Traffic Engineering Extensions to OSPF for GMPLS Control of Beyond-100G G.709 Optical Transport Networks
<draft-izh-ccamp-b100g-routing-00>

Abstract

This document describes Open Shortest Path First - Traffic Engineering (OSPF-TE) routing protocol extensions to support GMPLS control of Optical Transport Networks (OTNs) specified in ITU-T Recommendation G.709 published in 2016. The 2016 version of G.709 introduces support for higher rate OTU signals, termed OTUCn (which have a nominal rate of 100n Gbps). The newly introduced OTUCn represent a very powerful extension to the OTN capabilities, and one which naturally scales to transport any newer clients with bit rates in excess of 100G, as they are introduced. This document extends the mechanisms defined in [RFC7138].

Status of This Memo

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

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

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

This Internet-Draft will expire on September 8, 2017.
1. Introduction

The current GMPLS routing extensions RFC [RFC7138] includes coverage for all the OTN capabilities that were defined in the 2012 version of G.709 [ITU-T_G709_2012]. The 2016 version of G.709 [ITU-T_G709_2016] introduces the following key extensions:

a. OTUCn signals with bandwidth larger than 100G (n*100G)
b. ODUcn signals with bandwidth larger than 100G.

c. ODUflex signals with bandwidth larger than 100G

d. mapping client signals with bandwidth larger than 100G into the corresponding ODUflex containers.

e. Tributary Slot Granularity of 5G

This document provides extensions required in GMPLS OSPF-TE for B100G OTN technology. For a short overview of OTN evolution and implications of B100G on GMPLS routing, please refer to [I-D.zih-ccamp-otn-b100g-fwk].

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

As discussed in [I-D.zih-ccamp-otn-b100g-fwk], OSPF-TE must be extended to advertise the termination, Switching and multiplexing Capabilities for ODUcn and OTUCn (Optical Transport Unit) links. These capabilities are carried in the Switching Capability specific information field of the Interface Switching Capability Descriptor (ISCD) using formats defined in this document.

3. TE-Link Representation

G.709 ODUcn/OTUCn 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. Figure 1 below provides an illustration of one-hop OTUCn TE-Links.

```
+-------+    +-------+    +-------+
 |  OTN  |  OTN  |  OTN  |
 |Switch |<- OTUCn Link ->|Switch |<- OTUCn Link ->|Switch |
 |  A    |       |  B    |       |  C    |
 +-------+       +-------+       +-------+
 |<-- TE-Link --|    |<-- TE-Link --|

Figure 1: OTUCn TE-Links
```
4. ISCD Format Extensions

The ISCD describes the Switching Capability of an interface and is defined in [RFC4203]. This document resues the switching capability defined in [RFC7138] but introduces a new encoding type (to be assigned) as follows:

- G.709-2106 ODUCn (Digital Section): One codepoint (applicable to all values of n) needs to be defined in the signaling extensions [TBD]. The same value is used for advertising fixed rate ODUs, as well as ODUflex signals supported by an ODUCn link. When the Switching Capability and Encoding fields are set to values as stated above, the Interface Switching Capability Descriptor MUST be interpreted as defined in [RFC7138].

The MAX LSP Bandwidth field is used according to [RFC4203], i.e., 0 \leq MAX LSP Bandwidth \leq rate (ODUCn). The bandwidth is expressed in bytes/second and the encoding MUST be in IEEE floating point format. The discrete rates for new ODUs introduced in G709-2016 are shown in Table 1.

<table>
<thead>
<tr>
<th>ODU Type</th>
<th>ODU Bit Rate</th>
<th>IEEE encoding of bw (bytes/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODUflex for IMP mapped packet traffic</td>
<td>s x 239/238 x 5 156 250 kbit/s: s=2,8,5*n, n \geq 1</td>
<td>TBD</td>
</tr>
<tr>
<td>ODUflex for FlexE aware transport</td>
<td>103 125 000 x 240/238 x n/20 kbit/s, where n is total number of available tributary slots among all PHYs which have been crunched and combined.</td>
<td>TBD</td>
</tr>
</tbody>
</table>

Note that this table doesn’t include ODUCn -- since it cannot be generated by mapping a non-OTN signal. An ODUCn is always formed by multiplexing multiple LO-ODUs.

Table 1: Types and rates of ODUs usable for client mappings

ISCD advertisement and processing rules are exactly as specified in [RFC7138].
4.1. Switching Capability Specific Information

The technology-specific part of the OTN-TDM ISCD may include a variable number of sub-TLVs called Bandwidth sub-TLVs. Each sub-TLV is encoded with the sub-TLV header as defined in [RFC7138]. The muxing hierarchy tree MUST be encoded as an order-independent list. In addition to the sub-TLVs of types 1 and 2 defined in [RFC7138], Section 4.1.1 introduces a new sub-TLV type 3 to advertise ODUCn Information.

The Switching Capability specific information (SCSI) for OTUCn links MUST include a Type 3 TLV at the beginning, followed by Type 1 and/or Type 2 sub-TLVs as defined in [RFC7138].

With respect to ODUflex, new Signal Types need to be defined for the new ODUflex signals introduced in Table 1:

- 23 - ODUflex (IMP)
- 24 - ODUflex (FlexE)

Each ODUflex signal MUST always be advertised in a separate Type 2 sub-TLV as per [RFC7138].

4.1.1. Switching Capability Specific Information for ODUCn containers

The format of the Bandwidth sub-TLV for ODUCn signals is depicted in the following figure:

```
| 0 | 1 | 2 | 3 |
+---+---+---+---+
| 0 | 1 | 2 | 3 |
| 0 | 1 | 2 | 3 |
| 0 | 1 | 2 | 3 |
| Type = 3 (Unres-ODUC, TBA) | Length |
| Sig Type=ODUCn | N Value | T | S | TSG | Res | Priority |
```

Figure 2: Bandwidth Sub-TLV -- Type 3

The values of the fields in the Bandwidth sub-TLV shown in Figure 2 are explained below.

- Signal Type (8 bits): Indicates the ODU type being advertised. For this sub-TLV type, a new signal type needs to be defined for ODUCn signals. Rather than define a unique signal type for each value of the parameter ‘n’, this draft proposes that we allocate a
single signal type for the ODUCn signal family, and encode the value of ‘n’ as a separate field. The first row after Type and Length MUST be followed by ODUCn information as shown.

- **N-Value (8 bits):** Indicates the value of ‘n’ in ODUCn field. The value of this field is an integer in the range 1...256 as per [ITU-T_G709_2016].

- **Flags (8 bits):**
  - **T Flag (bit 17):** Indicates whether the advertised bandwidth can be terminated per [RFC7138]. Since an ODUCn MUST be advertised as non-switchable and terminated, the T field MUST be set to 1.
  - **S Flag (bit 18):** Indicates whether the advertised bandwidth can be switched. Since an ODUCn MUST be advertised as non-switchable and terminated, the S field MUST be set to 0.

- **TSG (3 bits):** Tributary Slot Granularity. Used for the advertisement of the supported tributary slot granularity. This document defines a new value for 5 Gbps time slots – which MUST be used when advertising OTUCn links. The values in the range 0-3 MUST be interpreted as defined in [RFC7138].
  - 0 - Ignored
  - 1 - 1.25 Gbps / 2.5 Gbps
  - 2 - 2.5 Gbps only
  - 3 - 1.25 Gbps only
  - 4 - 5.0 Gbps only [TBA by IANA]
  - 5-7 - Reserved

- **Priority (8 bits):** The meaning and usage of priority field MUST same as in [RFC7138].

5. Examples

The examples in the following pages are not normative and are not intended to imply or mandate any specific implementation.
5.1. MAX LSP Bandwidth Fields in the ISCD

This example shows how the MAX LSP Bandwidth fields of the ISCD are filled according to TE-Link bandwidth occupancy. In this example, an OTUC4 link is considered, with (a) supported priorities 0, 2, 4, 7 (b) 300G of bandwidth already consumed (c) 100G bandwidth available, and able to support an ODU4 LSP.

At time T0, the advertisement would be as shown in Figure 3:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SwCap=OTN_TDM | Encoding = TBA | Reserved (all zeros) | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 0 = 100 Gpbs + | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 1 = 0 | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 2 = 100 Gpbs | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 3 = 0 | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 4 = 100 Gpbs | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 5 = 0 | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 6 = 0 | +---------------------------------------------------+
| MAX LSP Bandwidth at priority 7 = 100 Gpbs | +---------------------------------------------------+
| Switching Capability Specific Information | +---------------------------------------------------+
| (variable length) | +---------------------------------------------------+
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: MAX LSP Bandwidth Fields in the ISCD at T0

At time T1, an ODU3 at priority 2 is set up. Once the ODU3 is carried over the ODUC4, the unreserved bandwidth reduces to 60G and consequently MAX LSP Bandwidth is advertised as ODU3, since no more ODU4s are available and the next supported ODUj in the hierarchy is ODU3. The updated advertisement is as shown in Figure 4:
At time $T_2$, an ODU2 at priority 4 is set up. The Max LSP bandwidth is still advertised as ODU3 as in Figure 4 since the remaining bandwidth is 50G. When the available BW drops below 40G, the max LSP BW is advertised as 10G. The advertisement is updated as shown in Figure 5:
5.2. Example of T, S, and TS Granularity Utilization

To be added later.

5.3. Example of ODUflex Advertisement

To be added later.

5.4. Example of Single-Stage Muxing

Suppose there is 1 OTUC4 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 the sake of simplicity, we also assume that only priorities 0 and 3 are supported.

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

Figure 5: MAX LSP Bandwidth Fields in the ISCD at T2
The related SCsIs are as follows:

0                   1                   2                   3
+---------------------------------------------------------------+
| Type = 3 (Unres-fix) | Length = 8 |
+---------------------------------------------------------------+
| Sig type=ODUCn | N-value=4 | 1 | 4 | 0 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 0 0 0 |
+---------------------------------------------------------------+
| Type = 1 (Unres-fix) | Length = 12 |
+---------------------------------------------------------------+
| Sig type=ODU1 | #stages= 1 | X | X | X X X | 0 0 0 0 0 0 0 0 0 0 0 |
| Stage#1=ODUCn | Padding (all zeros) |
+---------------------------------------------------------------+
| Unres ODU1 at Prio 0 =160 | Unres ODU1 at Prio 3 =160 |
+---------------------------------------------------------------+
| Type = 1 (Unres-fix) | Length = 12 |
+---------------------------------------------------------------+
| Sig type=ODU2 | #stages= 1 | X | X | X X X | 0 0 0 0 0 0 0 0 0 0 0 |
| Stage#1=ODUCn | Padding (all zeros) |
+---------------------------------------------------------------+
| Unres ODU2 at Prio 0 =40 | Unres ODU2 at Prio 3 =40 |
+---------------------------------------------------------------+
| Type = 1 (Unres-fix) | Length = 12 |
+---------------------------------------------------------------+
| Sig type=ODU3 | #stages= 1 | X | X | X X X | 0 0 0 0 0 0 0 0 0 0 0 |
| Stage#1=ODUCn | Padding (all zeros) |
+---------------------------------------------------------------+
| Unres ODU3 at Prio 0 =10 | Unres ODU3 at Prio 3 =10 |
+---------------------------------------------------------------+
| Type = 2 (Unres/MAX-var) | Length = 24 |
+---------------------------------------------------------------+
| Sig type=ODUCn | N-value=4 | 1 | 4 | 0 0 0 0 0 0 0 0 0 0 0 |
+---------------------------------------------------------------+
| S. type=ODUflex | #stages= 1 | X | X | X X X | 0 0 0 0 0 0 0 0 0 0 0 |
| Stage#1=ODUCn | Padding (all zeros) |
+---------------------------------------------------------------+
| Unreserved Bandwidth at priority 0 =400 Gbps |
| Unreserved Bandwidth at priority 3 =400 Gbps |
| MAX LSP Bandwidth at priority 0 =400 Gbps |
MAX LSP Bandwidth at priority 3 = 400 Gbps

Figure 6: Single-Stage Muxing

5.5. Example of Multi-Stage Muxing -- Unbundled Link

Suppose there is 1 OTUC4 link with muxing capabilities as shown in the following figure:

```
  ODU2    ODU0    ODUflex    ODU0
 /      /      /      /
 |      |      |      |
 ODU3    ODU2
 /      /      /      /
 |      |      |      |
 ODU3    ODU2
 /      /      /      /
 |      |      |      |
 ODU3    ODU2
 /      /      /      /
 |      |      |      |
 ODU3    ODU2
 /      /      /      /
 |      |      |      |
 ODUC4
```

The ODUC4 is not a switchable entity. It is advertised with zero counts to show TSG information. Considering only supported priorities 0 and 3, the advertisement is composed by the following Bandwidth sub-TLVs:

```
<table>
<thead>
<tr>
<th>Type = 3 (Unres-fix)</th>
<th>Length = 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODUCn</td>
<td>N-value=4</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Unres ODU3 at Prio 0 =10</td>
<td>Unres ODU3 at Prio 3 =10</td>
</tr>
</tbody>
</table>
```

The ODUC4 is not a switchable entity. It is advertised with zero counts to show TSG information. Considering only supported priorities 0 and 3, the advertisement is composed by the following Bandwidth sub-TLVs:

```
<table>
<thead>
<tr>
<th>Type = 3 (Unres-fix)</th>
<th>Length = 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig type=ODUCn</td>
<td>N-value=4</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Unres ODU3 at Prio 0 =10</td>
<td>Unres ODU3 at Prio 3 =10</td>
</tr>
</tbody>
</table>
```

Figure 7: Multi-Stage Muxing -- Unbundled Link
6. Security Considerations

Please refer to [RFC5920] for details on security threats; defensive techniques; monitoring, detection, and reporting of security attacks; and requirements.

7. IANA Considerations

TBD

8. Contributors

9. Acknowledgements

10. References

10.1. Normative References

[I-D.zih-ccamp-otn-b100g-fwk]

[ITU-T_G709_2012]

[ITU-T_G709_2016]


10.2. Informative References


Authors' Addresses

Rajan Rao
Infinera Corporation
140 Caspian CT.
Sunnyvale, CA-94089
USA
EMail: rrao@infinera.com

Iftekhar Hussain
Infinera Corporation
140 Caspian CT.
Sunnyvale, CA-94089
USA
EMail: IHussain@infinera.com

Radha Valiveti
Infinera Corporation
140 Caspian CT.
Sunnyvale, CA-94089
USA
EMail: rvaliveti@infinera.com

Qilei Wang (editor)
ZTE
Nanjing
CN
EMail: wang.qilei@zte.com.cn
Yuanbin Zhang  
ZTE  
Beijing  
CN  
EMail: zhang.yuanbin@zte.com.cn

Huub van Helvoort  
Hai Gaoming B.V  
EMail: huubatwork@gmail.com