CCAMP Working Group
Internet-Draft
Intended status: Proposed Standard
Expires: September 15, 2011

Ashok Kunjidhapatham
Rajan Rao
Biao Lu
Snigdho Bardalai
Khuzema Pithewan
Infinera Corp
John E Drake
Juniper Networks
Steve Balls
Metswitch Networks
March 14, 2011

OSPF TE Extensions for Generalized MPLS (GMPLS) Control of G.709 Optical Transport Networks
draft-ashok-ccamp-gmpls-ospf-g709-03.txt

Status of this Memo

Copyright and License Notice

Abstract

As OTN network capabilities continue to evolve, there is an increased need to support GMPLS control for the same. [RFC4328] introduced GMPLS signaling extensions for supporting the early version of G.709 [G.709-v1]. The basic routing considerations from signaling perspective is also specified in [RFC4328].

The recent revision of ITU-T Recommendation G.709 [G.709-v3] and [GSUP.43] have introduced new ODU containers (both fixed and flexible) and additional ODU multiplexing capabilities, enabling support for optimal service aggregation.

This document describes OSPF protocol extensions to support Generalized MPLS (GMPLS) control for routing services over the standardized OTU/ODU containers in support of ODU based TDM switching. Routing support for Optical Channel Layer switching (Lambda switching) is not covered in this document.

Table of Contents

1. Introduction

Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] extends MPLS from supporting Packet Switching Capable (PSC) interfaces and switching to include support of four new classes of interfaces and switching: Layer-2 Switching (L2SC), Time-Division Multiplex (TDM), Lambda Switch (LSC), and Fiber-Switch (FSC) Capable. A functional description of the extensions to MPLS signaling that are needed to support these new classes of interfaces and switching is provided in [RFC3471]. OSPF extensions for supporting GMPLS are defined in [RFC4203].

ITU-T Recommendations G.709 and G.872 provide specifications for OTN interface and network architecture respectively. As OTN network capabilities continue to evolve; there is an increased need to support GMPLS control for the same.

GMPLS signaling extensions to support G.709 OTN interfaces are specified in [RFC4328]. The basic routing considerations from signaling perspective is specified. G.709 specifications evolved rapidly over the last few years. Following are the features added in OTN since the first version [G.709-v1].

(a) OTU Containers:
    Pre-existing Containers: OTU1, OTU2 and OTU3
    New Containers introduced in [G.709-v3]: OTU2e and OTU4
    New Containers introduced in [GSUP.43]: OTU1e, OTU3e1 and OTU3e2

(b) Fixed ODU Containers:
    Pre-existing Containers: ODU1, ODU2 and ODU3
    New Containers introduced in [G.709-v3]: ODU0, ODU2e and ODU4
    New Containers introduced in [GSUP.43]: ODU1e, ODU3e1 and ODU3e2

(c) Flexible ODU Containers:
    ODUflex for CBR and GFP-F mapped services. ODUflex uses 'n' number of OPU Tributary Slots where 'n' is different from the number of OPU Tributary Slots used by the Fixed ODU Containers.

(d) Tributary Slot Granularity:
    OPU2 and OPU3 support two Tributary Slot Granularities: (i) 1.25Gbps and (ii) 2.5Gbps.

(e) Multi-stage ODU Multiplexing:
    Multi-stage multiplexing of LO-ODUs into HO-ODU is supported. Also, multiplexing could be heterogeneous (meaning LO-ODUs of different rates can be multiplexed into a HO-ODU).

OTN networks support switching at two layers: (i) ODU Layer - TDM

Switching and (ii) OCH Layer - Lambda (LSC) Switching. The nodes on the network may support one or both the switching types. When multiple switching types are supported MRN/MLN based routing [RFC5212] and [RFC6001] is assumed.

This document covers OSPF extensions to support routing over the standardized OTU/ODU containers in support of ODU Layer based TDM switching as outlined in the framework document [G.709-FRAME]. The Interface Switch Capability Descriptor extensions for ODU Layer switching and bandwidth representation for ODU containers are defined in this document.

Routing support for Optical Channel Layer switching (LSC) is beyond the scope of this document. Refer to [WSON-FRAME] for further details.

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 is to be interpreted as described in RFC-2119 [RFC2119].

In addition, the reader is assumed to be familiar with the terminology used in ITU-T [G.709-v3], [G.872] and [GSUP.43], as well as [RFC4201] and [RFC4203]. Abbreviations used in this document is detailed in Appendix A.

3. OTU/ODU Link Representation

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

3.1. OTUk TE-Link

OTUk Link can be modeled as a TE-Link. Switching at ODUk layer and ODUj layer (including multi-stage multiplexing) can be managed on OTUk TE-Link. Figure-1 below provides an illustration of this link type.

When a LO-ODU layer being switched on an OTUk interface involves multi-stage multiplexing, all the HO-ODU layer(s) should necessarily terminate between the same pair of nodes as the OTUk layer in this case. For example, if ODU1 layer switching is configured on a OTU3 link via multiplexing hierarchy

ODU3<-ODU2<-ODU1, HO-ODUs (namely ODU3 & ODU2) should terminate between the same pair of nodes as OTU3 layer.

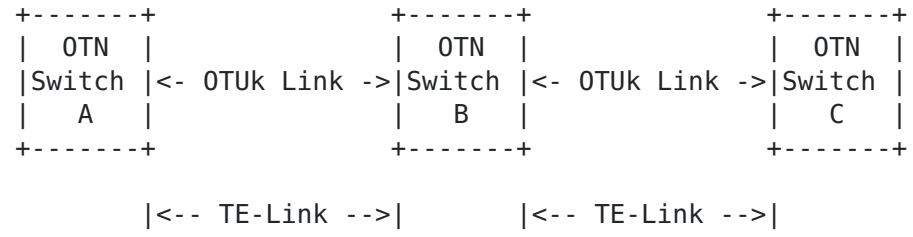


Figure-1: OTUk TE-Link

3.2. ODUk TE-Link

When ODUk layer does not terminate on the same pair of nodes as OTUk layer, ODUk link should be modeled as a TE-Link. As bandwidth is directly managed on the ODUk link, associated OTUk links are not significant in this case. Switching at ODUj layer (including multi-stage multiplexing) can be managed on ODUk TE-Link. Figure-2 below provides an illustration of this link type.

When a LO-ODU layer being switched on the ODUk interface involves multi-stage multiplexing, all the HO-ODU layer(s) should necessarily terminate between the same pair of nodes as ODUk in this case. For example, if ODU1 layer switching is configured on an ODU3 link via multiplexing hierarchy ODU3<-ODU2<-ODU1, HO-ODU (namely ODU2) should terminate between the same pair of nodes as ODU3.

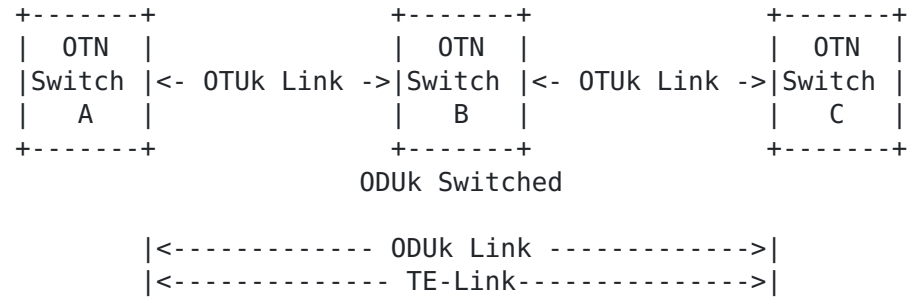


Figure-2: ODUk TE-Link

3.3. ODUj TE-Link

When a LO-ODUj within a HO-ODUk does not terminate on the same pair of nodes as HO-ODUk layer, Separate TE-Links needs to be modeled for ODUk link and ODUj link. Also, ODUk link shall no longer manage the bandwidth associated with the ODUj link. Switching at sub-ODUj layer (including multi-stage multiplexing)

can be supported on this ODUj TE-Link. Figure-3 below provides an illustration of this link type.

When a LO-ODU layer being switched on an ODUj interface involves multi-stage multiplexing, all the HO-ODU layer(s) should necessarily terminate between the same pair of nodes as ODUj in this case. For example, if ODU0 layer switching is configured on an ODU2 link via multiplexing hierarchy ODU2<-ODU1<-ODU0, HO-ODU (namely ODU1) should terminate between the same pair of nodes as ODU2.

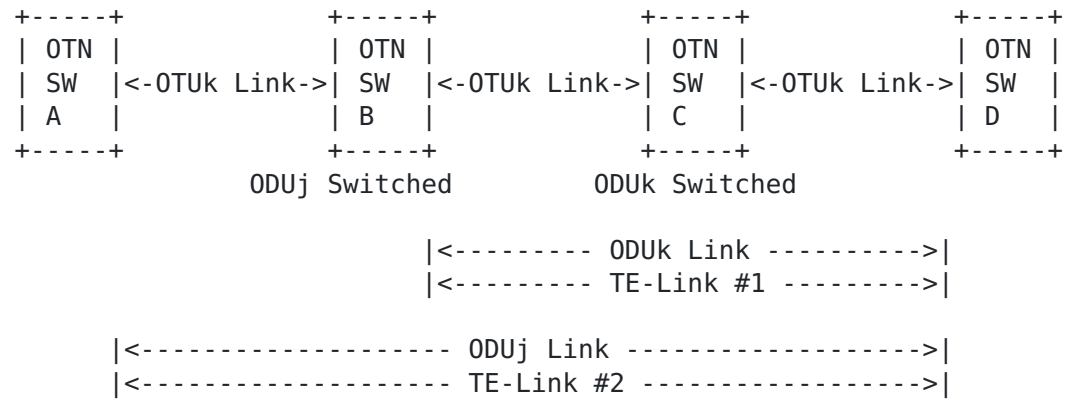


Figure-3: ODUj TE-Link

3.4. Bundled TE-Link

Any mix of OTU and ODU links of dissimilar rates that terminates on same pair of nodes and meets the entire bundling criterion specified in TE-Link Bundling specification [RFC4201] can be pulled together to form a Bundle TE-Link. This is required for better scalability.

3.5. OTU/ODU Link Property Agreement

The OTN interfaces (associated with peer nodes) participating in a TE-Link may not be fully compatible in terms of OTN interface properties. The lowest common denominator between the two links endpoints need to be used for forming the TE link. Some of OTN specific link properties that need to be agreed upon between the two link endpoints (on peer nodes) are:

(a) OPU Tributary Slot Granularity for OPU2 and OPU3.

(b) Multiplexing hierarchies supported - both number of stages and specific LO-ODUs supported in each stage. This includes both Fixed and Flexible ODU containers.

These link properties either can be configured or discovered through Link discovery mechanism. The details of such mechanism is beyond the scope of this document.

4. OTU/ODU Link Bandwidth Model

Bandwidth allocation/management on OTU/ODU links is done in terms of discrete units called OPU Tributary Slots. OPU Tributary Slots occurs in two granularities (1.25Gbps and 2.5Gbps) and the actual bit-rate of the OPU Tributary Slot slightly varies for different ODU container types (i.e., ODU1, ODU2, ODU3 and ODU4). As a result of this disparity, number of Tributary Slots required to map a LO-ODU on different HO-ODU container types could vary. For example, ODU2e requires 9 OPU TSs on ODU3 and 8 OPU TSs on ODU4.

The basic objectives of OTN interface bandwidth model are as follows:

(a) Support ODU multi-stage multiplexing hierarchy and yet not require advertisement of complete hierarchy tree.

(b) Account for bandwidth fragmentation that can result due to the restricted multiplexing hierarchy supported on an OTN interface. For example, assume that an ODU3 interface supports direct multiplexing of ODU2 only. Here, mapping of ODU1 and ODU0 is possible only through second stage multiplexing underneath ODU2. If two ODU1 are created under two different ODU2, only two ODU2 can be created further on the interface although 28 Tributary Slots (1.25Gbps) are available on the interface (ODU hierarchy).

(c) Hide the complexities in Tributary Slot Granularities (1.25Gbps and 2.5Gbps) from bandwidth model and thereby simplify the end-to-end path computation. As explained in the previous section, this needs to be negotiated as a part of link discovery or pre-configured locally on the either ends.

(d) Hide the complexities in Tributary Slot Size disparities (among ODU containers) and number of Tributary Slots required to map a LO-ODU. This can be achieved by advertising the number of LO-ODU containers that can be mapped on an OTN interface rather than number of Tributary Slots or absolute bandwidth in bytes/sec.

(e) For ODU-Flex service, Absolute bandwidth required (for CBR or GFP mapped service) needs to be mapped to 'n' Tributary Slots of certain bit rate. This needs Tributary Slot bit-rate and number of Tributary slots to be advertised.

5. OSPF TE-LSA Extension

This section describes the OSPF TE-LSA Extensions to support bandwidth encoding for OTU/ODU TE-Links.

5.1. Maximum Bandwidth

The format and interpretation of this attribute must be consistent with OSPF-TE Extension [RFC3630] and TE-Link Bundling Support [RFC4201] specifications. The OPUk payload nominal rate (in bytes per sec) as specified in [G.709-v3] shall be encoded in this attribute.

5.2. Maximum Reservable Bandwidth

The format and interpretation of this attribute must be consistent with OSPF-TE Extension [RFC3630] and TE-Link Bundling Support [RFC4201] specifications.

5.3. Unreserved Bandwidth

The format and interpretation of this attribute must be consistent with OSPF-TE Extension [RFC3630] and TE-Link Bundling Support [RFC4201] specifications.

Unreserved Bandwidth in bytes per second is not of much value for OTU/ODU interfaces. Unreserved Bandwidth per ODU rate is more appropriate and useful in this case. Implementations may choose to ignore this attribute and consider per ODU-rate Unreserved Bandwidth defined in Interface Switch Capability Descriptor for "G.709 ODUk" encoding type. See section 5.4.1 for details.

5.4. Interface Switch Capability Descriptor

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

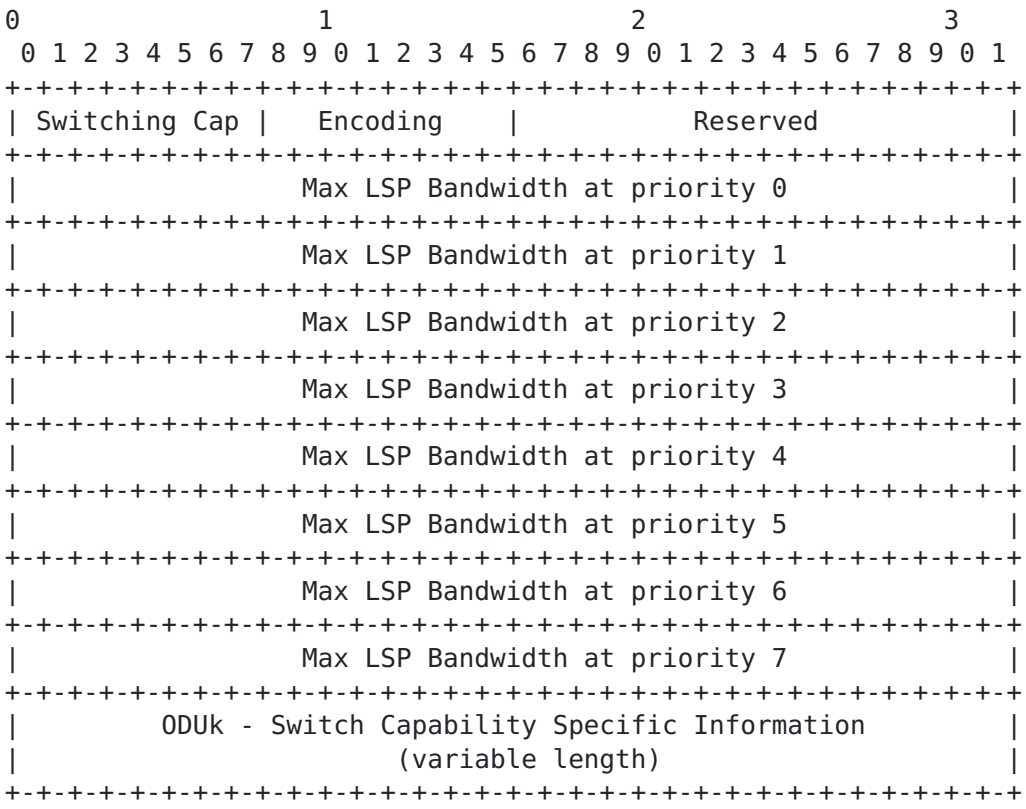
Value	Type
250	OTN-TDM capable (OTN-TDM)

Nodes advertising ODUk switching BW for for its links must use Switching Type and Encoding values as follows:

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

Both fixed ODUk (where k=0,1,2,3,4,1e,2e) and flexible ODUs (ODUflex) use same switching type and encoding values.

When Swithcing Type and Encoding fields are set to values as stated above, the Interface Switching Capability Descriptor should be interpreted as follows:



Maximum LSP Bandwidth

This field should be encoded with Nominal Rate of the ODUj (j<= k) for which Bandwidth is advertised. The bandwidth unit is in bytes per second & the encoding is in IEEE floating point format [RFC 3471]. The discrete values for varous ODUj(s) is shown in the table below.

For an unbundled link, the Maximum LSP Bandwidth at priority 'p' is set to Nominal rate of the ODUj for which bandwidth is advertised in Switch Capability Specific Information (SCSI).

For bundled link too, the Maximum LSP Bandwidth at priority 'p' is set to Nominal rate of the ODUj for which bandwidth is advertised in Switch Capability Specific Information (SCSI).

ODU type	Nomial Rate(bytes/s)	Value in Byes/Sec (IEEE format)
ODU0	15552000	
ODU1	312346890.75	
ODU2	1254659240.50	
ODU2e	1299940664.50	
ODU1e		
ODU3	5039902372.875	
ODU4	13099305726.875	
ODUflex	Any	

The Maximum LSP bandwidth field is used to identify the ODUj type.

5.4.1 ODU Switching

When Switching Capability is set to OTN-TDM, it means the node is capable of

- terminating OTUk layer
- Switching of HO-ODU (ODUk)
- switching of LO-ODU (ODUj) if HO-ODU supports mux/demux
  (termination of HO-ODU is required for mux/demux operation)

Multiple ISCDs would be advertised if an interface supports more than one type of ODUk switching. There would be one ISCD advertisement per ODUj independent of the OTN multiplexing branch it belongs to.

For e.g. If an OTU3 interface supports ODU0, ODU1 and ODU2 switching, there would be three ISCDs one for each ODU type.

Refer to examples in section 7.0.

5.4.2. ODUk Switch Capability Specific Information

This SCSI field contains bandwidth information for fixed ODUj(j=0,1,2,3,4,2e,1e) or ODUflex.

The type of ODUj/ODuflex is identified by Maximum LSP bandwidth field and BW sub TLV Type field as follows.

If bandwidth advertisement is for fixed size ODUj, then
- set BW sub TLV Type = 1
- Encode nominal rate of the ODUj in Max LSP BW field
- Encode avaialble number of ODUj(s) as shown below

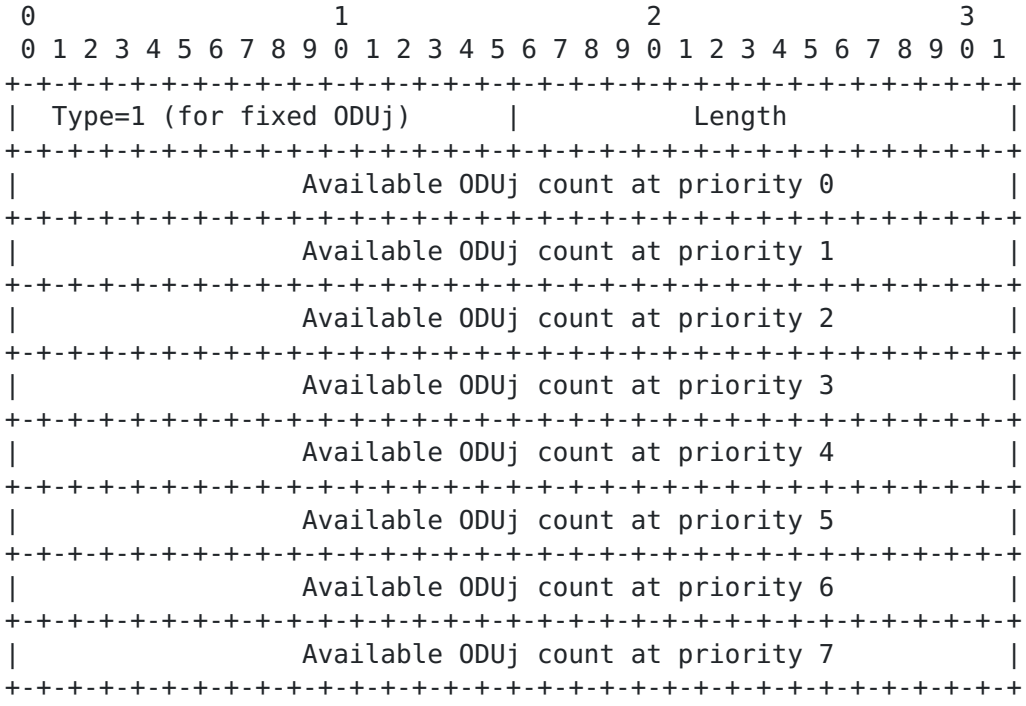
If bandwidth advertisment is for ODUflex, then
- set BW sub TLV Type = 2
- Encode available BW in Max LSP BW field

- Encode avaialble Bandwidth as shown below

The SCSI field must be included when Switching Capability is "OTN-TDM".

5.4.2.1 Bandwidth sub TLV for fixed ODUj

The format of Bandwidth sub TLV for fixed size ODUj is shown below; (j=0,1,2,3,4,2e,1e). The TLV Type must be set to 1 for fixed ODUs.



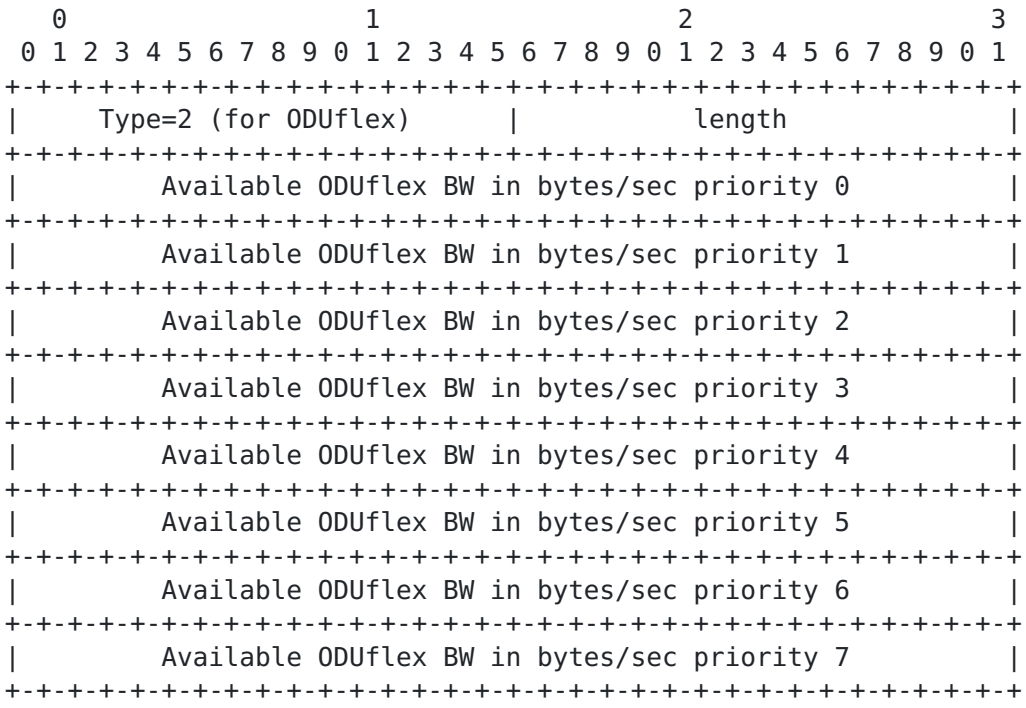
Available ODUj(s):

This field (32 bits) indicates the maximum number of Containers of a given ODUj-Type at priority 'p' available on this TE-Link.

The "Available ODUj(s)" of a bundled link at priority p is defined to be the sum of "Available ODUj(s)" at priority p of all of its component links.

5.4.2.2 Bandwidth sub-TLV for ODUflex

The format of Bandwidth sub TLV for ODUflex is shown below.
The TLV Type is set to 2 for flexible ODUs.



Available BW (in bytes/sec)

Available BW (in bytes/sec) is represented in IEEE float-point format similar to Max-Lsp-Bandwidth in ISCD.

The "Available BW" of a bundled link at priority p is defined to be the sum of "Available BW" at priority p of all of its component links.

This information may be used to route LSPs over links which have most bandwidth available.

5.5. Interface Multiplexing Capability Descriptor

The OTN multiplexing hierarchy involves one or more ODU layers. The server ODU layer is called the higher order ODU(HO-ODU) and the layer multiplexed into a server ODU layer is called lower order ODU (LO-ODU).

A HO-ODU can carry (mux/demux) one or more LO-ODUs as specified in G.709.
Termination of HO-ODU is necessary to mux/demux LO-ODUs. For e.g.

a) on a OTU2 interface with OTU2-ODU2-ODU0 muxing stack,
it is necessary to terminate ODU2(H) in order to mux/demux
contained ODU0s.

b) on a OTU2 interface with OTU2-ODU2-ODU1-ODU0 muxing stack,
it is necessary to terminate ODU2 and ODU1 layers to mux/demux
contained ODU0s.

An OTN interface supporting multi-stage multiplexing requires termination
of more than one HO-ODU to access one or more LO-ODUs for switching purposes.
For e.g. on an interface with OTU3-ODU3-ODU2-ODU0 multiplexing stack/
hierarchy,
ODU3 and ODU2 layers should be terminated to access ODU0s for switching
purposes.

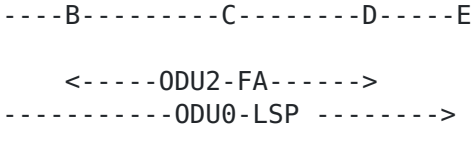
5.5.1 Multiplex Layers and Hierarchical LSP

It is possible to construct H-LSP(s) using different HO-ODU muxing layer(s).
While creation of H-LSP is optional, it becomes necessary in network
scenarios where switching restrictions exist for LO-ODUs.

Example #1:

- Nodes A, B, D & E are ODU0 and ODU2 switching capable;
- Nodes C is ODU2 switching capable only.

An ODU2-FA between nodes B & D is necessary to support E2E ODU0-LSP(s)



Example #2: ODU0-LSP over G.709-v1 capable node (legacy node)

- Nodes A, B, D & E are ODU0 & ODU1 switch capable nodes;
- Node C is ODU1 switching capable

An ODU1-FA between nodes B & D is necessary to support E2E ODU0-LSPs

In order to support identification of potential FA boundary points, it is

necessary to flood mux/demux information. This includes information about:

- the HO-ODU layer which can be terminated
- the LO-ODUs available upon HO-ODU termination (muxing hierarchy)

The multiplexing hierarchy provides information about specific branch(es) of the OTN muxing hierarchy. This includes

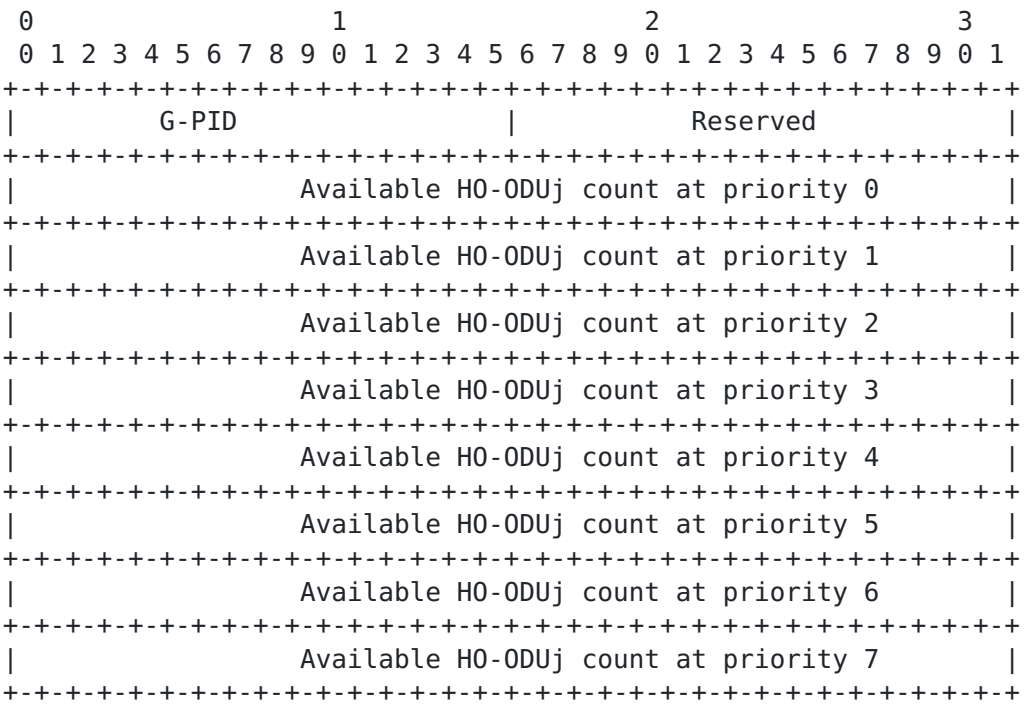
- one or more HO-ODU(s) which needs to be terminated and
- a LO-ODU layer which can be accessed after termination

The HO-ODUs which are terminate-able are potential FA end points. FA becomes necessary when switching bandwidth is not available at all nodes along the path for an LSP (specifically for LSPs at LO-ODU layers).

This draft proposes the use of IMCD (Interface Multiplexing Capability Descriptor) to distribute OTN mux/demux information of Te-end points.

5.5.2 IMCD format

The Interface Multiplexing Capability Descriptor (IMCD) describes "Multiplexing" capability of an interface. It is a sub-TLV of the Link TLV (Type TBD). The format of value field is as shown below:



5.5.2.1 G-PID

The G-PID field is a 16 bit field as defined in [RFC3471].
New G-PID values are defined in addition to those defined in [RFC3471].
Within OTN context, the new G-PID values identify multiplexing stack
supported by the Te-end point.

The table below shows newly defined values for G-PID:

Value	G-PID	Meaning
60	ODU1-ODU0	ODU1 termination required
61	ODU2-ODU0	ODU2 termination required
62	ODU2-ODU1	ODU2 termination required
63	ODU2-ODU1-ODU0	ODU2 & ODU1 termination required
64	ODU2-ODUflex	ODU2 termination required
65	ODU3-ODU0	ODU3 termination required
66	ODU3-ODU1	ODU3 termination required
67	ODU3-ODU1-ODU0	ODU3 & ODU1 termination required
68	ODU3-ODU2	ODU3 termination required
69	ODU3-ODU2-ODU0	ODU3 & ODU2 termination required
70	ODU3-ODU2-ODU1	ODU3 & ODU2 termination required
71	ODU3-ODU2-ODU1-ODU0	ODU3 & ODU2 & ODU1 termination required
72	ODU3-ODU2-ODUflex	ODU3 & ODU2 termination required
73	ODU3-ODUflex	ODU3 termination required
74	ODU3-ODU2e	ODU3 termination required
75	ODU4-ODU0	ODU4 termination required
76	ODU4-ODU1	ODU4 termination required
77	ODU4-ODU1-ODU0	ODU4 & ODU1 termination required
78	ODU4-ODU2	ODU4 termination required
79	ODU4-ODU2-ODU0	ODU4 & ODU2 termination required
80	ODU4-ODU2-ODU1	ODU4 & ODU2 termination required
81	ODU4-ODU2-ODU1-ODU0	ODU4 & ODU2 & ODU1 termination required
82	ODU4-ODU2-ODUflex	ODU4 & ODU2 termination required
83	ODU4-ODU3	ODU4 termination required
84	ODU4-ODU3-ODU0	ODU4 & ODU3 termination required
85	ODU4-ODU3-ODU1	ODU4 & ODU3 termination required
86	ODU4-ODU3-ODU1-ODU0	ODU4 & ODU3 & ODU1 termination required
87	ODU4-ODU3-ODU2	ODU4 & ODU3 termination required
88	ODU4-ODU3-ODU2-ODU0	ODU4 & ODU3 & ODU2 termination required
89	ODU4-ODU3-ODU2-ODU1	ODU4 & ODU3 & ODU2 termination required
90	ODU4-ODU3-ODU2-ODU1-ODU0	ODU4 & ODU3 & ODU2 & ODU1 termination required
91	ODU4-ODU3-ODU2-ODUflex	ODU4 & ODU3 & ODU2 termination required
92	ODU4-ODU3-ODUflex	ODU4 & ODU3 termination required
93	ODU4-ODU3-ODU2e	ODU4 & ODU3 termination required
94	ODU4-ODUflex	ODU4 termination required
95	ODU4-ODU2e	ODU4 termination required

96	ODU1	ODU1 termination required
97	ODU2	ODU2 termination required
98	ODU3	ODU3 termination required
99	ODU4	ODU4 termination required
100	ODU2-GFP-10GBE	ODU2 termination for Ethernet
101	ODU2e-10GBE	ODU2e termination for Ethernet
102	ODU2-OC192	ODU2 termination for SONET

5.5.2.2 Available Bandwidth

The available bandwidth advertised in "Available HO-ODUj" field indicates the number of "Terminations" possible at HO-ODUj layer. The HO-ODUj layer (Parent ODU) is identified by G-PID field.

This field (32 bits) indicates maximum number of Containers of a given HO-ODUj at priority 'p' available on the TE-Link; where {j=1,2,3,4}.

The "Available HO-ODUj(s)" of a bundled link at priority 'p' is defined to be the sum of "Available HO-ODUj(s)" at priority 'p' of all of its component links for that specific G-PID.

Example#1: Unbundled link with ODU2-ODU0 muxing hierarchy support

```
     A ---------- B
```

IMCD advertised would be as follows:
- o G-PID= ODU2-ODU0
- o Available HO-ODUj count = 1 (refers to ODU2 layer)

The ODU2 termination implies ability to mux/demux 8xODU0s.

Example#2: Bundled Te-link with ODU2-ODU0 muxing hierarchy support (3 links)

```
     A ========== B
```

IMCD advertised would be as follows:
- o G-PID= ODU2-ODU0
- o Available HO-ODUj count = 3 (refers to ODU2 layer)

The ODU2 termination implies ability to mux/demux 24xODU0s in total.

5.5.3 Controlling IMCD advertisement

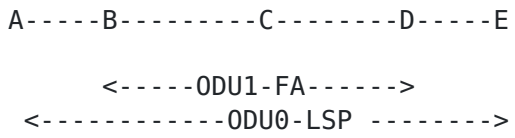
The IMCD advertisement is not mandatory and it is required only when FA support is needed.

The network operators can selectively enable IMCD advertisement for specific HO-ODU mux layer(s). This can be done on a link by link basis, node basis or network basis. The mechanism to achieve this is outside the scope of this document.

5.5.4 How to use IMCDs for FA creation

When computing path for an FA (induced or otherwise), the path computing node should look for matching G-PIDs at the FA boundary nodes. For example, to create ODU1-FA for ODU0 service, the path computation should look for matching G-PID = ODU1-ODU0 at nodes B & D

The need for FA is due to Node-C's ability to switch ODU1 only.



5.5.5 IMCD and non OTN services

In certain deployments it may be beneficial to advertise ODU termination bandwidth without the LO-ODU information. The intent is to allow signaling to decide non-OTN signal to adapt at the time of path establishment.

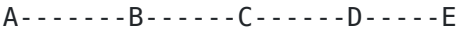
The G-PID values 96, 97, 98, 99 defined in section 5.5.2.1 are meant for this purpose.

The path computation can also be preformed for specific clients over an ODUj using G-PID values 100, 101 & 102 (e.g. 10GBE mapping to ODU2 using GFP).

6. Examples

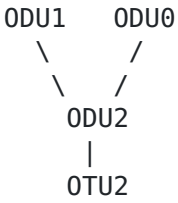
This sections presents some use-cases for bandwidth encoding and usage.

6.1. Network with no IMCD advertisement (no FA support)

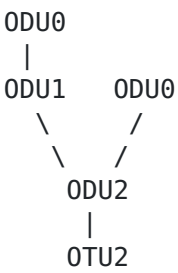


Suppose Muxing Hierarchy supported at the end points as shown:

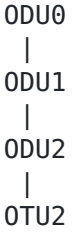
Link A-B: Mux hierarchy at A & B ends is as follows:



Link B-C: Mux hierarchy at B & C ends is as follows:



Link C-D: mux hierarchy at C & D ends is as follows:



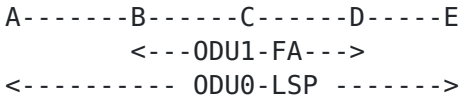
a) The ISCD advertisement by nodes A, B, C & D would be as follows

```
          ISCD1:
            Max LSP BW = ODU2 nominal rate in bytes/sec
            Available ODU2 count at priority 'p' = 1

          ISCD2:
            Max LSP BW = ODU1 nominal rate in bytes/sec
            Available ODU1 count at priority 'p' = 4

          ISCD3:
            Max LSP BW = ODU0 nominal rate in bytes/sec
            Available ODU0 count at priority 'p' = 8
```

b) BW advertisement after an ODU0-LSP creation from A to D.
   The bandwidth is no longer available at ODU2 rate.

```
          ISCD1:
            Max LSP BW = ODU1 nominal rate in bytes/sec
            Available ODU1 count at priority 'p' = 3

          ISCD2:
            Max LSP BW = ODU0 nominal rate in bytes/sec
            Available ODU0 count at priority 'p' = 7
```

6.2. Network with IMCD advertisement for FA support



The above network can support FA at ODU2 and ODU1 layers.
To support FA origination/termination, the IMCDs would be advertised
as follows. This is in addition to ISCD advertisement.

The ISCD1, ISCD2 & ISCD3 advertisement by A, B, C & D is same as in section 7.1

The IMCD advertisement by A & B for link A-B:

```
          IMCD1:
                G-PID = ODU2-ODU1
                Available HO-ODUj count at Pi = 1 (ODU2)

          IMCD2:
                G-PID = ODU2-ODU0
                Available HO-ODUj count at Pi = 1 (ODU2)
```

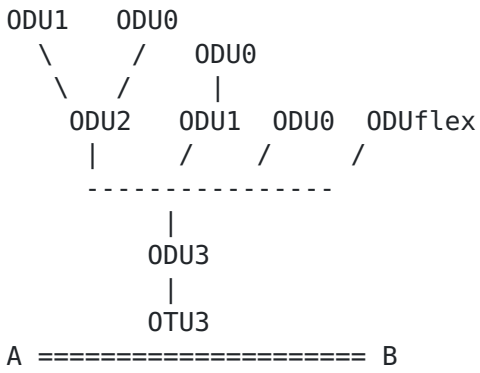
The IMCD advertisement by B & C for link B-C:

```
          IMCD1:
```

```
                G-PID = ODU2-ODU1
                Available HO-ODUj count at Pi = 1 (ODU2)

          IMCD2:
                G-PID = ODU2-ODU0
                Available HO-ODUj count at Pi = 1 (ODU2)

          IMCD3:
                G-PID = ODU1-ODU0
                Available HO-ODUj count at Pi = 4 (ODU1)

The IMCDs advertised by C & D for link C-D would be as follows:
          IMCD1:
                G-PID = ODU2-ODU1
                Available HO-ODUj count at Pi = 1 (ODU2)

          IMCD2:
                G-PID = ODU2-ODU1-ODU0
                Available HO-ODUj count at Pi = 1 (ODU2)

          IMCD3:
                G-PID = ODU1-ODU0
                Available HO-ODUj count at Pi = 4 (ODU1)


The IMCD advertisement by B & C for link B-C after ODU1-FA creation:
          IMCD1:
                G-PID = ODU1-ODU0
                Available HO-ODUj count at Pi = 3 (ODU1)

The IMCD advertisement by C & D for link C-D after ODU1-FA creation:
          IMCD1:
                G-PID = ODU1-ODU0
                Available HO-ODUj count at Pi = 3 (ODU1)
```

6.3. Link bundle with similar muxing capabilities

Consider a Bundled Te-link with 2xOTU3 links between Nodes A & B with multiplexing hierarchy as shown:



The ISCDs and IMCDs advertised by A & B would be as follows:

ISCD1:
  Max LSP BW = ODU3 nominal rate in bytes/sec
  Available ODU3 count at priority 'p' = 2

ISCD2:
  Max LSP BW = ODU2 nominal rate in bytes/sec
  Available ODU2 count at priority 'p' = 8

ISCD3:
  Max LSP BW = ODU1 nominal rate in bytes/sec
  Available ODU1 count at priority 'p' = 32

ISCD4:
  Max LSP BW = ODU0 nominal rate in bytes/sec
  Available ODU0 count at priority 'p' = 64

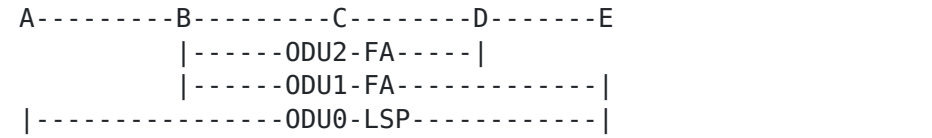
ISCD5:
  Max LSP BW = ODU3 nominal rate in bytes/sec
  Available ODUflex BW = 2xODU3 BW in byte/sec

To support FAs at ODU3, ODU2 & ODU1 rates, the following IMCDs are advertised

IMCD1:
  G-PID = ODU3-ODU2
  Available HO-ODUj count at Pi = 2 (ODU3)

```
IMCD2:
  G-PID = ODU3-ODU2-ODU1
  Available HO-ODUj count at Pi = 2 (ODU3)

IMCD3:
  G-PID = ODU3-ODU2-ODU0
  Available HO-ODUj count at Pi = 2 (ODU3)

IMCD4:
  G-PID = ODU3-ODU1
  Available HO-ODUj count at Pi = 2 (ODU3)

IMCD5:
  G-PID = ODU3-ODU1-ODU0
  Available HO-ODUj count at Pi = 2 (ODU3)

IMCD6:
  G-PID = ODU3-ODU0
  Available HO-ODUj count at Pi = 2 (ODU3)

IMCD7:
  G-PID = ODU2-ODU1
  Available HO-ODUj count at Pi = 8 (ODU2)

IMCD8:
  G-PID = ODU2-ODU0
  Available HO-ODUj count at Pi = 8 (ODU2)

IMCD9:
  G-PID = ODU1-ODU0
  Available HO-ODUj count at Pi = 32 (ODU1)

IMCD9:
  G-PID = ODU3-ODUflex
  Available HO-ODUj count at Pi = 3 (ODUflex)
```

6.4. Link bundle with dissimilar muxing capabilities: Layer relation



Link A-B: Hierarchy at both ends is OTU2-ODU2-ODU0
Link B-C: Is a bundled Te-link with 3 component links with multiplexing hierarchy at both ends as shown:

```
          Component link#1: OTU2 link with mux hierarchy: OTU2-ODU2-ODU1-ODU0
          Component link#2: OTU2 link with mux hierarchy: OTU2-ODU2-ODU1
          Component link#3: OTU1 link with mux hierarchy: OTU1-ODU1-ODU0
          Link C-D:
            - Hierarchy at C end is OTU2-ODU2
            - Hierarchy at D end is OTU2-ODU2-ODU1

          Link D-E:
            - Hierarchy at D end is OTU1-ODU1
            - Hierarchy at E end is OTU1-ODU1-ODU0



                    The IMCDs advertised for B-C would include the following:
                    IMCD1:
                      G-PID = ODU2-ODU1
                      Available HO-ODUj count at Pi = 2 (ODU2)

                    IMCD2:
                      G-PID = ODU1-ODU0
                      Available HO-ODUj count at Pi = 5 (ODU1)

                    IMCD3:
                      G-PID = ODU2-ODU1-ODU0
                      Available HO-ODUj count at Pi = 1 (ODU2)
```

In this example, we need two FAs to originate from the same point (at node-B). It is necessary to advertise IMCD3 as we can not conclude full mux relation from IMCD1 & IMCD2.

7. Backward Compatibility

If backwards compatibility is required with G.709-v1, then [RFC4328] based ISCDs should be a advertised in addition to ISCDs/IMCDs specified in this document.

8. Security Considerations

There are no additional security implications to OSPF routing protocol due to the extensions captured in this document.

9. IANA Considerations

The memo introduces two new sub-TLVs of the Interface Switch Capability Descriptor Sub-TLV of TE-LSA. [RFC3630] says that the sub-TLVs of the TE Link TLV in the range 10-32767 must be assigned by Expert Review, and must be registered with IANA.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels".

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)"

[RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)"

[RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.

[RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.

[RFC5212] Shiomoto, 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.

[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.

[RFC6001] D. Papadimitriou, et al, Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)

[G.709-v3] ITU-T, "Interfaces for the Optical Transport Network (OTN)", G.709 Recommendation, December 2009.

[GSUP.43] ITU-T, "Proposed revision of G.sup43 (for agreement)", December 2008.

10.2. Informative References

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

[G.709-v1] ITU-T, "Interface for the Optical Transport Network (OTN)," G.709 Recommendation (and Amendment 1), February 2001 (October 2001).

[G.872] ITU-T, "Architecture of optical transport networks", November 2001 (11 2001).

[G.709-FRAME] F. Zhang, D. Li, H. Li, S. Belotti, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", draft-zhang-ccamp-gmpls-g709-framework-02, work in progress.

[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.

11. Acknowledgements

Special thanks to Daniele Ceccarelli, Lyndon Ong, Sergio Belotti, Pietro Grandi, Jonathan Sadler, Remi Theillaud, Fatai Zhang and Diego Caviglia for discussions on various modeling options.

Authors would like to thank Lou Berger,Ping Pan, Radhakrishna Valiveti and Mohit Misra for review comments and suggestions.

Author's Addresses

Ashok Kunjidhapatham
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA
Email: akunjidhapatham@infinera.com

Rajan Rao
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA
Email: rrao@infinera.com

Snigdho Bardalai
Infinera Corporation

169, Java Drive
Sunnyvale, CA-94089
USA
Email: sbardalai@infinera.com

Khuzema Pithewan
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA
Email: kpithewan@infinera.com

Biao Lu
Infinera Corporation
169, Java Drive
Sunnyvale, CA-94089
USA
Email: blu@infinera.com

John Drake
Juniper Networks
USA
Email: jdrake@juniper.net

Steve Balls
Metaswitch Networks
100 Church Street
Enfield
EN2 6BQ U.K.
Email: steve.balls@metaswitch.com

Xihua Fu,
ZTE
China
fu.xihua@zte.com.cn

Appendix A: Abbreviations & Terminology

A.1 Abbreviations:

CBR	Constant Bit Rate
GFP	Generic Framing Procedure
HO-ODU	Higher Order ODU
LSC	Lambda Switch Capable
LSP	Label Switched Path
LO-ODU	Lower Order ODU
ISCD	Interface Switch Capability Descriptor
OCC	Optical Channel Carrier
OCG	Optical Carrier Group
OCh	Optical Channel (with full functionality)
OChr	Optical Channel (with reduced functionality)
ODTUG	Optical Date Tributary Unit Group
ODU	Optical Channel Data Unit
OMS	Optical Multiplex Section
OMU	Optical Multiplex Unit
OPS	Optical Physical Section
OPU	Optical Channel Payload Unit
OSC	Optical Supervisory Channel
OTH	Optical Transport Hierarchy
OTM	Optical Transport Module
OTN	Optical Transport Network
OTS	Optical Transmission Section
OTU	Optical Channel Transport Unit
OTUkV	Functionally Standardized OTUk
SCSI	Switch Capability Specific Information
TDM	Time Division Multiplex

A.2 Terminology

1. ODUk and ODUj

ODUk refers to the ODU container that is directly mapped to an OTU container. ODUj refers to the lower order ODU container that is mapped to an higher order ODU container via multiplexing.

2. LO-ODU and HO-ODU

LO-ODU refers to the ODU client layer of lower rate that is mapped to an ODU server layer of higher rate via multiplexing. HO-ODU refers to the ODU server layer of higher rate that supports mapping of one or more ODU client layers of lower rate.

In multi-stage multiplexing case, a given ODU layer can be a client for one stage (interpreted as LO-ODU) and at the same

time server for another stage (interpreted as HO-ODU). In this case, the notion of LO-ODU and HO-ODU needs to be interpreted in a recursive manner.

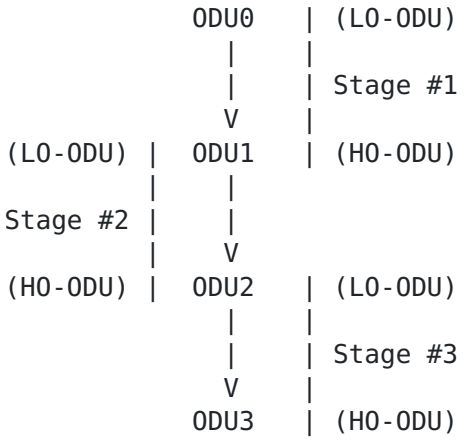


Figure-4 : LO-ODU and HO-ODU

Appendix B : Optimization Techniques

Optimization techniques can be used to reduce TE-LSA size. The following sub sections describe available options.

B.1 Multiple ISCDs Vs. Single ISCD

It is possible to encode ISCDs corresponding to different ODU layers into SCSI field of a single ISCD. This options was shown in previous version of this draft (draft-ashok-ccamp-gmpls-ospf-g709-02).

Doing so will reduce the LSA size by a factor of:
10 words x (#ODUjs - 1)

It is possible to reduce LSA size further by reducing the size of BW field to half word. Doing so will reduce LSA size by a factor of:
4 words x (#ODUjs)

B.1 Multiple IMCDs Vs. Single IMCD

This optimization doesn't save much. The shrinking of BW field to 1/2 word helps reduce LSA size to some extent. The size reduction depends on the number of ODUs supported.

4 words x (#ODUjs)

B.1 Eight priorities Vs. restricted number of priorities

It is possible to further optimize by advertising BW only for supported priorities. This can be easily achieved by having a bit vector as described in previous version of this draft.

Appendix C: Relation with MLN & MRN

The ISCD and IMCDs defined in this draft doesn't repalce IACDs. All three (ISCD, IMCD & IACD) can co-exist in a network and serve different purposes.

Appendix D : AMP, BMP & GMP Mapping

The G.709 defines various mapping schemes for LO-ODUs into HO-ODUs. From G.709 descriptions, the AMP & GMP mapping appears to be fixed for a given LO-ODU to HO-ODU based on the time slot granularity. Since the mapping is fixed we do not see value in advertising this information in TE-LSAs.