry.

This document introduces a new RRO sub-object:

Type	Name	Reference
------	------	-----------

-----	-----	-----
TBD (34)	SRLG sub-object	This I-D

7. Security Considerations

TBD.

8. Acknowledgements

The authors would like to thank Igor Bryskin and Ramon Casellas for their useful comments to the document.

9. References

- [RFC2119] S. Bradner, "Key words for use in RFCs to indicate requirements levels", RFC 2119, March 1997.
- [RFC2961] L. Berger, D. Gan, G. Swallow, P. Pan, F. Tommasi and S. Molendini, "RSVP Refresh Overhead Reduction Extensions", RFC 2961, April 2001.
- [RFC3477] K. Kompella, Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", rfc3477, January 2003.
- [RFC4206] K. Kompella, Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC4208] G. Swallow, J. Drake, Boeing, H. Ishimatsu, and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.
- [RFC4874] CY. Lee, A. Farrel, S. De Cnodder, " Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) ", RFC 4874, April 2007.

- [RFC5150] Ayyangar, A., Vasseur, J.P, and Farrel, A., "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.
- [RFC5420] A. Farrel, D. Papadimitriou, J.P, and A. Ayyangar, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.
- [RFC6107] K. Shiimoto, A. Farrel, " Procedures for Dynamically Signaled Hierarchical Label Switched Paths ", RFC 6107, February 2011.

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

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

Phone: +86-755-28970230
Email: danli@huawei.com

Oscar Gonzalez de Dios
Telefonica Investigacion y Desarrollo
Emilio Vargas 6
Madrid, 28045
Spain

Phone: +34 913374013
Email: ogondio@tid.es

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

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

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.

