CCAMP Working Group S. Belotti, Ed.
Internet-Draft P. Grandi
Intended status: Informational Alcatel-Lucent
Expires: June 1, 2013 D. Ceccarelli, Ed.
D. Caviglia
Ericsson
F. Zhang
D. Li
Huawei Technologies
November 28, 2012

Evaluation of existing GMPLS encoding against G.709v3 Optical Transport Networks (OTN)

draft-ietf-ccamp-otn-g709-info-model-05

Abstract

The recent revision of ITU-T recommendation G.709 [G.709-2012] has introduced new fixed and flexible Optical Data Unit (ODU) containers in Optical Transport Networks (OTNs), enabling optimized support for an increasingly abundant service mix.

This document provides an evaluation of existing Generalized Multiprotocol Label Switching (GMPLS) routing and signaling methods against the G.709-2012 OTN networks.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

GMPLS[RFC3945] extends MPLS to include Layer-2 Switching (L2SC), Time-Division Multiplexing (e.g., SONET/SDH, PDH, and OTN), Wavelength (OCh, Lambdas) Switching and Spatial Switching (e.g., incoming port or fiber to outgoing port or fiber).

The establishment of Label Switched Paths (LSPs) that span only interfaces recognizing packet/cell boundaries is defined in [RFC3036, RFC3212, RFC3209]. [RFC3471] presents a functional description of the extensions to MPLS signaling required to support GMPLS. ReSource reserVation Protocol-Traffic Engineering (RSVP-TE) -specific formats, mechanisms and technology specific details are defined in [RFC3473].

From a routing perspective, Open Shortest Path First-Traffic Engineering (OSPF-TE) generates Link State Advertisements (LSAs) carrying application-specific information and floods them to other nodes as defined in [RFC5250]. Three types of opaque LSA are defined, i.e. type 9 - link-local flooding scope, type 10 - area-local flooding scope, type 11 - AS flooding scope.

Type 10 LSAs are composed of a standard LSA header and a payload including one top-level TLV and possible several nested sub-TLVs. [RFC3630] defines two top-level TLVs: Router Address TLV and Link TLV; and nine possible sub-TLVs for the Link TLV, used to carry link related TE information. The Link type sub-TLVs are enhanced by [RFC4203] in order to support GMPLS networks and related specific link information. In GMPLS networks each node generates TE LSAs to advertise its TE information and capabilities (link-specific or node-specific)through the network. The TE information carried in the LSAs are collected by the other nodes of the network and stored into their local Traffic Engineering Databases (TED).

In a GMPLS enabled G.709-2012 OTNs, routing and signaling are fundamental in order to allow automatic calculation and establishment of routes for ODUk LSPs. The recent revision of ITU-T Recommendation G.709 [G.709-2012] has introduced new fixed and flexible ODU containers that augment those specified in [RFC4328]. As a result, it is necessary to provide OSPF-TE and RSVP-TE extensions to allow GMPLS control of all currently defined ODU containers.

This document provides an evaluation of GMPLS signaling and routing processes against [G.709-2012] requirements.

OSPF-TE and RSVP-TE requirements are defined in [OTN-FWK], while protocol extensions are defined in [OTN-OSPF] and [OTN-RSVP].

2. G.709 Mapping and Multiplexing Capabilities

The digital OTN layered structure is comprised of digital path layer networks (ODU) and digital section layer networks (OTU). An OTU (Optical Transport Unit) section layer supports one ODU path layer as client and provides monitoring capability for the OCh. An ODU path layer may transport a heterogeneous assembly of ODU clients. Some types of ODUs (i.e., ODU1, ODU2, ODU3, ODU4) may assume either a client or server role within the context of a particular networking domain. ITU-T G.872 amendment 2 recommendation [G.872-a2] provides two tables defining mapping and multiplexing capabilities of OTNs, which are reproduced below.

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

Figure 1: OTN mapping capability

ODU client	ODU server
1.25 Gbps client	ODU 0
-	ODU 0
2.5 Gbps client	ODU 1
ODU 0	ODU 1
10 Gbps client	ODU 2
ODU0,ODU1,ODUflex	ODU 2
10.3125 Gbps client	ODU 2e
-	ODU 2e
40 Gbps client	ODU 3
ODU0,ODU1,ODU2,ODU2e,ODUflex	ODU 3
100 Gbps client	ODU 4
ODU0,ODU1,ODU2,ODU2e,ODU3,ODUflex	ODU 4
CBR clients from greater than 2.5 Gbit/s to 100 Gbit/s: or GFP-F mapped packet clients from 1.25 Gbit/s to 100 Gbit/s.	ODUflex
-	ODUflex

Figure 2: OTN multiplexing capability

How an ODUk connection service is transported within an operator network is governed by operator policy. For example, the ODUk connection service might be transported over an ODUk path over an OTUk section, with the path and section being at the same rate as that of the connection service (see Table 1). In this case, an entire lambda of capacity is consumed in transporting the ODUk connection service. On the other hand, the operator might exploit different multiplexing capabilities in the network to improve infrastructure efficiencies within any given networking domain. In

this case, ODUk multiplexing may be performed prior to transport over various rate ODU servers (as per Table 2) over associated OTU sections.

From the perspective of multiplexing relationships, a given ODUk may play different roles as it traverses various networking domains.

As detailed in [OTN-FWK], client ODUk connection services can be transported over:

o Case A) one or more wavelength sub-networks connected by optical links or

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

o Case C) a mix of ODU links and wavelength sub-networks.

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

The following sections list and analyze, for each type of data that needs to be advertised and signaled, what is already there in GMPLS and what is missing.

3. Tributary Slot Granularity

ITU-T recommendation defines two type of Tributary Slot (TS) granularity. This TS granularity is defined per layer, meaning that both ends of a link can select proper TS granularity differently for each supported layer, based on the rules below:

- If both ends of a link are new cards supporting both 1.25Gbps TS and 2.5Gbps TS, then the link will work with 1.25Gbps TS.

- If one end is a new card supporting both the 1.25Gbps and 2,5Gbps TS, and the other end is an old card supporting just the 2.5Gbps TS, the link will work with 2.5Gbps TS.

3.1. Data Plane Considerations

3.1.1. Payload Type and TSG relationship

As defined in G.709-2012 an ODUk container consist of an Optical Payload Unit (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 payload type (PT) is used to indicate the composition of the OPUk signal. When an ODUj signal is multiplexed into an ODUk, the ODUj signal is first extended with frame alignment overhead and then mapped into an Optical channel Data Tributary Unit (ODTU). Two different types of ODTU are defined:

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

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

G.709-2012 introduces also a logical entity, called Optical Data Tributary Unit Group (ODTUGk), characterizing the multiplexing of the various ODTU. The ODTUGk is then mapped into OPUK. ODTUjk and ODTUk.ts signals are directly time-division multiplexed into the tributary slots of an HO OPUk.

When PT is assuming value 20 or 21,together with OPUk type (K= 1,2,3,4), it is used to discriminate two different ODU multiplex structure ODTUGx :

- Value 20: supporting ODTUjk only,

- Value 21: supporting ODTUk.ts or ODTUk.ts and ODTUjk.

The discrimination is needed for OPUk with K =2 or 3, since OPU2 and OPU3 are able to support both the different ODU multiplex structures. For OPU4 and OPU1, only one type of ODTUG is supported: ODTUG4 with PT=21 and ODTUG1 with PT=20. (see table Figure 6).The relationship between PT and TS granularity, is in the fact that the two different ODTUGk discriminated by PT and OPUk are characterized by two different TS granularities of the related OPUk, the former at 2.5 Gbps, the latter at 1.25Gbps.

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

```
                   +--------+      +------------+
        +----+     |        !------| ODTUjk     |-----Client
        |    |     | ODTUGk |      +-----.------+
        |    |-----| PT=21  |            .
        |    |     |        |      +-----.------+
        |    |     |        |------| ODTUk.TS   |-----Client
        |OPUk|     +--------+      +------------+
        |    |
        |    |     +--------+      +------------+
        |    |     |        |------| ODTUjk     |-----Client
        |    |-----|        |      +-----.------+
        +----+     | ODTUGk |            .
                   | PT=20  |      +-----.------+
                   |        |------| ODTUjk     |-----Client
                   +--------+      +------------+
```

Figure 3: OTN client multiplexing

3.1.2. Fall-back procedure

ITU-T G.798 recommendation Amendment 2 [G.798-a2] describes the so called PT=21-to-PT=20 interworking process that explains how two equipments with interfaces with different PayloadType, and hence different TS granularity (1.25Gbps vs. 2.5Gbps), can be coordinated so to permit the equipment with 1.25 TS granularity to adapt his TS allocation accordingly to the different TS granularity (2.5Gbps) of a neighbor.

Therefore, in order to let the NE change TS granularity accordingly to the nieghbour requirements, the AUTOpayloadtype needs to be set. When both the neighbors (link or trail) have been configured as structured, the payload type received in the overhead is compared to the transmitted PT. If they are different and the transmitted PT=21, the node must fallback to PT=20. In this case the fall-back process makes the system self consistent and the only reason for signaling the TS granularity is to provide the correct label (i.e. label for PT=21 has twice the TS number of PT=20). On the other side, if the AUTOpayloadtype is not configured, the RSVP-TE consequent actions in case of TS mismatch need to be defined.

3.2. Control Plane considerations

When setting up an ODUj over an ODUk, it is possible to identify two types of TSG, the server and the client one. The server TSG is used to map an end to end ODUj onto a server ODUk LSP or links. This parameter can not be influenced in any way from the ODUj LSP: ODUj LSP will be mapped on tributary slots available on the different

links/ODUk LSPs. When setting up an ODUj at a given rate, the fact that it is carried over a path composed by links/Forwarding Adjacencies(FAs) structured with 1.25Gbps or 2.5Gbps TS size is completely transparent to the end to end ODUj.

On the other side the client TSG is the tributary slot size that is exported towards the client layer. The client TSG information is one of the parameters needed to correctly select the adaptation towards the client layers at the end nodes and this is the only thing that the ODUj has to guarantee.

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

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

         ... Service LSP
         --- H-LSP
```

Figure 4: Client-Server TSG example

In this scenario, an ODU3 LSP is setup from node B to Z. Node B has an old interface able to support 2.5 TSG granularity, hence only client TSG equal to 2.5Gbps can be exported to ODU3 H-LSP possible clients. An ODU2 LSP is setup from node A to node Z with client TSG 1.25 signaled and exported towards clients. The ODU2 LSP is carried by ODU3 H-LSP from B to Z. Due to the limitations of old node B interface, the ODU2 LSP is mapped with 2.5Gbps TSG over the ODU3 H-LSP. Then an ODU1 LSP is setup from A to Z, carried by the ODU2 H-LSP and mapped over it using a 1.25Gbps TSG.

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

The TSG 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 TSG. In case they cannot, A will compute an alternate path from itself to Z (see figure 4).

Moreover, also TSG information needs to be signaled. Consider as example the setup of an ODU3 path that is going to carry an ODU0, hence the support of 1,25 GBps TS is needed. The information related to the TSG has to be carried in the signaling to permit node C (see figure 5) chosose the right one among the different interfaces (with different TSGs) towards D. In case the full ERO is provided in the signaling with explicit interface declaration, there is no need for C to choose the right interface towards D as it has been already decided by the ingress node or by the PCE.

```
                             ODU3
                            ________________________

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

Figure 5: TSG in signaling

In case an ODUk FA_LSP needs to be set up nesting another ODUj (as depicted in figure 4), there might be the need to know the hierarchy of nested LSPs in addition to TSG, to permit the penultimate hop choosing the correct interface towards the egress node. This is not needed in case we allow bundling only component links with homogeneous hierarchies. In case of specific implementation not specifying in the ERO the last hop interface, crank-back can be a solution.

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

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

	2.5G TS		1.25G TS			
	OPU2	OPU3	OPU1	OPU2	OPU3	OPU4
ODU0	-	-	AMP PT=20	GMP PT=21	GMP PT=21	GMP PT=21
ODU1	AMP PT=20	AMP PT=20	-	AMP PT=21	AMP PT=21	GMP PT=21
ODU2	-	AMP PT=20	-	-	AMP PT=21	GMP PT=21
ODU2e	-	-	-	-	GMP PT=21	GMP PT=21
ODU3	-	-	-	-	-	GMP PT=21
ODUfl	-	-	-	GMP PT=21	GMP PT=21	GMP PT=21

Figure 6: ODUj into OPUk mapping types (Source: Table 7-10 [G.709-2012])

The signaled TSGs information is not enough to have a complete choice since the penultimate hop node has to distinguish between interfaces with the same TSG (e.g. 1.25Gbps) whether the interface is able to support the right hierarchy, i.e. it is possible to have two interfaces both at 1.25 TSG but only one is supporting ODU0.

A dedicated optional object could be defined in order to carry the multiplexing hierarchy and adaptation information (i.e. TSG/PT, AMP/GMP) so to have a more precise choice capability. In this way, when the penultimate node receives such object, together with the Traffic Parameters Object, is allowed to choose the correct interface towards the egress node.

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

4. Tributary Port Number

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

As described in [G.709-2012] and [G.798-a2], the OPUk overhead in an OTUk frame contains n (n = the total number of TSs of the ODUk) MSI (Multiplex Structure Identifier) bytes (in the form of multi-frame), each of which is used to indicate the association between tributary port number and tributary slot of the ODUk.

The association between TPN and TS has to be configured by the control plane and checked by the data plane on each side of the link. (Please refer to [OTN-FWK] for further details). As a consequence, the RSVP-TE signaling needs to be extended to support the TPN assignment function.

5. Signal type

From a routing perspective, [RFC4203] allows advertising [RFC4328] interfaces (single TS type) without the capability of providing precise information about bandwidth specific allocation. For example, in case of link bundling, dividing the unreserved bandwidth by the MAX LSP bandwidth it is not possible to know the exact number of LSPs at MAX LSP bandwidth size that can be set up. (see example fig. 3)

The lack of spatial allocation heavily impacts the restoration process, because the lack of information of free resources highly increases the number of crank-backs affecting network convergence time.

Moreover actual tools provided by [RFC4203] only allow advertising signal types with fixed bandwidth and implicit hierarchy (e.g. SDH/ SONET networks) or variable bandwidth with no hierarchy (e.g. packet switching networks) but do not provide the means for advertising networks with mixed approach (e.g. ODUflex CBR and ODUflex packet).

For example, advertising ODU0 as MIN LSP bandwidth and ODU4 as MAX LSP bandwidth it is not possible to state whether the advertised link supports ODU4 and ODUflex or ODU4, ODU3, ODU2, ODU1, ODU0 and ODUflex. Such ambiguity is not present in SDH networks where the hierarchy is implicit and flexible containers like ODUFlex do not exist. The issue could be resolved by declaring 1 ISCD for each signal type actually supported by the link.

Supposing for example to have an equivalent ODU2 unreserved bandwidth in a TE-link (with bundling capability) distributed on 4 ODU1, it would be advertised via the ISCD in this way:

MAX LSP Bw: ODU1

MIN LSP Bw: ODU1

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

- Unreserved Bandwidth (of the bundle) set to ODU2

In conclusion, the OSPF-TE extensions defined in [RFC4203] require a different ISCD per signal type in order to advertise each supported container. This motivates attempting to look for a more optimized solution, without proliferations of the number of ISCD advertised. Per [RFC2328], OSPF messages are directly encapsulated in IP datagrams and depend on IP fragmentation when transmitting packets larger than the network MTU. [RFC2328] recommends that "IP fragmentation should be avoided whenever possible." This recommendation further constraints solutions as OSPF does not support any generic mechanism to fragment OSPF LSAs.

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 singaling point of view, [RFC4328] describes GMPLS signaling extensions to support the control for pre-G.709-2012 OTNs. 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 Packet can have variable bit rates and tolerances (please refer to [OTN-FWK] table 2); it is thus needed to upgrade the signaling traffic patameters so to specify requested bit rates and tolerance values during LSP setup.

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 not to know how many LSPs of a specific signal type can be restored. As example it is possible to consider the scenario depicted in the following figure.

```
                  +------+ component link 1 +------+
                  |      +------------------+      |
                  |      | component link 2 |      |
                  |  N1  +------------------+  N2  |
                  |      | component link 3 |      |
                  |      +------------------+      |
                  +------+                  +---+--+
```

Figure 7: Concurrent path computation

Suppose to have a TE link comprising 3 ODU3 component links with 32TSs available on the first one, 24TSs on the second, 24TSs on the third and supporting ODU2 and ODU3 signal types. The node would advertise a TE link unreserved bandwidth equal to 80 TSs and a MAX LSP bandwidth equal to 32 TSs. In case of restoration the network could try to restore 2 ODU3 (64TSs) in such TE-link while only a single ODU3 can be set up and a crank-back would be originated. In more complex network scenarios the number of crank-backs can be much higher.

8. Maximum LSP Bandwidth

Maximum LSP bandwidth is currently advertised in the common part of the ISCD and advertised per priority, while in OTN networks it is only required for ODUflex advertising. This leads to a significant waste of bits inside each LSA.

9. Distinction between terminating and switching capability

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

```
             MATRIX                   LINE INTERFACE
       +-----------------+          +-----------------+
       |    +-------+    |   ODU2   |                 |
      ----->| ODU-2 |----|----------|--------\        |
       |    +-------+    |          |      +----+     |
       |                 |          |       \__/      |
       |                 |          |        \/       |
       |    +-------+    |   ODU3   |         | ODU3  |
      ----->| ODU-3 |----|----------|------\  |       |
       |    +-------+    |          |       \ |       |
       |                 |          |        \|       |
       |                 |          |      +----+     |
       |                 |          |       \__/      |
       |                 |          |        \/       |
       |                 |          |         ---------> OTU-3
       +-----------------+          +-----------------+
```

Figure 8: Switching and Terminating capabilities

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

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

- Map an ODU3 coming from the switching matrix into an OTU3

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

This piece of information needs to be advertised together with the related unreserved bandwidth and signal type. As a consequence signaling must have the possibility to setup an LSP allowing the local selection of resources consistent with the limitations considered during the path computation.

In figures 9 and 10 there are two examples of the need of termination/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 path 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 ODU2 interface. Node A has to know that from E to D exists an available OTU2 link from which node D can extract the ODU0 service. This information is required in order to avoid that the OTU3 link is considered in the path computation.

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

```
               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

10. Priority Support

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

11. Multi-stage multiplexing

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

With respect to the routing, please note that in case of multi stage muxing hierarchy (e.g. ODU1->ODU2->ODU3), not only the ODUk/OTUk bandwidth (ODU3) and service layer bandwidth (ODU1) are needed, but also the intermediate one (ODU2). This is a typical case of spatial allocation problem.

Suppose in this scenario to have the following advertisement:

Hierarchy: ODU1->ODU2->ODU3

Number of ODU1==5

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

It is possible that 2 links are bundled together and 3 ODU1->ODU2->ODU3 are available on a component link and 2 on the other one, in such a case no ODU2 FA could be set up. The advertisement of the ODU2 is needed because in case of ODU1 spatial allocation (3+2), the ODU2 available bandwidth would be 0 (no ODU2 FA can be created), while in case of ODU1 spatial allocation (4+1) the ODU2 available bandwidth would be 1 (1 ODU2 FA can be created).

12. Generalized Label

The ODUk label format defined in [RFC4328] could be updated to support new signal types defined in [G.709-2012] but would hardly be further enhanced to support possible new signal types.

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

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

13. Security Considerations

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

New types of information to be conveyed regard OTN containers and hierarchies 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. IANA Considerations

This informational document does not make any requests for IANA action.

15. 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

16. Acknowledgements

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

17. References

17.1. Normative References

[OTN-OSPF]
D.Ceccarelli,D.Caviglia,F.Zhang,D.Li,Y.Xu,P.Grandi,S.Belotti, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolutive G.709 OTN Networks", work in

progress draft-ietf-ccamp-gmpls-ospf-g709v3-04, November 2012.

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

[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.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

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

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

[RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.

[RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

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

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

[RFC5339] Le Roux, JL. and D. Papadimitriou, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 5339, September 2008.

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

17.2. Informative References

[G.709-2012]
ITU-T, "Rec G.709, version 4", approved by ITU-T in 2012.

[G.798-a2]
ITU-T, "Amendment 2 of G.798 Characteristics of optical transport network hierarchy equipment functional blocks", consented by ITU-T on April 2012.

[G.872-a2]
ITU-T, "Amendment 2 of G.872 Architecture of optical transport networks for consent", consented by ITU-T on June 2010.

[OTN-FWK] F.Zhang, D.Li, H.Li, S.Belotti, D.Ceccarelli, "Framework for GMPLS and PCE Control of G.709 Optical Transport Networks", work in progress draft-ietf-ccamp-gmpls-g709-framework-11, November 2012.

[OTN-RSVP]
F.Zhang, G.Zhang, S.Belotti, D.Ceccarelli, K.Pithewan, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for the evolving G.709 Optical Transport Networks Control, work in progress draft-ietf-ccamp-gmpls-signaling-g709v3-05", November 2012.

Authors' Addresses

Sergio Belotti (editor)
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

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

Email: pietro_vittorio.grandi@alcatel-lucent.com

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

Email: daniele.ceccarelli@ericsson.com

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

Email: diego.caviglia@ericsson.com

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

Email: zhangfatai@huawei.com

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

Email: danli@huawei.com