                   Evaluation of Existing GMPLS Encoding
              against G.709v3 Optical Transport Networks (OTNs)

Abstract

   ITU-T recommendation G.709-2012 has introduced new fixed and flexible
   Optical channel Data Unit (ODU) containers in Optical Transport
   Networks (OTNs).

   This document provides an evaluation of existing Generalized
   Multiprotocol Label Switching (GMPLS) routing and signaling protocols
   against the G.709 OTNs.

Status of This Memo

Table of Contents

1. Introduction

    GMPLS routing [RFC4203] [RFC5307] and signaling [RFC3473] [RFC4328]
    provide the mechanisms for basic GMPLS control of Optical Transport
    Networks (OTNs) based on the 2001 revision of the G.709 specification
    [G.709-2001]. The 2012 revision of the G.709 specification
    [G.709-2012] includes new OTN features that are not supported by
    GMPLS.

    This document provides an evaluation of exiting GMPLS signaling and
    routing protocols against G.709 requirements. Background information
    and a framework for the GMPLS protocol extensions needed to support
    G.709 is provided in [RFC7062]. Specific routing and signaling
    extensions defined in [OTN-OSPF] and [OTN-RSVP] specifically address
    the gaps identified in this document.

2. G.709 Mapping and Multiplexing Capabilities

   The digital OTN-layered structure is comprised of the digital path
   layer (ODU) and the digital section layer (OTU). An OTU (Optical
   channel Transport Unit) section layer supports one ODU path layer as
   a client and provides monitoring capability for the Optical Channel
   (OCh), which is the optical path carrying the digital OTN structure.
   An ODU path layer may transport a heterogeneous assembly of ODU
   clients. Some types of ODUs (i.e., ODU1, ODU2, ODU3, and ODU4) may
   assume either a client or server role within the context of a
   particular networking domain. The terms ODU1, ODU2, ODU3, ODU4, and
   flexible ODU (ODUflex) are explained in G.709. G.872 [G.872]
   provides two tables defining mapping and multiplexing capabilities of
   OTNs, which are reported below.

```
+-------------------+-------------------+
|     ODU client    |     OTU server    |
+-------------------+-------------------+
|       ODU0        |         -         |
+-------------------+-------------------+
|       ODU1        |       OTU 1       |
+-------------------+-------------------+
|       ODU2        |       OTU 2       |
+-------------------+-------------------+
|       ODU2e       |         -         |
+-------------------+-------------------+
|       ODU3        |       OTU 3       |
+-------------------+-------------------+
|       ODU4        |       OTU 4       |
+-------------------+-------------------+
|      ODUflex      |         -         |
+-------------------+-------------------+
```

                  Figure 1: OTN Mapping Capability

```
+========================================+=======================+
|              ODU client                |      ODU server       |
+----------------------------------------+-----------------------+
|          1.25 Gbit/s client            |                       |
+--------------------------------+       |         ODU0          |
|               -                |       |                       |
+========================================+=======================+
|           2.5 Gbit/s client            |                       |
+--------------------------------+       |         ODU1          |
|              ODU0              |        |                       |
+========================================+=======================+
|           10 Gbit/s client             |                       |
+--------------------------------+       |         ODU2          |
|        ODU0,ODU1,ODUflex       |       |                       |
+========================================+=======================+
|         10.3125 Gbit/s client          |                       |
+--------------------------------+       |         ODU2e         |
|               -                |       |                       |
+========================================+=======================+
|           40 Gbit/s client             |                       |
+--------------------------------+       |         ODU3          |
|   ODU0,ODU1,ODU2,ODU2e,ODUflex  |      |                       |
+========================================+=======================+
|          100 Gbit/s client             |                       |
+--------------------------------+       |         ODU4          |
|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   |       |        ODUflex        |
|from 1.25 Gbit/s to 100 Gbit/s. |       |                       |
+--------------------------------+       |                       |
|               -                |       |                       |
+========================================+=======================+
(*) - Constant Bit Rate
(**) - Generic Framing Procedure - Framed (GFP-F)
```

Figure 2: OTN Multiplexing Capability

In the following, the terms Optical channel Data Unit-j (ODUj) and
Optical channel Data Unit-k (ODUk) are used in a multiplexing
scenario to identify the lower order signal (ODUj) and the higher
order signal (ODUk). 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 Optical channel Transport Unit-k (OTUk) section,
with the same path and section rates as that of the connection

service (see Figure 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 Figure 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 [RFC7062], client ODUk connection services can be
transported over:

Case A: one or more wavelength subnetworks connected by optical
         links, or

Case B: one or more ODU links (having sub-lambda and/or lambda
         bandwidth granularity), or

Case C: a mix of ODU links and wavelength subnetworks.

This document considers the Traffic Engineering (TE) information
needed for ODU path computation and the parameters needed to be
signaled for Label Switched Path (LSP) setup.

The following sections list and analyze what GMPLS already has and
what it is missing with regard to each type of data that needs to be
advertised and signaled.

3. Tributary Slot Granularity

   G.709 defines two types of Tributary Slot (TS) granularities. This
   TS granularity is defined per layer, meaning that both ends of a link
   can select proper TS granularity differently for each supported
   layer, based on the rules below:

   o If both ends of a link are new cards supporting both 1.25 Gbit/s
      TS and 2.5 Gbit/s TS, then the link will work with 1.25 Gbit/s TS.

   o If one end of a link is a new card supporting both the 1.25 Gbit/s
      and 2.5 Gbit/s TS granularities, and the other end is an old card
      supporting just the 2.5 Gbit/s TS granularity, the link will work
      with 2.5 Gbit/s TS granularity.

3.1. Data-Plane Considerations

3.1.1. Payload Type and TS Granularity Relationship

   As defined in G.709, an ODUk container consists of an Optical channel
   Payload Unit-k (OPUk) plus a specific ODUk Overhead (OH). OPUk OH
   information is added to the OPUk information payload to create an
   OPUk. It includes information to support the adaptation of client
   signals. Within the OPUk overhead, there is the payload structure
   identifier (PSI) that includes the payload type (PT). The PT is used
   to indicate the composition of the OPUk signal. When an ODUj signal
   is multiplexed into an ODUk, the ODUj signal is first extended with
   the frame alignment overhead and then mapped into an Optical channel
   Data Tributary Unit (ODTU). Two different types of ODTUs are
   defined:

   o ODTUjk ((j,k) = {(0,1), (1,2), (1,3), (2,3)}; ODTU01, ODTU12,
      ODTU13, and ODTU23) in which an ODUj signal is mapped via the
      Asynchronous Mapping Procedure (AMP), as defined in Section 19.5
      of [G.709-2012].

   o ODTUk.ts ((k,ts) = (2,1..8), (3,1..32), (4,1..80)) in which a
      lower order ODU (ODU0, ODU1, ODU2, ODU2e, ODU3, and ODUflex)
      signal is mapped via the Generic Mapping Procedure (GMP), as
      defined in Section 19.6 of [G.709-2012].

   G.709 also introduces a logical entity, called Optical channel Data
   Tributary Unit Group (ODTUGk), characterizing the multiplexing of the
   various ODTU. The ODTUGk is then mapped into OPUk. Optical channel
   Data Tributary Unit j into k (ODTUjk) and Optical channel Data
   Tributary Unit k with ts tributary slots (ODTUk.ts) are directly
   time-division multiplexed into the tributary slots of an OH OPUk.

   When PT is assuming values 0x20 or 0x21, together with OPUk type
   (k=1, 2, 3, 4), it is used to discriminate two different ODU
   multiplex structures for ODTUGx:

   o Value 0x20: supporting ODTUjk only

   o Value 0x21: supporting ODTUk.ts or ODTUk.ts and ODTUjk

   The distinction is needed for OPUk with k=2 or 3 since OPU2 and OPU3
   are able to support both the different ODU multiplex structures. For
   OPU4 and OPU1, only one type of ODTUG is supported: ODTUG4 with
   PT=0x21 and ODTUG1 with PT=0x20 (see Figure 6). The relationship
   between PT and TS granularity is due to the fact that the two

   different ODTUGk types discriminated by PT and OPUk are characterized
   by two different TS granularities of the related OPUk, the former at
   2.5 Gbit/s and the latter at 1.25 Gbit/s.

   In order to complete the picture, in the PSI OH, there is also the
   Multiplex Structure Identifier (MSI) that provides the information on
   which tributary slots of the different ODTUjk or ODTUk.ts are mapped
   into the related OPUk. The following figure shows how the client
   traffic is multiplexed till the OPUk layer.

```
                    +--------+       +-----------+
        +----+      |        !------| ODTUjk     |-----Client
        |    |      | ODTUGk |      +-----.------+
        |    |-----| PT=0x21|            .
        |    |      |        |      +-----.------+
        |    |      |        |------| ODTUk.ts   |-----Client
        |OPUk|      +--------+      +-----------+
        |    |
        |    |      +--------+       +-----------+
        |    |      |        |------| ODTUjk     |-----Client
        |    |-----|        |      +-----.------+
        +----+      | ODTUGk |            .
                    | PT=0x20|      +-----.------+
                    |        |------| ODTUjk     |-----Client
                    +--------+      +-----------+
```

                   Figure 3: OTN Client Multiplexing

3.1.2. Fallback Procedure

   G.798 [G.798] describes the so-called PT=0x21-to-PT=0x20 interworking
   process that explains how two nodes with interfaces that have
   different payload types and, hence, different TS granularity (1.25
   Gbit/s vs. 2.5 Gbit/s), can be coordinated to permit the equipment
   with 1.25 Gbit/s TS granularity to adapt the TS allocation according
   to the different TS granularity (2.5 Gbit/s) of a neighbor.

   Therefore, in order to let the Network Element (NE) change TS
   granularity accordingly to the neighbor requirements, the
   AUTOpayloadtype [G.798] needs to be set. When both the neighbors
   (link or trail) have been configured as structured, the payload type
   received in the overhead is compared to the transmitted PT. If they
   are different and the transmitted one is PT=0x21, the node must fall
   back to PT=0x20. In this case, the fallback process makes the system
   self-consistent, and the only reason for signaling the TS granularity
   is to provide the correct label (i.e., the label for PT=0x21 has
   twice the TS number of PT=0x20). On the other side, if the

AUTOpayloadtype is not configured, the Resource Reservation Protocol-
Traffic Engineering (RSVP-TE) consequent actions need to be defined
in case of a TS mismatch.

3.2. Control-Plane Considerations

When setting up an ODUj over an ODUk, it is possible to identify two
types of TS granularity (TSG): the server and the client. The server
TS granularity is used to map an end-to-end ODUj onto a server ODUk
LSP or links. This parameter cannot be influenced in any way from
the ODUj LSP: the ODUj LSP will be mapped on tributary slots
available on the different links / ODUk LSPs. When setting up an
ODUj at a given rate, the fact that it is carried over a path
composed by links / Forwarding Adjacencies (FAs) structured with 1.25
Gbit/s or 2.5 Gbit/s TS granularity is completely transparent to the
end-to-end ODUj.

The client TS granularity information is one of the parameters needed
to correctly select the adaptation towards the client layers at the
end nodes, and this is the only thing that the ODUj has to guarantee.

In Figure 4, an example of client and server TS granularity
utilization in a scenario with mixed OTN [RFC4328] and OTN interfaces
[G.709-2012] is shown.

```
                     ODU1-LSP
         ...........................................
    TSG-C|                                         |TSG-C
     1.25|              ODU2-H-LSP                  |1.25 Gbit/s
  Gbit/s+-----------X------------------------+
       |        TSG-S|                       |TSG-S
       |         2.5|                        |2.5 Gbit/s
       |        Gbit/s|     ODU3-H-LSP        |
       |         |-----------X------------|
       |         |                        |
       |         |                        |
    +--+--+    +--+--+                  +--+-+
    |     |    |     |      +-+   +-+    |    |
    |  A  +------+  B +-----+ +***+ +-----+  Z  |
    | V.3 | OTU2 | V.1 |OTU3 +-+   +-+ OTU3| V.3 |
    +-----+    +-----+                  +-----+

       ...  Service LSP
       ---  Hierarchical-LSP (H-LSP)
```
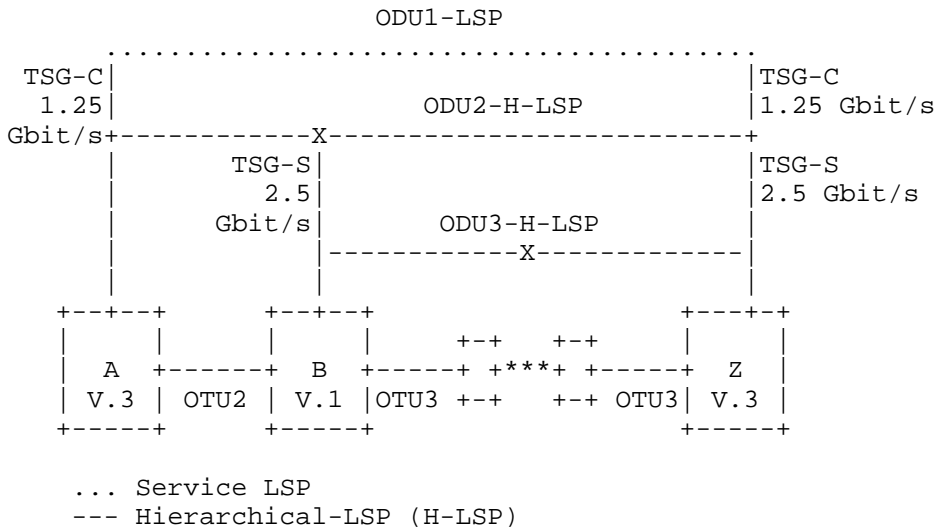
          Figure 4: Client-Server TS Granularity Example

In this scenario, an ODU3 LSP is set up from nodes B to Z. Node B
has an old interface that is able to support 2.5 Gbit/s TS
granularity; hence, only client TS granularity equal to 2.5 Gbit/s
can be exported to ODU3 H-LSP-possible clients. An ODU2 LSP is set
up from nodes A to Z with client TS granularity 1.25 Gbit/s signaled
and exported towards clients. The ODU2 LSP is carried by ODU3 H-LSP
from nodes B to Z. Due to the limitations of the old node B
interface, the ODU2 LSP is mapped with 2.5 Gbit/s TS granularity over
the ODU3 H-LSP. Then, an ODU1 LSP is set up from nodes A to Z, which
is carried by the ODU2 H-LSP and mapped over it using 1.25 Gbit/s TS
granularity.

What is shown in the example is that the TS granularity processing is
a per-layer issue: even if the ODU3 H-LSP is created with the TS
granularity client at 2.5 Gbit/s, the ODU2 H-LSP must guarantee a
1.25 Gbit/s TS granularity client. The ODU3 H-LSP is eligible from
an ODU2 LSP perspective since it is known from the routing that this
ODU3 interface at node Z supports an ODU2 termination exporting a TS
granularity at 1.25 Gbit/s / 2.5 Gbit/s.

The TS granularity information is needed in the routing protocol as
the ingress node (A in the previous example) needs to know if the
interfaces at the last hop can support the required TS granularity.
In case they cannot, A will compute an alternate path from itself to
Z (see Figure 4).

Moreover, TS granularity information also needs to be signaled. As
an example, consider the setup of an ODU3 forwarding adjacency that
is going to carry an ODU0; hence, the support of 1.25 Gbit/s TS is
needed. The information related to the TS granularity has to be
carried in the signaling to permit node C (see Figure 5) to choose
the right one among the different interfaces (with different TS
granularities) towards D. In case the full Explicit Route Object
(ERO) is provided in the signaling with explicit interface
declaration, there is no need for C to choose the right interface
towards D as it has been already decided by the ingress node or by
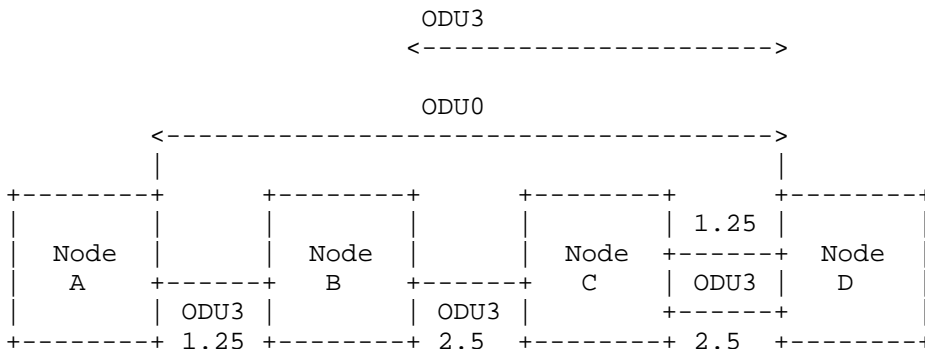the Path Computation Element (PCE).

```
                              ODU3
                     <--------------------->

                          ODU0
              <--------------------------------------->
              |                                       |
     +--------+     +--------+     +--------+     +--------+
     |        |     |        |     |        | 1.25 |        |
     | Node   |     | Node   |     | Node   +------+ Node   |
     |  A     +------+  B     +------+  C     | ODU3 |  D     |
     |        | ODU3 |        | ODU3 |        +------+        |
     +--------+ 1.25 +--------+ 2.5  +--------+ 2.5  +--------+
```

                 Figure 5: TS Granularity in Signaling

In case an ODUk FA_LSP needs to be set up as nesting another ODUj (as
depicted in Figure 5), there might be the need to know the hierarchy
of nested LSPs in addition to TS granularity to permit the
penultimate hop (i.e., C) to choose the correct interface towards the
egress node or any intermediate node (i.e., B) to choose the right
path when performing the ERO expansion. This is not needed in case
we allow bundling only component links with homogeneous hierarchies.
In the case in which a specific implementation does not specify the
last hop interface in the ERO, crankback can be a solution.

In a multi-stage multiplexing environment, any layer can have a
different TS granularity structure; for example, in a multiplexing
hierarchy such as ODU0->ODU2->ODU3, the ODU3 can be structured at TS
granularity = 2.5 Gbit/s in order to support an ODU2 connection, but
this ODU2 connection can be a tunnel for ODU0 and, hence, structured
with 1.25 Gbit/s TS granularity. Therefore, any multiplexing level
has to advertise its TS granularity capabilities in order to allow a
correct path computation by the end nodes (both the ODUk trail and
the H-LSP/FA).

The following table shows the different mapping possibilities
depending on the TS granularity types. The client types are shown in
the left column, while the different OPUk server and related TS
granularities are listed in the top row. The table also shows the
relationship between the TS granularity and the payload type.

| | 2.5 Gbit/s TS | | 1.25 Gbit/s TS | | | |
	OPU2	OPU3	OPU1	OPU2	OPU3	OPU4
ODU0	-	-	AMP PT=0x20	GMP PT=0x21	GMP PT=0x21	GMP PT=0x21
ODU1	AMP PT=0x20	AMP PT=0x20	-	AMP PT=0x21	AMP PT=0x21	GMP PT=0x21
ODU2	-	AMP PT=0x20	-	-	AMP PT=0x21	GMP PT=0x21
ODU2e	-	-	-	-	GMP PT=0x21	GMP PT=0x21
ODU3	-	-	-	-	-	GMP PT=0x21
ODUfl	-	-	-	GMP PT=0x21	GMP PT=0x21	GMP PT=0x21

Figure 6: ODUj into OPUk Mapping Types
(Source: [G.709-2012], Tables7-10)

Specific information could be defined in order to carry the
multiplexing hierarchy and adaptation information (i.e., TS
granularity / PT and AMP / GMP) to enable precise path selection.
That way, when the penultimate node (or the intermediate node
performing the ERO expansion) receives such an object, together with
the Traffic Parameters Object, it is possible to choose the correct
interface towards the egress node.

In conclusion, both routing and signaling need to be extended to
appropriately represent the TS granularity/PT information. Routing
needs to represent a link's TS granularity and PT capabilities as
well as the supported multiplexing hierarchy. Signaling needs to
represent the TS granularity/PT and multiplexing hierarchy encoding.

4. Tributary Port Number

   [RFC4328] supports only the deprecated auto-MSI mode, which assumes
   that the Tributary Port Number (TPN) is automatically assigned in the
   transmit direction and is not checked in the receive direction.

   As described in [G.709-2012] and [G.798], the OPUk overhead in an
   OTUk frame contains n (n = the total number of TSs of the ODUk) MSI
   bytes (in the form of multiframe), each of which is used to indicate
   the association between the TPN and TS of the ODUk.

   The association between the 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 [RFC7062] for further details.) As a consequence,
   the RSVP-TE signaling needs to be extended to support the TPN
   assignment function.

5. Signal Type

   From a routing perspective, GMPLS OSPF [RFC4203] and GMPLS IS-IS
   [RFC5307] only allow advertising interfaces [RFC4328] (the single TS
   type) without the capability of providing precise information about
   bandwidth-specific allocation. For example, in case of link
   bundling, when 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 the example in Figure 3).

   The lack of spatial allocation heavily impacts the restoration
   process because the lack of information on free resources highly
   increases the number of crankbacks affecting network convergence
   time.

   Moreover, actual tools provided by [RFC4203] and [RFC5307] only allow
   advertising signal types with fixed bandwidth and implicit hierarchy
   (e.g., Synchronous Digital Hierarchy (SDH) networks / Synchronous
   Optical Networks (SONETs)) or variable bandwidth with no hierarchy
   (e.g., packet switching networks); but, they do not provide the means
   for advertising networks with a mixed approach (e.g., ODUflex
   Constant Bit Rate (CBR) and ODUflex packet).

   For example, when 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 Interface
   Switching Capability Descriptor (ISCD) for each signal type actually
   supported by the link.

Suppose, for example, there is 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 Bandwidth: ODU1

   MIN LSP Bandwidth: ODU1

   - Maximum Reservable Bandwidth (of the bundle) set to ODU2

   - Unreserved Bandwidth (of the bundle) set to ODU2

In conclusion, the routing extensions defined in [RFC4203] and
[RFC5307] require a different ISCD per signal type in order to
advertise each supported container. This motivates an attempt to
look for a more optimized solution without proliferation of the
number of ISCDs advertised.

Per [RFC2328], OSPF messages are directly encapsulated in IP
datagrams and depend on IP fragmentation when transmitting packets
larger than the network's MTU. [RFC2328] recommends that "IP
fragmentation should be avoided whenever possible". This
recommendation further constrains solutions since OSPF does not
support any generic mechanism to fragment OSPF Link State
Advertisements (LSAs). Even when used in IP environments, IS-IS
[RFC1195] does not support message sizes larger than a link's maximum
frame size.

With respect to link bundling [RFC4201], 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 signaling point of view, [RFC4328]
describes GMPLS signaling extensions to support the control of G.709
OTNs defined before 2011 [G.709-2001]. 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.

6. Bit Rate and Tolerance

   In the current traffic parameters signaling, bit rate and tolerance
   are implicitly defined by the signal type. ODUflex CBR and ODUflex
   packet can have variable bit rates (please refer to [RFC7062],
   Table 2); hence, signaling traffic parameters need to be upgraded.
   With respect to tolerance, there is no need to upgrade GMPLS
   protocols as a fixed value (+/-100 parts per million (ppm) or +/-20
   ppm depending on the signal type) is defined for each signal type.

7. 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 to uncertainty about how many LSPs
   of a specific signal type can be restored. As an example, consider
   the scenario depicted in the following figure.

```
             +------+ component link 1 +------+
             |      +-----------------+        |
             |      | component link 2 |       |
             |  N1  +-----------------+  N2   |
             |      | component link 3 |       |
             |      +-----------------+        |
             +------+                 +---+--+
```
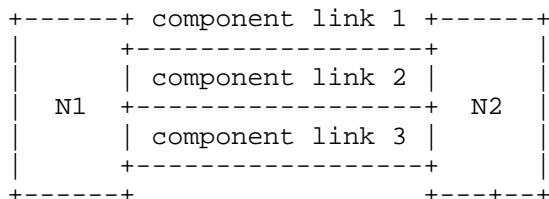
                Figure 7: Concurrent Path Computation

   Consider the case where a TE link is composed of three ODU3 component
   links with 32 TSs available on the first one, 24 TSs on the second,
   and 24 TSs on the third and is supporting ODU2 and ODU3 signal types.
   The node would advertise a TE link with 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 two ODU3s (64 TSs) in
   such a TE link while only a single ODU3 can be set up, and a
   crankback would be originated. In more complex network scenarios,
   the number of crankbacks can be much higher.

8. Maximum LSP Bandwidth

   Maximum LSP bandwidth is currently advertised per priority in the
   common part of the ISCD. Section 5 reviews some of the implications
   of advertising OTN information using ISCDs and identifies the need
   for a more optimized solution. While strictly not required, such an
   optimization effort should also consider the optimization of the per-
   priority maximum LSP bandwidth advertisement of both fixed and
   variable ODU types.

9. Distinction between Terminating and Switching Capabilities

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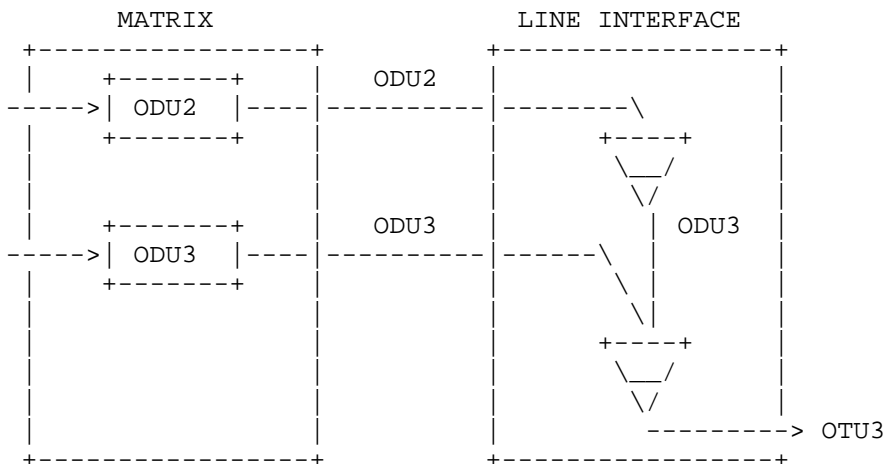   explain the switching and terminating capabilities.

```
              MATRIX                      LINE INTERFACE
        +----------------+          +----------------+
        |    +-------+   |   ODU2   |                |
   ----->| ODU2  |----|----------|--------\          |
        |    +-------+   |          |       +----+    |
        |                |          |       \___/     |
        |                |          |        \/       |
        |    +-------+   |   ODU3   |                 | ODU3  |
   ----->| ODU3  |----|----------|------\   |         |
        |    +-------+   |          |      \  |        |
        |                |          |       \ |        |
        |                |          |       \|        |
        |                |          |      +----+     |
        |                |          |      \___/      |
        |                |          |       \/        |
        |                |          |        ---------> OTU3
        +----------------+          +----------------+
```

           Figure 8: Switching and Terminating Capabilities

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

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

   o 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 capability to set up an LSP, allowing the
   local selection of resources to be consistent with the limitations
   considered during the path computation.

In Figure 9 and Figure 10, there are two examples of the terminating/
switching capability differentiation. In both examples, all nodes
only support single-stage capability. Figure 9 represents a scenario
in which a failure on link B-C forces node A to calculate another
ODU2 LSP carrying ODU0 service along the nodes B-E-D. As node D is a
single stage capable node, it is able to extract ODU0 service only
from the 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 the OTU3
link being considered in the path computation.

```
                    ODU0 Transparently Transported
         ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
         |               ODU2 LSP Carrying ODU0 Service             |
         |      |''''''''''''''''''''''''''''''''''''''''''|        |
         |      |              |           |               |        |
         |   +----++  OTU2  +-----+  OTU2  +-----+  OTU2  ++-----+  |
  ODU0   |   |    |  Link   |     |  Link  |     |  Link  |     |  | ODU0
  ---->| A |_____| B |_____| C |_____| D |---->
         |   |    |         |     |        |     |        |     |  |
         +-----+         +--+--+        +-----+        ++--+-+
                            |                             |  |
                        OTU3|                             |  |
                        Link|    +-----+_____|  |
                            |    |     |     OTU3 Link     |
                            |____|  E  |                   |
                            |    |     |_____|
                            +-----+      OTU2 Link
```
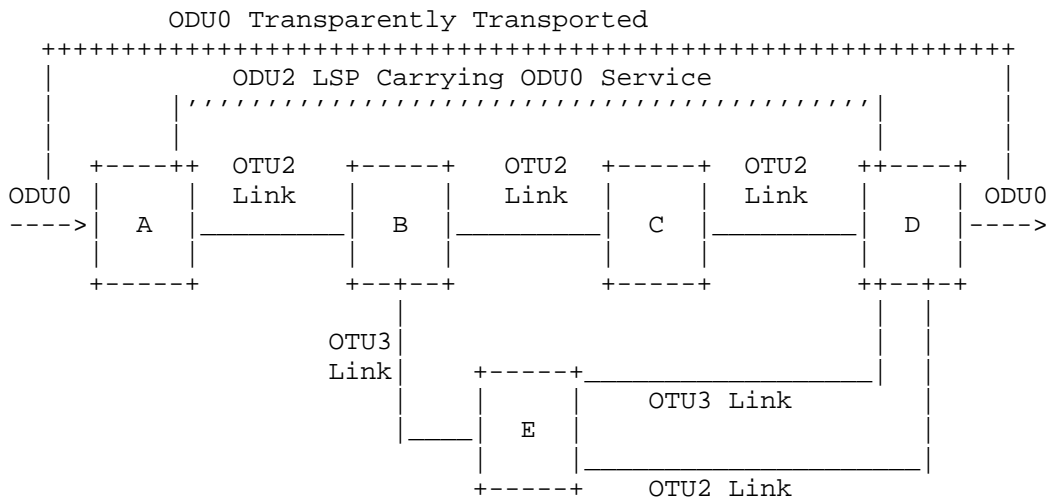
              Figure 9: Switching and Terminating Capabilities - Example 1

   Figure 10 addresses the scenario in which the restoration of the ODU2
   LSP (A-B-C-D) 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 (A-B-C-D). However, such ODU2
   unreserved bandwidth must be advertised since it can be used for a
   different ODU2 LSP terminating on E, e.g., F-B-E. Node A has to know
   that the ODU2 capability on the OTU2 link can only be terminated, and
   that the restoration of A-B-C-D can only be performed using the ODU2
   bandwidth available on the OTU3 link.

```
              ODU0 Transparently Transported
       ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
       |            ODU2 LSP Carrying ODU0 Service               |
       |         |'''''''''''''''''''''''''''''''''''''''''|      |
       |         |                                         |      |
       |  +----++  OTU2   +-----+  OTU2   +-----+  OTU2   ++-----+ |
  ODU0 |  |    |  Link   |     |  Link   |     |  Link   |    | ODU0
  ---->|  A |_____| B |_____| C |_____| D |---->
       |  |    |         |     |         |     |         |    |
       |  +-----+         ++-+-++         +-----+         +--+-+ |
                          | | |                           |
                     OTU2| | |                            |
          +-----+    Link| | |    OTU3    +-----+         |
          |     |        | | |    Link    |     |         |
          |  F  |_____| | |_____|  E  |         |
          |     |        | |_|            |     |_____|
          +-----+        |_____|  +-----+  OTU2 Link
                            OTU2 Link
```
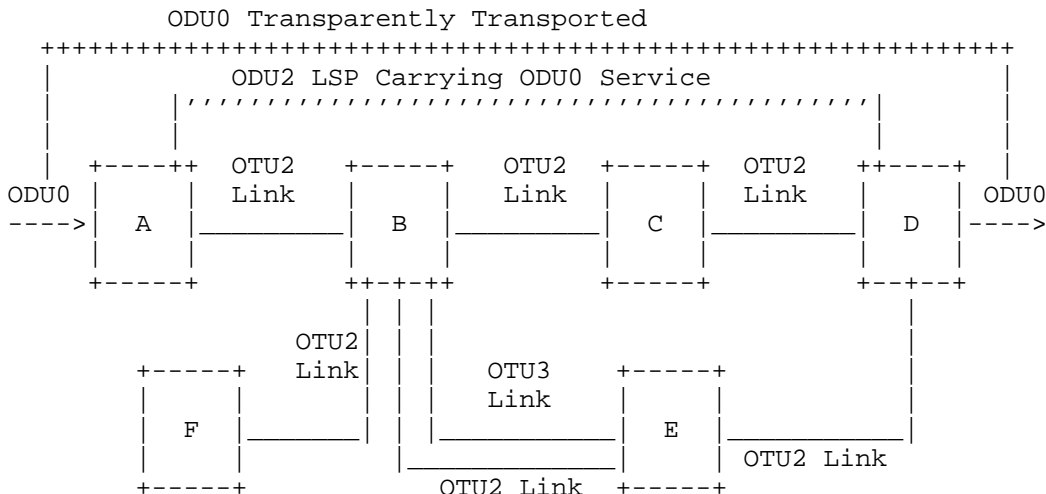
             Figure 10: Switching and Terminating Capabilities - Example 2

   The issue shown above is analyzed in an OTN context, but it is a
   general technology-independent GMPLS limitation.

10. Priority Support

   [RFC4202] defines eight priorities for resource availability and
   usage. As defined, each is advertised independent of the number of
   priorities supported by a network, and even unsupported priorities
   are included. As is the case in Section 8, addressing any
   inefficiency with such advertisements is not required to support
   OTNs. But, any such inefficiency should also be considered as part
   of the optimization effort identified in Section 5.

11. Multi-stage Multiplexing

   With reference to [RFC7062], the introduction of multi-stage
   multiplexing implies the advertisement of cascaded adaptation
   capabilities together with the matrix access constraints. The
   structure defined by the IETF for the advertisement of adaptation
   capabilities is the Interface Adaptation Capability Descriptor
   (IACD), as defined in [RFC6001].

   With respect to routing, please note that in case of multi-stage
   multiplexing 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 a
   spatial allocation problem.

In this scenario, suppose the following advertisement:

   Hierarchy: ODU1->ODU2->ODU3

   Number of ODU1==5

The number of ODU1 suggests that it is possible to have an ODU2 FA,
but it depends on the spatial allocation of such ODU1s.

It is possible that two links are bundled together and three
ODU1->ODU2->ODU3 are available on a component link and two on the
other one; in such a case, the ODU2 FA could not 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 (ODU2 FA
cannot be created), while in case of ODU1 spatial allocation (4+1),
the ODU2 available bandwidth would be 1 (1 ODU2 FA can be created).

The information stated above implies augmenting both the ISCD and the
IACD.

12. Generalized Label

The ODUk label format defined in [RFC4328] could be updated to
support new signal types as defined in [G.709-2012], but it would be
difficult to further enhance it to support possible new signal types.

Furthermore, such a 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 1.25 Gbit/s
tributary slots, it would require the utilization of 31 labels
(31*4*8=992 bits) to be allocated, while an ODUflex into an ODU4 may
need up to 80 labels (80*4*8=2560 bits).

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

13. Security Considerations

This document provides an evaluation of OTN requirements against
actual routing ([RFC4202], [RFC4203], and [RFC5307]) and signaling
mechanisms ([RFC3471], [RFC3473], and [RFC4328]) in GMPLS.

This document defines new types of information to be carried that
describes OTN containers and hierarchies. It does not define any new
protocol elements, and from a security standpoint, this memo does not
introduce further risks with respect to the information that can be
currently conveyed via GMPLS protocols. For a general discussion on
MPLS and GMPLS-related security issues, see the MPLS/GMPLS security
framework [RFC5920].

14. Contributors

   Jonathan Sadler
   Tellabs
   EMail: jonathan.sadler@tellabs.com

   John Drake
   Juniper
   EMail: jdrake@juniper.net

   Francesco Fondelli
   Ericsson
   Via Moruzzi 1
   Pisa - 56100
   EMail: francesco.fondelli@ericsson.com

15. Acknowledgements

   The authors would like to thank Lou Berger, Eve Varma, and Sergio
   Lanzone for their precious collaboration and review.

16. References

16.1. Normative References

   [G.709-2001] ITU-T, "Interfaces for the Optical Transport Network
                 (OTN)", G.709/Y.1331 Recommendation, February 2001.

   [G.709-2012] ITU-T, "Interfaces for the Optical Transport Network
                 (OTN)", G.709/Y.1331 Recommendation, February 2012.

   [G.798]   ITU-T, "Characteristics of Optical Transport Network
                 Hierarchy Equipment Functional Blocks", G.798
                 Recommendation, December 2012.

   [G.872]   ITU-T, "Architecture of Optical Transport Networks",
                 G.872 Recommendation, October 2012.

   [RFC1195]   Callon, R., "Use of OSI IS-IS for routing in TCP/IP and
                 dual environments", RFC 1195, December 1990.

   [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.

   [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.

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

   [RFC6001]   Papadimitriou, D., Vigoureux, M., Shiomoto, K.,
                   Brungard, D., and JL. Le Roux, "Generalized MPLS
                   (GMPLS) Protocol Extensions for Multi-Layer and
                   Multi-Region Networks (MLN/ MRN)", RFC 6001, October
                   2010.

16.2. Informative References

   [OTN-OSPF]   Ceccarelli, D., Ed., Zhang, F., Belotti, S., Rao, R.,
                   and J. Drake, "Traffic Engineering Extensions to OSPF
                   for Generalized MPLS (GMPLS) Control of Evolving G.709
                   OTN Networks", Work in Progress, December 2013.

   [OTN-RSVP]   Zhang, F., Ed., Zhang, G., Belotti, S., Ceccarelli, D.,
                   and K. Pithewan, "Generalized Multi-Protocol Label
                   Switching (GMPLS) Signaling Extensions for the evolving
                   G.709 Optical Transport Networks Control", Work in
                   Progress, September 2013.

   [RFC2328]   Moy, J., "OSPF Version 2", STD 54, RFC 2328, April
                   1998.

   [RFC4201]   Kompella, K., Rekhter, Y., and L. Berger, "Link
                   Bundling in MPLS Traffic Engineering (TE)", RFC 4201,
                   October 2005.

   [RFC5920]   Fang, L., "Security Framework for MPLS and GMPLS
                   Networks", RFC 5920, July 2010.

   [RFC7062]   Zhang, F., Li, D., Li, H., Belotti, S., and D.
                   Ceccarelli, "Framework for GMPLS and PCE Control of
                   G.709 Optical Transport Networks", RFC 7062, November
                   2013.

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
   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: +86-755-28973237
   EMail: danli@huawei.com