CCAMP Working Group
Internet Draft
Intended status: Informational
Expires: January 2017

Italo Busi
Huawei
Sergio Belotti
Nokia
Victor Lopez
Oscar Gonzalez de Dios
Telefonica
Anurag Sharma
Infinera
Yan Shi
China Unicom
Ricard Vilalta
CTTC
Karthik Sethuraman
NEC

July 7, 2016

Path Computation API
draft-busibel-ccamp-path-computation-api-00.txt

Status of this Memo

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

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

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

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

This Internet-Draft will expire on January 7, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Abstract

There are scenarios, typically in a hierarchical SDN context, in which an orchestrator may not have detailed information to be able to perform an end-to-end path computation and would need to request lower layer/domain controllers to calculate some (partial) feasible paths.

Multiple protocol solutions can be used for communication between different controller hierarchical levels. This document assumes that the controllers are communicating using YANG-based Application Programming Interface (APIs).

This document describes some use cases for an Application Programming Interface for path computation. A related yang model will be proposed in a next version or in another document.

Table of Contents

1. Introduction

There are scenarios, typically in a hierarchical SDN context, in which an orchestrator may not have detailed information to be able to perform an end-to-end path computation and would need to request lower layer/domain controllers to calculate some (partial) feasible paths.

Multiple protocol solutions can be used for communication between different controller hierarchical levels. This document assumes that the controllers are communicating using YANG-based Application Programming Interface (APIs).

Path Computation Elements, Controllers and Orchestrators perform their operations based on Traffic Engineering Databases (TED). Such TEDs can be described, in a technology agnostic way, with the YANG Data Model for TE Topologies [TE-TOPO]. Furthermore, the technology specific details of the TED are modeled in the augmented TE topology models (e.g. [L1-TOPO] for Layer-1 ODU technologies).

The availability of such topology models allows providing the TED via Netconf or Restconf API. Furthermore, it enables that a PCE/Controller performs the necessary abstractions or modifications and offer this customized topology to another PCE/Controller or high level orchestrator.

The tunnels that can be provided over the networks described with the topology models can be also set-up, deleted and modified via Netconf or Restconf API using the TE-Tunnel Yang model [TE-TUNNEL].

This document describes some use cases where a path computation function, also using Netconf or Restconf API, can be needed. A related yang model will be proposed in a next version or in another document.

2. Use Cases

This document presents different use cases, where an API for path computation is required. The presented uses cases have been grouped, depending on the different underlying topologies: a) IP-Optical integration; b) Multi-domain Optical Networks; and c) Data center interconnections.

2.1. IP-Optical integration

In these use cases, there is an Optical domain which is used to provide connectivity between IP routers which are connected with the Optical domains using access links (see Figure 1).

```
--------------------------------------------------------------------
I                                                                  I
I                                                                  I
I                                                                  I
I                      IP+Optical Use Cases                        I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                     (only in PDF version)                        I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
--------------------------------------------------------------------
```

Figure 1 - IP+Optical Use Cases

It is assumed that the Optical domain controller provides to the orchestrator an abstracted view of the Optical network. A possible abstraction shall be representing the optical domain as one "virtual node" with "virtual ports" connected to the access links.

The path computation request helps the orchestrator to know which are the real connections that can be provided at the optical domain.

```
------------------------------------------------------------------
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I               IP+Optical Topology Abstraction                  I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                     (only in PDF version)                      I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
------------------------------------------------------------------
```

Figure 2 - IP+Optical Topology Abstraction

2.1.1. Inter-layer path computation

In this use case the orchestrator needs to setup an optimal path between two IP routers R1 and R2.

As depicted in Figure 2, the Orchestrator has only an "abstracted view" of the physical network, and it does not know the feasibility or the cost of the possible optical paths (e.g., VP1-VP4 and VP2-VP5), which depend from the current status of the physical resources within the optical network and on vendor-specific optical attributes.

However, the orchestrator can ask the underlying Optical domain controller to compute a set of potential optimal paths, taking into account optical constraints Then, based on its own constraints, policy and knowledge (e.g. cost of the access links), it can choose which one of these potential paths to use to setup the optimal e2e path crossing optical network.

```
--------------------------------------------------------------------
I                                                                  I
I              IP+Optical Path Computation Example                 I
I                                                                  I
I                                                                  I
I                     (only in PDF version)                        I
I                                                                  I
--------------------------------------------------------------------
```

Figure 3 - IP+Optical Path Computation Example

For example, in Figure 3, the Orchestrator can request the Optical domain controller to compute the paths between VP1-VP4 and VP2-VP5 and then decide to setup the optimal end-to-end path which passes through the VP2-VP5 Optical path even this is not the optimal path from the Optical domain perspective.

An alternative approach could be to have the Optical domain controller making the information shown in Figure 3 available to the Orchestrator.

One possibility, under discussion within the TEAS WG, is to provide a "detailed connectivity matrix" which extends the "connectivity matrix" defined in [RFC7446] and describes not only the valid inbound-outbound TE link switching combinations, but also specifies a vector of various costs (in terms of delay, OSNR, intra-node SRLGs and summary TE metrics) a potential TE path associated with the connectivity matrix entry.

The information provided by the "detailed abstract connectivity matrix" would be equivalent to the information that should be provided by "virtual link model" as defined in [TE-INTERCONNECT].

In this case, the Path Computation Element (PCE) within the Orchestrator could use this information to calculate by its own the optimal path between routers R1 and R2, without requesting any additional information to the Optical Domain Controller.

However, there is a tradeoff between the accuracy (i.e., providing "all" the information that might be needed by the Orchestrator’s PCE) and scalability to be considered when designing the amount of information to provide within the "detailed abstract connectivity matrix".

Figure 4 below shows another example, similar to the one in Figure 3, but where there are two possible Optical paths between VP1 and VP4 with different properties (e.g., available bandwidth and cost).

```
--------------------------------------------------------------------
I                                                                  I
I             IP+Optical Path Computation Example                  I
I                    with multiple choices                         I
I                                                                  I
I                                                                  I
I                     (only in PDF version)                        I
I                                                                  I
--------------------------------------------------------------------
```

Figure 4 - IP+Optical Path Computation Example with multiple choices

Reporting all the information, as in Figure 4, using the "detailed abstract connectivity matrix" is quite challenging from a scalability perspective since the amount of this information is not just based on number of end points (which would scale as N-square), but also on many other parameters, including client rate, user constraints / policies for the service, e.g. max latency < N ms, max cost, etc., exclusion policies to route around busy links, min OSNR margin, max preFEC BER etc. All these constraints could be different based on connectivity requirements.

It is also worth noting that the "connectivity matrix" has been originally defined in WSON, [RFC7446] to report the connectivity constrains of a physical node within the WDM network: the information it contains is pretty "static" and therefore, once taken and stored in the TE data base, it can be always being considered valid and up-to-date in path computation request.

Using the "connectivity matrix" with an abstract node to abstract the information regarding the connectivity constraints of an Optical domain, would make this information more "dynamic" since the connectivity constraints of an Optical domain can change over time because some optical paths that are feasible at a given time may become unfeasible at a later time when e.g., another optical path is established. The information in the "detailed abstract connectivity matrix" is even more dynamic since the establishment of another optical path may change some of the parameters (e.g., delay or available bandwidth) in the "detailed abstract connectivity matrix" while not changing the feasibility of the path.

"Connectivity matrix" is sometimes confused with optical reach table that contain multiple (e.g. k-shortest) regen-free reachable paths for every A-Z node combination in the network. Optical reach tables can be calculated offline, utilizing vendor optical design and planning tools,and periodically uploaded to the Controller: these optical path reach tables are fairly static. However, to get the connectivity matrix, between any two sites, either a regen free path can be used, if one is available, or multiple regen free paths are concatenated to get from src to dest, which can be a very large combination. Additionally, when the optical path within optical domain needs to be computed, it can result in different paths based on input objective, constraints, and network conditions. In summary, even though "optical reachability table" is fairly static, which regen free paths to build the connectivity matrix between any source and destination is very dynamic, and is done using very sophisticated routing algorithms.

There is therefore the need to keep the information in the "connectivity matrix" updated which means that there another tradeoff between the accuracy (i.e., providing "all" the information that might be needed by the Orchestrator’s PCE) and having up-to-date information. The more the information is provided and the longer it takes to keep it up-to-date which increases the likelihood that the Orchestrator’s PCE computes paths using not updated information.

It seems therefore quite challenging to have a "detailed abstract connectivity matrix" that provides accurate, scalable and updated information to allow the Orchestrator’s PCE to take optimal decisions by its own.

If the information in the "detailed abstract connectivity matrix" is not complete/accurate, we can have the following drawbacks considering for example the case in Figure 4:

- o If only the VP1-VP4 path with available bandwidth of 2 Gb/s and cost 50 is reported, the Orchestrator's PCE will fail to compute a 5 Gb/s path between routers R1 and R2, although this would be feasible;

- o If only the VP1-VP4 path with available bandwidth of 10 Gb/s and cost 60 is reported, the Orchestrator's PCE will compute, as optimal, the 1 Gb/s path between R1 and R2 going through the VP2-VP5 path within the Optical domain while the optimal path would actually be the one going thought the VP1-VP4 sub-path (with cost 50) within the Optical domain.

Instead, using the approach proposed in this document, the Orchestrator, when it needs to setup an end-to-end path, it can request the Optical domain controller to compute a set of optimal paths (e.g., for VP1-VP4 and VP2-VP5) and take decisions based on the information received:

- o When setting up a 5 Gb/s path between routers R1 and R2, the Optical domain controller may report only the VP1-VP4 path as the only feasible path: the Orchestrator can successfully setup the end-to-end path passing though this Optical path;

- o When setting up a 1 Gb/s path between routers R1 and R2, the Optical domain controller (knowing that the path requires only 1 Gb/s) can report both the VP1-VP4 path, with cost 50, and the VP2-VP5 path, with cost 65. The Orchestrator can then compute the optimal path which is passing thought the VP1-VP4 sub-path (with cost 50) within the Optical domain.

Considering the dynamicity of the connectivity constraints of an Optical domain, it is possible that a path computed by the Optical domain controller when requested by the Orchestrator is no longer valid when the Orchestrator requests it to be setup up.

It is worth noting that with the approach proposed in this document, the likelihood for this issue to happen can be quite small since the time window between the path computation request and the path setup request should be quite short (especially if compared with the time that would be needed to update the information of a very detailed abstract connectivity matrix).

If this risk is still not acceptable, the Orchestrator may also optionally request the Optical domain controller not only to compute the path but also to keep track of its resources (e.g., these resources can be reserved to avoid being used by any other connection). In this case, some mechanism (e.g., a timeout) needs to be defined to avoid having stranded resources within the Optical domain.

These issues and solutions can be fine-tuned during the design of the Path Computation API.

2.1.2. Route Diverse IP Services

This is for further study.

2.2. Multi-domain Optical Networks

In this use case there are two optical domains which are interconnected together by multiple inter-domains links.

```
--------------------------------------------------------------------
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I             Multi-domain multi-link interconnection              I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                    (only in PDF version)                         I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
I                                                                  I
--------------------------------------------------------------------
```

Figure 5 - Multi-domain multi-link interconnection

In order to setup an end-to-end multi-domain Optical path (e.g., between nodes A and H), the orchestrator needs to know the feasibility or the cost of the possible optical paths within the two optical domains, which depend from the current status of the physical resources within each optical network and on vendor-specific optical attributes (which may be different in the two domains if they are provided by different vendors).

There is a trade-off between having the Orchestrator's PCE being able to take path computation decisions by its own versus having the Orchestrator being able to ask the Domain Controllers to provide a set of feasible optimal optical paths.

Orchestrator could want to select/optimize end-to-end path based on abstract topology information provided by the domain controllers. For example:

o Need to compute a path between A and H

o That path can go through inter-domain link C-E or through inter-domain link D-F

o Orchestrator's PCE, based on its own information, can compute the optimal multi-domain path being A-B-C-E-G-H

o But, during path setup, the domain controller may find out that A-B-C is not optically feasible, while only the path A-B-D is feasible

o So what the hierarchical controller computed is not good and need to re-start the path computation from scratch

As discussed in section 3.1, providing more extensive abstract information from the Optical domain controllers to the multi-domain Orchestator may lead to scalability problems.

Alternatively the Orchestrator can request the Optical domain controllers to compute a set of optimal paths and take decisions based on the information received. For example:

o Need to compute a path between A and H

o The Orchestrator asks Optical domain controllers to provide set of paths between A-C, A-D, E-H and F-H

o Optical domain controllers return a set of feasible paths with the associated costs: the path A-C would not be part of this set

- o The Orchestrator will select the path A-B-D-F-G-H since it is the only feasible path and then request the Optical domain controllers to setup the A-B-D and F-G-H paths

- o If there are multiple feasible paths, the Orchestrator can select the optimal path knowing the cost of the intra-domain paths (provided by the Optical domain controllers) and the cost of the inter-domain links (known by the Orchestrator)

In a sense this is similar to the problem of routing and wavelength assignment within an Optical domain. It is possible to do first routing (step 1) and then wavelength assignment (step 2), but the chances of ending up with a good path is low. Alternatively, it is possible to do combined routing and wavelength assignment, which is known to be a more optimal and effective way for Optical path setup. Similarly, it is possible to first compute an abstract end-to-end path within the multi-domain Orchestrator (step 1) and then compute an intra-domain path within each Optical domain (step 2), but there are more chances not to find a path or to get a suboptimal path that performing per-domain path computation and then stitch them.

The approach to request each Optical domain controllers to compute a set of optimal paths and take decisions based on the information received may still have some scalability issues when the number of Optical domains is quite big (e.g. 20).

In this case, it would be worthwhile combining the two approaches and use the abstract topology information provided by the domain controllers to limit the number of potential optimal end-to-end paths and then the Path Computation to decide what is the optimal path within this limited set.

```
  --------------------------------------------------------------------
  I                                                                  I
  I                                                                  I
  I                                                                  I
  I                Multi-domain with many domains                    I
  I                      (Topology information)                      I
  I                                                                  I
  I                                                                  I
  I                     (only in PDF version)                        I
  I                                                                  I
  I                                                                  I
  I                                                                  I
  --------------------------------------------------------------------
```

Figure 6 - Multi-domain with many domains (Topology information)

An example can be described considering multi-domain abstract topology shown in Figure 6. In this example an end-to-end Optical path between domains A and F needs to be setup. The transit domain should be selected between domains B, C, D and E.

The actual cost of each intra-domain path is not known a priori from the abstract topology information. The Orchestrator only knows the feasibility of some intra-domain paths and some upper-bound and/or lower-bound cost information. With this information, together with the cost of inter-domain links, the Orchestrator can decide that:

o Domain B cannot be selected as the path connecting domains A and E is not feasible;

o Domain E cannot be selected as a transit domain since it is know from the abstract topology information provided by domain controllers that the cost of the multi-domain path A-E-F (which is 100, in the best case) will be always be higher than the cost of the multi-domain paths A-D-F (which is 90, in the worst case) and A-E-F (which is 80, in the worst case)

Therefore, the Orchestrator can decide by its own that the optimal multi-domain path could be either A-D-F or A-E-F.

The Orchestrator can therefore request only the Optical domain controllers A, D, E and F to provide a set of optimal paths.

```
   --------------------------------------------------------------------
   I                                                                  I
   I                                                                  I
   I                                                                  I
   I                Multi-domain with many domains                    I
   I                (Path Computation information)                    I
   I                                                                  I
   I                                                                  I
   I                     (only in PDF version)                        I
   I                                                                  I
   I                                                                  I
   I                                                                  I
   --------------------------------------------------------------------
```

Figure 7 - Multi-domain with many domains (Path Computation information)

Based on these requests, the Orchestrator can know the actual cost of each intra-domain paths which belongs to potential optimal end-to-end paths, as shown in Figure 7, and then compute the optimal end-to-end path (e.g., A-D-F, having total cost of 50, instead of A-C-F having a total cost of 70).

2.3. Data center interconnections

In these use case, there is an Optical domain which is used to provide connectivity between data centers which are connected with the Optical domains using access links.

```
------------------------------------------------------------------
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                Data Center Interconnection Use Case            I
I                                                                I
I                                                                I
I                                                                I
I                     (only in PDF version)                      I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
I                                                                I
------------------------------------------------------------------
```

Figure 8 - Data Center Interconnection Use Case

In this use case, a virtual machine within Data Center 1 (DC1) needs to transfer data to another virtual machine that can reside either in DC2 or in DC3.

The optimal decision depends both on the cost of the optical path (DC1-DC2 or DC1-DC3) and of the computing power (data center resources) within DC2 or DC3.

The Cloud Orchestrator may not be able to make this decision because it has only an abstract view of the optical network (as in use case in 3.1).

The cloud orchestrator can request to the Optical domain controller to compute the cost of the possible optical paths (e.g., DC1-DC2 and DC1-DC3) and to the DC controller to compute the cost of the computing power (DC resources) within DC2 and DC3 and then it can take the decision about the optimal solution based on this information and its policy.

3. Security Considerations

This is for further study

4. IANA Considerations

This document requires no IANA actions.

5. References

5.1. Normative References

[RFC7446] Lee, Y. et al., "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", RFC 7446, February 2015.

5.2. Informative References

[TE-TOPO] Liu, X. et al., "YANG Data Model for TE Topologies", draft-ietf-teas-yang-te-topo, work in progress.

[L1-TOPO] Zhang, X. et al., "A YANG Data Model for Layer 1 (ODU) Network Topology", draft-zhang-ccamp-l1-topo-yang, work in progress.

[TE-TUNNEL] Xhang, X. et al., "A YANG Data Model for Traffic Engineering Tunnels and Interfaces", draft-ietf-teas-yang-te, work in progress.

[TE-INTERCONNECT] Farrel, A. et al., "Problem Statement and Architecture for Information Exchange Between Interconnected Traffic Engineered Networks", draft-ietf-teas-interconnected-te-info-exchange, work in progress.

6. Acknowledgments

The authors would like to thank Igor Bryskin and Xian Zhang for participating in discussions and providing valuable insights.

This document was prepared using 2-Word-v2.0.template.dot.

Contributors

Dieter Beller
Nokia
Email: dieter.beller@nokia.com

Authors' Addresses

Italo Busi
Huawei
Email: italo.busi@huawei.com

Sergio Belotti
Nokia
Email: sergio.belotti@nokia.com

Victor Lopez
Telefonica
Email: victor.lopezalvarez@telefonica.com

Oscar Gonzalez de Dios
Telefonica
Email: oscar.gonzalezdedios@telefonica.com

Anurag Sharma
Infinera
Email: AnSharma@infinera.com

Yan Shi
China Unicom
Email: shiyan49@chinaunicom.cn

Ricard Vilalta
CTTC
Email: ricard.vilalta@cttc.es

Karthik Sethuraman
NEC
Email: karthik.sethuraman@necam.com

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 8, 2017

I. Hussain
R. Valiveti
K. Pithewan
Infinera Corp
July 7, 2016

FlexE GMPLS Signaling Extensions
draft-hussain-ccamp-flexe-signaling-extensions-00

Abstract

This document describes GMPLS signaling extensions for configuring a FlexE group and adding or removing FlexE client(s) to a FlexE group [OIFFLEXE1].

Status of This Memo

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

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

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

This Internet-Draft will expire on January 8, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

This document describes GMPLS signaling extensions for configuring a FlexE group and adding or removing FlexE client(s) to a FlexE group [OIFFLEXE1]. The various usecases that arise when transporting Flexible Rate Ethernet signals in Optical transport networks are described in [FLEXEUSECASES]. The routing extensions in support of carrying link state information for a FlexE group are described in [FLEXEROUTING].

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

a. Ethernet PHY: an entity representing 100G-R Physical Coding Sublayer (PCS), Physical Media Attachment (PMA), and Physical Media Dependent (PMD) layers.

b. FlexE Group: a group of from 1 to 254 bonded Ethernet PHYs.

c. FlexE Client: an Ethernet flow based on a MAC data rate that may or may not correspond to any Ethernet PHY rate (e.g., 10, 40, m x 25 Gb/s).

d. FlexE Shim: the layer that maps or demaps the FlexE clients carried over a FlexE group.

e. FlexE Calendar: Representation of a FlexE group of n PHYs as a calendar of 20n slots logical length with 20 slots per PHY for scheduling of slots (i.e., a PHY bandwidth) among the FlexE clients.

3. Protocol Extensions

This section describes extensions to RSVP-TE signaling for GMPLS [RFC3473] to support FlexE.

3.1. Generalized Label

Figure 1 shows the proposed FlexE generalized label format to be carried in the Generalized Label Request [RFC3471]. This document proposes LSP Encoding type = Flexible Ethernet (FlexE) (a new value of 15 as defined in [FLEXEROUTING]), Switching type = Layer-2 Switch Capable (L2SC) (a value of 51 as defined in [RFC3471]) and Generalized PID (G-PID) = FlexE (a new value of 71 as defined in this document). A FlexE Group consists of 1 to n 100GBASE-R Ethernet PHYs. The label lists all PHY numbers (1 to 254) that are members of the FlexE group. For a client, the label also lists calendar slots in each member PHY that are assigned to the client.

```
 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
+---------------------------------------+-----------------------+
|         FlexE Group Number            |        Reserved       |
+---------------------------------------+-----------------------+
| Client (being added or removed) |  Flags                      |
+----------------+----------------+----------------+------------+
|  PHY Number    |      Rate      | Granularity    | Unav. Slots |
+----------------+----------------+-------------+-+-------------+
|  Slot Map (0 to 19 slot for 100G PHY)         |   Reserved    |
+-----------------------------------------------+---------------+
|                           ......                              |
+----------------+----------------+---------------+-------------+
|   PHY Number   |      Rate      |  Granularity  | Unav.  Slots|
+----------------+----------------+-------------+-+-------------+
|  Slot Map (0 to 19 slot for 100G PHY)         |   Reserved    |
+-----------------------------------------------+---------------+
|                           ......                              |
+---------------------------------------------------------------+
```

Figure 1: FlexE Generalized Label

FlexE Group Number (20 bits) fields allows to check that the correct PHY is being received from the correct group number [OIFFLEXE1].

Client (16 bits) field indicates which of the FlexE clients is mapped into a given calendar slot in the A and B calendar configurations for the sub-calendar carried over that PHY [OIFFLEXE1].

Flags (8 bits) field is reserved for future use. [OIFFLEXE1].

PHY Number (8 bits) field is used to identify PHY by a number in the 1-254 range [OIFFLEXE1].

Rate (8 bits) field is used to specify rate of the given PHY number. Currently [OIFFLEXE1] has defined a 100G PHY rate. In the future, this field may be used to indicate other PHY rates (e.g., 400G).

Granularity (8 bits) field is used to indicate granularity of the FlexE calendar. Currently [OIFFLEXE1] has defined 5G granularity. In future, this field can have additional values, as further granularity are defined.

Slot Map (20 bits) field is used to indicate which calendar slots of the associated PHY number is assigned to a given FlexE client. For a PHY with a rate of 100G and granularity of 5G, the slot map consists of 20 slots (0-19 range). In the future, when other PHY rates and/or calender granularities are defined, the slot map size for a PHY can be derived based on the Rate and Granularity fields values.

Unavailable Slots (8 bits) field is used to indicate the number of unavailable calendar slots (0-19 range) for example due to transport network constraints (i.e., no FlexE client should be assigned to the unused slots). Unavailable slots are placed at the end of each relevant sub-calendar (i.e., the highest numbered slots) [OIFFLEXE1].

3.2. FlexE Group Initial Setup

Suppose it is desired to establish a FlexE group containing two 100G PHYs between node A and B. This can be accomplished by having node A send a RSVP-TE message containing a FlexE generalized label to node B with the following field values:

a. FlexE Group Number = 100 (say), Client = 0x0000 (i.e., no client)

b. First PHY Number = 5 (say), Rate= 100G, Granularity=5G, Unavailable Slots=0, Slot Map = 0-19 bit set to 0 (i.e., all slots available)

c. Second PHY Number = 7 (say), Rate-100G, Granularity=5G, Unavailable Slots=0, Slot Map = 0-19 bit set to 0 (i.e., all slots available)

Thus both ends will have the same FlexE group configuration and the FlexE group can be brought in service.

3.3. FlexE Client Setup

Suppose it is desired to establish a FlexE client of rate 50G node A and B to the FlexE group created in the Section 3.2. This can be accomplished by having node A send a RSVP-TE message containing a FlexE generalized label to node B with the following field values:

a. FlexE Group Number = 100, Client = 0x0001 (i.e., client id = 1)

b. First PHY Number = 5 , Rate= 100G, Granularity=5G, Unavailable Slots=0, Slot Map = 0 to 4 bit set to 1 (i.e., 25G on this PHY)

c. Second PHY Number = 7, Rate-100G, Granularity=5G, Unavailable Slots=0, Slot Map = 0 to 4 bit set to 1 (i.e., 25G on this PHY)

3.4. Related Work

The generalized label described in [FLEXESIGNAL] is limited to 100G PHY only. In contrast, the generalized label proposed in this document is extendible to PHY rates beyond 100G. Specifically, the label proposed in this document introduces additional per PHY fields, namely, Rate and Granularity. This enables to drive per PHY calendar size information in the face of calendar granularity and/or calendar size changes that might be required for PHY rates beyond 100G (such as 400G).

4. Acknowledgements

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

None.

7. References

7.1. Normative References

[FLEXEROUTING]
IETF, "FlexE GMPLS Routing Extension, draft-pithewan-ccamp-flexe-routing-extensions", June 2016.

[FLEXESIGNAL]
IETF, "RSVP-TE Signaling Extensions in support of Flexible Ethernet networks, draft-wang-ccamp-flexe-signaling-00", March 2016.

[FLEXEUSECASES]
IETF, "FlexE Usecases, draft-hussain-ccamp-flexe-usecases", June 2016.

[OIFFLEXE1]
OIF, "FLex Ethernet Implementation Agreement Version 1.0 (OIF-FLEXE-01.0)", March 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3471] IETF, "G Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description, RFC3471", January 2003.

[RFC3473] IETF, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions, RFC3473", January 2003.

7.2. Informative References

[OIFMLG3] OIF, "Multi-Lane Gearbox Implementation Agreement Version 3.0 (OIF-MLG-3.0)", April 2016.

Appendix A. Additional Stuff

This becomes an Appendix.

Authors' Addresses

Iftekhar Hussain
Infinera Corp
169 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1-408-572-5200
Email: IHussain@infinera.com

Radha Valiveti
Infinera Corp
169 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1-408-572-5200
Email: rvaliveti@infinera.com

Khuzema Pithewan
Infinera Corp
169 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1-408-572-5200
Email: kpithewan@infinera.com

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 9, 2017

R. Kunze, Ed.
Deutsche Telekom
G. Grammel, Ed.
Juniper
D. Beller, Ed.
Nokia
G. Galimberti, Ed.
Cisco

July 8, 2016

A framework for Management and Control of DWDM optical interface parameters

draft-ietf-ccamp-dwdm-if-mng-ctrl-fwk-02

Abstract

To ensure an efficient data transport, meeting the requirements requested by today’s IP-services the control and management of DWDM interfaces is a precondition for enhanced multilayer networking and for an further automation of network provisioning and operation. This document describes use cases and requirements for the control and management of optical interfaces parameters according to different types of single channel DWDM interfaces. The focus is on automating the network provisioning process irrespective on how it is triggered i.e. by EMS, NMS or GMPLS. This document covers management as well as control plane considerations in different management cases of a single channel DWDM interface. The purpose is to identify the necessary information elements and processes to be used by control or management systems for further processing.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 9, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

The usage of the single channel DWDM interfaces in client nodes (e.g. routers) connected to a DWDM Network (which include ROADMs and optical amplifiers) adds a further networking option for operators opening to new scenarios and requiring more control/management plane integration.

Carriers deploy their networks today as a combination of transport and packet infrastructures to ensure high availability and flexible data transport. Both network technologies are usually managed by different operational units using different management concepts. This is the status quo in many carrier networks today. In the case of deployments, where the optical transport interface moves into the client device (e.g. , router), it is necessary to coordinate the management of the optical interface at the client domain with the optical transport domain. There are different levels of coordination, which are specified in this framework.

The objective of this document is to provide a framework that describes the solution space for the control and management of single channel interfaces and give use cases on how to manage the solutions. In particular, it examines topological elements and related network management measures. From an architectural point of view, the network can be considered as a set of pre- configured/qualified unidirectional, single-fiber, network connections between reference points S and R shown in figure 2. The optical transport network is managed and controlled in order to provide optical connections at the intended centre frequencies and the optical interfaces are managed and controlled to generate signals of the intended centre frequencies and further parameters as specified for example in ITU-T Recommendations G.698.2 and G.798. The management or control plane of the client and DWDM network be aware of the parameters of the interfaces to properly set up the optical link. This knowledge can be used furthermore, to support fast fault detection.

Optical routing and wavelength assignment based on WSON is out of scope although can benefit of the way the optical parameters are exchanged between the Client and the DWDM Network.

Additionally, the wavelength ordering process and the process how to determine the demand for a new wavelength from A to Z is out of scope.

Note that the Control and Management Planes are two separate entities that are handling the same information in different ways. This document covers management as well as control plane considerations in different management cases of single channel DWDM interfaces.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology and Definitions

Current generation WDM netwoks are single vendor networks where the optical line system and the transponders are tightly integrated. The DWDM interfaces migration from the Transponders to the Client interfaces changes this scenario, by introducing a standardized interface at the level of OCh between the Client DWDM interface and the DWDM network.

Black Link: The Black Link [ITU.G698.2] allows supporting an optical transmitter/receiver pair of a single vendor or from different vendors to provide a single optical channel interface and transport it over an optical network composed of amplifiers, filters, add-drop multiplexers which may be from a different vendor. Therefore the standard defines the ingress and egress parameters for the optical interfaces at the reference points Ss and Rs.

Single Channel DWDM Interface: The single channel interfaces to DWDM systems defined in G.698.2, which currently include the following features: channel frequency spacing: 50 GHz and wider (defined in [ITU-T G.694.1]); bit rate of single channel: Up to 10 Gbit/s. Future revisions are expected to include application codes for bit rates up to 40 Gb/s.

Forward error correction (FEC): FEC is a way of improving the performance of high-capacity optical transmission systems. Employing FEC in optical transmission systems yields system designs that can accept relatively large BER (much more than 10-12) in the optical transmission line (before decoding).

Administrative domain [G.805]: For the purposes of this Recommendation an administrative domain represents the extent of resources which belong to a single player such as a network operator, a service provider or an end-user. Administrative domains of different players do not overlap amongst themselves.

Intra-domain interface (IaDI) [G.872]: A physical interface within an administrative domain.

Inter-domain interface (IrDI) [G.872]: A physical interface that represents the boundary between two administrative domains.

Management Plane [G.8081]: The management plane performs management functions for the transport plane, the control plane and the system as a whole. It also provides coordination between all the planes. The following management functional areas are performed in the management plane: performance management; fault management; configuration management; accounting management and security management.

Control Plane[G.8081]: The control plane performs neighbour discovery, call control and connection control functions. Through signalling, the control plane sets up and releases connections, and may restore a connection in case of a failure. The control plane also performs other functions in support of call and connection control, such as neighbour discovery and routing information dissemination.

Transponder: A Transponder is a network element that performs O/E/O (Optical /Electrical/Optical) conversion. In this document it is referred only transponders with 3R (rather than 2R or 1R regeneration) as defined in [ITU.G.872].

Client DWDM interface: A Transceiver element that performs E/O (Electrical/Optical) conversion. In this document it is referred as the DWDM side of a transponder as defined in [ITU.G.872].

3. Solution Space Client DWDM interface

The management of optical interfaces using the Black Link approach deals with aspects related to the management of single-channel optical interface parameters of physical point-to-point and ring DWDM applications on single-mode optical fibres.

The solution allows the direct connection of a wide variety of equipments using a DWDM link, for example:

1. A digital cross-connect with multiple optical interfaces, supplied by a different vendor from the line system

2. Multiple optical client devices, each from a different vendor, supplying one channel each

3. A combination of the above

Table 1 provides a list of management tasks regarding the configuration of optical parameters.

Task	Domain	a	b	c	d
determination of centre frequency	optical	R	R	R	R
configuration of centre frequency at optical IF	client	NR	NR	R	R
path computation of wavelength	optical	NR	NR	R	R
routing of wavelength	optical	NR	NR	R	R
wavelength setup across optical network	optical	?	?	R	R
detection of wavelength fault	client	R	R	R	R
fault isolation, identification of root failure	optical	NR	R	R	R
repair actions within optical network	optical	R	R	R	R
protection switching of wavelength	optical	NR	NR	R	R
restoration of wavelength	optical	NR	NR	R	R

Note: R = relevant, NR = not relevant

Table 1: List of tasks related to Client - Network interconnection management

Furthermore the following deployment cases will be considered:

a. Passive WDM

b. P2P WDM systems

c. WDM systems with OADMs

d. Transparent optical networks supporting specific functions, nterfaces, protocols etc.

Case a) is added for illustration only, since passive WDM is specified in ITU-T Recommendations G.695 and G.698.1.

Case b) and case c)are motivated by the usage of legacy equipment using the traditional connection as described in Figure 1 DWDM interface integration on the client side.

3.1. Comparison of approaches for transverse compatibility

3.1.1. Multivendor DWDM line system with transponders

As illustrated in Figure 1, for this approach interoperability is achieved via the use of optical transponders providing OEO (allowing conversion to appropriate parameters). The optical interfaces can then be any short reach standardized optical interface that both vendors support, such as those found in [ITU-T G.957] [ITU-T G.691], [ITU-T G.693], [ITU-T G.959.1], etc.

```
                IrDI                              IaDI
                 |                                 |
                 .                                 .
                 |   +-----------------------------|----+
                 .   |      +    WDM Domain     +  .    |
                 |   |      |\                 /|  |    |
   +------+      .   |      | \     |\        / |  .    |      +------+
   |  TX/ |-->--+---+--T/-|OM|----|/-------|OD|--+-\T-+---->--| RX/  |
   |  RX  |--<--+---+--T/-|  |----- /|-----|  |--.-\T-+----<--| TX   |
   +------+      |   |      | /       \|      \ |  |    |      +------+
                 .   |      |/                 \|  .    |
                 |   |      +                   +  |    |
                 .   +-----------------------------.----+
                 |                                 |


         TX/RX = Single channel non-DWDM interfaces
         T/ = Transponder
         OM = Optical Mux
         OD = Optical Demux
```

Figure 1: Inter and Intra-Domain Interface Identification

In the scenario of Figure 1 the administrative domain is defined by the Interdomain Interface (IrDI). This interface terminates the DWDM domain. The line side is characterized by the IaDI. This interface specifies the internal parameter set of the optical administrative domain. In the case of a client DWDM interface deployment this interface moves into the client device and extends the optical and administrative domain towards the client node. ITU-T G.698.2 for example specifies the parameter set for a certain set of applications.

This document elaborates only the IaDI Interface as shown in Figure 1 as transversely compatible and multi-vendor interface within one administrative domain controlled by the network operator.

3.1.2. Integrated single channel DWDM deployments on the client site

In case of a deployment as shown in Figure 2, through the use of single channel DWDM interfaces, multi-vendor interconnection can also be achieved while removing the need for one short reach transmitter and receiver pair per channel (eliminating the transponders).

Figure 2 shows a set of reference points, for single-channel connection (Ss and Rs) between transmitters (Tx) and receivers (Rx). Here the DWDM network elements include an optical multiplexer (OM) and an optical demultiplexer (OD) (which are used as a pair with the peer element), one or more optical amplifiers and may also include one or more OADMs.

```
        |==================== Black Link =======================|


           +-------------------------------------------------+
        Ss |  |           DWDM Network Elements           |  | Rs
   +---+ | |  | \                                       / |  | | +---+
   Tx L1---|->|   \    +------+            +------+   /   |--|-->Rx L1
   +---+   |  |    |   |      |  +------+  |      |  |    |  |   +---+
   +---+   |  |    |   |      |  |      |  |      |  |    |  |   +---+
   Tx L2---|->| OM |-|>|------|->| OADM |--|------|->| OD |--|-->Rx L2
   +---+   |  |    |   |      |  |      |  |      |  |    |  |   +---+
   +---+   |  |    |   |      |  +------+  |      |  |    |  |   +---+
   Tx L3---|->|   /    | DWDM |    |  ^    | DWDM |   \   |--|-->Rx L3
   +---+   |  | /      | Link +----|--|----+ Link |     \ |  |   +---+
           +-----------+           |  |           +----------+
                                +--+  +--+
                                |        |
                                v        |
                              +---+    +---+
                               RxLx     TxLx
                              +---+    +---+

       Ss = Reference point at the DWDM network element tributary output
       Rs = Reference point at the DWDM network element tributary input
       Lx = Lambda x
       OM = Optical Mux
       OD = Optical Demux
       OADM = Optical Add Drop Mux
```

Linear DWDM network as per ITU-T G.698.2

Figure 2: Linear Black Link

As shown in Figure 2, the administrative domain may consists of several vendor domains. Even a in that case a common north bound management interface is required to ensure a consistent management of the entire connection.

The following documents[DWDM-interface-MIB], [YANG], [LMP] define such a protocol- FIX-THE-REFERENCE specific information using SNMP/ SMI, Yang models and LMP TLV to support the direct exchange of information between the client and the network control plane.

4. Solutions for managing and controlling the optical interface

Operation and management of WDM systems is traditionally seen as a homogenous group of tasks that could be carried out best when a single management system or an umbrella management system is used. Currently each WDM vendor provides an Element Management System (EMS) that also administers the wavelengths.

Therefore from the operational point of view the following approaches will be considered to manage and operate optical interfaces.

<vspace>:

1. Separate operation and management of client device and the transport network whereas the single channel interface of the client belongs to the administrative domain of the transport network and will be managed by the transport group. This results in two different approaches to send information to the management system

   a.Direct connection from the client to the management system,ensuring a management of the single channel of the optical network (e.g. EMS, NMS)

   b.Indirect connection to the management system of the optical network using a protocol (LMP) between the client device and the directly connected WDM system node to exchange management information with the optical domain

2. Common operation and management of client device including the single channel DWDM part and the Transport network

The first option keeps the status quo in large carrier networks as mentioned above. In that case it must be ensured that the full FCAPS Management (Fault, Configuration, Accounting, Performance and Security) capabilities are supported. This means from the management staff point of view nothing changes. The transceiver/receiver optical interface will be part of the optical management domain and will be managed from the transport management staff.

The second solution addresses the case where underlying WDM transport network is mainly used to interconnect a homogeneous set of client nodes (e.g. IP routers or digital crossconnects). Since the service creation and restoration could be done by the higher layers (e.g. IP), this may lead to an efficient network operation and a higher level of integration.

4.1. Separate Operation and Management Approaches

4.1.1. Direct connection to the management system

As depicted in Figure 3 (case 1a) one possibility to manage the optical interface within the client domain is a direct connection to the management system of the optical domain. This ensures manageability as usual.

```
                        +-----+
                        | NMS |
                        |_____|
                        /_____/
                           |
                           |
                           |
                       +---+---+
                +----->+  EMS  |
               /       |       |
              /        +-------+
             /             | MI SNMP or Yang
       SNMP / or Yang      |                DCN Network
       --------------------+--------------------------------
          /         +------+-----------------------+
         /          |     +|     WDM Domain   +    |
        /           |     |\                 /|    |
   +---+--+         |     | \     |\        / |    |          +------+
   |  CL  |-/C------+-----|OM|----|/-------|OD|----+-------/C-|  CL  |
   |      |-/C------+-----|  |----- /|-----|  |----+-------/C-|      |
   +------+         |     | /       \|      \ |    |          +------+
                    |     |/                 \|    |
                    |     +                   +    |
                    +------------------------------+
```

CL = Client Device
/C = Single Channel Optical Interface
OM = Optical Mux
OD = Optical Demux
EMS = Element Management System
MI= Management Interface

Figure 3: Connecting Single Channel optical interfaces to the Transport Management system

The exchange of management information between client device and the management system assumes that some form of a direct management communication link exists between the client device and the DWDM management system (e.g. EMS). This may be an Ethernet Link or a DCN connection (management communication channel MCC).

It must be ensured that the optical network interface can be managed in a standardised way to enable interoperable solutions between different optical interface vendors and vendors of the optical network management application. RFC 3591 [RFC3591] defines managed objects for the optical interface type but needs further extension to cover the optical parameters required by this framework document. Therefore an extension to this MIB for the optical interface has been drafted in [DWDM-interface-MIB]. SNMP is used to read parameters and get notifications and alarms, netconf and Yang models are needed to easily provision the interface with the right parameter set as described in [YANG]

Note that a software update of the optical interface components of the client nodes must not lead obligatory to an update of the software of the EMS and vice versa.

4.1.2. Direct connection to the DWDM management system

An alternative as shown in Figure 4 can be used in cases where a more integrated relationship between transport node (e.g. OM or OD) and client device is aspired. In that case a combination of control plane features and manual management will be used.

```
                        +-----+
                        | NMS |
                        |_____|
                        /_____/
                           |
                           |
                       +---+---+
                       | EMS   |
                       |       |
                       +-------+
                           | MI  SNMP or Yang
                           |
           LMP      +------+-----------------------+
       +------------+---+ +|                  +    |
       |            |   | |\                 /|    |
   +---+--+         |   +-+ \     |\        / |    |          +------+
   |  CL  |-/C------+--- -|OM|----|/-------|OD|--- +-------/C-|  CL  |
   |      |-/C------+--- -|  |----- /|-----|  |----+-------/C-|      |
   +------+         |     | /       \|      \ |    |          +------+
                    |     |/                 \|    |
                    |     +                   +    |
                    +------------------------------+

           CL = Client Device
           /C = Single Channel optical Interface
           OM = Optical Mux
           OD = Optical Demux
           EMS= Element Management System
           MI= Management Interface
```

Figure 4: Direct connection between peer node and first optical network node

For information exchange between the client node and the direct connected node of the optical transport network LMP as specified in RFC 4209 [RFC4209] should be used. This extension of LMP may be used between a peer node and an adjacent optical network node as depicted in Figure 4.

The LMP based on RFC 4209 does not yet support the transmission of configuration data (information). This functionality must be added to the existing extensions of the protocol. The use of LMP-WDM assumes that some form of a control channel exists between the client node and the WDM equipment. This may be a dedicated lambda, an Ethernet Link, or other signalling communication channel (SCC or IPCC).

4.2. Control Plane Considerations

The concept of integrated single channel DWDM interfaces equally applies to management and control plane mechanisms. The general GMPLS control plane for wavelength switched optical networks is work under definition in the scope of WSON. One important aspect of the BL is the fact that it includes the wavelength that is supported by the given link. Thus a BL can logically be considered as a fiber that is transparent only for a single wavelength. In other words, the wavelength becomes a characteristic of the link itself. Nevertheless the procedure to light up the fiber may vary depending on the implementation. Since the implementation is unknown a priori, different sequences to light up a wavelength need to be considered:

1. Interface first, interface tuning: The transmitter is switched on and the link is immediately transparent to its wavelength. This requires the transmitter to carefully tune power and frequency not overload the line system or to create transients.

2. Interface first, OLS tuning: The transmitter is switched on first and can immediately go to the max power allowed since the OLS performs the power tuning. This leads to an intermediate state where the receiver does not receive a valid signal while the transmitter is sending out one. Alarm suppression mechanisms shall be employed to overcome that condition.

3. OLS first, interface tuning: At first the OLS is tuned to be transparent for a given wavelength, then transponders need to be tuned up. Since the OLS in general requires the presence of a wavelength to fine-tune it is internal facilities there may be a period of time where a valid signal is transmitted but the receiver is unable to detect it. This equally need to be covered by alarm suppression mechanisms.

4. OLS first, OLS tuning: The OLS is programmed to be transparent for a given wavelength, then the interfaces need to be switched on and further power tuning takes place. The sequencing of enabling the link needs to be covered as well.

The preferred way to address these in a Control Plane enabled network is neighbour discovery including exchange of link characteristics and link property correlation. The general mechanisms are covered in RFC4209 [LMP-WDM] and RFC 4204[LMP] which provides the necessary protocol framework to exchange those characteristics between client and black link. LMP-WDM is not intended for exchanging routing or signalling information but covers:

1. Control channel management

2. Link property correlation

3. Link verification

4. Fault management

Extensions to LMP/LMP-WDM covering the code points of the BL definition are needed. Additionally when client and server side are managed by different operational entities, Link state exchange is required to align the management systems.

4.2.1. Considerations using GMPLS UNI

The deployment of single channel optical interfaces is leading to some functional changes related to the control plane models and has therefore some impact on the existing interfaces especially in the case of an overlay model where the edge node requests resources from the core node and the edges node do not participate in the routing protocol instance that runs among the core nodes. RFC 4208 [RFC4208] defines the GMPLS UNI that will be used between edge and core node. In case of integrated interfaces deployment additional functionalities are needed to setup a connection.

It is necessary to differentiate between topology/signalling information and configuration parameters that are needed to setup a wavelength path. RSVP-TE could be used for the signalling and the reservation of the wavelength path. But there are additional information needed before RSVP-TE can start the signalling process. There are three possibilities to proceed:

a. Using RSVP-TE only for the signalling and LMP as described above to exchange information to configure the optical interface within the edge node or

b. RSVP-TE will be used to transport additional information

c. Leaking IGP information instead of exchanging this information needed from the optical network to the edge node (overlay will be transformed to a border-peer model)

Furthermore following issues should be addressed:

a) The Communication between peering edge nodes using an out of band control channel. The two nodes have to exchange their optical capabilities. An extended version of LMP is needed to exchange FEC Modulation scheme, etc. that must be the same. It would be helpful to define some common profiles that will be supported. Only if the profiles match with both interface capabilities it is possible start signalling.

b) Due to the bidirectional wavelength path that must be setup it is obligatory that the upstream edge node inserts a wavelength value into the path message for the wavelength path towards the upstream node itself. But in the case of an overlay model the client device may not have full information which wavelength must/should be selectedand this information must be exchanged between the edge and the core node.

5. Use cases

A Comparison with the traditional operation scenarios provides an insight of similarities and distinctions in operation and management of single channel optical interfaces. The following use cases provide an overview about operation and maintenance processes.

5.1. Service Setup

It is necessary to differentiate between two operational issues for setting up a light path (a DWDM connection is specific in having defined maximum impairments) within an operational network. The first step is the preparation of the connection if no optical signal is applied. Therefore it is necessary to define the path of the connection.

The second step is to setup the connection between the client DWDM interface and the ROADM port. This is done using the NMS of the optical transport network. From the operation point of view the task is similar in a Black Link scenario and in a traditional WDM environment. The Black Link connection is measured by using BER tester which use optical interfaces according to G.698.2. These measurements are carried out in accordance with [ITU-TG.692]. When needed further connections for resilience are brought into service in the same way.

In addition some other parameters like the transmit optical power, the received optical power, the frequency, etc. must be considered.

If the optical interface moves into a client device some of changes from the operational point of view have to be considered. The centre frequency of the Optical Channel was determined by the setup process.

The optical interfaces at both terminals are set to the centre frequency before interconnected with the dedicated ports of the WDM network. Optical monitoring is activated in the WDM network after the terminals are interconnected with the dedicated ports in order to monitor the status of the connection. The monitor functions of the optical interfaces at the terminals are also activated in order to monitor the end to end connection.

Furthermore it should be possible to automate this last step. After connecting the client device towards the first control plane managed transport node a control connection may e.g. be automatically established using LMP to exchange configuration information.

If tunable interfaces are used in the scenario it would be possible to define a series of backup wavelength routes for restoration that could be tested and stored in backup profile. In fault cases this wavelength routes can be used to recover the service.

5.2. Link monitoring Use Cases

The use cases described below are assuming that power monitoring functions are available in the ingress and egress network element of the DWDM network, respectively. By performing link property correlation it would be beneficial to include the current transmit power value at reference point Ss and the current received power value at reference point Rs. For example if the Client transmitter power (OXC1) has a value of 0dBm and the ROADM interface measured power (at OLS1) is -6dBm the fiber patch cord connecting the two nodes may be pinched or the connectors are dirty. More, the interface characteristics can be used by the OLS network Control Plane in order to check the Optical Channels feasibility. Finally the OXC1 transceivers parameters (Application Code) can be shared with OXC2 using the LMP protocol to verify the transceivers compatibility. The actual route selection of a specific wavelength within the allowed set is outside the scope of LMP. In GMPLS, the parameter selection (e.g. central frequency) is performed by RSVP-TE.

G.698.2 defines a single channel optical interface for DWDM systems that allows interconnecting network-external optical transponders across a DWDM network. The optical transponders are considered to be external to the DWDM network. This so-called 'black link' approach illustrated in Figure 5-1 of G.698.2 and a copy of this figure is provided below. The single channel fiber link between the Ss/Rs reference points and the ingress/egress port of the network element on the domain boundary of the DWDM network (DWDM border NE) is called access link in this contribution. Based on the definition in G.698.2 it is considered to be part of the DWDM network. The access link typically is realized as a passive fiber link that has a specific optical attenuation (insertion loss). As the access link is an integral part of the DWDM network, it is desirable to monitor its attenuation. Therefore, it is useful to detect an increase of the access link attenuation, for example, when the access link fiber has been disconnected and reconnected (maintenance) and a bad patch panel connection (connector) resulted in a significantly higher access link attenuation (loss of signal in the extreme case of an open connector or a fiber cut). In the following section, two use cases are presented and discussed:

1) pure access link monitoring
2) access link monitoring with a power control loop

These use cases require a power monitor as described in G.697 (see section 6.1.2), that is capable to measure the optical power of the incoming or outgoing single channel signal. The use case where a power control loop is in place could even be used to compensate an increased attenuation as long as the optical transmitter can still be operated within its output power range defined by its application code.

   Figure 5 Access Link Power Monitoring

```
                                       +-------------------------+
                                       |    P(in) = P(Tx) - a(Tx) |
                                       |    ___                   |
       +----------+                    |    \ /   Power Monitor   |
       |          | P(Tx)              |     V    P(in)           |
       |  +----+  | Ss    //\\         |     |    |\              |
       |  | TX |----|-----\\//------------------->| \             |
       |  +----+  | Access Link (AL-T) |       .  |  |            |
       |          | attenuation a(Tx)  |       .  |  |==============>
       |          |                    |       .  |  |            |
       | External |                    |      --->| /             |
       | Optical  |                    |          |/              |
       |Transpond.|                    |    P(out)                |
       |          |                    |    ___                   |
       |          |                    |    \ /   Power Monitor   |
       |          | P(Rx)              |     V    P(out)          |
       |  +----+  | Rs    //\\         |     |    |\              |
       |  | RX |<---|-----\\//--------------------| \             |
       |  +----+  | Access Link (AL-R) |       .  |  |            |
       |          | Attenuation a(Rx)  |       .  |  |<==============
       +----------+                    |       .  |  |            |
                                       |      <---| /             |
       P(Rx) = P(out) - a(Rx)          |          |/              |
                                       |                          |
                                       |         ROADM            |
                                       +--------------------------+
```

- For AL-T monitoring: P(Tx) and a(Tx) must be known
- For AL-R monitoring: P(RX) and a(Rx) must be known

An alarm shall be raised if P(in) or P(Rx) drops below a configured threshold (t [dB]):

- P(in) < P(Tx) - a(Tx) - t (Tx direction)
- P(Rx) < P(out) - a(Rx) - t (Rx direction)
- a(Tx) =| a(Rx)

Figure 5: Extended LMP Model

5.2.1. Pure Access Link (AL) Monitoring Use Case

Figure 6 illustrates the access link monitoring use case and the different physical properties involved that are defined below:

- Ss, Rs: Single Channel reference points
- P(Tx): current optical output power of transmitter Tx
- a(Tx): access link attenuation in Tx direction (external transponder point of view)
- P(in): measured current optical input power at the input port of border DWDM NE
- t: user defined threshold (tolerance)
- P(out): measured current optical output power at the output port of border DWDM NE
- a(Rx): access link attenuation in Rx direction (external transponder point of view)
- P(Rx): current optical input power of receiver Rx

Description:
- The access link attenuation in both directions (a(Tx), a(Rx)) is known or can be determined as part of the commissioning process. Typically, both values are the same.
- A threshold value t has been configured by the operator. This should also be done during commissioning.
- A control plane protocol (e.g. this draft) is in place that allows to periodically send the optical power values P(Tx) and P(Rx) to the control plane protocol instance on the DWDM border NE. This is llustrated in Figure 3.
- The DWDM border NE is capable to periodically measure the optical power Pin and Pout as defined in G.697 by power monitoring points depicted as yellow triangles in the figures below.

Access Link monitoring process:
- Tx direction: the measured optical input power Pin is compared with the expected optical input power P(Tx) - a(Tx). If the measured optical input power P(in) drops below the value (P(Tx) - a(Tx) - t) a low power alarm shall be raised indicating that the access link attenuation has exceeded a(Tx) + t.
- Rx direction: the measured optical input power P(Rx) is compared with the expected optical input power P(out) - a(Rx). If the measured optical input power P(Rx) drops below the value (P(out) - a(Rx) - t) a low power alarm shall be raised indicating that the access link attenuation has exceeded a(Rx) + t.
- to avoid toggling errors, the low power alarm threshold shall be lower than the alarm clear threshold.

Figure 6 Use case 1: Access Link monitoring

```
     +----------+                    +---------------------------+
     | +------+ |    P(Tx), P(Rx)    |  +-------+                |
     | |      | | =================> |  |       |                |
     | | LMP  | |    P(in), P(out)   |  |  LMP  |                |
     | |      | | <================= |  |       |                |
     | +------+ |                    |  +-------+                |
     |          |                    |                           |
     |          |                    |    P(in) - P(Tx) - a(Tx)  |
     |          |                    |    ___                    |
     |          |                    |    \ /  Power Monitor     |
     |          | P(Tx)              |     V                     |
     |  +----+  | Ss    //\\         |     |     |\              |
     |  | TX |----|-----\\//------------------->| \             |
     |  +----+  | Access Link (AL-T) |       .   |  |            |
     |          | attenuation a(Tx)  |       .   |  |==============>
     |          |                    |       .   |  |            |
     | External |                    |      --->| /             |
     | Optical  |                    |           |/              |
     |Transpond.|                    |    P(out)                 |
     |          |                    |    ___                    |
     |          |                    |    \ /  Power Monitor     |
     |          | P(Rx)              |     V                     |
     |  +----+  | Rs    //\\         |     |     |\              |
     |  | RX |<---|-----\\//--------------------| \             |
     |  +----+  | Access Link (AL-R) |       .   |  |            |
     |          | Attenuation a(Rx)  |       .   |  |<==============
     +----------+                    |       .   |  |            |
                                     |      <---| /             |
     P(Rx) = P(out) - a(Rx)          |           |/              |
                                     |                           |
                                     |          ROADM            |
                                     +---------------------------+

      - For AL-T monitoring: P(Tx) and a(Tx) must be known
      - For AL-R monitoring: P(RX) and a(Rx) must be known
      An alarm shall be raised if P(in) or P(Rx) drops below a
      configured threshold  (t [dB]):
      -  P(in) < P(Tx) - a(Tx) - t (Tx direction)
      -  P(Rx) < P(out) - a(Rx) - t (Rx direction)
      -  a(Tx) = a(Rx)
```

Figure 6: Extended LMP Model

5.2.2. Power Control Loop Use Case

This use case is based on the access link monitoring use case as described above. In addition, the border NE is running a power control application that is capable to control the optical output power of the single channel tributary signal at the output port of the border DWDM NE (towards the external receiver Rx) and the optical output power of the single channel tributary signal at the external transmitter Tx within their known operating range. The time scale of this control loop is typically relatively slow (e.g. some 10s or minutes) because the access link attenuation is not expected to vary much over time (the attenuation only changes when re-cabling occurs).
From a data plane perspective, this use case does not require additional data plane extensions. It does only require a protocol extension in the control plane (e.g. this LMP draft) that allows the power control application residing in the DWDM border NE to modify the optical output power of the DWDM domain-external transmitter Tx within the range of the currently used application code. Figure 5 below illustrates this use case utilizing the LMP protocol with extensions defined in this draft.

Figure 7 Use case 2: Power Control Loop

```
     +----------+                       +---------------------------+
     | +------+ | P(Tx),P(Rx),Set(Pout) |  +-------+   +--------+  |
     | |      | | ====================> |  |       |   | Power  |  |
     | | LMP  | | P(in),P(out),Set(PTx) |  |  LMP  |   |Control |  |
     | |      | | <==================== |  |       |   | Loop   |  |
     | +------+ |                       |  +-------+   +--------+  |
     |    |     |                       |                           |
     | +------+ |                       |    P(in) = P(Tx) - a(Tx) |
     | |C.Loop| |                       |    ___                    |
     | +------+ |                       |    \ /  Power Monitor     |
     |    |     | P(Tx)                 |     V                     |
     | +------+ | Ss    //\\            |     |     |\              |
     | | TX |>----|-----\\//--------------------->| \             |
     | +------+ | Access Link (AL-T)    |       .  |  |            |
     |  VOA(Tx) | attenuation a(Tx)     |       .  |  |==============>
     |          |                       |       .  |  |            |
     | External |                       |      --->| /             |
     | Optical  |                       |          |/              |
     |Transpond.|                       |    P(out)                 |
     |          |                       |    ___                    |
     |          |                       |    \ /  Power Monitor     |
     |          | P(Rx)                 |     V                     |
     |  +----+  | Rs    //\\            |     |  VOA(out) |\        |
     |  | RX |<---|-----\\//--------------------<|-------| \       |
     |  +----+  | Access Link (AL-R)    |             .  |  |       |
     |          | attenuation a(Rx)     |             .  |  |<=======
     +----------+                       |        VOA(out) |  |      |
                                        |     <--<|-------| /       |
     P(Rx) = P(out) - a(Rx)             |               |/          |
                                        |                           |
                                        |          ROADM            |
                                        +---------------------------+
```

Figure 7: Power control loop

- The Power Control Loops in Transponder and ROADM controls the Variable Optical Attenuators (VOA) to adjust the proper power in base of the ROADM and Receiver caracteristics and the Access Link attenuation

6. Requirements

Even if network architectures becomes more complex the management and operation as well as the provisioning process should have a higher degree of automation or should be fully automated. Simplifying and automating the entire management and provisioning process of the network in combination with a higher link utilization and faster restoration times will be the major requirements that has been addressed in this section.

Data Plane interoperability as defined for example in [ITU.G698.2] is a precondition to ensure plain solutions and allow the usage of standardized interfaces between network and control/management plane.

The following requirements are focusing on the usage of standardised integrated single channel interfaces but also valid in other environments.

1 To ensure a lean management and provisioning process of single channel interfaces management and control plane of the client and DWDM network must be aware of the parameters of the interfaces and the optical network to properly setup the optical connection.

2 A standardized northbound API (to network management system)must be supported based on SNMP and Netconf.

3 A standardized data model for single channel interfaces must be supported to exchange optical parameters with control/ management plane.

4..Netconf should be used also for configuration of the single channel interfaces including the setting of the power.

5 LMP should be extended and used in cases where optical parameters need to be exchanged between peer nodes to correlate link characteristics and adopt the working mode of the single channel interface.

6 Legacy operational models should be supported (parameters must be exchanged with the DWDM transport EMS to manage the configuration and the transmission of alarms and other FCAPS messages.

7 LMP should be used to adjust the output power of the single channel DWDM interface to ensure that the interface works in the right range defined by the application code.

8 Parameters e.g. PRE-FEC BER could be used to trigger a FRR

mechanism on the IP control plane to reroute traffic before the link breaks.

9 LMP should be used to automate the end to end connection setup of the optical connection.

10 Power monitoring functions at both ends of the DWDM connection should be implemented to further automate the setup and shoutdown process of the optical interfaces.

11 A standardized procedure to setup an optical connection must be defined and implemented in DWDM and client devices (containing the single channel optical interface).LMP should be used to ensure that the process follows the right order.

12 Pre-tested and configured backup paths should be stored in so called backup profiles. In fault cases this wavelength routes can be used to recover the service.

13 LMP should be used to monitor and observe the access link.

7. Acknowledgements

The authors would like to thank all who supported the work with fruitful discussions and contributions.

8. IANA Considerations

This memo includes no request to IANA.

9. Security Considerations

The architecture and solution space in scope of this framework imposes no additional requirements to the security models already defined in RFC5920 for packet/optical networks using GMPLS, covering also Control Plane and Management interfaces. Respective security mechanisms of the components and protocols, e.g. LMP security models, can be applied unchanged.

As this framework is focusing on the single operator use case, the security concerns can be relaxed to a subset compared to a setup where information is exchanged between external parties and over external interfaces.

Concerning the access control to Management interfaces, security issues can be generally addressed by authentication techniques providing origin verification, integrity and confidentiality.

Additionally, access to Management interfaces can be physically or logically isolated, by configuring them to be only accessible out-of-band, through a system that is physically or logically separated from the rest of the network infrastructure. In case where management interfaces are accessible in-band at the client device or within the optical transport netork domain, filtering or firewalling techniques can be used to restrict unauthorized in-band traffic. Authentication techniques may be additionally used in all cases.

10. Contributors

Arnold Mattheus
Deutsche Telekom
Darmstadt
Germany
email arnold.Mattheus@telekom.de

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

Josef Roese
Deutsche Telekom
Darmstadt
Germany
email j.roese@telekom.de

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

11. References

11.1. Normative References

[RFC2863] McCloghrie, K. and F. Kastenholz, "The Interfaces Group MIB", RFC 2863, DOI 10.17487/RFC2863, June 2000, <http://www.rfc-editor.org/info/rfc2863>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, DOI 10.17487/RFC2578, April 1999, <http://www.rfc-editor.org/info/rfc2578>.

[RFC2579] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Textual Conventions for SMIv2", STD 58, RFC 2579, DOI 10.17487/RFC2579, April 1999, <http://www.rfc-editor.org/info/rfc2579>.

[RFC2580] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Conformance Statements for SMIv2", STD 58, RFC 2580, DOI 10.17487/RFC2580, April 1999, <http://www.rfc-editor.org/info/rfc2580>.

[RFC3591] Lam, H-K., Stewart, M., and A. Huynh, "Definitions of Managed Objects for the Optical Interface Type", RFC 3591, DOI 10.17487/RFC3591, September 2003, <http://www.rfc-editor.org/info/rfc3591>.

[RFC6205] Otani, T., Ed. and D. Li, Ed., "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, DOI 10.17487/RFC6205, March 2011, <http://www.rfc-editor.org/info/rfc6205>.

[RFC4209] Fredette, A., Ed. and J. Lang, Ed., "Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems", RFC 4209, DOI 10.17487/RFC4209, October 2005, <http://www.rfc-editor.org/info/rfc4209>.

[ITU.G698.2] International Telecommunications Union, "Amplified multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.

[ITU.G709] International Telecommunications Union, "Interface for the Optical Transport Network (OTN)", ITU-T Recommendation G.709, March 2003.

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

11.2. Informative References

[DWDM-interface-MIB]
Internet Engineering Task Force, "A SNMP MIB to manage the DWDM optical interface parameters of DWDM applications", draft-galimkunze-ccamp-dwdm-if-snmp-mib draft-galimkunze-ccamp-dwdm-if-snmp-mib, July 2011.

[ITU-TG.691]
ITU-T, "Optical interfaces for single channel STM-64 and other SDH systems with optical amplifiers", ITU-T Recommendation ITU-T G.691, 2008.

[ITU-TG.693]
ITU-T, "Optical interfaces for intra-office systems", ITU-T Recommendation ITU-T G.693, 2009.

[ITU-TG.957]
ITU-T, "Optical interfaces for equipments and systems relating to the synchronous digital hierarchy", ITU-T Recommendation ITU-T G.957, 2006.

[ITU-TG.692]
ITU-T, "Transmission media characteristics - Characteristics of optical components and sub-systems", ITU-T Recommendation ITU-T G.692, 1998.

[ITU-TG.959.1]
ITU-T, "Optical transport network physical layer interfaces", ITU-T Recommendation ITU-T G.959.1, 2009.

[ITU-TG.8081]
ITU-T, "Terms and definitions for Automatically Switched Optical Networks (ASON)", ITU-T Recommendation G.8081", ITU-T Recommendation ITU-T G.8081, September 2010.

Authors' Addresses

Ruediger Kunze (editor)
Deutsche Telekom
Winterfeldtstr. 21-27
10781 Berlin
Germany

Phone: +491702275321
Email: RKunze@telekom.de

Gert Grammel (editor)
Juniper
Oskar-Schlemmer Str. 15
80807 Muenchen
Germany

Phone: +49 1725186386
Email: ggrammel@juniper.net

Dieter Beller (editor)
Nokia
Lorenzstrasse, 10
70435 Stuttgart
Germany

Phone: +4971182143125
Email: Dieter.Beller@nokia.com

Gabriele Galimberti (editor)
Cisco
Via S. Maria Molgora, 48
20871 - Vimercate
Italy

Phone: +390392091462
Email: ggalimbe@cisco.com

CCAMP Working Group Y. Lee (Editor)
D. Dhody
X. Zhang
Internet Draft Huawei
Intended status: Standard Track A. Guo
Huawei
V. Lopez
Telefonica
D. King
U. of Lancaster
B. Yoon
ETRI

Expires: January 8, 2017 July 8, 2016

A Yang Data Model for WSON Optical Networks

draft-ietf-ccamp-wson-yang-02.txt

Abstract

This document provides a YANG data model for the routing and wavelength assignment (RWA) TE topology in wavelength switched optical networks (WSONs).

Status of this Memo

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

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

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

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

This Internet-Draft will expire on October 8, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

This document provides a YANG data model for the routing and wavelength assignment (RWA) Traffic Engineering (TE) topology in wavelength switched optical networks (WSONs). The YANG model described in this document is a WSON technology-specific Yang model based on the information model developed in [RFC7446] and the two encoding documents [RFC7581] and [RFC7579] that developed protocol independent encodings based on [RFC7446]. This document augments the generic TE topology draft [TE-TOPO].

What is not in scope of this document is both impairment-aware WSON and flex-grid.

2. YANG Model (Tree Structure)

```
module: ietf-wson-topology
augment /nd:networks/nd:network/nd:network-types/tet:te-topology:
   +--rw wson-topology!
augment /nd:networks/nd:network/nd:node/tet:te/tet:config/tet:te-node-attributes
/tet:connectivity-matrix:
   +--rw wson-matrix
      +--rw device-type?        devicetype
      +--rw dir?                directionality
      +--rw matrix-interface* [in-port-id]
         +--rw in-port-id     wson-interface-ref
         +--rw out-port-id?   wson-interface-ref
augment /nd:networks/nd:network/lnk:link/tet:te/tet:config:
   +--rw wavelength-available-bitmap*   boolean
augment /nd:networks/nd:network/nd:node/tet:te/tet:config:
   +--rw resource-pool* [resource-pool-id]
      +--rw resource-pool-id    uint32
      +--rw pool-state?         boolean
      +--rw matrix-interface* [in-port-id]
         +--rw in-port-id     wson-interface-ref
         +--rw out-port-id?   wson-interface-ref
```

3. WSON-RWA YANG Model

```
<CODE BEGINS> file "ietf-wson-topology@2016-07-08.yang"

module   ietf-wson-topology {
   namespace "urn:ietf:params:xml:ns:yang:ietf-wson-topology";

   prefix wson;

   import ietf-network {
   prefix "nd";
   }

   import ietf-network-topology {
   prefix "lnk";
   }

   import ietf-inet-types {
   prefix "inet";
   }

   import ietf-te-topology {
   prefix "tet";
   }

   organization
      "IETF CCAMP Working Group";

   contact
      "Editor:   Young Lee  <leeyoung@huawei.com>";

   description
      "This module contains a collection of YANG definitions for
      RWA WSON.

      Copyright (c) 2015 IETF Trust and the persons identified as
      authors of the code.  All rights reserved.
```

```
      Redistribution and use in source and binary forms, with or
      without modification, is permitted pursuant to, and subject
      to the license terms contained in, the Simplified BSD
      License set forth in Section 4.c of the IETF Trust's Legal
      Provisions Relating to IETF Documents
      (http://trustee.ietf.org/license-info).";

   revision 2015-10-14 {
      description
         "version 2.";

   reference
      "RFC XXX: A Yang Data Model for WSON Optical Networks ";
   }

   typedef wson-topology-id {
      type inet:uri;
      description
         "The WSON Topology ID";
   }

   typedef wson-node-id {
      type inet:ip-address;
      description
         "The WSON Node ID";
   }

   typedef devicetype {
      type enumeration {
         enum adm {
            value 1;
            description
               "Device is ADM";
         }

         enum roadm {
            value 2;
            description
               "Device is ROAMD/OXC";
         }
```

```
         }
         description
            "device type: fixed (ADM) or switched (ROADM/OXC)";
      }

      typedef directionality {
         type enumeration {
            enum bidir {
               value 0;
               description
                  "bi-directional";
            }
            enum input {
               value 1;
               description
                  "input direction";
            }
            enum output {
               value 2;
               description
                  "output direction";
            }
         }
         description
            "The directionality of link set";
      }



      typedef wson-interface-ref {
         type leafref {
         path "/nd:networks/nd:network/lnk:link/lnk:link-id";
         }
         description
            "This type is used by data models that need to
            reference WSON interface.";
      }

       grouping wson-topology-type {
          description
            "Identifies the wson topology type.";
```

```
          container wson-topology {
            presence "Indicates WSON topology.";
            description
              "Its presence identifies the WSON topology type.";
          }
        } // wson-topology-type

      augment "/nd:networks/nd:network/nd:network-types/tet:te-
   topology" {

         description "Augment network-types to include WSON topology";
         uses wson-topology-type {
            description
               "An empty WSON container to identify
               the topology type.";
         }
      }

      augment "/nd:networks/nd:network/nd:node/tet:te/tet:config"
            +"/tet:te-node-attributes"
            +"/tet:connectivity-matrix" {
         when "nd:network-types/tet:te-topology/wson-topology" {
            description
               "This augment is only valid for WSON.";
         }
         description "WSON Connectivity Matrix augmentation.";
         container wson-matrix{
            description "WSON specific Matrix.";
            leaf device-type {
               type devicetype;
               description
                  "device type: fixed (ADM) or switched
                  (ROADM/OXC)";
            }
            leaf dir {
               type directionality;
               description
                  "bi-directionality or input or output
                  of link set";
            }
```

```
            list matrix-interface {
               key "in-port-id";

               description
                  "matrix-interface describes input-ports
                  and out-ports around a connectivity
                  matrix";

               leaf in-port-id {
                  type wson-interface-ref;
                  description
                     "The reference to in-port";
               }

               leaf out-port-id {
                  type wson-interface-ref;
                  description
                     "The reference to out-port";
               }
            }
         }
      }

      augment "/nd:networks/nd:network/lnk:link/tet:te/tet:config" {

         when "nd:network-types/tet:te-topology/wson-topology" {

         description
            "This augment is only valid for WSON.";
         }
         description "WSON Link augmentation.";

         leaf-list wavelength-available-bitmap {
            type boolean;
            description
               "array of bits (i.e., bitmap) that indicates
               if a wavelength is available or not on each
               channel.";
         }
      }
```

```
      augment "/nd:networks/nd:network/nd:node/tet:te/tet:config" {
         when "nd:network-types/tet:te-topology/wson-topology" {
         description
            "This augment is only valid for WSON.";
         }
         description "WSON Node augmentation.";

         list resource-pool {
            key "resource-pool-id";
            description
               "The resource pool list";

            leaf resource-pool-id {
               type uint32;
               description
                  "The resource pool ID";
            }

            leaf pool-state {
               type boolean;
               description
                  "TRUE is state UP; FALSE is state down";
            }

            list matrix-interface {
               key "in-port-id";

               description
                  "pool is described as matrix-interface
                  with input-ports and output-ports
                  around the pool";

               leaf in-port-id {
                  type wson-interface-ref;
                  description
                     "The reference to in-interface";
               }

               leaf out-port-id {
                  type wson-interface-ref;
                  description
```

```
                    "The reference to out-interface";
              }
           }
        }
     }
  }

  <CODE ENDS>
```

4. Security Considerations

TDB

5. IANA Considerations

TDB

6. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

7. References

7.1. Normative References

[RFC7446] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", RFC 7446, Feburary 2015.

[RFC7579] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", RFC 7579, June 2015.

[RFC7581] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", RFC 7581, June 2015.

[TE-TOPO] X. Liu, et al., "YANG Data Model for TE Topologies", work in progress: draft-ietf-teas-yang-te-topo.

7.2. Informative References

8. Contributors

Authors' Addresses

Young Lee (ed.)
Huawei Technologies
5340 Legacy Drive, Building 3
Plano, TX 75023
USA

Phone: (469) 277-5838
Email: leeyoung@huawei.com

Dhruv Dhody
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area, Kundalahalli Village, Whitefield,
Bangalore - 560 037 [H1-2A-245]

Email: dhruv.dhody@huawei.com

Xian Zhang
Huawei Technologies

Email: zhang.xian@huawei.com

Aihua Guo
Huawei Technologies
Email: aihuaguo@huawei.com

Victor Lopez
Telefonica
Email: victor.lopezalvarez@telefonica.com

Daniel King
University of Lancaster
Email: d.king@lancaster.ac.uk

Bin Yeong Yoon
ETRI
218 Gaijeongro, Yuseong-gu
Daejeon, Korea
Email: byyun@etri.re.kr

CCAMP WG J. Ahlberg
Internet-Draft Ericsson AB
Intended status: Informational LM. Contreras
Expires: January 9, 2017 TID
A. Ye Min
Huawei
M. Vaupotic
Aviat Networks
J. Tantsura
Individual
K. Kawada
NEC Corporation
X. Li
NEC Laboratories Europe
I. Akiyoshi
NEC
July 8, 2016

A framework for Management and Control of microwave and millimeter wave interface parameters - problem statement
draft-mwdt-ccamp-problem-statement-00

Abstract

To ensure an efficient data transport, meeting the requirements requested by today’s transport services, the unification of control and management of microwave and millimeter wave radio link interfaces is a precondition for seamless multilayer networking and automated network wide provisioning and operation.

This document describes the required characteristics and use cases for control and management of radio link interface parameters using a YANG Data Model. It focuses on the benefits of a standardized management model that is aligned with how other packet technology interfaces in a microwave/millimeter wave node are modeled, the need to support core parameters and at the same time allow for optional product/feature specific parameters supporting new, unique innovative features until they have become mature enough to be included in the standardized model.

The purpose is to create a framework for identification of the necessary information elements and definition of a YANG Data Model for control and management of the radio link interfaces in a microwave/millimeter wave node.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 9, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

Network requirements vary between operators globally as well as within individual countries. The overall goal is however the same - to deliver the best possible network performance and quality of experience in a cost-efficient way.

Microwave/millimeter wave (hereafter referred to as microwave, but including the frequency bands represented by millimeter wave) are important technologies to fulfill this goal today, but also in the future when demands on capacity and packet features increases. Microwave is already today able to fully support the capacity needs of a backhaul in a radio access network and will evolve to support multiple gigabits in traditional frequency bands and beyond 10 gigabit in the millimeter wave. L2 packet features are normally an integrated part of microwave nodes and more advanced L2 and L3 features will over time be introduced to support the evolution of the transport services to be provided by a backhaul/transport network.

The main application for microwave is backhaul for mobile broadband. Those networks will continue to be modernized using a combination of microwave and fiber technologies. The choice of technology is a question about fiber presence and cost of ownership, not about capacity limitations in microwave. In 2020, more than 65% of all cell sites are expected to be connected with microwave solutions.

Open and standardized interfaces are a pre-requisite for efficient management of equipment from multiple vendors. This framework addresses management and control of the radio link interface(s) and the relationship to other packet interfaces, typically to Ethernet interfaces, in a microwave node. A radio link provides the transport over the air, using one or several carriers in aggregated or protected configurations. Managing and controlling a transport service over a microwave node involves both radio link and packet functionality.

Already today there are numerous IETF data models, RFCs and drafts, with technology specific extensions that cover a large part of the packet domain. Examples are IP Management [RFC7277], Routing Management [I-D.ietf-netmod-routing-cfg] and Provider Bridge [PB-YANG]. They are based on RFC 7223 [RFC7223], which is the IETF YANG model for Interface Management, and is an evolution of the SNMP IF-MIB [RFC 2863].

Since microwave nodes will contain more and more packet functionality and the interfaces for those technologies are expected to be managed using those models, there are advantages if radio link interfaces can be modeled and be managed using the same structure and the same approach, specifically for use cases in which a microwave node are managed as one common entity including both the radio link and the packet functionality, e.g. at installation, initial setup, trouble shooting, upgrade and maintenance. All interfaces in a node, irrespective of technology, would then be accessed from the same core model, i.e. RFC 7223, and could be extended with technology specific parameters in models augmenting that core model. The relationship/connectivity between interfaces would be given by the equipment configuration and slot positions or be configured via management systems or controllers.

```
+------------------------------------------------------------------+
| Interface [RFC7223]                                              |
|                +------------------+                              |
|                |Ethernet Port     |                              |
|                +------------------+                              |
|                      \                                           |
|                    +-----------------------+                     |
|                    |Radio link terminal    |                     |
|                    +-----------------------+                     |
|                               \                                  |
|                            +------------------------+            |
|                            |Carrier termination     |            |
|                            +------------------------+            |
+------------------------------------------------------------------+
```

Figure 1: Relationship between interfaces in a node

There will always be certain implementations that differ among products and it is therefore practically impossible to achieve industry consensus on every design detail. It is therefore important to focus on the parameters that are required to support the use cases applicable for centralized, unified, multi-vendor management and to allow other parameters to be optional or to be covered by extensions to the standardized model. Furthermore, a standard that allows for a certain degree of freedom encourages innovation and competition which is something that benefits the entire industry. It is therefore important that a radio link management model covers all relevant functions but also leaves room for product/feature-specific parameters.

For microwave radio link functionality work has been initiated (ONF: Microwave Modeling [ONF-model], IETF: Radio Link Model [I-D.ahlberg-ccamp-microwave-radio-link]). This effort is expected to take these initiatives into consideration and complement them on the gaps identified with the ambition to reach consensus within the industry around one common approach.

2. Conventions used in this document

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

3. Terminology and Definitions

Software Defined Networking (SDN) is an emerging architecture that decouples the network control and forwarding functions enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services. This results in an extremely dynamic, manageable, cost-effective, and adaptable architecture that gives administrators unprecedented programmability, automation, and control. The SDN concept is widely applied for network management, the adoption of SDN framework to manage and control the microwave and millimeter wave interface is one of the key applications of this work.

Microwave is a band of spectrum with wavelengths ranging from 1 meter to 1 millimeter and with frequencies ranging between 300 MHz and 300 GHz. Microwave radio technology is widely used for point-to-point telecommunications because of their small wavelength that allows conveniently-sized antennas to direct them in narrow beams, and their comparatively higher frequencies that allows broad bandwidth and high data transmission rates.

Millimeter wave is also known as extremely high frequency (EHF) or very high frequency (VHF) by the International Telecommunications Union (ITU), which can be used for high-speed wireless broadband communications. Millimeter wave is an undeveloped band of spectrum that can be used for a broad range of services on mobile and wireless networks including high-speed, point-to-point wireless local area networks (WLANs) and broadband access. This band has short wavelengths that range from 10 millimeters to 1 millimeter, namely millimeter band or millimeter wave. The 71 - 76 GHz, 81 - 86 GHz and 92-95 GHz bands are used for point-to-point high-bandwidth communication links, which allows for higher data rates up to 10 Gbit/s but requires a license. Unlicensed short-range data links can be used on 60 GHz millimeter wave. For instance, the upcoming IEEE Wi-Fi standard 802.11ad will run on the 60 GHz spectrum with data transfer rates of up to 7 Gbit/s.

Carrier Termination is an interface for the capacity provided over the air by a single carrier. It is typically defined by its transmitting and receiving frequencies.

Radio Link Terminal is an interface providing packet capacity and/or TDM capacity to the associated Ethernet and/or TDM interfaces in a node and used for the capacity required for setting up a transport service over a microwave/millimeter wave link.

4. Application

This framework addresses the definition of an open and standardized interface for the radio link functionality in a microwave/millimeter wave node. The application of such an interface used for management and control of nodes and networks typically vary from one operator to another, in terms of the systems used, what they are called and how they interact.

4.1. Network Management Solutions

The classic network management solutions, with vendor specific domain management combined with cross domain functionality for service management and analytics, still dominates the market.

These solutions are expected to evolve and benefit from an increased focus on standardization by simplifying multi-vendor management, remove the need for vendor/domain specific management, and enabling use of open source systems.

4.2. Software Defined Networking

SDN is another application emerging on the market. The main drivers for SDN introduction from an operator perspective is simplification and automation of network provisioning as well as E2E network services. The vision is to have a global view of the network conditions spanning across different vendors' equipment and multiple technologies. A variety of different SDN architectures and functions are being discussed and proposed within the industry, but they all have in common a very clear relationship to network management. In some proposals the SDN controller completely replaces the role of a network management system while in some cases the SDN controller is seen as an entity managed by the NMS. In any case the SDN functions shall be seen as part of an overall NMS strategy, no matter how the functionality is partitioned across units and no matter what terminology is used.

If nodes from different vendors shall be managed by the same SDN functions, without the extra effort of introducing mediation layers, all nodes must align their interfaces. Hence, open and standardized node interfaces are closely associated with the introduction of SDN.

4.2.1. SDN example of radio link configuration

Radio link terminals comprising a group of carriers are widely used in microwave technology. There are several kinds of groups: aggregation/bonding, 1+1 protection/redundancy, etc. To avoid configuration on each carrier termination, a logical control provides flexible management without configuring parameters on each carrier termination directly. An operator using SDN manages radio links by selecting an allowed operation mode. Alternatively, an operator can prefer the complete configuration of all the parameters of interest. Such case is not described in the document for simplicity). The operation mode is a set of logical metrics or parameters describing a complete radio link configuration, such as capacity, availability, priority and power consumption.

```
+------+       +-----+                        +-----+       +-------+
|      |-------| CT1 |   ~~~~~~~~~~~          | CT1 |-------|       |
|      |       +-----+                        +-----+       |       |
| RLT1 |                                                    | RLT2  |
|      |       +-----+                        +-----+       |       |
|      |-------| CT2 |   ~~~~~~~~~~~          | CT2 |-------|       |
+------+       +-----+                        +-----+       +-------+
```

Figure 2: Radio link example

Example of an operation mode table is shown Table 1. One mode (ID = 1) provides high capacity but requires high power consumption; another mode (ID = 2) provides low power consumption but low capacity. SDN controller selects one of the operation modes dynamically, according to the actual capacity need.

ID	Description	Capacity	Availability	Priority	Power
1	High capacity	400	99.9%	Low	High
2	High avail- ability	100	99.999%	High	Low

Figure 3: Example of an operation mode table

The procedure of logical control is as following:

1/ SDN controller gets the supported operation modes by reporting from the node NE or by other means, e.g., NMS.

2/ According to its operation policy, SDN controller selects one operation mode from the table and translates that into the required configuration of the individual parameters for the radio link terminals and the associated carrier terminations.

The operation mode could be selected based on power consumption.

Nowadays, more and more operators have strong requirement to decrease the node power consumption. The SDN controller can monitor the real time traffic distribution, and generated corresponding policy: to set the operation mode to high capacity on the radio links with heavy traffic load; to set the operation mode to low power consumption on the radio links with light traffic load

Radio link aggregation/bonding is widely used in microwave: multiple carriers are used to carry the traffic in cases where the needed capacity is beyond the capability of single carriers. During night time, the actual traffic load may be 1/3 of the peak day time. The operation mode of low power consumption could be set to turn off some of the radios within the group to save power consumption.

The operation mode could also be configured based on traffic priority.

For high priority traffic, it is necessary to assure high availability. On the other hand, low priority traffic is often high volume and requires high capacity. The SDN controller can generate corresponding policy based on the priority of traffic: to set the operation mode to high availability on the radio links with high priority traffic; to set the operation mode to high capacity on the radio links with low priority traffic.

Radio protection/redundancy like a 1+1 Hot Standby is used to increase overall availability of links between nodes while radio link aggregation is used to achieve higher capacity. If traffic is predominately TDM traffic, high availability is required on radio links. In this case, SDN controller set operation mode as high availability. On the other hand, if traffic is packet traffic, e.g., LTE traffic, capacity is more of a concern. In this case, SDN controller set the operation mode as high capacity.

5. Use cases

The use cases described should be the basis for identification and definition of the parameters to be supported by a YANG Data model for management of radio links, applicable for centralized, unified, multi-vendor management.

Use cases addressing installation, on-site trouble shooting, fault resolution and other activities not performed from a centralized system and which in most cases are product specific are outside the scope of this framework.

5.1. Configuration Management

Configuration of a radio link terminal, the constituent carrier terminations and when applicable the relationship to packet/Ethernet interfaces.

5.1.1. Understand the capabilities and limitations

Exchange of information between a manager and a device about the capabilities supported and specific limitations in the parameter values and enumerations that can be used.

Support for the XPIC feature or not and the maximum modulation supported are two examples on information that could be exchanged.

5.1.2. Initial Configuration

Initial configuration of a radio link terminal, enough to establish L1 connectivity over the hop to an associated radio link terminal on a device at far end. It MAY also include configuration of the relationship to a packet/Ethernet interface, unless that is given by the equipment configuration.

Frequency, modulation, coding and output power are examples of parameters typically configured for a carrier termination and type of aggregation/bonding or protection configurations expected for a radio link terminal.

5.1.3. Radio link re-configuration and optimization

Re-configuration, update or optimization of an existing radio link terminal. Output power and modulation for a carrier termination and protection schemas and activation/de-activation of carriers in a radio link terminal are examples on parameters that can be re-configured and used for optimization of the performance of a network.

5.2. Inventory

5.2.1. Retrieve logical inventory and configuration from device

Request from manager and response by device with information about radio interfaces, their constitution and configuration.

5.2.2. Retrieve logical inventory and configuration from device

Request from manager about physical and/or equipment inventory associated with the radio link terminals and carrier terminations

5.3. Status and statistics

5.3.1. Actual status and performance of a radio link interface

Manager requests and device responds with information about actual status and statistics of configured radio link interfaces and their constituent parts.

5.4. Performance management

5.4.1. Configuration of historical measurements to be performed

Configuration of historical measurements to be performed on a radio link interface and/or its constituent parts is a subset of the configuration use case to be supported. See 5.1 above.

5.4.2. Collection of historical performance data

Collection of historical performance data in bulk by the manager is a general use case for a device and not specific to a radio link interface.

Collection of an individual counter for a specific interval is in same cases required as a complement to the retrieval in bulk as described above.

5.5. Fault Management

5.5.1. Configuration of alarm reporting

Configuration of alarm reporting associated specifically with radio interfaces, e.g. configuration of alarm severity, is a subset of the configuration use case to be supported. See 5.1 above.

5.5.2. Alarm management

Alarm synchronization, visualization and handling, and notifications and events are generic use cases for a device and not specific to a radio link interface an should be supported accordingly.

5.5.3. Troubleshooting and Root Cause Analysis

Information and actions required by a manager/operator to investigate and understand the underlying issue to a problem in the performance and/or functionality of a radio link terminal and the associated carrier terminations.

6. Requirements

This is a placeholder for a separate chapter about requirements.

7. Security Considerations

TBD.

8. IANA Considerations

N/A.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, DOI 10.17487/RFC2234, November 1997, <http://www.rfc-editor.org/info/rfc2234>.

[RFC2863] McCloghrie, K. and F. Kastenholz, "The Interfaces Group MIB", RFC 2863, DOI 10.17487/RFC2863, June 2000, <http://www.rfc-editor.org/info/rfc2863>.

[RFC7223] Bjorklund, M., "A YANG Data Model for Interface Management", RFC 7223, DOI 10.17487/RFC7223, May 2014, <http://www.rfc-editor.org/info/rfc7223>.

9.2. Informative References

[I-D.ahlberg-ccamp-microwave-radio-link] Ahlberg, J., Carlson, J., Lund, H., Olausson, T., Ye, M., and M. Vaupotic, "Microwave Radio Link YANG Data Models", draft-ahlberg-ccamp-microwave-radio-link-01 (work in progress), May 2016.

[I-D.ietf-netmod-routing-cfg] Lhotka, L. and A. Lindem, "A YANG Data Model for Routing Management", draft-ietf-netmod-routing-cfg-22 (work in progress), July 2016.

[ONF-model]
"Microwave Modeling - ONF Wireless Transport Group", May 2016.

[PB-YANG] "IEEE 802.1X and 802.1Q YANG models, Marc,H.", October 2015.

[RFC7227] Hankins, D., Mrugalski, T., Siodelski, M., Jiang, S., and S. Krishnan, "Guidelines for Creating New DHCPv6 Options", BCP 187, RFC 7227, DOI 10.17487/RFC7227, May 2014, <http://www.rfc-editor.org/info/rfc7227>.

Authors' Addresses

Jonas Ahlberg
Ericsson AB
Lindholmspiren 11
Goeteborg 417 56
Sweden

Email: jonas.ahlberg@ericsson.com

Luis M. Contreras
Telefonica I+D
Ronda de la Comunicacion, S/N
Madrid 28050
Spain

Email: luismiguel.contrerasmurillo@telefonica.com

Ye Min (Amy)
Huawei Technologies CO., Ltd
No.1899, Xiyuan Avenue
Chengdu 611731
P.R.China

Email: amy.yemin@huawei.com

Marko Vaupotic
Aviat Networks
Motnica 9
Trzin-Ljubljana 1236
Slovenia

Email: Marko.Vaupotic@aviatnet.com

Jeff Tantsura
Individual

Email: jefftant.ietf@gmail.com

Koji Kawada
NEC Corporation
1753, Shimonumabe Nakahara-ku
Kawasaki, Kanagawa 211-8666
Japan

Email: k-kawada@ah.jp.nec.com

Xi Li
NEC Laboratories Europe
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Email: Xi.Li@neclab.eu

Ippei Akiyoshi
NEC
1753, Shimonumabe Nakahara-ku
Kawasaki, Kanagawa 211-8666
Japan

Email: i-akiyoshi@ah.jp.nec.com

K. Pithewan
I. Hussain
R. Valiveti
Infinera

Intended Status: Proposed Standard
Expires: January 7, 2017
July 6, 2016

FlexE GMPLS Routing Extension
draft-pithewan-ccamp-flexe-routing-extensions-00

Abstract

This document specifies routing extensions in support of carrying link state information for Generalized Multi-Protocol Label Switching (GMPLS) for Flex Ethernet.

Status of this Memo

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

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

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

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

Copyright and License Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1 Introduction

FlexE implementation agreement version 1.0 as specified by OIF supports flexible size Ethernet interfaces supported by one or more Ethernet PHY(s). FlexE interface represents the aggregate bandwidth of underlying PHY(s). In order to compute a path for a FlexE LSP that spans over one or more FlexE links, the bandwidth of FlexE interface needs to be advertised and flooded in routing domain. This document specifies the OSPF routing extension for FlexE interfaces and enables GMPLS control plane to flood FlexE link bandwidth in routing domain.

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 RFC 2119 [RFC2119].

2. FlexE TE Link

FlexE mechanism operates using calendar that assigns 66bit block positions to sub-calendars on each PHY of the FlexE group. The calendar has a granularity of 5G, and has a length of 20 slots per 100G of FlexE group capacity. Two calendar configurations are supported, referred as "A" and "B". At any given time, one of the calendar configurations is used for mapping the FlexE clients into the FlexE group. The two calendar configurations are provided to facilitate reconfiguration. To add or remove FlexE clients from the FlexE group, the clients are added to (or removed from) the inactive calendar and configuration is switched to make it effective. The effective bandwidth is represented by active calendar.

The effective bandwidth between two FlexE group (or interfaces) is represented by FlexE TE Link.

```
+-------+               +-------+               +-------+
|FlexE  |               |FlexE  |               |FlexE  |
|Switch |<- FlexE Link->|Switch |<- FlexE Link->|Switch |
|   A   |               |   B   |               |   C   |
+-------+               +-------+               +-------+

        |<-- TE-Link -->|       |<-- TE-Link -->|
```

Figure-1: FlexE TE-Link

In Figure-1 FlexE Switch A,B,C operates in FlexE aware mode (i.e., these switches terminate FlexE shim) and are capable of multiplexing and de-multiplexing FlexE group. In other words, the FlexE switches instantiate the FlexE Shim functions (as specified in the OIF IA) at each FlexE interface. The mux-demux capability of FlexE switch allows it to select time-slots for FlexE LSP at each hop. FlexE Link (TE-Link) exist between each FlexE aware switch.

```
+-------+ <------------ FlexE Link------------->+-------+
|FlexE  |               +-------+               |FlexE  |
|Switch |               |unaware|               |Switch |
|   A   |               |  (B)  |               |   C   |
+-------+               +-------+               +-------+

        |<-------------- TE-Link -------------->|
```

Figure-2: FlexE TE-Link

In Figure-2 A and C are FlexE aware switches and B is unaware. In this case, FlexE group is terminated by the switch A and C. Therefore, TE-link will also exist between A and C. BW modeled by TE-link and switch capability of FlexE aware switch is used for end-to-end path calculation for FleXE LSP.

3. Flex link property

FlexE group realized by the FlexE SHIM needs to have same number of components (i.e., grouped PHYs), number of slots and slot granularity on both ends.

The bandwidth of a TE Link is represented by number of slots and size of the slots. Theoretically, it is conceivable that a FlexE TE link could be constructed out of heterogeneous collection of Ethernet PHY(s), with different rates, which may result in a FlexE TE Link bandwidth realized by multiple sets of slots where each set may have different granularity/size. However, OIF agreement doesn’t support different granularity/size, hence, it is out of scope for this document.

4. OSPF TE-LSA Extension

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

4.1. Interface Switch Capability Descriptor

The Interface Switching Capability Descriptor (ISCD) describes switching capability of an interface [RFC 4202]. The Switching capability is essentially described by

- Interface Switching Capability
- Encoding
- Reservable Bandwidth

For FlexE interfaces this proposal uses L2 Switching Capability (L2SC) for Switching Type and proposes new encoding type

Value	Type
15	Flex Ethernet (FlexE)

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

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

```
      Switching Type :
                L2SC [as defined in RFC4202]
      Encoding Type  :
                FlexE [as defined in this document]
      Max LSP Bandwidth :
               The FlexE SHIM nominal rate (in kbps)
```

Reservable Bandwidth is advertised as part of Switch Capability Specific Information which is variable field in ISCD. Unused or available or reservable bandwidth is expressed as combination of [Number of Slots, Granularity]

As per OIF agreement Granularity has only one value i.e. 5G. In future additional values may be defined.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type=1             |           Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Granularity        |           Reserved                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 0                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 1                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 2                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 3                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 4                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 5                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 6                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Available Slots at priority 7                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Granularity : This is 8 bit field and takes values like Enum. Current OIF agreement only allow 5G granularity. In future, this field can have more values, as further granularity are defined.

Value	- Granularity
1	- 5G

Reserved : This field is reserved and marked as 0.

Available Slots : This field (32 bits) indicates the maximum number of slots available at priority 'p' on active calendar of the TE Link.

3 Security Considerations

TBD

4 IANA Considerations

New Encoding type is defined in this document for Flex Ethernet. Proposed value is 15

5 References

5.1 Normative References

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

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

[OIFFLEXE1] OIF, "FLex Ethernet Implementation Agreement Version 1.0 (OIF-FLEXE-01.0)", March 2016.

5.2 Informative References

Authors' Addresses

Khuzema Pithewan
169 W Java Dr, Sunnyvale, CA, USA
EMail: kpithewan@infinera.com

Iftekhar Hussain
169 W Java Dr, Sunnyvale, CA, USA
EMail: ihussain@infinera.com

Radha Valiveti
169 W Java Dr, Sunnyvale, CA, USA
EMail: rvaliveti@infinera.com

CCAMP Working Group A. Sharma
Internet-Draft R. Rao
Intended status: Standards Track Infinera Corp
Expires: January 7, 2017 X. Zhang
Huawei Technologies
July 6, 2016

OTN Service YANG Model
draft-sharma-ccamp-otn-service-model-00

Abstract

This document describes the YANG data model for OTN Services.

Status of this Memo

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

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

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

This Internet-Draft will expire on January 7, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

OTN transport networks can carry various types of client services. In many cases, the client service is an OTN service across connected domains in a multi-domain network. These OTN services can either be transported or switched in the OTN network. If an OTN service is switched then additional parameters need to be provided to create a Mux OTN service.

This document provides YANG model for creating OTN service. The model augments the TE Tunnel model, which is an abstract model to create TE Tunnels.

2. Model Overview

This section provides an overview of the OTN Service Model.

2.1. OTN Mux Service

```
                              OTN Mux Service
                  <--------------------------------->

  XXXXXXXXXXX          XXXXXXXXXXXXXXXXXXXXXXXXX         XXXXXXXXX
 XX         XXX       XX                       XXX      XX        XX
XX   +---+    X      XX +---+          +---+    X     XX  +---+   XX
X    |NE1+--------+-----+NE2+-----------+NE3+--------+-X---+NE4|    X
X    +---+   XX   ^  X  +---+          +---+   XX   ^ XX  +---+    X
 XX         XX    |   XX                       XX   |  XX        XXX
   XXXXXXXXXX     |     XXXXXXXXXXXXXXXXXXXXXXXX    |   XXXXXXXXXX
    Domain-1      |            Domain-2             |    Domain-3
                  +                                 +
         Same OTN Service attributes:      Same OTN Service attributes:
         1. Client Signal                  1. Client Signal
         2. Tributary Port Number          2. Tributary Port Number
         3. Tributary Slot Granularity     3. Tributary Slot Granularity
         4. Timeslots                      4. Timeslots
```

Figure 1: OTN Mux Service in a multi-domain network topology

Figure 1 shows a multi-domain OTN network with three domains. In this example, user wants to setup an end-to-end OTN service that passes through Domain-2. In order to create an OTN mux service in Domain-2, user will need to specify the exact details of the client side LO-ODU on NE2 and NE3, so that these service endpoints can be paired with the LO-ODU endpoints on NE1 and NE4, respectively.

Let’s assume that ODU4 is the client side HO-ODU on NE2 and NE3, and the client signal is ODU2. User will need to specify the OTN client signal (ODU2 in this example), the Tributary Port Number (TPN), Tributary Slot Granularities (TSG) and timeslots to be used. As shown in the figure above, these service parameters must be the same between NE1 and NE2, and NE3 and NE4.

Once the OTN Mux service is setup in Domain-2, the incoming signal from either NE1 and/or NE4 will be switched inside Domain-2, and delivered to NE at the other end.

2.2. Model Tree

```
module: ietf-otn-service
augment /te:te/te:tunnels/te:tunnel/te:config:
   +--rw payload-treatment?    enumeration
   +--rw src-client-signal?    identityref
   +--rw src-tpn?              uint16
   +--rw src-tsg?              identityref
   +--rw src-timeslot-count?   uint16
   +--rw src-timeslots
   |  +--rw values*   uint8
   +--rw dst-client-signal?    identityref
   +--rw dst-tpn?              uint16
   +--rw dst-tsg?              identityref
   +--rw dst-timeslot-count?   uint16
   +--rw dst-timeslots
      +--rw values*   uint8
augment /te:te/te:tunnels/te:tunnel/te:state:
   +--ro payload-treatment?    enumeration
   +--ro src-client-signal?    identityref
   +--ro src-tpn?              uint16
   +--ro src-tsg?              identityref
   +--ro src-timeslot-count?   uint16
   +--ro src-timeslots
   |  +--ro values*   uint8
   +--ro dst-client-signal?    identityref
   +--ro dst-tpn?              uint16
   +--ro dst-tsg?              identityref
   +--ro dst-timeslot-count?   uint16
   +--ro dst-timeslots
      +--ro values*   uint8
```

2.3. OTN Service YANG Model

```
<CODE BEGINS> file "ietf-otn-service@2016-06-24.yang"

module ietf-otn-service {
```

```
    yang-version 1;
    namespace "urn:ietf:params:xml:ns:yang:ietf-otn-service";
    prefix "otn-svc";

    import ietf-te { prefix "te"; }
    import ietf-transport-types { prefix "tran-types"; }
    import yang-ext { prefix ext; revision-date 2013-07-09; }

    organization
        "IETF CCAMP Working Group";

    contact
        "WG Web: <http://tools.ietf.org/wg/ccamp/>
        WG List: <mailto:ccamp@ietf.org>

        Editor: Anurag Sharma
                <mailto:AnSharma@infinera.com>

        Editor: Rajan Rao
                <mailto:rrao@infinera.com>

        Editor: Xian Zhang
                <mailto:zhang.xian@huawei.com>";

    description
        "This module defines a model for OTN Services.";

    revision "2016-06-24" {
        description "Initial revision";
        reference "TBD";
    }

    grouping otn-tunnel-endpoint {
        description "Parameters for OTN service.";

                leaf payload-treatment {
            type enumeration {
                enum switching;
                enum transport;
            }
            default switching;
            description
                "Treatment of the incoming payload. Payload can
                either be switched, or transported as is.";
        }

        leaf src-client-signal {
                type identityref {
```

```
                        base tran-types:client-signal;
                }
                description
                        "Client signal at the source endpoint of
                        the tunnel.";
        }

        leaf src-tpn {
            type uint16 {
                range "0..4095";
            }
            description
                "Tributary Port Number. Applicable in case of mux
                services.";
            reference
                "RFC7139: GMPLS Signaling Extensions for Control of
                Evolving G.709 Optical Transport Networks.";
        }

        leaf src-tsg {
            type identityref {
                base tran-types:tributary-slot-granularity;
            }
            description
                "Tributary slot granularity. Applicable in case of mux
                services.";
            reference
                "G.709/Y.1331, February 2012: Interfaces for the
                Optical Transport Network (OTN)";
        }

        leaf src-timeslot-count {
             type uint16;
             description
                 "Number of timeslots used at the source.";
        }

        container src-timeslots {
            description
                 "A list of tributary timeslots used by the
                 client service. Applicable in case of mux
                 services.";
            leaf-list values {
                type uint8;
                description
                        "Tributary timeslot value.";
                reference
                    "G.709/Y.1331, February 2012: Interfaces for the
```

```
                        Optical Transport Network (OTN)";
                }
            }

            leaf dst-client-signal {
                    type identityref {
                            base tran-types:client-signal;
                    }
                    description
                            "Client signal at the destination endpoint of
                            the tunnel.";
            }

            leaf dst-tpn {
                type uint16 {
                    range "0..4095";
                }
                description
                    "Tributary Port Number. Applicable in case of mux
                    services.";
                reference
                    "RFC7139: GMPLS Signaling Extensions for Control of
                    Evolving G.709 Optical Transport Networks.";
            }

            leaf dst-tsg {
                type identityref {
                    base tran-types:tributary-slot-granularity;
                }
                description
                    "Tributary slot granularity. Applicable in case of mux
                    services.";
                reference
                    "G.709/Y.1331, February 2012: Interfaces for the
                    Optical Transport Network (OTN)";
            }

            leaf dst-timeslot-count {
                 type uint16;
                 description
                     "Number of timeslots used at the destination.";
            }

            container dst-timeslots {
                 description
                    "A list of tributary timeslots used by the
                    client service. Applicable in case of mux
                    services.";
```

```
            leaf-list values {
                type uint8;
                description
                        "Tributary timeslot value.";
                reference
                    "G.709/Y.1331, February 2012: Interfaces for the
                        Optical Transport Network (OTN)";
            }
        }
    }

    /*
    Note: Comment has been given to authors of TE Tunnel model to add
    tunnel-types to the model in order to identify the technology
    type of the service.

    grouping otn-service-type {
        description
          "Identifies the OTN Service type.";
        container otn-service {
          presence "Indicates OTN Service.";
          description
            "Its presence identifies the OTN Service type.";
        }
      } // otn-service-type

    augment "/te:te/te:tunnels/te:tunnel/te:tunnel-types" {
        description
                "Introduce OTN service type for tunnel.";
        ext:augment-identifier otn-service-type-augment;
        uses otn-service-type;
    }
    */

    /*
    Note: Comment has been given to authors of TE Tunnel model to add
    list of endpoints under config to support P2MP tunnel.
    */
    augment "/te:te/te:tunnels/te:tunnel/te:config" {
        description
                "Augment with additional parameters required for OTN
                service.";
        ext:augment-identifier otn-tunnel-endpoint-config-augment;
        uses otn-tunnel-endpoint;
    }

    augment "/te:te/te:tunnels/te:tunnel/te:state" {
        description
```

```
                  "Augment with additional parameters required for OTN
                  service.";
          ext:augment-identifier otn-tunnel-endpoint-state-augment;
          uses otn-tunnel-endpoint;
      }

      /*
      Note: Comment has been given to authors of TE Tunnel model to add
      tunnel-lifecycle-event to the model. This notification is reported
      for all lifecycle changes (create, delete, and update) to the
      tunnel or lsp.
      augment "/te:tunnel-lifecycle-event" {
          description
            "OTN service event";
          uses otn-service-type;
          uses otn-tunnel-params;
          list endpoint {
            key
              "endpoint-address tp-id";
            description
              "List of Tunnel Endpoints.";
            uses te:tunnel-endpoint;
            uses otn-tunnel-params;
          }
        }
      */
  }

  <CODE ENDS>
```

2.4. Transport Types YANG Model

```
   <CODE BEGINS> file "ietf-transport-types@2016-06-24.yang"

    module ietf-transport-types {
       yang-version 1;
       namespace "urn:ietf:params:xml:ns:yang:ietf-transport-types";
       prefix "tran-types";

       organization
           "IETF CCAMP Working Group";
       contact
           "WG Web: <http://tools.ietf.org/wg/ccamp/>
           WG List: <mailto:ccamp@ietf.org>

           Editor: Anurag Sharma
                   <mailto:AnSharma@infinera.com>
```

```
        Editor: Rajan Rao
                <mailto:rrao@infinera.com>

        Editor: Xian Zhang
                <mailto:zhang.xian@huawei.com>";

    description
        "This module defines transport types.";

    revision "2016-06-24" {
        description "Initial revision";
        reference "TBD";
    }

    identity tributary-slot-granularity {
        description
                "Tributary slot granularity.";
        reference
                "G.709/Y.1331, February 2012: Interfaces for the
                        Optical Transport Network (OTN)";
    }

    identity tsg-1.25G {
        base tributary-slot-granularity;
        description
                "1.25G tributary slot granularity.";
    }

    identity tsg-2.5G {
        base tributary-slot-granularity;
        description
                "2.5G tributary slot granularity.";
    }

    identity tributary-protocol-type {
        description
                "Base identity for protocol framing used by
                tributary signals.";
    }

    identity prot-OTU1 {
        base tributary-protocol-type;
        description
                "OTU1 protocol (2.66G)";
    }

    identity prot-OTU1e {
        base tributary-protocol-type;
```

```
           description
                   "OTU1e protocol (11.04G)";
       }


       identity prot-OTU2 {
           base tributary-protocol-type;
           description
                   "OTU2 protocol (10.70G)";
       }

       identity prot-OTU2e {
           base tributary-protocol-type;
           description
                   "OTU2e protocol (11.09G) for 10G LAN PHY";
       }

       identity prot-OTU2f {
           base tributary-protocol-type;
           description
                   "OTU2f protocol (11.32G) for transporting a 10
                   fiber channel.";
       }

       identity prot-OTU3 {
           base tributary-protocol-type;
           description
                   "OTU3 protocol (43.01G)";
       }

       identity prot-OTU3e {
           base tributary-protocol-type;
           description
                   "OTU3e protocol (44.57G) for transporting four OTU2e
                   signals.";
       }

       identity prot-OTU3e2 {
           base tributary-protocol-type;
           description
                   "OTU3e2 protocol (44.58G).";
       }


       identity prot-OTU4 {
           base tributary-protocol-type;
           description
                   "OTU4 protocol (112G) for transporting 100GE
```

```
                   signal.";
       }

       identity prot-OTUCn {
           base tributary-protocol-type;
           description
                   "OTUCn protocol (beyond 100G) for transporting
                   more than 100G signals.";
       }

       identity prot-ODU0 {
           base tributary-protocol-type;
           description
                   "ODU0 protocol (1.24G).";
       }

       identity prot-ODU1 {
           base tributary-protocol-type;
           description
                   "ODU1 protocol (2.49G).";
       }

       identity prot-ODU1e {
           base tributary-protocol-type;
           description
                   "ODU1e protocol (10.35G).";
       }

       identity prot-ODU2 {
           base tributary-protocol-type;
           description
                   "ODU2 protocol (10.03G).";
       }

       identity prot-ODU2e {
           base tributary-protocol-type;
           description
                    "ODU2e protocol (10.39G).";
       }

       identity prot-ODU3 {
           base tributary-protocol-type;
           description
                   "ODU 3 protocol (40.31G).";
       }

       identity prot-ODU3e2 {
           base tributary-protocol-type;
```

```
           description
                   "ODU3e2 protocol (41.78G).";
       }

       identity prot-ODU4 {
           base tributary-protocol-type;
           description
                   "ODU4 protocol (104.79G).";
       }

       identity prot-ODUFlex-cbr {
           base tributary-protocol-type;
           description
                   "ODU Flex CBR protocol for transporting constant bit
                   rate signal.";
       }

       identity prot-ODUFlex-gfp {
           base tributary-protocol-type;
           description
                   "ODU Flex GFP protocol for transporting stream
                   of packets using Generic Framing Procedure.";
       }

       identity prot-ODUCn {
           base tributary-protocol-type;
           description
                   "ODUCn protocol (beyond 100G).";
       }

       identity prot-1GbE {
           base tributary-protocol-type;
           description
                   "1G Ethernet protocol";
       }

       identity prot-10GbE-LAN {
           base tributary-protocol-type;
           description
                   "10G Ethernet LAN protocol";
       }

       identity prot-40GbE {
           base tributary-protocol-type;
           description
                   "40G Ethernet protocol";
       }
```

```
       identity prot-100GbE {
           base tributary-protocol-type;
           description
                   "100G Ethernet protocol";
       }

       identity client-signal {
           description
                   "Base identity from which specific client signals
                   for the tunnel are derived.";
         }

       identity client-signal-1GbE {
           base client-signal;
           description
                   "Client signal type of 1GbE";
         }

       identity client-signal-10GbE-LAN {
           base client-signal;
           description
                   "Client signal type of 10GbE LAN";
         }

       identity client-signal-10GbE-WAN {
           base client-signal;
           description
                   "Client signal type of 10GbE WAN";
         }

       identity client-signal-40GbE {
           base client-signal;
           description
                   "Client signal type of 40GbE";
         }

       identity client-signal-100GbE {
           base client-signal;
           description
                   "Client signal type of 100GbE";
         }

       identity client-signal-OC3_STM1 {
           base client-signal;
           description
                   "Client signal type of OC3 and STM1";
         }
```

```
       identity client-signal-OC12_STM4 {
           base client-signal;
           description
                   "Client signal type of OC12 and STM4";
         }

       identity client-signal-OC48_STM16 {
           base client-signal;
           description
                   "Client signal type of OC48 and STM16";
         }

       identity client-signal-OC192_STM64 {
           base client-signal;
           description
                   "Client signal type of OC192 and STM64";
         }

       identity client-signal-OC768_STM256 {
           base client-signal;
           description
                   "Client signal type of OC768 and STM256";
         }

       identity client-signal-OTU1 {
           base client-signal;
           description
                   "Client signal type of OTU1 (2.66G)";
       }

       identity client-signal-OTU2 {
           base client-signal;
           description
                   "Client signal type of OTU2 (10.70G)";
       }

       identity client-signal-OTU2e {
           base client-signal;
           description
                   "Client signal type of OTU2e (11.09G)";
       }

       identity client-signal-OTU2f {
           base client-signal;
           description
                   "Client signal type of OTU2f (11.32G)";
       }
```

```
       identity client-signal-OTU3 {
           base client-signal;
           description
                   "Client signal type of OTU3 (43.01G)";
       }

       identity client-signal-OTU3e {
           base client-signal;
           description
                   "Client signal type of OTU3e (44.58G)";
       }

       identity client-signal-OTU4 {
           base client-signal;
           description
                   "Client signal type of OTU4 (112G)";
       }

       identity client-signal-OTUCn {
           base client-signal;
           description
                   "Client signal type of OTUCn (beyond 100G)";
       }

       identity client-signal-ODU0 {
           base client-signal;
           description
                   "Client signal type of ODU0 (1.24G)";
       }

       identity client-signal-ODU1 {
           base client-signal;
           description
                   "ODU1 protocol (2.49G)";
       }

       identity client-signal-ODU2 {
           base client-signal;
           description
                   "Client signal type of ODU2 (10.03G)";
       }

       identity client-signal-ODU2e {
           base client-signal;
           description
                   "Client signal type of ODU2e (10.39G)";
       }
```

```
    identity client-signal-ODU3 {
        base client-signal;
        description
                "Client signal type of ODU 3 (40.31G)";
    }

    identity client-signal-ODU3e2 {
        base client-signal;
        description
                "Client signal type of ODU3e2 (41.78G)";
    }

    identity client-signal-ODU4 {
        base client-signal;
        description
                "Client signal type of ODU4 (104.79G)";
    }

    identity client-signal-ODUFlex-cbr {
        base client-signal;
        description
                "Client signal type of ODU Flex CBR";
    }

    identity client-signal-ODUFlex-gfp {
        base client-signal;
        description
                "Client signal type of ODU Flex GFP";
    }

    identity client-signal-ODUCn {
        base client-signal;
        description
                "Client signal type of ODUCn (beyond 100G).";
    }

    identity client-signal-FC400 {
        base client-signal;
        description
                "Client signal type of Fibre Channel FC400.";
    }

    identity client-signal-FC800 {
        base client-signal;
        description
                "Client signal type of Fibre Channel FC800.";
    }
```

```
    identity client-signal-FICON-4G {
        base client-signal;
        description
                "Client signal type of Fibre Connection 4G.";
    }

    identity client-signal-FICON-8G {
        base client-signal;
        description
                "Client signal type of Fibre Connection 8G.";
    }
}

<CODE ENDS>
```

3. Security Considerations

TBD

4. IANA Considerations

TBD

5. Acknowledgements

6. Normative References

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

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <http://www.rfc-editor.org/info/rfc6020>.

[RFC7139] Zhang, F., Ed., Zhang, G., Belotti, S., Ceccarelli, D., and K. Pithewan, "GMPLS Signaling Extensions for Control of Evolving G.709 Optical Transport Networks", RFC 7139, DOI 10.17487/RFC7139, March 2014, <http://www.rfc-editor.org/info/rfc7139>.

Authors' Addresses

Anurag Sharma
Infinera Corp
169 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1-408-572-5365
Email: AnSharma@infinera.com

Rajan Rao
Infinera Corp
169 Java Drive
Sunnyvale, CA 94089
USA

Phone: +1-408-543-7755
Email: rrao@infinera.com

Xian Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Industrial Base, Bantian, Longgang District
Shenzhen, Guangdong 518129
P.R.China

Email: zhang.xian@huawei.com

J.E. Lopez de Vergara
Universidad Autonoma de Madrid
V. Lopez
O. Gonzalez de Dios
Telefonica I+D/GCTO
D. King
Lancaster University
Y. Lee
Huawei
Z. Ali
Cisco Systems
July 8, 2016

YANG data model for Flexi-Grid Optical Networks
draft-vergara-ccamp-flexigrid-yang-03.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

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

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

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

This Internet-Draft will expire on January 7, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Abstract

This document defines a YANG model for managing flexi-grid optical Networks. The model described in this document is composed of two submodels: one to define a flexi-grid traffic engineering database, and other one to describe the flexi-grid paths or media channels.

Table of Contents

1. Introduction

Internet-based traffic is dramatically increasing every year. Moreover, such traffic is also becoming more dynamic. Thus, transport networks need to evolve from current DWDM systems towards elastic optical networks, based on flexi-grid transmission and switching technologies. This technology aims at increasing both transport network scalability and flexibility, allowing the optimization of bandwidth usage.

This document presents a YANG model for flexi-grid objects in the dynamic optical network, including the nodes, transponders and links between them, as well as how such links interconnect nodes and transponders.

The YANG model for flexi-grid [RFC7698] networks allows the representation of the flexi-grid optical layer of a network, combined with the underlying physical layer. The model is defined in two YANG modules:

- o Flexi-grid-TED (Traffic Engineering Database): This module defines all the information needed to represent the flexi-grid optical node, transponder and link.
- o Media-channel: This module defines the whole path from a source transponder to the destination through a number of intermediate nodes in the flexi-grid optical network.

This document identifies the flexi-grid components, parameters and their values, characterizes the features and the performances of the flexi-grid elements. An application example is provided towards the end of the document to better understand their utility.

2. Conventions used in this document

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

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

In this document, the characters ">>" preceding an indented line(s) indicates a compliance requirement statement using the key words listed above. This convention aids reviewers in quickly identifying or finding the explicit compliance requirements of this RFC.

3. Flexi-grid network topology model overview

YANG is a data modeling language used to model configuration data manipulated by the NETCONF protocol. Several YANG models have already been specified for network configurations. For instance, the work in [I-D.draft-ietf-i2rs-yang-network-topo] has proposed a generic YANG model for network/service topologies and inventories. The work in [I-D.draft-ietf-teas-yang-te-topo] presents a data model to represent, retrieve and manipulate Traffic Engineering (TE) Topologies. These models serve as base models that other technology specific models can augment. A YANG model has also been proposed in [I-D.draft-dharini-ccamp-dwdm-if-yang] to manage single channel optical interface parameters of DWDM applications, and in

[I-D.draft-ietf-ccamp-wson-yang] another model has been specified for the routing and wavelength assignment TE topology in wavelength switched optical networks (WSONs). None of them are specific for flexi-grid technology.

Then, as stated before, we propose a model to describe a flexi-grid topology that is split in two YANG sub-modules:

- o Flexi-grid-TED: In order to be compatible with existing proposals, we augment the definitions contained in [I-D.draft-ietf-i2rs-yang-network-topo] and [I-D.draft-ietf-teas-yang-te-topo], by defining the different elements we find in a flexi-grid network: a node, a transponder and a link. For that, each of those elements is defined as a container that includes a group of attributes. References to the elements are provided to be later used in the definition of a media channel. It also includes the data types for the type of modulation, the flexi-grid technology, the FEC, etc.
- o Media-channel: This module defines the whole path from a source transponder to the destination through a number of intermediate nodes and links. For this, it takes the information defined before in the flexi-grid TED.

The following section provides a detailed view of each module.

4. Main building blocks

Subsections below detail each of the defined YANG modules. They are listed in Appendix A.

4.1. Flexi-grid TED

The description of the three main components, flexi-grid-node, flexi-grid-transponder and flexi-grid-link is provided below. flexi-grid-sliceable-transponders are also defined.

<flexi-grid-node> ::= <config> <state>

<flexi-grid-node>: This element designates a node in the network.

<config> ::= <flexi-grid-node-attributes-config>

<config>: Contains the configuration of a node.
<flexi-grid-node-attributes-config> ::= <list-interface>
<connectivity_matrix>

<flexi-grid-node-attributes-config>: Contains all the attributes related to the node configuration, such as its interfaces or its management addresses.

```
                   <list-interface> ::= <name> <port-number>
                   <input-port> <output-port> <description>
                   <interface-type>
                   [<numbered-interface> / <unnumbered-interface>]

                       <list-interface>: The list containing all the
                       information of the interfaces.

                       <name>: Determines the interface name.

                       <port-number>: Port number of the interface.

                       <input-port>: Boolean value that defines
                       whether the interface is input or not.

                       <output-port>: Boolean value that defines
                       whether the interface is output or not.

                       <description>: Description of the usage of
                       the interface.

                       <interface-type>: Determines if the interface
                       is numbered or unnumbered.

                       <numbered-interface> ::= <n-i-ip-address>

                           <numbered-interface>: An interface with
                           its own IP address.

                           <n-i-ip-address>: Only available if
                           <interface-type> is "numbered-interface".
                           Determines the IP address of the interface.

                       <unnumbered-interface> ::= <u-i-ip-address>
                       <label>

                           <unnumbered-interface>: A interface that
                           needs a label to be unique.

                           <u-i-ip-address>: Only available if
                           <interface-type> is "numbered-interface".
                           Determines the node IP address, which with
                           the label defines the interface.

                           <label>: Label that determines the
                           interface, joint with the node IP address.

                   <connectivity-matrix> ::= <connections>

                       <connectivity-matrix>: Determines whether a
                       connection port in/port out exists.

                       <connections> ::= <input-port-id>
                       <output-port-id>
```

```
           <state> ::= <flexi-grid-node-attributes-config>
           <flexi-grid-node-attributes-state>

               <state>: Contains the state of a node.

               <flexi-grid-node-attributes-config>: See above.

               <flexi-grid-node-attributes-state>: Contain all the
               attributes related to the state of a node.

   <flexi-grid-transponder> ::= <config> <state>
       <config> ::= <flexi-grid-transponder-attributes-config>

           <config>: Contains the configuration of a transponder.

           <flexi-grid-transponder-attributes-config> ::=
           <available-modulation> <modulation-type>
           <available-FEC> <FEC-enabled> [<FEC-type>]

               <flexi-grid-transponder-attributes>: Contains all the
               attributes related to the transponder, such as whether
               it has FEC enabled or not, or its modulation type.

                  <available-modulation>: It provides a list of the
                  modulations available at this transponder.

                  <modulation-type>: Determines the type of modulation
                  in use: QPSK, QAM16, QAM64...

                  <available-FEC>: It provides a list of the FEC
                  algorithms available at this transponder.

                  <FEC-enabled>: Boolean value that determines whether
                  is the FEC enabled or not.

                <FEC-type>: Determines the type of FEC in use:
                reed-solomon, hamming-code, enum golay, BCH...

        <state> ::= <flexi-grid-transponder-attributes-config>
        <flexi-grid-transponder-attributes-state>

            <state>: Contains the state of a transponder.

            <flexi-grid-transponder-attributes-config>: See above.

            <flexi-grid-transponder-attributes-state>: Contains the
            state of a transponder.
```

```
   <flexi-grid-sliceable-transponder> ::= <config> <state>

       <flexi-grid-sliceable-transponder>: A list of transponders.

       <config> ::= <flexi-grid-transponder-attributes-config>
       <flexi-grid-sliceable-transponder-attributes-config>

           <flexi-grid-transponder-attributes-config>: See above.

           <flexi-grid-sliceable-transponder-attributes-config> ::=
           <transponder-list>

               <flexi-grid-sliceable-transponder-attributes-config>:
               Contains the configuration of a sliceable transponder

               <transponder-list> ::= <carrier-id>
                   <transponder-list>: A list of transponders.

                   <carrier-id>: An identifier for each one of the
                   transponders in the list.

       <state> ::= <flexi-grid-transponder-attributes-state>
       <flexi-grid-sliceable-transponder-attributes-state>
       <flexi-grid-transponder-attributes-config>
       <flexi-grid-sliceable-transponder-attributes-config>

           <state>: Contains the state of a sliceable transponder.

           <flexi-grid-transponder-attributes-state>: See above.

           <flexi-grid-sliceable-transponder-attributes-state>:
           Contains the state attributes of a sliceable transponders.

           <flexi-grid-transponder-attributes-config>: See above.

           <flexi-grid-sliceable-transponder-attributes-config>: See
           above.


    <link> ::= <config> <state>

        <link>: This element describes all the information of a link.

        <config> ::= <flexi-grid-link-attributes-config>

            <config>: Contains the configuration of a link.

            <flexi-grid-link-attributes-config> ::= <technology-type>
            <available-label-flexi-grid> <N-max> <base-frequency>
            <nominal-central-frequency-granularity>
            <slot-width-granularity>
```

<flexi-grid-link-attributes>: Contains all the attributes related to the link, such as its unique id, its N value, its latency, etc.

<link-id>: Unique id of the link.

<available-label-flexi-grid>: Array of bits that determines, with each bit, the availability of each interface for flexi-grid technology.

<N-max>: The max value of N in this link, being N the number of slots.

<base-frequency>: The default central frequency used in the link.

<nominal-central-frequency-granularity>: It is the spacing between allowed nominal central frequencies and it is set to 6.25 GHz (note: sometimes referred to as 0.00625 THz).

<slot-width-granularity>: 12.5 GHz, as defined in G.694.1.

<state> ::= <flexi-grid-link-attributes-config>
<flexi-grid-link-attributes-state>

<state>: Contains the state of a link.

<flexi-grid-link-attributes-config>: See above.

<flexi-grid-link-attributes-state>: Contains all the the information related to the state of a link.

4.2. Media-channel/network-media-channel

The model defines two types of media channels, following the terminology summarized in [RFC7698]:
media-channel, which represents a (effective) frequency slot supported by a concatenation of media elements (fibers, amplifiers, filters, switching matrices...);
network-media-channel: It is a media channel that transports an Optical Tributary Signal. In the model, the network media channel has as end-points transponders, which are the source and destination of the optical signal. The description of these components is provided below:

<media-channel> ::= <source> <destination> <link-channel> <effective-freq-slot>

<media-channel>: Determines a media-channel and its components.

<source > ::= <source-node> <source-port>

<source>: In a media-channel, the source is a node and a port.

<source-node>: Reference to the source node of the media channel.

<source-port>: Reference to the source port in the source <node.

<destination> ::= <destination-node> <destination-port>

<destination>: In a media-channel, the destination is a node and a port.

<destination-node>: Reference to the destination node of the media channel.

<destination-port>: Reference to the destination port in the destination node.

<link-channel> ::= <link-id> <N> <M> <source-node> <source-port> <destination-node> <destination-port> <link> <bidirectional>

<link-channel>: Defines a list with each of the links between elements in the media channel.

<link-id>: Unique identifier for the link channel

<N>: N used for this link channel.

<M>: M used for this link channel.

<source-node>: Reference to the source node of this link channel.

<source-port>: Reference to the source port of this link channel.

<destination-node>: Reference to the destination node of this link channel.

<destination-port>: Reference to the destination port of this link channel.

<link>: Reference to the link of this link channel.

<bidirectional>: Indicates if this link is bidirectional or not.

<effective-freq-slot> ::= <N> <M>

<effective-freq-slot>: Defines the effective frequency slot of the media channel, which could be different from the one defined in the link channels.

<N>: Defines the effective N for this media channel.

<M>: Defines the effective M for this media channel.

<network-media-channel> ::= <source> <destination> <link-channel> <effective-freq-slot>

<network-media-channel>: Determines a network media-channel and its components.

<source > ::= <source-node> <source-transponder>

<source>: In a network media channel, the source is defined by a node and a transponder.

<source-node>: Reference to the source node of the media channel.

<source-transponder>: Reference to the source transponder in the source node.

<destination> ::= <destination-node> <destination-transponder>

<destination>: In a network media channel, the destination is defined by a node and a transponder

<destination-node>: Reference to the destination node of the media channel.

<destination-port>: Reference to the destination port in the destination node.

<link-channel>: See above, the information is reused for both types of media channels.

<effective-freq-slot>: See above, this information is reused for both types of media channels.

5. Example of use

In order to explain how this model is used, we provide the following example. An optical network usually has multiple transponders, switches (nodes) and links between them. Figure 1 shows a simple topology, where two physical paths interconnect two optical transponders.

```
                               Media channel
            <===================================================>
                                  Path x
            <--------------------------------------------------->

                     +----------+          +----------+
              Link 1 |Flexi-grid|  Link 2  |Flexi-grid| Link 3
                .--->|   node   |<-------->|   node   |<---.
                |    |    B     |          |    C     |    |
                |    +----------+          +----------+    |
                v                                          v
            +-------------+                      +-------------+
            | Flexi-grid  |                      | Flexi-grid  |
            | transponder |                      | transponder |
            |      A      |                      |      E      |
            +-------------+                      +-------------+
                  ^                                     ^
                  |             +----------+            |
                  |    Link 4   |Flexi-grid|   Link 5   |
                  '------------>|   node   |<-----------'
                                |    D     |
                                +----------+
            <--------------------------------------------------->
                                  Path y
```

Figure 1. Topology example.

In order to configure a media channel to interconnect transponders A and E, first of all we have to populate the flexi-grid TED YANG model with all elements in the network:

1. We define the transponders A and E, including their FEC type, if enabled, and modulation type. We also provide node identifiers and addresses for the transponders, as well as interfaces included in the transponders. Sliceable transponders can also be defined if needed.

2. We do the same for the nodes B, C and D, providing their identifiers, addresses and interfaces, as well as the internal connectivity matrix between interfaces.

3. Then, we also define the links 1 to 5 that interconnect nodes and transponders, indicating which flexi-grid labels are available. Other information, such as the slot frequency and granularity are also provided.

Next, we can configure the media channel from the information we have stored in the flexi-grid TED, by querying which elements are available, and planning the resources that have to be provided on each situation. Note that every element in the flexi-grid TED has a reference, and this is the way in which they are called in the media channel.

4. Depending on the case, it is possible to define either the source and destination node ports, or the source and destination node and transponder. In our case, we would define a network media channel, with source transponder A and source node B, and destination transponder E and destination node C. Thus, we are going to follow path x.

5. Then, for each link in the path x, we indicate which channel we are going to use, providing information about the slots, and what nodes are connected.

Finally, the flexi-grid TED has to be updated with each element usage status each time a media channel is created or torn down.

6. Formal Syntax

The following syntax specification uses the augmented Backus-Naur Form (BNF) as described in [RFC5234].

7. Security Considerations

The transport protocol used for sending the managed information MUST support authentication and SHOULD support encryption.

The defined data-model by itself does not create any security implications.

8. IANA Considerations

The namespace used in the defined models is currently based on the IDEALIST project URI. Future versions of this document could register a URI in the IETF XML registry [RFC3688], as well as in the YANG Module Names registry [RFC6020].

9. References

9.1. Normative References

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

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008, <http:/www.rfc-editor.org/info/rfc5234>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.

[RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688, January 2004.

9.2. Informative References

[RFC7698] Gonzalez de Dios, O., Casellas, R., "Framework and Requirements for GMPLS-Based Control of Flexi-Grid Dense Wavelength Division Multiplexing (DWDM) Networks", RFC7698, November 2015.

[I-D.draft-ietf-i2rs-yang-network-topo] Clemm, A., Medved, J., Varga, R., Tkacik, T., Bahadur, N., Ananthakrishnan, H., Liu, X., "A Data Model for Network Topologies", Internet Draft draft-ietf-i2rs-yang-network-topo-03.txt, 2016.

[I-D.draft-ietf-teas-yang-te-topo] Liu, X., Bryskin, I., Pavan Beeram, V., Saad, T., Shah, H., Gonzalez De Dios, O., "YANG Data Model for TE Topologies", Internet Draft draft-ietf-teas-yang-te-topo-04.txt, 2016

[I-D.draft-dharini-ccamp-dwdm-if-yang] Galimberti, G., Kunze, R., Lam, K., Hiremagalur, D., Grammel, G., Fang, L., Ratterree, G., Eds., "A YANG model to manage the optical interface parameters for an external transponder in a WDM network", Internet Draft, draft-dharini-ccamp-dwdm-if-yang-00.txt, 2016.

[I-D.draft-ietf-ccamp-wson-yang] Lee, Y. Dhody, D., Zhang, X., Guo, A., Lopez, V., King, D., Yoon, B.,"A Yang Data Model for WSON Optical Networks", Internet Draft, draft-ietf-ccamp-wson-yang-01.txt

10. Contributors

The model presented in this paper was contributed to by more people than can be listed on the author list. Additional contributors include:

o Daniel Michaud Vallinoto, Universidad Autonoma de Madrid
o Daniel Perdices Burrero, Universidad Autonoma de Madrid

11. Acknowledgments

The work presented in this Internet-Draft has been partially funded by the European Commission under the project Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Networks (IDEALIST) of the Seventh Framework Program, with Grant Agreement Number: 317999, and by the Spanish Ministry of Economy and Competitiveness under the project TRAFICA (MINECO/FEDER TEC2015-69417-C2-1-R).

Appendix A. YANG models

A.1. Flexi-grid TED YANG Model

A.1.1. Yang Model - Tree Structure

```
module: ietf-flexi-grid-topology
flexi-grid-network-type
   augment /nd:networks/nd:network/nd:network-types:
      +--rw flexi-grid-network!
flexi-grid-link-attributes-config
   augment /nd:networks/nd:network/lnk:link/tet:te/tet:config:
      +--rw available-label-flexi-grid*             bits
      +--rw N-max?                                  int32
      +--rw base-frequency?                         decimal64
      +--rw nominal-central-frequency-granularity?   decimal64
      +--rw slot-width-granularity?                 decimal64
flexi-grid-link-attributes-state
   augment /nd:networks/nd:network/lnk:link/tet:te/tet:state:
      +--ro available-label-flexi-grid*             bits
      +--ro N-max?                                  int32
      +--ro base-frequency?                         decimal64
      +--ro nominal-central-frequency-granularity?   decimal64
      +--ro slot-width-granularity?                 decimal64
flexi-grid-node-attributes-config
   augment /nd:networks/nd:network/nd:node/tet:te/tet:config:
      +--rw interfaces* [name]
         +--rw name                    string
         +--rw port-number?            uint32
         +--rw input-port?             boolean
         +--rw output-port?            boolean
         +--rw description?            string
         +--rw type?                   interface-type
         +--rw numbered-interface
         |  +--rw n-i-ip-address?   inet:ip-address
         +--rw unnumbered-interface
            +--rw u-i-ip-address?   inet:ip-address
            +--rw label?            uint32
```

```
   flexi-grid-node-attributes-state
      augment /nd:networks/nd:network/nd:node/tet:te/tet:state:
         +--ro interfaces* [name]
            +--ro name                    string
            +--ro port-number?            uint32
            +--ro input-port?             boolean
            +--ro output-port?            boolean
            +--ro description?            string
            +--ro type?                   interface-type
            +--ro numbered-interface
            |  +--ro n-i-ip-address?   inet:ip-address
            +--ro unnumbered-interface
               +--ro u-i-ip-address?   inet:ip-address
               +--ro label?            uint32
   flexi-grid-connectivity-matrix-attributes
      augment /nd:networks/nd:network/nd:node/tet:te/tet:config/
      tet:te-node-attributes/tet:connectivity-matrix:
         +--rw connections* [input-port-id]
            +--rw input-port-id     flexi-grid-node-port-ref
            +--rw output-port-id?   flexi-grid-node-port-ref
   flexi-grid-connectivity-matrix-attributes
      augment /nd:networks/nd:network/nd:node/tet:te/tet:state/
      tet:te-node-attributes/tet:connectivity-matrix:
         +--ro connections* [input-port-id]
            +--ro input-port-id     flexi-grid-node-port-ref
            +--ro output-port-id?   flexi-grid-node-port-ref
   flexi-grid-transponder
      augment /nd:networks/nd:network/nd:node:
         +--rw config
         |  +--rw available-modulation*   modulation
         |  +--rw modulation-type?        modulation
         |  +--rw available-FEC*          FEC
         |  +--rw FEC-enabled?            boolean
         |  +--rw FEC-type?               FEC
         +--rw state
         |  +--rw available-modulation*   modulation
         |  +--rw modulation-type?        modulation
         |  +--rw available-FEC*          FEC
         |  +--rw FEC-enabled?            boolean
         |  +--rw FEC-type?               FEC
         +--rw node-type?   flexi-grid-node-type
```

```
   flexi-grid-sliceable-transponder
      augment /nd:networks/nd:network/nd:node:
         +--rw config
         |  +--rw available-modulation*   modulation
         |  +--rw modulation-type?        modulation
         |  +--rw available-FEC*          FEC
         |  +--rw FEC-enabled?            boolean
         |  +--rw FEC-type?               FEC
         |  +--rw transponder-list* [carrier-id]
         |     +--rw carrier-id           uint32
         +--rw state
         |  +--rw available-modulation*   modulation
         |  +--rw modulation-type?        modulation
         |  +--rw available-FEC*          FEC
         |  +--rw FEC-enabled?            boolean
         |  +--rw FEC-type?               FEC
         +--rw node-type?   flexi-grid-node-type
```

A.1.2. YANG Model - Code

```
   <CODE BEGINS> file "ietf-flexi-grid-ted.yang"

module ietf-flexi-grid-ted {
  yang-version 1.1;

  namespace "urn:ietf:params:xml:ns:yang:ietf-flexi-grid-ted";
  prefix "fg-topo";

  import ietf-network {
    prefix "nd";
  }

  import ietf-network-topology {
    prefix "lnk";
  }

  import ietf-te-topology {
    prefix "tet";
  }

   import ietf-inet-types {
      prefix "inet";
   }
```

```
  organization
         "IETF CCAMP Working Group";

      contact
         "Editor: Jorge Lopez de Vergara
                  <jorge.lopez_vergara@uam.es>";

      description
         "This module contains a collection of YANG definitions for
         a Flexi-Grid Traffic Engineering Database (TED).

         Copyright (c) 2016 IETF Trust and the persons identified as
         authors of the code.  All rights reserved.

         Redistribution and use in source and binary forms, with or
         without modification, is permitted pursuant to, and subject
         to the license terms contained in, the Simplified BSD
         License set forth in Section 4.c of the IETF Trust's Legal
         Provisions Relating to IETF Documents
         (http://trustee.ietf.org/license-info).";

      revision 2016-07-08 {
         description
            "version 4.";

         reference
             "RFC XXX: A Yang Data Model for
             Flexi-Grid Optical Networks ";
      }

  /*
  typedef
  */

   typedef flexi-grid-node-type {
         type enumeration {
             enum flexi-grid-node {
                description
                   "Flexi-grid node";
             }
             enum flexi-grid-transponder {
                description
                   "Flexi-grid transponder";
             }
             enum flexi-grid-sliceable-transponder {
                description
                   "Flexi-grid sliceable transponder";
             }
         }
```

```
         description "Determines the node type:
             flexi-grid-node,
             flexi-grid-transponder or
             flexi-grid-sliceable-transponder";
   }

  typedef modulation {
         type enumeration {
            enum QPSK {
                description
                "QPSK (Quadrature Phase Shift Keying) modulation";
                }
            enum DP_QPSK {
                description "DP-QPSK (Dual Polarization Quadrature
                Phase Shift Keying) modulation";
                }
            enum QAM16 {
                description "QAM16 (Quadrature Amplitude Modulation
                - 4 bits per symbol) modulation";
                }
            enum DP_QAM16 {
                description "DP-QAM16 (Dual Polarization
                Quadrature Amplitude Modulation - 4 bits per
                symbol) modulation";
                }
            enum DC_DP_QAM16 {
                description "DC DP-QAM16 (Dual Polarization
                Quadrature Amplitude Modulation - 4 bits per
                symbol) modulation";
                }
         }
         description
            "Enumeration that defines the type of wave modulation";
   }

   typedef FEC {
         type enumeration {
            enum reed-solomon {
                description "Reed-Solomon error correction";
            }
            enum hamming-code{
                description "Hamming Code error correction";
            }
            enum golay{
                description "Golay error correction";
            }
         }
         description "Enumeration that defines the type of
         Forward Error Correction";
   }
```

```
   typedef interface-type {
         type enumeration {
            enum numbered-interface {
                description "The interface is numbered";
            }
            enum unnumbered-interface {
                description "The interface is unnumbered";
            }
         }
         description
            "Enumeration that defines if an interface is numbered or
            unnumbered";
   }

   /*
      Typedef related to references
   */

   typedef flexi-grid-transponder-ref {
         type leafref {
            path
             "/nd:networks/nd:network/nd:node/nd:node-id";
         }
         description
            "This type is used by data models that need to reference
            a flexi-grid optical transponder.";
      }

      typedef flexi-grid-node-ref {
         type leafref {
            path
             "/nd:networks/nd:network/nd:node/tet:te/tet:te-node-id";
         }
         description
            "This type is used by data models that need to reference
            a flexi-grid optical node.";
      }

      typedef flexi-grid-link-ref {
         type leafref {
            path
               "/nd:networks/nd:network/lnk:link/lnk:link-id";
         }

         description
            "This type is used by data models that need to reference
            a flexi-grid optical link.";
      }
```

```
      typedef flexi-grid-node-port-ref {
         type leafref {
            path "/nd:networks/nd:network/nd:node/tet:te/tet:config/"
            +"fg-topo:interfaces/fg-topo:port-number";
         }
         description
            "This type is used by data models that need to reference
            a flexi-grid optical link.";
      }

   /*
   Groupings
   */

    grouping flexi-grid-network-type {
         container flexi-grid-network {
            presence "indicates a flexi-grid optical network";
            description "flexi-grid optical network";
         }
         description "If present, it indicates a flexi-grid
         optical TED network";
    }

    grouping flexi-grid-node-attributes-config {
         description "Set of attributes of an optical node.";

         list interfaces {
            key "name";
            unique "port-number";
            description "List of interfaces contained in the node";
            leaf name {
               type string;
               description "Interface name";
            }
            leaf port-number {
               type uint32;
               description "Number of the port used by the interface";
            }
            leaf input-port {
               type boolean;
               description "Determines if the port is an input port";
            }
            leaf output-port {
               type boolean;
               description
               "Determines if the port is an output port";
            }
            leaf description {
               type string;
               description "Description of the interface";
            }
```

```
            leaf type {
               type interface-type;
               description "Determines the type of the interface";
            }
            container numbered-interface {
               when "type == numbered-interface" {
                 description
                 "If the interface is a numbered interface";
               }
               description "Container that defines an numbered
               interface with an ip-address";
               leaf n-i-ip-address{
                  type inet:ip-address;
                  description "IP address of the numbered interface";
               }
            }
            container unnumbered-interface {
               when "type == unnumbered-interface" {
                 description
                   "If the interface is an unnumbered interface";
               }
               description "Container that defines an unnumbered
               interface with an ip-address and a label";
               leaf u-i-ip-address{
                  type inet:ip-address;
                  description "IP address of the interface";
               }
               leaf label {
                  type uint32;
                  description "Number as label for the interface";
               }
            }
         }
      }

   grouping flexi-grid-node-attributes-state {
         description "Flexigrid node attributes (state).";
   }
```

```
   grouping flexi-grid-link-attributes-config {
         description "Set of attributes of an optical link";
            leaf-list available-label-flexi-grid {
               type bits {
                  bit is-available{
                    description "Set to 1 when it is available";
                  }
               }
               description
                  "Array of bits that determines whether a spectral
                  slot is available or not.";
            }

            leaf N-max {
               type int32;
               description "Maximum number of channels available.";
            }

            leaf base-frequency {
               type decimal64 {
                  fraction-digits 5;
               }
               units THz;
               default 193.1;
               description "Default central frequency";
               reference "rfc7698";
            }

            leaf nominal-central-frequency-granularity {
               type decimal64 {
                  fraction-digits 5;
               }
               units GHz;
               default 6.25;
               description
                  "It is the spacing between allowed nominal central
                  frequencies and it is set to 6.25 GHz";
               reference "rfc7698";
            }

            leaf slot-width-granularity {
               type decimal64 {
                  fraction-digits 5;
               }
               units GHz;
               description "Minimum space between slot widths";
               reference "rfc7698";
            }

   }
```

```
   grouping flexi-grid-link-attributes-state {
      description "Flexigrid link attributes (state)";
   }

   grouping flexi-grid-transponder-attributes-config {
      description "Configuration of an optical transponder";
      leaf-list available-modulation {
         type modulation;
         description
            "List determining all the available modulations";
      }
      leaf modulation-type {
         type modulation;
         description "Modulation type of the wave";
      }
      leaf-list available-FEC {
         type FEC;
         description "List determining all the available FEC";
      }
      leaf FEC-enabled {
         type boolean;
         description
            "Determines whether the FEC is enabled or not";
      }
      leaf FEC-type {
         type FEC;
         description "FEC type of the transponder";
      }
   }

   grouping flexi-grid-transponder-attributes-state {
      description "State of an optical transponder";
   }

   grouping flexi-grid-sliceable-transponder-attributes-config {
      description
         "Configuration of a sliceable transponder.";
      list transponder-list {
         key "carrier-id";
         description "List of carriers";
         leaf carrier-id {
            type uint32;
            description "Identifier of the carrier";
         }
      }
   }

   grouping flexi-grid-sliceable-transponder-attributes-state {
      description "State of a sliceable transponder.";
      uses flexi-grid-transponder-attributes-state;
   }
```

```
   grouping flexi-grid-connectivity-matrix-attributes {
      description "Connectivity matrix between the input and
          output ports";
      list connections {
         key "input-port-id";
         leaf input-port-id {
            type flexi-grid-node-port-ref;
            description "Identifier of the input port";
         }
         leaf output-port-id {
            type flexi-grid-node-port-ref;
            description "Identifier of the output port";
         }
         description "List of connections between input and
          output ports";
      }
   }

  /*
   * Data nodes
   */

   augment "/nd:networks/nd:network/nd:network-types" {
            uses flexi-grid-network-type;
             description "Augment network-types including flexi-grid
             topology";
   }

  augment "/nd:networks/nd:network/lnk:link/tet:te/tet:config" {
    when "/nd:network-types/tet:te-topology/flexi-grid-network" {
      description "Augment only for Flexigrid network.";
    }
    description "Augment link configuration";

    uses flexi-grid-link-attributes-config;
  }

  augment "/nd:networks/nd:network/lnk:link/tet:te/tet:state" {
    when "nd:network-types/tet:te-topology/flexi-grid-network" {
      description "Augment only for Flexigrid network.";
    }
    description "Augment link state";

    uses flexi-grid-link-attributes-config;
    uses flexi-grid-link-attributes-state;
  }
```

```
   augment "/nd:networks/nd:network/nd:node/tet:te/tet:config"  {
      when "nd:network-types/tet:te-topology/flexi-grid-network" {
         description "Augment only for Flexigrid network.";
      }
      uses flexi-grid-node-attributes-config;
      description "Augment node config with flexi-grid attributes";
   }

   augment "/nd:networks/nd:network/nd:node/tet:te/tet:state"  {
      when "nd:network-types/tet:te-topology/flexi-grid-network" {
         description "Augment only for Flexigrid network.";
      }
      uses flexi-grid-node-attributes-config;
      uses flexi-grid-node-attributes-state;
      description "Augment node config with flexi-grid attributes";
   }

   augment "/nd:networks/nd:network/nd:node/tet:te/tet:config"+
         "/tet:te-node-attributes/tet:connectivity-matrix" {
       when "nd:network-types/tet:te-topology/flexi-grid-network"{
          description "Augment only for Flexigrid network.";
       }

      uses flexi-grid-connectivity-matrix-attributes;
      description "Augment node connectivity-matrix for node config";
    }

    augment "/nd:networks/nd:network/nd:node/tet:te/tet:state"+
         "/tet:te-node-attributes/tet:connectivity-matrix" {
       when "nd:network-types/tet:te-topology/flexi-grid-network"{
          description "Augment only for Flexigrid network.";
       }

      uses flexi-grid-connectivity-matrix-attributes;
      description "Augment node connectivity-matrix for node config";
    }
```

```
   /*
   Elements out of TE
   */

   augment "/nd:networks/nd:network/nd:node" {
      when "nd:network-types/tet:te-topology/flexi-grid-network"{
          description "Augment only for Flexigrid network.";
       }

       container config {
          description "Configuration of either a transponder or
          a sliceable transponder";

       }

       container state {
        description "State of either a transponder or
        a sliceable transponder";

       }

       leaf node-type {
            type flexi-grid-node-type;
            description "Type of flexi-grid node";
      }

       description "Augment node with configuration and state
       for transponder";
   }

   augment "/nd:networks/nd:network/nd:node/fg-topo:config" {
      when "/nd:networks/nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-transponder|/nd:networks/
              nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-sliceable-transponder"{
          description "When it is either a flexi-grid transponder
           or a sliceable transponder";
      }
      uses flexi-grid-transponder-attributes-config;
      description "Augment node state with transponder attributes";
   }
```

```
   augment "/nd:networks/nd:network/nd:node/fg-topo:state" {
      when "/nd:networks/nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-transponder|/nd:networks/
              nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-sliceable-transponder"{
          description "When it is either a flexi-grid transponder
           or a sliceable transponder";
      }
      uses flexi-grid-transponder-attributes-state;
      uses flexi-grid-transponder-attributes-config;
      description "Augment node state with transponder attributes";
   }

   augment "/nd:networks/nd:network/nd:node/fg-topo:config" {
      when "/nd:networks/nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-sliceable-transponder"{
          description
              "When it is a flexi-grid sliceable transponder";
      }
      uses flexi-grid-sliceable-transponder-attributes-config;
      description "Augment node with sliceable transponder
          attributes";
   }

   augment "/nd:networks/nd:network/nd:node/fg-topo:state" {
      when "/nd:networks/nd:network/nd:node/fg-topo:node-type/
              fg-topo:flexi-grid-sliceable-transponder"{
          description
              "When it is a flexi-grid sliceable transponder";
    }
    uses flexi-grid-sliceable-transponder-attributes-state;
    uses flexi-grid-sliceable-transponder-attributes-config;
    description "Augment node with sliceable transponder
        attributes";
   }
}

   <CODE ENDS>
```

A.2. Media Channel YANG Model

A.2.1. YANG Model - Tree

```
module: ietf-flexi-grid-media-channel
   +--rw media-channel
   |  +--rw source
   |  |  +--rw source-node?   fg-ted:flexi-grid-node-ref
   |  |  +--rw source-port?   fg-ted:flexi-grid-node-port-ref
   |  +--rw destination
   |  |  +--rw destination-node?   fg-ted:flexi-grid-node-ref
   |  |  +--rw destination-port?   fg-ted:flexi-grid-node-port-ref
   |  +--rw effective-freq-slot
   |  |  +--rw N?   int32
   |  |  +--rw M?   int32
   |  +--rw link-channel* [link-id]
   |     +--rw link-id            int32
   |     +--rw N?                 int32
   |     +--rw M?                 int32
   |     +--rw source-node?       fg-ted:flexi-grid-node-ref
   |     +--rw source-port?       fg-ted:flexi-grid-node-port-ref
   |     +--rw destination-node?   fg-ted:flexi-grid-node-ref
   |     +--rw destination-port?   fg-ted:flexi-grid-node-port-ref
   |     +--rw link?              fg-ted:flexi-grid-link-ref
   |     +--rw bidireccional?     boolean
   +--rw network-media-channel
      +--rw source
      |  +--rw source-node?          fg-ted:flexi-grid-node-ref
      |  +--rw source-transponder?   fg-ted:flexi-grid-transponder-ref
      +--rw destination
      |  +--rw destination-node?          fg-ted:flexi-grid-node-ref
      |  +--rw destination-transponder?
      |             fg-ted:flexi-grid-transponder-ref
      +--rw effective-freq-slot
      |  +--rw N?   int32
      |  +--rw M?   int32
      +--rw link-channel* [link-id]
         +--rw link-id            int32
         +--rw N?                 int32
         +--rw M?                 int32
         +--rw source-node?       fg-ted:flexi-grid-node-ref
         +--rw source-port?       fg-ted:flexi-grid-node-port-ref
         +--rw destination-node?   fg-ted:flexi-grid-node-ref
         +--rw destination-port?   fg-ted:flexi-grid-node-port-ref
         +--rw link?              fg-ted:flexi-grid-link-ref
         +--rw bidireccional?     boolean
```

A.2.2. YANG Model - Code

```
   <CODE BEGINS> file "ietf-flexi-grid-media-channel.yang"

module ietf-flexi-grid-media-channel {
   yang-version 1;

   namespace
     "urn:ietf:params:xml:ns:yang:ietf-flexi-grid-media-channel";
   prefix fg-mc;

   import ietf-flexi-grid-ted {
      prefix fg-ted;
   }

   organization
      "IETF CCAMP Working Group";

   contact
      "Editor: Jorge Lopez de Vergara
                 <jorge.lopez_vergara@uam.es>";

   description
      "This module contains a collection of YANG definitions for
      a Flexi-Grid media channel.

      Copyright (c) 2016 IETF Trust and the persons identified as
      authors of the code.  All rights reserved.

      Redistribution and use in source and binary forms, with or
      without modification, is permitted pursuant to, and subject
      to the license terms contained in, the Simplified BSD
      License set forth in Section 4.c of the IETF Trust's Legal
      Provisions Relating to IETF Documents
      (http://trustee.ietf.org/license-info).";

   revision 2016-07-08 {
      description
         "version 4.";

      reference
         "RFC XXX: A Yang Data Model for Flexi-Grid Optical
         Networks ";
   }
```

```
   container media-channel {
      description
         "Media association that represents both the topology
         (i.e., path through the media) and the resource
         (frequency slot) that it occupies.  As a topological
         construct, it represents a (effective) frequency slot
         supported by a concatenation of media elements (fibers,
         amplifiers, filters, switching matrices...).  This term
         is used to identify the end-to-end physical layer entity
         with its corresponding (one or more) frequency slots
         local at each link filters.";
      reference "rfc7698";
      container source {
         description "Source of the media channel";
         leaf source-node {
            type fg-ted:flexi-grid-node-ref;
            description "Source node";
         }
         leaf source-port {
            type fg-ted:flexi-grid-node-port-ref;
            description "Source port";
          }
      }
      container destination {
         description "Destination of the media channel";
         leaf destination-node {
            type fg-ted:flexi-grid-node-ref;
            description "Destination node";
         }
         leaf destination-port {
            type fg-ted:flexi-grid-node-port-ref;
            description "Destination port";
         }
      }
      uses media-channel-attributes;
   }
   container network-media-channel {
      description
         "It is a media channel that transports an Optical
         Tributary Signal ";
      reference "rfc7698";
      container source {
         description "Source of the network media channel";
         leaf source-node {
            type fg-ted:flexi-grid-node-ref;
            description "Source node";
         }
         leaf source-transponder {
            type fg-ted:flexi-grid-transponder-ref;
            description "Source transponder";
         }
      }
```

```
      container destination {
         description "Destination of the network media channel";
         leaf destination-node {
            type fg-ted:flexi-grid-node-ref;
            description "Destination node";
         }
         leaf destination-transponder {
            type fg-ted:flexi-grid-transponder-ref;
            description "Destination transponder";
         }
      }
      uses media-channel-attributes;
   }

   grouping media-channel-attributes {
      description "Set of attributes of a media channel";
      container effective-freq-slot {
         description
            "The effective frequency slot is an attribute of
            a media channel and, being a frequency slot, it is
            described by its nominal central frequency and slot
            width";
         reference "rfc7698";
         leaf N {
            type int32;
            description
               "Is used to determine the Nominal Central
               Frequency. The set of nominal central frequencies
               can be built using the following expression:
                  f = 193.1 THz + n x 0.00625 THz,
               where 193.1 THz is ITU-T ''anchor frequency'' for
               transmission over the C band, n is a positive or
               negative integer including 0.";
            reference "rfc7698";
         }
         leaf M {
            type int32;
            description
               "Is used to determine the slot width. A slot width
               is constrained to be M x SWG (that is, M x 12.5 GHz),
               where M is an integer greater than or equal to 1.";
            reference "rfc7698";
         }
      }
```

```
      list link-channel {
         key "link-id";
         description
             "A list of the concatenated elements of the media
             channel.";
         leaf link-id {
            type int32;
            description "Identifier of the link";
         }
         uses link-channel-attributes;
      }
   }

   grouping link-channel-attributes {
      description
         "A link channel is one of the concatenated elements of
         the media channel.";
      leaf N {
         type int32;
         description
            "Is used to determine the Nominal Central Frequency.
            The set of nominal central frequencies can be built
            using the following expression:
               f = 193.1 THz + n x 0.00625 THz,
            where 193.1 THz is ITU-T ''anchor frequency'' for
            transmission over the C band, n is a positive or
            negative integer including 0.";
         reference "rfc7698";
      }
      leaf M {
         type int32;
            description
               "Is used to determine the slot width. A slot
               width is constrained to be M x SWG (that is,
               M x 12.5 GHz), where M is an integer greater than
               or equal to 1.";
            reference "rfc7698";
      }
      leaf source-node {
         type fg-ted:flexi-grid-node-ref;
         description "Source node of the link channel";
      }
      leaf source-port {
         type fg-ted:flexi-grid-node-port-ref;
         description "Source port of the link channel";
      }
      leaf destination-node {
         type fg-ted:flexi-grid-node-ref;
         description "Destination node of the link channel";
      }
```

```
      leaf destination-port {
         type fg-ted:flexi-grid-node-port-ref;
         description "Destination port of the link channel";
      }
      leaf link {
         type fg-ted:flexi-grid-link-ref;
         description "Link of the link channel";
      }
      leaf bidireccional {
         type boolean;
         description
            "Determines whether the link is bidireccional or
            not";
      }
   }
}
```

<CODE ENDS>

A.3. License

Copyright (c) 2016 IETF Trust and the persons identified as authors of the code. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- o Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

- o Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

- o Neither the name of Internet Society, IETF or IETF Trust, nor the names of specific contributors, may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Authors' Addresses

Jorge E. Lopez de Vergara
Universidad Autonoma de Madrid
Escuela Politecnica Superior
C/Francisco Tomas y Valiente, 11
E-28049 Madrid, Spain

Email: jorge.lopez_vergara@uam.es

Victor Lopez
Telefonica I+D/GCTO
Distrito Telefonica
E-28050 Madrid, Spain

Email: victor.lopezalvarez@telefonica.com

Oscar Gonzalez de Dios
Telefonica I+D/GCTO
Distrito Telefonica
E-28050 Madrid, Spain

Email: oscar.gonzalezdedios@telefonica.com

Daniel King
Lancaster University

Email: d.king@lancaster.ac.uk

Young Lee
Huawei Technologies

Email: leeyoung@huawei.com

Zafar Ali
Cisco Systems

Email: zali@cisco.com

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 9, 2017

Q. Wang, Ed.
ZTE
I. Hussain, Ed.
K. Pithewan
R. Valiveti
Infinera

July 8, 2016

Framework and Requirements for GMPLS-based Control of Flexible Ethernet Network

draft-wang-ccamp-flexe-fwk-02

Abstract

Flex Ethernet (FlexE) technology, which is defined by Optical Internetworking Forum (OIF), is a new kind of data plane technology and can be used to provides a generic mechanism for supporting a variety of Ethernet MAC rates that may or may not correspond to any existing Ethernet PHY rate. This includes MAC rates that are greater than (through bonding) and less than (through sub-rate and channelization) the standard Ethernet PHY rates.

Based on the applications/use cases given in the Flex Ethernet Implementation Agreement [FlexE-IA], this document defines a framework and control plane requirements for the application of existing GMPLS architecture and control plane protocols to the control of flexible Ethernet network. The actual extensions to the GMPLS protocols will be defined in companion documents.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 9, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

Traditionally, Ethernet MAC rates were constrained to match the rates of the Ethernet PHY(s). OIF recently approved Flex Ethernet (FlexE) Implementation Agreement [FlexE-IA], which can be used to provide a generic mechanism to support a variety of Ethernet MAC rates that may or may not correspond to any existing Ethernet PHY rate. In this kind of network scenario, FlexE uses more than one Ethernet PHYs as server layer and these Ethernet PHYs are bonded together as a FlexE group to carry FlexE client signal. The general capabilities supported by FlexE implementation includes:

Bonding of Ethernet PHYs, e.g., supporting a 200G MAC over two bonded 100GBASE-R PHYs.

Sub-rates of Ethernet PHYs, e.g., supporting a 50G MAC over a 100GBASE-R PHY.

Channelization within a PHY or a group of bonded PHYs, e.g, support a 150G and two 25G MACs over two bonded 100GBASE-R PHYs.

Note that hybrids are also possible, for example a sub-rate of a group of bonded PHYs, for example, a 250G MAC over three bonded 100GBASE-R PHYs. For more use cases, you can refer to [draft-hussain-ccamp-flexe-usecases].

In order to operate on the Ethernet PHYs, FlexE capable nodes uses a calendar to assign slot positions on sub-calendars on each PHY of the FlexE group for each of the FlexE clients. The calendar has a granularity of 5G, and has a length of 20 slots per 100G Ethernet PHYs.

Based on the FlexE Implementation Agreement [FlexE-IA], this document defines the framework for GMPLS-based control of flexible Ethernet network to depict the layer model of Flex Ethernet as well as a set of associated control-plane requirements.

Note: currently, ITU-T already include this part of work, such as, [ITU-T G.709] already include the content of mapping of FlexE Client signals into OPUflex using a new mapping method, and mapping of FlexE Aware signals into OPUflex. Also [ITU-T G.872] is going to include a description of FlexE, such as the layer model in different cases.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Overview of FlexE Networks

2.1. FlexE Terminology

This section describes the definitions of the terms used in FlexE networks. More details about these terms can be found in OIF Flex Ethernet (FlexE) Implementation Agreement [FlexE-IA].

FlexE group: the FlexE Group refers to a group of from 1 to n bonded Ethernet PHYs. This version of the Implementation Agreement supports FlexE groups composed of one or more bonded 100GBASE-R PHYs.

FlexE Client: a FlexE Client is an Ethernet flow based on a MAC data rate that may or may not correspond to any Ethernet PHY rate. The FlexE client MAC rates supported by this implementation agreement are 10, 40, and m * 25 Gb/s.

FlexE Shim: the FlexE Shim is the layer that maps or demaps the FlexE clients carried over a FlexE group. The FlexE mux refers to the transmit direction which maps the FlexE clients over the FlexE group. The FlexE demux refers to the receive direction which demaps the FlexE clients from the FlexE group.

2.2. Scenarios Supported by FlexE

According to the FlexE Implementation Agreement [FlexE-IA], FlexE can support a variety of cases. A non-exhaustive list includes:

One case of router to transport connection is where the transport network is unaware of FlexE. This may be used with legacy transport equipment that provides PCS-codeword transparent transport of 100GbE, but provides no special support for FlexE. In this case, all PHYs of the FlexE group are carried independently, but over the same fiber route, over the transport network.

Another case of router to transport connection is where the transport network equipment terminates the FlexE group. In the FlexE terminating case, FlexE group is terminated before crossing the transport network and FlexE client is extracted from the FlexE signal and then carried over the transport network.

The final router to transport example described is one where the transport network is aware that it is carrying FlexE PHYs (as opposed to 100GbE), but the FlexE group is not terminated on the transport equipment. Transport network equipment may "crunch" the PHY of the FlexE group by allowing bits or bytes to be discarded from the unavailable calendar slots at the transport network ingress and these bits or bytes re-inserted with fixed values at the transport network egress. This may be used to support cases where the Ethernet PHY rate is be greater than the wavelength rate, the wavelength rate is not an integral multiple of the PHY rate, or there is a reason (for example, wavelengths terminated on different transponder line cards) that it is not possible to terminate the FlexE group in the transport equipment. This kind of equipment is a kind of special transport equipment which can support partial-rate transport.

2.3. FlexE Layer Model

Based on the cases addressed in section 2.2, FlexE has different kinds of mapping hierarchy accordingly. This section gives some description of FlexE layer model in different cases. Figure 1 depicts a FlexE layered network scenario. In this network, B and E are FlexE capable nodes, C and D are OTN ODUflex/ODU4 capable nodes. Node B, C are mainly used to encapsulate the client layer signal into the server layer, while node D, E are mainly used to extract the client layer signal from the server layer signal.

As defined in FlexE Implementation Agreement, a FlexE client may be generated internally within a system, received from an Ethernet PHY or from another FlexE shim. In this network scenario, we suppose the FlexE client is generated in router B.

Feature of cases can be found in section 3.2.

In all the following cases, we suppose FlexE client at node B has a path setup request from source B to destination E.

```
+---+                       +---+
| B |                       | E |
+---+                       +---+
    \                     /
     \                   /
     +---+         +---+
     | C |---------| D |
     +---+         +---+
```

Figure 1: FlexE Layer Network

2.3.1. Layer Model in FlexE Unaware Case

In this case which is depicted in Figure 2, there exist four network layers. FlexE client layer represents an end-to-end connection, which is from the source B to destination E. When the FlexE client signal is generated inside node B, the FlexE client signal is first mapped into the slots of FlexE, then the FlexE signal is carried by Ethernet PHYs towards the destination E. When the Ethernet PHYs arrive at node C, each PHY will be mapped into a separate ODU4 connection and then forwarded across the OTN network towards the ODU layer connection destination D.

Note: in this case, more than one FlexE clients can be carried by FlexE layer.

Four layers exist in this case, and the mapping hierarchy between node C and node D can be seen in Figure 3.

```
             Ethernet Client Connections
     |------------------------------------------|
                  FlexE Connection
      |----------------------------------------|
    +---+                                    +---+
    | B |          PHY Connections           | E |
    +---+|--------------------------------|+---+
         \                                /
          \                              /
           +---+  ODU4 Connections  +---+
           | C |--------------------| D |
           +---+                    +---+
```

Figure 2: FlexE Unaware Layer Network

```
+-------------------+
| Ethernet Client   |
+-------------------+
|       FlexE       |
+-------------------+
|  PHY    |  PHY    |
+-------------------+
| ODU4    |  ODU4   |
+-------------------+
```

Figure 3: FlexE Unaware Layer Hierarchy

2.3.2. Layer Model in FlexE Terminating Case

In this case, FlexE client layer represents an end-to-end connection, which is from the source B to destination E. When the FlexE client signal is generated inside node B,, the Ethernet signal is first mapped into the slots of FlexE, then the FlexE signal is carried by Ethernet PHYs towards the destination C. When the FlexE signal arrives at node C, node C first extracts the FlexE client signal, then maps the Ethernet client signal into ODU signal and forwards across the OTN network towards destination node D. Node D will first extract the FlexE client signal from the ODU signal, then map the Ethernet client signal into FlexE signal, which will then be carried by Ethernet PHYs towards destination node E.

Two segments of FlexE connection exist in this case. one is from node B to node C, and the other is from node D to node E.

```
                 Ethernet Client Connection
           |----------------------------------------|
          +---+                                    +---+
          | B |                                    | E |
          +---+                                    +---+
PHY Connections\FlexE Connection                  /
                \                                /
                 +---+   ODU Connection   +---+
                 | C |--------------------| D |
                 +---+                    +---+
```

Figure 4: FlexE Terminating Layer Network

Two kinds of mapping hierarchy exist. For the signal transferred on the links between B and C, D and E, the mapping hierarchy is depicted in Figure 5. For the signal transferred on the links between C and D, the mapping hierarchy is depicted in Figure 6.

```
+-------------------+
| Ethernet Client   |
+-------------------+
|      FlexE        |
+-------------------+
|  PHY    |  PHY    |
+-------------------+
```

Figure 5: Mapping Hierarchy of FlexE

```
+-------------------+
| Ethernet Client   |
+-------------------+
|       ODU         |
+-------------------+
```

Figure 6: Mapping Hierarchy on OTN Link

2.3.3. Layer Model in FlexE Aware Case

FlexE client layer represents an end-to-end connection, which is from the source B to destination E. When the FlexE client signal is generated inside node B,, the Ethernet signal is first mapped into the slots of FlexE, then the FlexE signal is carried by Ethernet PHYs towards the destination E. When the FlexE signal arrives at node C, node C will first discards unavailable slots, then transfers the remaining FlexE slots to ODU Connection. According to the description in [ITU-T G.709], these FlexE slots are carried across the OTN network via other ODUflex signals carried in other ODUCn/OTUCn/OTSiA signals.

In this scenario, Ethernet PHYs connection exist between node B and node C, node D and node E.

```
                       Ethernet Client Connection
                 |-----------------------------------------|
                           FlexE Connection
                  |----------------------------------------|
                +---+                                    +---+
                | B |                                    | E |
                +---+                                    +---+
      PHY Connections\                                  /
                      \                                /
                       +---+ ODUFlex Connection +---+
                       | C |--------------------| D |
                       +---+                    +---+
```

Figure 7: FlexE Aware Layer Network

Two kinds of mapping hierarchy exist. For the signal transferred on the links between B and C, D and E, the mapping hierarchy can be seen in Figure 8. For the signal transferred on the links between C and D, the mapping hierarchy can be seen in Figure 9.

```
                       +-------------------+
                       | Ethernet Client   |
                       +-------------------+
                       |      FlexE        |
                       +-------------------+
                       |  PHY    |  PHY    |
                       +-------------------+
```

Figure 8: Mapping Hierarchy of FlexE

```
+------------------+
| Ethernet Client  |
+------------------+
|      FlexE       |
+------------------+
|     ODUFlex      |
+------------------+
```

Figure 9: Mapping Hierarchy on OTN Link

3. GMPLS Applicability

The goal of this section is to provide an insight into the application of GMPLS as a control mechanism in FlexE networks. Specific control-plane requirements for the support of FlexE networks are covered in Section 4. This section aims to describe the modelling of controlling the FlexE shim layer specific attributes in different network scenario based on the capability of FlexE described in OIF Flex Ethernet (FlexE) Implementation Agreement. [FlexE-IA]

3.1. General Considerations

The GMPLS control of the FlexE layer deals with the establishment of FlexE connections that are transferred in FlexE capable nodes. GMPLS labels are used to locally represent the FlexE connections and its associated slots assignment information for client.

3.2. Consideration of FlexE LSPs

The FlexE LSP is a control-plane representation of a FlexE Connection and MUST be carried by Ethernet PHYs LSP or ODU LSP in the network.

Figure 4 depicts a scenario that the FlexE LSP is carried over Ethernet PHYs LSP between node B and node C, node D and node E. When the Ethernet client signal arrives at node B, node B first check if there are enough Ethernet PHYs available for setting up FlexE LSP. If no, node B will first set up Ethernet PHYs LSP from node B to node C, and then set up the FlexE LSP over the Ethernet PHYs LSP. This process involves two signalling procedures, one is to set up Ethernet PHYs, and the other is used to set up FlexE LSP over the Ethernet PHYs. The set-up signalling of FlexE LSP includes the allocation of resource for Ethernet client.

Figure 7 depicts a scenario that the FlexE LSP is carried over ODU LSP between node C and node D. This scenario is different, and is used to support cases where the Ethernet PHY rate is be greater than the wavelength rate, the wavelength rate is not an integral multiple of the PHY rate. Node C and node D support the partial-rate ability.

When the FlexE LSP over Ethernet PHYs arrives at node C, node C first check if there is enough resource for carrying the FlexE LSP signal across the transport network. If no, node C will check if there is enough resource for carrying FlexE LSP signal after discarding the unavailable slots. If yes, node C will first set up the ODUFlex LSP to node D, and then continue the signalling process of FlexE LSP across the transport network.

3.3. Control-Plane Modelling of FlexE Network Elements

FlexE is a new kinds of transport technology, which has many new constraints. These constraints are listed below:

Unavailable slots: this is different from "unused" slot, in that it is known, due to transport network constraints, that not all of the calendar slots generated from the FlexE mux will reach the FlexE demux and therefore no FlexE client should be assigned to those slots. As defined in the Flex Ethernet Implementation Agreement, unavailable slots are always at the end of the sub-calendar configuration for the respective PHY.

Unused slots: unused slots can be allocated to Ethernet client as available resource.

Partial-rate capability: the partial-rate capability is usually supported by an OTN access equipment. If an equipment supports partial-rate, it means this equipment has the capability of discarding unavailable slots and transfers the left slots across OTN transport network.

Slot granularity: currently, only one kinds of 5G slot granularity is defined in OIF Flex Ethernet (FlexE) Implementation Agreement.

3.4. FlexE Layer Resource Allocation Considerations

FlexE LSP transfers based on the slot information, so it SHOULD be able to expose the unused slot resource information towards the client layer. Besides the slot information, there are also some other attributes that need to be specified when allocating resource. In GMPLS-controlled system, these information should be taken into consideration as a label when forwarding.

FlexE group number: a bunch of Ethernet PHYs can be bounded together and used as a whole as one FlexE LSP. FlexE LSPs between the same source and destination equipment SHOULD NOT have the same FlexE group number. Source equipment and destination equipment SHOULD be aware of the existing of different FlexE groups and which Ethernet PHYs are in which FlexE group.

PHY Number: it’s a dynamic and logical number that is assigned through control plane or management plane, which is unique within the context of (source, destination), and has a one-to-one correlation with physical port. This information will also be carried in the FlexE overhead. Source equipment and destination equipment SHOULD negotiate a value for every Ethernet PHYs within one FlexE group.

Slot Assignment information: the FlexE LSP transfers based on the slot positions, so the equipment SHOULD be able to tell which slot is assigned to which client.

Partial-rate: during the process of resource allocation, where the partial-rate would happen should be indicated.

Granularity: currently, only one kinds of 5G slot granularity is defined in OIF Flex Ethernet (FlexE) Implementation Agreement [FlexE-IA].

3.5. Neighbour Discovery and Link Property Correlation

There are potential interworking problems between different FlexE capable equipment. Devices or equipments might not be able to support the interworking of every slot due to the constraints of transport network equipment or other constraints. In this case, two directly connected FlexE capable equipments SHOULD run the neighbour discovery process and correlate the link property to make sure which slots are unavailable, which slots can be used by the client. Neighbour discovery protocol can be communicated in in-band FlexE section management channel, and also can be communicated through out-of-band management channel.

3.6. Routing and Topology Dissemination

The topology and routing information is used by the path computation entity to compute an end-to-end path. Besides the basic interconnected information, there are also some FlexE specific attributes that should be taken into consideration.

Partial-rate: partial-rate capability is a special feature which allows an equipment to discard unavailable slots and transfers the left slots across OTN transport network. Path computation entity is more likely to compute a feasible path if this capability is taken into consideration when computing path.

Unavailable slot information: this information is used to indicate certain slots SHOULD not be considered when computing an end-to-end path. The unavailable slots can not be used to forward signal because of the transport constraints.

Unused slot information: unused slot can be allocated to the path as available resource.

3.7. Resizing of FlexE

Currently, OIF has many contributions about the FlexE at the PHY level and intends to do the resizing process through data plane negotiation scheme. This framework draft will include this requirement later when the work is stable in OIF.

4. Control-Plane Requirements

The control of FlexE networks brings some new additional requirements to the GMPLS protocols. This section summarizes those requirements for signalling,routing and Link management protocol.

4.1. Support for Signalling of FlexE

Aim of the signaling is to set up an end-to-end LSP for FlexE signal.

The signalling procedures shall be able to assign FlexE releated attributes for an LSP, which include FlexE group number for a FlexE LSP. This FlexE group number is unique and can be used to indicate a group of Ethernet PHYs bonded together.

The signalling procedures shall be able to assign an unique PHY number for each bonded Ethernet PHY, and a correlation relationship SHOULD also be indicated between the assigned PHY number and real physical port number when signalling.

The signalling procedures shall be able to configure the slots information allocated for a FlexE LSP.

The Signalling procedures shall be able to indicate the palace where partial-rate mapping happens.

4.2. Support for Routing of FlexE

The routing protocol extensions are mainly based on the functionality that is described in [RFC4202] and these extensions are made to fit into FlexE network.

The routing protocol SHALL distribute sufficient information to compute paths to enable the signalling procedure to establish LSPs as described in the previous sections.

The routing protocol SHALL update its advertisements of available resources and capabilities to include the partial-rate support information and unused slot information on each Ethernet PHY port.

4.3. Support for Neighbour Discovery and Link Property and Link Correlation

The control plane MAY include support for neighbour discovery such that a FlexE network can be constructed in a "plug-and-play" manner.

The control plane SHOULD allow the nodes at opposite ends of a link to correlate the properties that they will apply to the link. Such a correlation SHOULD include at least the identities of the nodes and the identities that they apply to the link. Other FlexE specific properties, such as the link characteristics of unavailable slot information, SHOULD also be correlated. Such neighbour discovery and link property correlation, if provided, MUST be able to operate in both in-band and out-of-band manner.

5. Security Considerations

TBD

6. Manageability Considerations

TBD

7. References

7.1. Normative References

[FlexE-IA]
Stephen, J. and David. R. Stauffer, "FlexE Implementation Agreement", 2016.

[I-D.hussain-ccamp-flexe-usecases]
Hussain, I., Valiveti, R., and K. Pithewan, "FlexE Usecases", draft-hussain-ccamp-flexe-usecases-01 (work in progress), July 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <http://www.rfc-editor.org/info/rfc3209>.

[RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <http://www.rfc-editor.org/info/rfc3471>.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <http://www.rfc-editor.org/info/rfc3473>.

[RFC3603] Marshall, W., Ed. and F. Andreasen, Ed., "Private Session Initiation Protocol (SIP) Proxy-to-Proxy Extensions for Supporting the PacketCable Distributed Call Signaling Architecture", RFC 3603, DOI 10.17487/RFC3603, October 2003, <http://www.rfc-editor.org/info/rfc3603>.

[RFC4202] Kompella, K., Ed. and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, DOI 10.17487/RFC4202, October 2005, <http://www.rfc-editor.org/info/rfc4202>.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <http://www.rfc-editor.org/info/rfc4203>.

[RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, DOI 10.17487/RFC4204, October 2005, <http://www.rfc-editor.org/info/rfc4204>.

7.2. Informative References

[RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, DOI 10.17487/RFC3945, October 2004, <http://www.rfc-editor.org/info/rfc3945>.

Authors’ Addresses

Qilei Wang (editor)
ZTE
Nanjing
CN

Email: wang.qilei@zte.com.cn

Iftekhar Hussain (editor)
Infinera
Sunnyvale
USA

Email: IHussain@infinera.com

Khuzema Pithewan
Infinera
Sunnyvale
USA

Email: kpithewan@infinera.com

Radha Valiveti
Infinera
Sunnyvale
USA

Email: rvaliveti@infinera.com

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 9, 2017
July 8, 2016

Q. Wang, Ed.
ZTE

RSVP-TE Signaling Extensions in support of Flexible Ethernet networks
draft-wang-ccamp-flexe-signaling-02

Abstract

This draft describes the extensions to the Resource Reservation Protocol Traffic Engineering (RSVP-TE) signalling protocol to support Label Switched Paths (LSPs) in a GMPLS-controlled flexible Ethernet network.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 9, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

Traditionally, Ethernet MAC rates were constrained to match the rates of the Ethernet PHY(s). OIF recently approved Flex Ethernet (FlexE) Implementation Agreement [FlexE-IA], which can be used to provide a generic mechanism to support a variety of Ethernet MAC rates that may or may not correspond to any existing standard Ethernet PHY rate. In this kind of network scenario, FlexE uses more than one Ethernet PHYs as server layer through inverse multiplexing and these Ethernet PHYs are bonded together as a FlexE group to carry FlexE client signal. The general capabilities supported by FlexE implementation includes:

Bonding of Ethernet PHYs, e.g., supporting a 200G MAC over two bonded 100GBASE-R PHYs.

Sub-rates of Ethernet PHYs, e.g., supporting a 50G MAC over a 100GBASE-R PHY.

Channelization within a PHY or a group of bonded PHYs, e.g, support a 150G and two 25G MACs over two bonded 100GBASE-R PHYs.

Note that hybrids are also possible, for example a sub-rate of a group of bonded PHYs, for example, a 250G MAC over three bonded 100GBASE-R PHYs. For more use cases, you can refer to [draft-hussain-ccamp-flexe-usecases].

In order to operate on the Ethernet PHYs, FlexE capable nodes uses a calendar to assign slot positions on sub-calendars on each PHY of the FlexE group for each of the FlexE clients. The calendar has a granularity of 5G, and has a length of 20 slots per 100G of Ethernet PHYs.

[FLEXE-FWK] defines a framework and the associated control plane requirements for the Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] based control of FlexE networks.

Based on the requirements described in FlexE framework document, this document defines additional requirements and protocol extensions to Resource Reservation Protocol-Traffic Engineering (RSVP-TE) [RFC3473] to set up FlexE LSPs.

1.1. FlexE Terminology

For terminology related to flexible Ethernet, please refer to [FLEXE-FWK] and FlexE Implementation Agreement [FlexE-IA].

1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Requirements for FlexE Signalling

The requirements for establishing LSPs in a FlexE network is described in [FLEXE-FWK].

The FlexE LSP is a control-plane representation of a FlexE Connection and MUST be carried by standard Ethernet PHY LSPs or ODU LSPs in the network. So in order to set up an end-to-end FlexE LSP, Ethernet PHY LSPs or ODU LSPs MUST be set up in advance. [FLEXE-FWK] gives a description of FlexE layer resource information needed to be reserved when signalling an end-to-end LSP, which include the identifier information used for indicate one specific FlexE LSP and resource assignment information for FlexE client. This section gives a detailed description of these identifier and resource information. Based on these information, source equipment and destination equipment will be able to map and demap the FlexE clients from the FlexE group properly.

2.1. FlexE Group

As defined in FlexE Implementation Agreement, the FlexE Group refers to a group of from 1 to n bonded 100GBASE-R Ethernet PHYs. Ethernet PHYs are bounded together and used as a whole through inverse multiplexing by the FlexE LSP. FlexE LSPs between the same source and destination equipment SHOULD NOT have the same FlexE group number. Source equipment and destination equipment SHOULD be aware of the existing of different FlexE group and which Ethernet PHYs are in which FlexE group.

This FlexE group information MUST be carried in the generalized label of signalling message during LSP establishment if FlexE group is needed.

2.2. PHY Number

The PHY number is a dynamic and logical number that is assigned by control plane or management plane, which is unique within a pair of source and destination, and has a one-to-one correlation with physical port. This PHY number information will also be carried in the FlexE overhead. Source equipment and destination equipment SHOULD negotiate a value for every Ethernet PHYs within one FlexE group.

The PHY number information MUST be carried in the generalized label of signalling message during LSP establishment. Besides the PHY number carried in the generalized label, RSVP_HOP object MUST also be used to indicate the correlation between PHY number and physical port number. The sequence of the PHY numbers listed in the generalized label SHOULD be in accordance with the physical ports carried in RSVP_HOP object.

2.3. Partial-rate

The partial-rate capability is usually supported by an OTN access equipment. If an equipment supports partial-rate, it means this equipment has the capability of discarding unavailable slots and forwards the remaining slots across OTN transport network. During the process of resource allocation, where the partial-rate would happen should be indicated.

2.4. Slot Assignment

The FlexE LSP forwards client signal based on the slot positions, so the equipment SHOULD be able to tell which slot is assigned to which client according to the generalized label carried in the signalling message. This attribute SHOULD also take the unavailable slots information into consideration.

2.5. Granularity

Currently, only one kinds of 5G slot granularity is defined in OIF Flex Ethernet (FlexE) Implementation Agreement [FlexE-IA]. During signalling process, this information can be inferred through the bandwidth parameters and slot number information within one Ethernet PHY.

3. Protocol Extensions

This section defines the extensions to RSVP-TE signalling for GMPLS [RFC3473] to support FlexE networks.

3.1. Traffic Parameters

In RSVP-TE, the SENDER_TSPEC object in the Path message characterizes the traffic parameters for the data flow from the corresponding sender. The FLOWSPEC object in the Resv message indicates the actual resource reservation. As defined in [RFC3473], bandwidth encodings are carried in the SENDER_TSPEC and FLOWSPEC objects, and these values are set in the Peak Data Rate field of Int-Serv objects, see [RFC2210]. Other bandwidth/service related parameters in the object are ignored and carried transparently. Signalling procedure of RSVP-TE used in FlexE network can also reuse the SENDER_TSPEC object defined in [RFC2210] to describe the traffic parameters for the data flow of the sender.

3.2. Generalized Label

In the case of FlexE network, the GMPLS labels are used to locally represent the FlexE connections and its associated slots transferring. Parameters defined in section 3 are needed to be carried in the generalized label to represent the FlexE connections.

The following is the GENERALIZED_LABEL object format that is used with the TDM Switching Type:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         FlexE Group Number            |        Reserved       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Flags             |       Client Indicator        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  PHY Number   | Slots Assignment information (bitmap) |   U   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  PHY Number   | Slots Assignment information (bitmap) |   U   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               ......                                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: FlexE Label

FlexE Group Number: 20 bits

The FlexE Group refers to a group of from 1 to n bonded 100GBASE-R Ethernet PHYs.

Flags:

Currently, only one flag is allocated to indicate the configuration of "A" calendar or "B" calendar.

Client Indicator:

This field is used to represent a value which can be filled in the client calendar of the FlexE overhead, and the client calendar is used to indicate which of the FlexE clients is mapped into a given calendar slot in the A and B calendar configurations for the sub-calendar carried over that PHY. For slots in any PHYs with the same value, they belong to the same FlexE client.

PHY number: 8 bits

The PHY number is a dynamic and logical number that is assigned through control plane or management plane, which is unique within the context of (source, destination), and has a one-to-one correlation with physical port. This information will also be carried in the FlexE overhead.

Slots Assignment information (bitmap) : 20 bits

This attribute is used to indicate slots assignment information, including slots assigned for the client, which are indicated by the bit set to"1" and slots unused, which are indicated by the bit set to "0".

U field: 4 bits

This field is used to indicate the number of unavailable slot. As defined in OIF FlexE Implementation Agreement, unavailable slots are always at the end of the sub-calendar configuration for the respective PHY, so this draft uses a specific field to describe them.

Note: the number of the Ethernet PHY used by FlexE can be referred from the number of the "PHY number" in the generalized label.

3.3. Flag extensions in Hop Attributes TLVs

The Attribute Flags TLV defined in [RFC5420] is carried in an ERO Hop Attributes subobject. Flags set in the Attribute Flags TLV [RFC5420] carried in an ERO Hop Attributes subobject SHALL be interpreted in the context of the received ERO. Only a subset of defined flags are defined as valid for use in Attribute Flags TLV carried in an ERO Hop Attributes subobject. Invalid flags SHALL be silently ignored.

A new bit in the Attribute Flags TLV is assigned to indicate the partial-rate mux and demux. This ERO Hop Attributes subobject MUST come in pairs. The node which do the partial-rate mux MUST check the existence of partial-rate demux flag in the ERO Hop Attributes subobject of the path message. If it does not exist, path will not be set up successfully.

3.4. FlexE Link Available Slot Number TLV

As indicated in [FlexE-IA], the number of PHY used by FlexE at both ends of the OTN network SHOULD be the same in aware case, or the OTN may encounter demux problem due to the format of FlexE 66 bits overhead blocks. So in this link available slot number TLV, there should have a number of PHY field to collect the number of the PHY needed at both ends of the OTN network.

Also a new FlexE Link Available Slot Number TLV is extended in the LSP_ATTRIBUTES object to calculate the number of end-to-end available slots. This TLV contains several 1 byte fields which correspond to a Ethernet PHY, and it is used to collect the number of available slots can be used at most along an end-to-end LSP.

3.5. Signalling Procedure

This section gives description of FlexE layer model in different cases. Figure 2 depicts a FlexE layered network scenario. In this network, B and E are FlexE capable nodes, C and D are OTN ODUflex/ODU4 switch capable nodes. Node B, C are mainly used to encapsulate the client layer signal into the server layer, while node D, E are mainly used to extract the client layer signal from the server layer signal.

```
+---+                       +---+
| B |                       | E |
    +---+                       +---+
     \                       /
      \                     /
       +---+          +---+
       | C |----------| D |
       +---+          +---+
```

Figure 2: FlexE Layer Network

3.5.1. Procedure for Setting up FlexE LSP in Unaware Case

In this case, node C and node D are unaware of the FlexE signal.

Suppose node B receives a FlexE client path set up request from node B to node E and the bandwidth of this path is 150G. There are three Ethernet PHYs between B and C, D and E. Also there exist enough ODU4 connections between node C and node D to bear the traffic from node B. Following is the signalling procedure.

a. Node B intends to send the RSVP-TE path message to set up an end-to-end path towards the destination node E. The bandwidth requirement is 150G.

b. Before node B begins to forward the FlexE client path message, it will first determine a new FlexE path needs to be set up from node B to node E to carry the Ethernet client traffic by comparing the switching capability carried in the FlexE client path message and the switching capability that node B can support. Node B first blocks the Ethernet client path message, then initiate another new path message to set up the FlexE path from node B to node E. The requested switching capability of the FlexE path is set to TDM and the encoding type is set to "FlexE-LSP".

Before node B send the FlexE path message towards the next hop, node B first check if there are enough Ethernet PHYs to carry the FlexE traffic. Then node B will start the set up of two Ethernet PHYs LSP from node B to node E. The Ethernet PHY path messages are then sent towards the next hop node C. Two Ethernet PHY LSPs will be set up.

c. When node C receives the Ethernet PHY path messages from node B, it will first determine two new ODU4 path needs to be set up first from node C to node D to carry the Ethernet PHY traffic by comparing the switching capability carried in the path message and the switching capability that node C can support. Node C first blocks the Ethernet PHY path message, then initiate another new path messages to set up the ODU4 path from node C to node D.

Considering the set up of ODU4 path is not the focus of this draft, procedure of setting up ODU4 paths are not going to be described in this draft. Node C will receive the Resv message from node D and confirm the successful set up of ODU4 paths. The ODU4 LSPs behave as the server layer of Ethernet PHY paths.

d. Node C will then continue sending the Ethernet PHY path messages towards node E to finish the set up of Ethernet PHY LSPs. The Ethernet PHY LSPs from node B to node E behave as the a link with only one hop.

e. After node B receives the Resv message from node E and confirm the successful set up of Ethernet PHY paths, node B will continue sending the FlexE path message towards the next hop node E. The bandwidth requirement is 150G.

f. Node E receives the FlexE path message from node B. Considering there are already two Ethernet PHYs from node B to node E, node E determines to set up the FlexE path over the two Ethernet PHYs by carrying the assigned FlexE group number, dynamic PHY number and slot positions for the Ethernet client in the generalized label.

Besides the generalized label, RSVP_HOP object MUST also be used to indicate the correlation between PHY number and physical port number. The sequence of the PHY numbers listed in the generalized label SHOULD be in accordance with the physical ports number carried in RSVP_HOP object.

Node E then send the Resv message toward the FlexE path source, node B, to finish the set up of FlexE path.

g. After node B receive the Resv message from node E and confirm the successful set up of FlexE path, node B will continue sending the blocked FlexE client path message towards the destination node E to finish the set up of the client path.

FlexE LSP appears as an Ethernet client link once it was set up. Unused bandwidth on the FlexE LSP can be used further by another FlexE client.

3.5.2. Procedure for Setting up FlexE LSP in Aware Case

In this case, node C and node D are aware of the FlexE signal.

Suppose node B receives a FlexE client path set up request from node B to node E and the bandwidth of this path is 150G. There are three Ethernet PHYs between B and C, D and E. Also there exist 180G ODUFlex connections between node C and node D to bear the traffic from node B. The number of unavailable slots between B and C is 4, between D and E is 5. Following is the signalling procedure.

a. Node B intends to send the RSVP-TE path message to set up an end-to-end path towards the destination node E. The bandwidth requirement is 150G.

b. Before node B begins to transfer the FlexE client path message, it will first determine a new FlexE path needs to be set up from node B to node E to carry the FlexE client traffic by comparing the switching capability carried in the FlexE client path message and the switching capability that node B can support. As node B is aware of the existence of unavailable slots between node B and node C, node D and node E, Node B’s computation element will compute an end-to-end partial-rate supported path B-C-D-E. Node B then first blocks the Ethernet client path message, then initiate another new path message to set up the FlexE path from node B to node E. The requested switching capability of the FlexE path is set to TDM and the encoding type is set to "Partial-rate FlexE-LSP".

c. Before node B send the FlexE path message towards the next hop, node B first check if there are enough Ethernet PHYs to carry the FlexE traffic. Then node B will start the set up of two Ethernet PHYs LSP from node B to node E with the G-PID field set to "Partial-rate FlexE-LSP". The Ethernet PHY path messages are then sent towards the next hop node C. Two Ethernet PHY LSPs will be set up.

d. When node C receives the Ethernet PHY path messages which are used to support partial-rate FlexE-LSP from node B, it will first check if itself support partial-rate capability as a result of the "Partial-rate FlexE-LSP" carried in G-PID. Node C then sends the Resv message to node B to finish the set up of Ethernet PHYs LSP if it confirms the support of partial-rate capability.

e. When node B receives the Ethernet PHYs resv message from node C, it will first finish the set up of Ethernet PHYs LSP, then continue sending the FlexE path message towards next hop node C carrying the FlexE Link Available Slot Number TLV to record the number of PHYs and available slots on the link between node B and node C. Before the sending of FlexE path message, neighbor discovery process needs to be run to negotiate the number of unavailable slots. These Ethernet PHY LSPs behave as a one hop FlexE link in the FlexE network.

f. When node C receives the FlexE path message, node B first block the FlexE path message, then checks if there are ODUFlex LSP available to carry the FlexE traffic. If no, node B first initiate the ODUFlex LSP set up and the bandwidth requested does not take the unavailable slots into consideration.

Considering the set up of ODUFlex path is not the focus of this draft, procedure of setting up ODUFlex paths are not going to be described in this draft. Node C will receive the Resv message from node D and confirm the successful set up of ODUFlex paths. The ODUFlex LSPs behave as a one hop FlexE link in the FlexE network.

g. When node C receives the ODUFlex Resv message from node D. It will first finish the set up of ODUFlex LSP, then continue sending the FlexE path message towards next hop node D. Node C first compares the currently number carried in the path message with the number of the available slots supported on link between node C and node D, then fills in the FlexE Link Available Slot Number TLV with the smaller one. These ODUFlex LSPs behave as a one hop FlexE link in the FlexE network and this link carries all FlexE slots except unavailable slots.

h. When node D receives the FlexE path message, it will repeat the procedure in c, d, e. Node E will receive the FlexE path message sent from node B and it will construct the Resv message sent back to node B to finish the FlexE client resource reservation along the path.

Note: in this case, if node D is aware of the number of PHY needed between node B, node C and node D, node E is different, it will trigger a path error message and send back to node B to indicate the larger one.

i. After node B receives the FlexE Resv message, it first finish the path set up and then continue the blocked FlexE client signalling procedure.

4. IANA Considerations

In this draft, two new encoding types "FlexE-LSP" and "Partial-rate FlexE-LSP" are assigned for FlexE LSP.

5. Security Considerations

TBD

6. Manageability Considerations

TBD

7. References

7.1. Normative References

[FlexE] Stephen, J. and David. R. Stauffer, "FlexE Implementation Agreement", 2016.

[I-D.hussain-ccamp-flexe-usecases] Hussain, I., Valiveti, R., and K. Pithewan, "FlexE Usecases", draft-hussain-ccamp-flexe-usecases-01 (work in progress), July 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <http://www.rfc-editor.org/info/rfc3209>.

[RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <http://www.rfc-editor.org/info/rfc3471>.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <http://www.rfc-editor.org/info/rfc3473>.

7.2. Informative References

[RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, DOI 10.17487/RFC3945, October 2004, <http://www.rfc-editor.org/info/rfc3945>.

Author's Address

Qilei Wang (editor)
ZTE
Nanjing
CN

Email: wang.qilei@zte.com.cn

CCAMP Working Group   Xian Zhang
Internet-Draft   Baoquan Rao
Intended status: Standards Track   Huawei
                                                           Anurag Sharma
                                                                Infinera
                                                             Xufeng Liu
                                                                Ericsson

Expires: January 06, 2017   July 08, 2016

A YANG Data Model for Layer 1 (ODU) Network Topology

draft-zhang-ccamp-l1-topo-yang-03.txt

Abstract

A transport network is a server-layer network designed to provide connectivity services for a client-layer network to carry the client traffic opaquely across the server-layer network resources. A transport network may be constructed from equipment utilizing any of a number of different transport technologies such as the evolving optical transport infrastructure (Synchronous Optical Networking (SONET) / Synchronous Digital Hierarchy (SDH) and Optical Transport Network (OTN)) or packet transport as epitomized by the MPLS Transport Profile (MPLS-TP).

This draft describes a YANG data model to manipulate the topologies of a layer 1 Optical Transport Network (OTN). It is independent of control plane protocols and captures topology related information pertaining to OTN and also enables manipulation (e.g., obtaining) of an OTN network.

Status of this Memo

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

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

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

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

This Internet-Draft will expire on Janruary 6th, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

A transport network is a server-layer network designed to provide connectivity services for a client-layer network to carry the client traffic opaquely across the server-layer network resources. A transport network may be constructed from equipment utilizing any of a number of different transport technologies such as the evolving optical transport infrastructure (Synchronous Optical Networking (SONET) / Synchronous Digital Hierarchy (SDH) and Optical Transport Network (OTN)) or packet transport as epitomized by the MPLS Transport Profile (MPLS-TP).

This document defines a data model of a layer one network topology, using YANG [RFC6020]. The model can be used by an application exposing to a management system. Moreover, it can also be used by an application in the following ways (but not limited to):

- o to obtain a whole view of the network topology information of its interest e.g., via a network element or maybe a controller or a network management system within the network ;

- o to receive notifications with regard to the information of the change of the network topology of its interest;

- o to enforce the establishment/update of a network topology with the characteristic specified in the data model, e.g., by a network controller or a client controller to manipulate the network provided by the provider for flexible control and management;

The YANG model defined in this draft is independent of control plane protocols and captures topology related information pertaining to an Optical Transport Networks (OTN)-electronic layer and also enables manipulation of an OTN network. Other network layers, such as fixed Dense Wavelength Switched Optical Network (WSON) and flexible optical networks (a.k.a., flexi-grid networks) are covered in [WSON-YANG] and [Flexi-YANG], respectively.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

3. Terminology and Notations

A simplified graphical representation of the data model is used in this document. The meaning of the symbols in the YANG data tree presented later in this draft is defined in [ietf-netmod-rfc6087bis]. They are provided below for reference.

o Brackets "[" and "]" enclose list keys.

o Abbreviations before data node names: "rw" means configuration (read-write) and "ro" state data (read-only).

o Symbols after data node names: "?" means an optional node, "!" means a presence container, and "*" denotes a list and leaf-list.

o Parentheses enclose choice and case nodes, and case nodes are also marked with a colon (":").

o Ellipsis ("...") stands for contents of subtrees that are not shown.

4. YANG Data Model for Layer 1 Topology

4.1. YANG Tree

```
module: ietf-odu-topology
augment /nd:networks/nd:network/nd:network-types/tet:te-topology:
   +--rw l1-network!
augment /nd:networks/nd:network:
   +--rw name?   string
augment /nd:networks/nd:network/nd:node:
   +--rw name?   string
augment /nd:networks/nd:network/nd:node/lnk:termination-
point/tet:te/tet:config:
   +--rw client-facing?         empty
   +--rw tpn?                   uint16
   +--rw tsg?                   identityref
   +--rw protocol-type?         identityref
   +--rw fec-enabled?           boolean
   +--rw adaptation-type?       adaptation-type
   +--rw sink-adapt-active?     boolean
   +--rw source-adapt-active?   boolean
   +--rw timeslots
      +--rw values*   uint8
augment /nd:networks/nd:network/nd:node/lnk:termination-
point/tet:te/tet:state:
   +--ro client-facing?         empty
   +--ro tpn?                   uint16
   +--ro tsg?                   identityref
   +--ro protocol-type?         identityref
```

```
   +--ro fec-enabled?           boolean
   +--ro adaptation-type?       adaptation-type
   +--ro sink-adapt-active?     boolean
   +--ro source-adapt-active?   boolean
   +--ro timeslots
      +--ro values*   uint8
augment /nd:networks/nd:network/lnk:link/tet:te/tet:config:
   +--rw odu-type?   identityref
   +--rw distance?   uint32
augment /nd:networks/nd:network/lnk:link/tet:te/tet:state:
   +--ro odu-type?   identityref
   +--ro distance?   uint32
augment /nd:networks/nd:network/nd:node/tet:te/tet:tunnel-termination-
point/tet:state:
   +--ro odu-Type?   identityref
augment /nd:networks/nd:network/lnk:link/tet:te/tet:config/tet:te-link-
attributes/tet:schedules/tet:schedule:
   +--rw odu-type?     identityref
   +--rw oduflex-bw?   uint32
augment /nd:networks/nd:network/lnk:link/tet:te/tet:state/tet:te-link-
attributes/tet:schedules/tet:schedule:
   +--ro odu-type?     identityref
   +--ro oduflex-bw?   uint32
```

4.1.1. Augmentation

As can be seen, from the tree shown in Section 4,1, the YANG module presented augments from a more generic Traffic Engineered (TE) network topology model, i.e., the ietf-te-topology YANG module as specified in [draft-ietf-teas-yang-te-topo-04].

Note the model in this draft uses the YANG model named ietf-transport-types.yang defined in [draft-sharma-ccamp-otn-service-model-00].

4.1.2. The node and link list

The module presented in this draft contains all the nodes and links information pertaining to a layer one network. As specified in the ietf-network YANG module, a node is identified by the node-id, which is unique within the network. Within the nodes, all the interfaces pertaining to this node and their potential capabilities/constraints SHOULD be present.

Similarly, a link is identified by the link-id, which is unique within a node. It includes the association with nodes as well as interfaces. Moreover, it includes information that is of interest to the client, for purposes, such as path computation, monitoring etc.

Since for an optical transport network, its client interface attributes has technology-specific attributes that need to be captured to give the client all the necessary information for its use. For full control purpose, this attributes and information are also captured and listed in this YANG module.

4.2. YANG Code

```
<CODE BEGINS> file " ietf-odu-topology@2016-07-07.yang"

module ietf-odu-topology {
 yang-version 1;

 namespace "urn:ietf:params:xml:ns:yang:ietf-odu-topology";
 prefix "l1topo";

 import ietf-network {
   prefix "nd";
 }

 import ietf-network-topology {
   prefix "lnk";
 }

 import ietf-te-topology {
   prefix "tet";
 }

 import ietf-transport-types {
   prefix "tran-types";
 }

 organization
   "Internet Engineering Task Force (IETF) CCAMP WG";
 contact
   "
   WG List: <mailto:ccamp@ietf.org>

   ID-draft editor:
   Xian ZHANG (zhang.xian@huawei.com);
   Anurag Sharma (AnSharma@infinera.com);
   ";
```

```
   description
     "This module defines a protocol independent Layer 1/ODU
     topology data model.";

   revision 2016-07-07 {
     description
       "Initial version.";
     reference
       "draft-zhang-ccamp-l1-topo-yang-01.txt";
   }

   /*
   typedef
   */

   typedef adaptation-type {
     type enumeration {
       enum CBR {
         description "Constant Bit Rate.";
       }
       enum ATMvp {
         description "ATM VP.";
       }
       enum GFP {
         description "Generic Framing Procedure.";
       }
       enum NULL {
         description "NULL";
       }
       enum PRBS {
         description "Pseudo Random Binary Sequence";
       }
       enum RSn {
         description "TBD";
       }
     }

     description
       "Defines a type representing the adaptation type
       on the termination point.";
   }


   /*
   Groupings
   */
```

```
   grouping l1-network-type {
     container l1-network {
       presence "indicates a L1 network, i.e., Optical
       Transport Network (OTN).";
       description "l1 network type";
     }
     description "l1-network-type";
   }

   grouping l1-network-attributes {
     leaf name {
       type string;
       description "the network name";
     }
     description "name attribute for l1 network";
   }

   grouping l1-node-attributes {
     description "l1-node-attributes";
     leaf name {
       type string;
       description "a name for this node.";
     }
   }

   grouping l1-link-attributes {
     description "l1 link attributes";

     leaf odu-type {
       type identityref{
         base tran-types:tributary-protocol-type;
       }
       description "the ODU type supported by this link";
     }

     leaf distance {
       type uint32;
       description "distance in the unit of kilometers";
     }
   }

   grouping ttp-odu-attributes {
     description "ttp ODU attributes";

     leaf odu-Type {
       type identityref{
         base tran-types:tributary-protocol-type;
```

```
        }
        description "the ODU type supported by this ttp";
      }
    }

    grouping sch-odu-attributes {
      description "additional odu attributes for scheduled link";

      leaf odu-type {
        type identityref{
          base tran-types:tributary-protocol-type;
        }
        description "the ODU type scedhuled";
      }

      leaf oduflex-bw {
        type uint32;
        description "banwidth for ODUflex type";
      }
    }

    grouping l1-tp-attributes {
      description "l1-tp-attributes";

      leaf client-facing {
        type empty;
        description
          "if present, it means this tp is a client-facing tp";
      }

      leaf tpn {
        type uint16 {
          range "0..4095";
        }
        description
        "Tributary Port Number. Applicable in case of mux services.";
        reference
        "RFC7139: GMPLS Signaling Extensions for Control of Evolving
         G.709 Optical Transport Networks.";
      }

      leaf tsg {
        type identityref {
          base tran-types:tributary-slot-granularity;
        }
        description "Tributary slot granularity.";
        reference
```

```
          "G.709/Y.1331, February 2012: Interfaces for the Optical
           Transport Network (OTN)";
        }

        leaf protocol-type {
          type identityref {
            base tran-types:tributary-protocol-type;
          }
          description "Protocol type for the Termination Point.";
        }

        leaf fec-enabled {
          type boolean;
          description
           "This attribute is optional and indicates whether Forward
           Error Correction (FEC) is enabled or not for the
           Termination Point.";
        }

        leaf adaptation-type {
          type adaptation-type;
          description
           "This attribute indicates the type of the supported
           adaptation function at the termination point.";
          reference
          "G.874.1, January 2002: Optical transport network (OTN):
          Protocol-neutral management information model for the
          network element view.";
        }

        leaf sink-adapt-active {
          type boolean;
          description
          "This attribute allows for activation or deactivation of
         the sink adaptation function. The value of TRUE means active.";
          reference
          "G.874.1, January 2002: Optical transport network (OTN):
          Protocol-neutral management information model for the
          network element view ";
        }

        leaf source-adapt-active {
          type boolean;
          description
          "This attribute allows for activation or deactivation of
          the sink adaptation function. The value of TRUE
          means active.";
```

```
       reference
       "G.874.1, January 2002: Optical transport network (OTN):
       Protocol-neutral management information model for
       the network element view ";
     }

     container timeslots {
       description
        "A list of tributary timeslots used by the ODU
        Termination Point.";
       leaf-list values {
         type uint8;
         description
           "Tributary timeslot value.";
         reference
          "G.709/Y.1331, February 2012: Interfaces for the
          Optical Transport Network (OTN)";
       }
     }
   }


    /*
    * Data nodes
    */
   augment "/nd:networks/nd:network/nd:network-types/tet:te-topology" {
     uses l1-network-type;
     description "augment network types to include L1 newtork";
   }

   augment "/nd:networks/nd:network" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network";
     }
     uses l1-network-attributes;
     description "Augment network configuration";
   }

   augment "/nd:networks/nd:network/nd:node" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network";
     }
     description "Augment node configuration";
     uses l1-node-attributes;
   }

   augment "/nd:networks/nd:network/nd:node/"
```

```
      +"lnk:termination-point/tet:te/tet:config" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network";
     }
     description "OTN TP attributes config in a ODU topology.";
     uses l1-tp-attributes;
   }

   augment "/nd:networks/nd:network/nd:node" +
           "/lnk:termination-point/tet:te/tet:state" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network";
     }
     description "OTN TP attributes state in a ODU topology.";
     uses l1-tp-attributes;
   }

   augment "/nd:networks/nd:network/lnk:link/tet:te/tet:config" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network.";
     }
     description "Augment link configuration";

     uses l1-link-attributes;
   }

   augment "/nd:networks/nd:network/lnk:link/tet:te/tet:state" {
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network.";
     }
     description "Augment link state";

     uses l1-link-attributes;
   }

   augment "/nd:networks/nd:network/nd:node/tet:te" +
       "/tet:tunnel-termination-point/tet:state"{
     when "nd:network-types/tet:te-topology/l1-network" {
       description "Augment only for L1 network";
     }
     description "Augment ttp state";
     uses ttp-odu-attributes;
   }

   augment "/nd:networks/nd:network/lnk:link/tet:te/tet:config"
       +"/tet:te-link-attributes/tet:schedules/tet:schedule"{
     when "nd:network-types/tet:te-topology/l1-network" {
```

```
        description "Augment only for L1 network.";
      }
      description "Augment link schedule config";

      uses sch-odu-attributes;
    }

    augment "/nd:networks/nd:network/lnk:link/tet:te/tet:state"
        +"/tet:te-link-attributes/tet:schedules/tet:schedule"{
      when "nd:network-types/tet:te-topology/l1-network" {
        description "Augment link schedule state.";
      }
      description "Augment link schedule config";

      uses sch-odu-attributes;
    }
  }

  <CODE ENDS>
```

5. Security Considerations

Since the data model defined in this draft is manipulated via, for example, the interface between an orchestrator and a transport network controller. The security concerns mentioned in [draft-ietf-teas-yang-te-topo-04] also applies to this draft.

The YANG module defined in this memo is designed to be accessed via the RESTCONF protocol defined in [draft-ietf-netconf-restconf-13] , or maybe via the NETCONF protocol [RFC6241].

There are a number of data nodes defined in the YANG module which are writable/creatable/deletable (i.e., config true, which is the default). These data nodes may be considered sensitive or vulnerable in some network environments. Write operations (e.g., <POST>) to these data nodes without proper protection can have a negative effect on network operations.

[Editor’s note: to List specific subtrees and data nodes and their sensitivity/vulnerability.]

6. Manageability Considerations

TBD.

7. IANA Considerations

TBD.

8. Acknowledgements

The initial YANG model specified in this draft is based on draft-clemm-i2rs-yang-network-topo but it is modified according to the features of the layer one networks.

We would like to thank the authors of the above mentioned draft for their helpful discussion during the creation of this draft.

9. References

9.1. Normative References

[RFC2119] S. Bradner, "Key words for use in RFCs to indicate requirements levels", RFC 2119, March 1997.

[RFC6020] Bjorklund, M., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.

[draft-ietf-i2rs-yang-network-topo] Clemm A., Medved J., Tkacik T., Varga R., et al, "A YANG Data Model for Network Topologies", draft-ietf-i2rs-yang-network-topo-01, work in progress, June 2015;

[draft-ietf-netconf-restconf] Bierman A., Bjorklund M., Watsen K., "RESTCONF Protocol", draft-ietf-netconf-restconf-13, work in progress, April 2016;

[draft-sharma-ccamp-otn-service-model] Sharma A., Rao R., Zhang X., "OTN Service YANG Model", draft-sharma-ccamp-otn-service-model-00, work in progress, July 2016;

9.2. Informative References

[draft-ietf-i2rs-architecture] Atlas, A., Halpern, J., Hares, S., Ward, D., Nadeau T., "An Architecture for the Interface to the Routing System", draft-ietf-i2rs-architecture-08, work in progress, January 2015;

[ietf-netmod-rfc6087bis] Bierman, A., "Guidelines for Authors and Reviewers of YANG Data Model Documents", draft-ietf-netmod-rfc6087bis-01, work in progress, October 2014.

[RFC6241] Enns, R., Bjorklund, M., Schoenwaelder, J., and A. Bierman, "Network Configuration Protocol (NETCONF)", RFC6241, June 2011.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, June 2011.

[RFC6536] Bierman, A. and M. Bjorklund, "Network Configuration Protocol (NETCONF) Access Control Model", RFC 6536, March 2012.

[WSON-YANG] Lee, Y., et al, " A Yang Data Model for WSON Optical Networks", draft-lee-ccamp-wson-yang-02, work in progress, July 2015.

[Flexi-YANG] Lopez de Varga, J.E., et al, "YANG data model for Flexi-Grid Optical Networks", draft-vergara-ccamp-flexigrid-yang-01, work in progress, July 2015.

10. Contributors' Address

Sergio Belotti
Alcatel-Lucent
Sergio.belotti@alcatel-lucent.com

Authors' Addresses

Xian Zhang
Huawei Technologies
Email: zhang.xian@huawei.com

Baoquan Rao
Huawei Technologies
raobaoquan@huawei.com

Anurag Sharma
Infinera

ansharma@infinera.com

Xufeng Liu
Ericsson
Xufeng.liu@ericsson.com

CCAMP Working Group
Internet-Draft
Intended status: Informational
Expires: January 08, 2017

X. Zhang, Ed.
Huawei Technologies
A. Sharma, Ed.
Infinera
S. Belotti
Nokia
T. Cummings
Ericsson
July 07, 2016

YANG Models for the Northbound Interface of a Transport Network Controller: Requirements and Gap Analysis
draft-zhang-ccamp-transport-yang-gap-analysis-00

Abstract

A transport network is a lower-layer network designed to provide connectivity services for a higher-layer network to carry the traffic opaquely across the lower-layer network resources. A transport network may be constructed from equipment utilizing any of a number of different transport technologies such as the evolving optical transport infrastructure (Synchronous Optical Networking (SONET) / Synchronous Digital Hierarchy (SDH) and Optical Transport Network (OTN)) or packet transport as epitomized by the MPLS Transport Profile (MPLS-TP).

All transport networks have high benchmarks for reliability and operational simplicity. This suggests a common, technology-independent management/control paradigm that can be extended to represent and configure specific technology attributes.

This document describes the high-level requirements facing transport networks in order to provide open interfaces for resource programmability and control/management automation. Furtheremore, gap analysis against existing models are also provided so that it can used as the guidance to separate efforts/drafts proposing new models or augmentation models based on existing models.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 8, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction

A transport network is a server-layer network designed to provide connectivity services, or more advanced services like Virtual Private Networks (VPN) for a client-layer network to carry the client traffic opaquely across the server-layer network resources. It acts as a pipe provider for upper-layer networks, such as IP network and mobile networks.

Transport networks, such as Synchronous Optical Networking (SONET) / Synchronous Digital Hierarchy (SDH), Optical Transport Network (OTN), Wavelength Division Multiplexing (WDM), and flexi-grid networks, are often built using equipment from a single vendor and are managed using proprietary interfaces to dedicated Element Management Systems (EMS) / Network Management Systems (NMS). All transport networks have high benchmarks for reliability and operational simplicity. This suggests a common, technology-independent management/control paradigm that is extended to represent and configure specific technology attributes.

Network providers need a common way to manage multi-vendor and multi-domain transport networks (where each domain is an island of equipment from a single supplier) and this requirement has been further stressed by the expansion in network size. At the same time, applications such as data center interconnection require larger and more dynamic connectivities. Therefore, transport networks face new challenges going beyond automatic provisioning of tunnel setup enabled by GMPLS (Generalized Multi-Protocol Label Switching) protocols to achieve automatic service provisioning, as well as address opportunities enabled by partitioning the transport network through the process of resource slicing. With a reduction in operational expenditure (OPEX) and capital expenditure (CAPEX) as the usual objectives, a common interface to transport network controllers are considered by network providers as a way to meet the requirements. The concept of Software Defined Networking (SDN) leverages these ideas.

The YANG language [RFC6020] is currently the data modeling language of choice within the IETF and has been adopted by a number of industry-wide open management and control initiatives. YANG may be used to model both configuration and operational states; it is vendor-neutral and supports extensible APIs for control and management of elements.

This document first specifies the scope and provides high-level requirements for transport network open interface modelling. Furthermore, detailed gap analysis of the typical scenarios with the existing model are provided. Thus, this document can used as a reference of existing models, and provides information of the missing ones which suggest further work.

2. Scope

For this draft, we use the domain controller as the reference point, with South Bound Interface (SBI) to the transport devices and North Bound Interface (NBI) to the orchestrator.

Transport networks have been evolving and deploying for decades, making them very heterogeneous. New and legacy transport devices support many of interface protocols like Path Computation Element Protocol (PCEP), TL1, SNMP, CLI, XML, NETCONF, Openflow etc. Domain controllers interfacing with transport devices need to support these protocols on its SBI, making the southbound fragmented. Domain controllers abstract the fragmented southbound view for its northbound clients by normalizing the NBI across various technologies, protocols, and vendors. The focus of this document is not to go into various southbound protocols to interface with the transport devices. Instead, this document focuses on the models that can be used by the domain controller and the orchestrator for various use cases identified in later sections of this document. This document analyzes IETF models for various use cases, such as single-domain network, multi-domain network, multi-layer network, etc. to identify any modeling gaps.

YANG models are currently developed not only in IETF, but also in other Standard Development Organizations (SDO) such as ONF and MEF, which can be used on the interfaces of a domain controller and an orchestrator. Each domain controller and orchestrator can use models developed by different SDOs. Therefore it is important to ensure that deployment use cases and related funcionalities are supported by all models to allow a seamless translation/mediation between systems using different models.

If the Abstraction and Control of Traffic-Engineered Networks (ACTN) defined in [I-D.ceccarelli-teas-actn-framework] is used as a reference architcture, then the focus is equivalent to MPI (MDSC-PNC Interface) and CMI (CNC-MDSC Interface).

3. High-level Modeling Requirements

This section covers various high-level modeling requirements for transport networks.

3.1. Generic Requirements

The following are generic requirements for Transport models:

- User Intent: Transport models should maintain separation between high level user intent and the operational state of the network. For e.g., maintain separation between user service request, including all constraints, and the actual service and connection state in the network.

- State Management: Network and service objects should support the following states: administrative state, operational state, and lifecycle state. Administrative state and operational states are well understood. Lifecycle state is defined in the ONF to model the following entity lifecylce states: planned state, potential state, installed state, in conflict state, and pending removal state.

- Identifiers: Network and service objects should support the following identifier:

  * ID: A unique entity ID provided by the controller. The identifier SHOULD be chosen such that the same entity in a real network topology will always be identified through the same ID, even if the model is instantiated in separate datastores. Controller may choose to capture semantics in the identifier, for example to indicate the type of entity and/or the type of the parent identity.

  * Name: A unique name provided by the client for the entity. The name can be modified, if required, by the client.

  * User Labels: A list of freeform strings that can be used as alias for the entity by the client. Multiple user labels are permitted for the entity, and client can edit these user labels. User labels do not need to be unique.

3.2. Transport Network and TE Topology Requirements

3.2.1. Topological Link Requirements

The model should support the following Topological Links:

- o Physical Links
- o Abstract Links [I-D.ietf-teas-interconnected-te-info-exchange]
- o Compound Link which are are internally aggregated lower level links
- o Access Links which connect the router port to the client port of the transport system
- o Transitional Links which provide adaptation capability between layers within a network element

The Link should support various link related attributes like cost, latency, capacity, risk characteristics (including shared risk). The model should provide clear association between Link and its topology (including virtual topology), nodes and termination points.

The model should provide association between the Link and any underlay circuit / service supporting the Link.

3.2.2. Topology Node Requirements

The model should support the following Topology Node:

- o Physical Node
- o Abstract Node
- o Chassis / Forwarding Domain

[Editors’ note: more details will be added later.]

3.2.3. Termination Point Requirements

[Editors’ note: this will be added later.]

3.3. Transport Service Requirements

[Editors’ note: this will be added later.]

3.4. Tunnel/LSP Requirements

[Editors' note: this will be added later.]

4. Scenarios

There are several scenarios (a.k.a., use cases) where an open interface via domain controller to access server-layer (transport) network resources would be useful. Three scenarios are provided and can be used for model instantiation exercise to identify missing pieces of existing models.

4.1. Single-domain Scenario

The first scenario is depicted as below (Figure 1):

```
        /--\   +------+                  +------+     /--\
       | 1  ~~~| A    +------------------| B    |~~~~~  3 |
        \--/   +-----++                  +--+---+     \--/
                     |                      |
                     |                      |
                     |                      |
                    ++-----+            +---+--+
                    |   F  +------------+  C   |
                    ++-----+            +--+---+
                     |                     |
                     |                     |
                     |                     |
                     |                     |
                     |                     |
                 +---+-+                +---+-+    /---\
                 |  E  +---------------+  D  |~~~~  4  |
        /--\~~~~~+-----+               +-----+    \---/
       | 2  |
        \--/

          +----+                      /--\
          |    |  Transport NE       |    |   DC
          +----+                      \--/

              -----   Transport Link      ~~~   Transport-DC link

           (a)     Data Centers interconnected via a transport network

                  +---------+----------+
                  | Data Center Network |
                  |   Controller        |
                  +---------+-----------+
                            |
                            |
                            |
                            | Open Interface
                            |
                            |
                  +---------+-----------+
                  |  Transport Network  |
                  |    Controller       |
                  +---------------------+
      (b)  The controller architecture for data center interconnection
```

Figure 1: Scenario 1: Data centers interconnected via a transport network and the controller architecture

For the data center operator, as a client of the transport network, assuming the objective is to trigger the transport network to provide connectivity on demand, the following capabilities, at a minimum, would be required on the common interface between the two controllers illustrated in Figure 1:

- The ability to obtain information about a set of access points of the transport network, including information such as access point identifiers, capabilities, etc.; for instance, transport-network-side end point identifiers related to the access link between DC1 and Transport NE A.

- The capability to send a request for a service using the aforementioned access point information, as well as the ability to retrieve a list of service requests and their statuses. In this request, it should at least be possible to include source node, destination node, and requested bandwidth to request the transport network to set up tunnels/paths so as to provide the requested connectivitiy for the service request.

- Note that in this case, the acquisition of the topology, be it physical or logical, of the transport network is not a compulsory requirement, but it may indeed be able to give data center providers more control over the transport resource usage. Furtheremore, the client controller can impose a virtual network of its own choice by requesting a slice of network resource with its choice of network parameters (such as network topology type, bandwidth etc.).

4.2. Multi-domain Scenario

The second scenario, more complicated than the first, is depicted as below (Figure 2). In this example, we focus on the management and control via common interfaces for multi-domain networks with homogeneous technologies (such as OTN), but it can be extended further to multi-domain networks with heterogeneous technologies with higher complexity.

```
                 +-------------------------------------------------+
                 |              Orchestrator/Coordinator           |
                 +---------+--------------+-------------------+----+
                                   |              |                   |
                                   |              |                   |
          +------------+--+           |        +----------+----------+
          | Controller 1  |           |        |   Controller 2      |
          +---------+-----+           |        +-------+-------------+
                                #        +----------------+        #
                                #Qx      | Controller 3   |        #
                                #        +----------------+        #TL1
                                #                #                 #
                        ----+-----           #             ----+-----
           ____/          \____      #        ____/          \____
          |                    |     #       |                    |
         |                      |    #      |                      |
         |    Network Domain    +***********+   Network Domain     |
         |          1          |     #      |          2           |
          |____            ___|      #       |____             ___|
                   \           /         #PCEP        \           /
                        -----------          #            *----------
                          *       *          #           *       *
                           *       *         #          *       *
                                *       *        #         *       *
                                 *       *       #        *       *
                                  *       *      #       *       *
                                   *       *     #      *       *
                                        *       *----+-----*       *
                                         *  ____/          \____  *
                                          *|                   |*
                                          |                     |
                                          |    Network Domain    |
                                          |          3          |
                                           |____            ___|
                                                        \           /
                                                         ----------

                                                 *****  inter-domain links
                                                 -----  Open Interfaces
                                                 #####  Controller-device interf
aces
```

Figure 2: Scenario 2: Multi-domain network control and management

For the second scenario, the orchestrator controls and manages three distinct network domains, each controlled/managed by their domain controller. This scenario is of interest not only to transport-only networks, but also to heretegenous network orchestration such as coordinating the transport, the radio (5G) and packet core domains. But to keep the functions explanation later accurate, only transport-only multi-domain networks are considered.

In order to orchestrate across domains/layers, besides the capabilities mentioned for the first scenario, the orchestrator needs its interface between domain controllers to be equipped with the following additional functions:

o Access to the topologies reported by each domain controller, including cross-domain links for the purpose of planning and requesting the paths of end-to-end tunnels. Multiple technologies within a domain (i.e., a multi-layer network), this might be reflected in the reported topology. Depending on the abstraction level of the reported topology, the orchestrator has different control granularities.

o Alterntively, the capability for the orchestrator to request "path computation" to a domain controller in order to create an end-to-end tunnel stitched together by different connection contribution obtained by consulting to each domain controller.

o The ability to set up, delete and modify tunnels, be it within one domain or across multiple domains. Furthermore, it should have the abilty to view the tunnels created within each domain as well as those that cross domains as reported by each domain controller.

4.3. Multi-layer Scenario

[Editors’ note: to be added later.]

4.4. Function Summary and Related YANG Models

For the common interface of a transport controller towards a northbound client, five functions are derived from the scenarios explained in the last section. They are summarized in the table below and we also match these functions with YANG models that are being developed in existing drafts.

Functions	Description	Related Existing YANG Models
Obtaining Access Point Info	Getting the necessary access points info	[TE-Topo]
Obtaining Topology	Getting the topology info	[TE-Topo], [WDM-Topo] [ODU-Topo]
Tunnel Operations	Tunnel Setup, Deletion Modification and Info Retrieval	[TE-Tunnel]
Service Request	Requesting connectivity service and retrieval the list of service request	NONE
Path Comp.	Path Computation pre service provisioning	NONE
Virtual Network Operations	Requesting a virtual network and related control operations, (e.g.,update, deletion)	[TE-Topo], [WDM topo] [ODU-Topo]

Analysis and descriptions of whether and how these functions are supported by the YANG models are provided in more detail in Section 5.

5. Function Gap Analysis on YANG Model Level

5.1. Topology Related Functions

As shown in the previous section, the functions of obtaining access point information, obtaining topology, and imposing virtual network operations can take advantages of the same set of topology YANG models. These functions are briefly explained further in the following sub-sections.

5.1.1. Obtaining Access Point Info

For cases such as scenario 1, a client may have no interest in directly controlling network resources, but might want an automated common control interface for initiating service requests. In this case, a transport domain controller may provide the access point information. This information can then be used in service request sent over the common interface.

The TE Topology YANG model provided in [TE-topo] [I-D.ietf-teas-yang-te-topo] can be used to provide a list of links. If the remote node and termination point information is unknown, it is omitted from the reported information. If the client-side node and termination point information is obtained via configuration or a distributed discovery mechanism, then it can also be added into the reported information. Technology-specific details might also be needed to further express the constraints/attributes associated with the access points. Note that all of this information is usually read only.

5.1.2. Obtaining Topology

Refer to [I-D.ietf-teas-yang-te-topo] for explanations and examples on how to obtain the topology. For technology specific topology information, other models such as those provided in [WDM-Topo] [I-D.ietf-ccamp-wson-yang] and [ODU-Topo] [I-D.zhang-ccamp-l1-topo-yang] may be used.

5.1.3. Virtual Network Operations

There are two ways to request the creation of a virtual network. One is to define the topology explicitly using the model provided in the topology YANG drafts listed in previous section. The other way is to provide an estimated traffic information (a traffic matrix) and ask for a domain controller of the provider network to provide a virtual network that can fulfill the demand. This second approach does not have a supporting model and need further work.

5.2. Tunnel Operations

The current [TE-Tunnel] [I-D.ietf-teas-yang-te] provides a technology agnostic Traffic-Engineered (TE) device tunnel. The model included in that draft is currently being developed to make it generic for both controller and device usage. In the latest version, it already provides such a generic TE tunnel model that can cater to the base requirementss for tunnel operations but it may need to be augmented to support controller-specific operations.

Furthermore, technology-specific augmentations of the base generic TE tunnel models are needed. For example, for Optical Channel (OCh) (note: ITU is updating this term as OTSi.) tunnels in WDM networks, information such as the lambda resource usage is needed. Similarly, for ODU tunnels, information such as ODU-specific client signal, tributary slot information etc. is needed.

5.3. Service Requests

The service model is an important model that enables automated operations between a client controller or an orchestrator and a domain controller. The transport connectivity service model is different from the model of a tunnel since the transport connectivity service model hides technical details from a client.

6. IANA Considerations

This document requests no IANA actions.

7. Security Considerations

Clearly modifying server-layer resources will have a significant impact on network infrastructure. More specifically they will provide the services and applications running across client-layers, which the server-layer is supporting. Therefore, security must be an important consideration when implementing the architecture, models and protocol mechanisms discussed in this document.

Communicating service and network information (including access point identifiers, capabilities, topologies, etc.) across external interfaces represents a security risk. Thus, mechanisms to encrypt or preserve the domain topology confidentiality should be used.

A key consideration are the external protocols (those shown as entering or leaving the orchestrator and controllers shown in Figure 2 (Scenario 2: Multi-domain network control and management)) which must be appropriately secured. This security should include authentication and authorization to control access to different functions that the orchestrator may perform to modify or create state in the server-layer, and the establishment and management of the orchestrator to controller relationship.

The orchestrator will contain significant data about the network domains, the services carried by each domain, and customer type information. Therefore, access to information held in the orchestrator must be secured. Since such access will be largely through external mechanisms, it may be pertinent to apply policy-based controls to restrict access and functions.

8. Manageability Considerations

The core objectives of this document are to assist in the deployment and operation of transport services across server-layer network infrastructure. The model-driven management/control principles, which are vendor-neutral and supported by extensible APIs, should be utilized.

The open models described in this document are based on YANG [RFC6020] and the RESTCONF [RESTCONF] messaging protocol, a REST-like protocol running over HTTP for accessing data defined in YANG, may also be used.

9. Acknowledgements

We would like to thank Young Lee, Igor Bryskin and Aihua Guo for their comments and discussions.

10. Contributing Authors

The following people all contributed to this document and are listed below:

Ruiquan Jing
China Telecom
Email: jingrq@ctbri.com.cn

Yan Shi
China Unicom
Email: shiyan49@chinaunicom.cn

Jeong-dong Ryoo
ETRI
Email: ryoo@etri.re.kr

Yunbin Xu
CAICT
Email: xuyunbin@ritt.cn

Daniel King
Lancaster University
Email: d.king@lancaster.ac.uk

11. References

11.1. Normative References

[I-D.ietf-netconf-restconf]
Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", draft-ietf-netconf-restconf-13 (work in progress), April 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <http://www.rfc-editor.org/info/rfc6020>.

11.2. Informative References

[I-D.ceccarelli-teas-actn-framework]
Ceccarelli, D. and Y. Lee, "Framework for Abstraction and Control of Traffic Engineered Networks", draft-ceccarelli-teas-actn-framework-02 (work in progress), April 2016.

[I-D.ietf-ccamp-wson-yang]
Lee, Y., Dhody, D., Zhang, X., Guo, A., Lopezalvarez, V., King, D., and B. Yoon, "A Yang Data Model for WSON Optical Networks", draft-ietf-ccamp-wson-yang-01 (work in progress), April 2016.

[I-D.ietf-teas-interconnected-te-info-exchange]
Farrel, A., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D., and X. Zhang, "Problem Statement and Architecture for Information Exchange Between Interconnected Traffic Engineered Networks", draft-ietf-teas-interconnected-te-info-exchange-07 (work in progress), May 2016.

[I-D.ietf-teas-yang-te]
Saad, T., Gandhi, R., Liu, X., Beeram, V., Shah, H., Chen, X., Jones, R., and B. Wen, "A YANG Data Model for Traffic Engineering Tunnels and Interfaces", draft-ietf-teas-yang-te-02 (work in progress), October 2015.

[I-D.ietf-teas-yang-te-topo]
Liu, X., Bryskin, I., Beeram, V., Saad, T., Shah, H., and O. Dios, "YANG Data Model for TE Topologies", draft-ietf-teas-yang-te-topo-04 (work in progress), March 2016.

[I-D.zhang-ccamp-l1-topo-yang]
Zhang, X., Rao, B., and X. Liu, "A YANG Data Model for Layer 1 Network Topology", draft-zhang-ccamp-l1-topo-yang-01 (work in progress), December 2015.

Authors' Addresses

Xian Zhang (editor)
Huawei Technologies
F3-5-B R&D Center, Huawei Industrial Base, Bantian, Longgang District
Shenzhen, Guangdong 518129
P.R.China

Email: zhang.xian@huawei.com

Anurag Sharma (editor)
Infinera
US

Email: ansharma@infinera.com

Sergio Belotti
Nokia
Italy

Email: sergio.belotti@nokia.com

Tara Cummings
Ericsson

Email: tara.cummings@ericsson.com