CCAMP Working Group   Vishnu Pavan Beeram (Ed)
Internet Draft   Juniper Networks
Intended status: Standards Track   Igor Bryskin (Ed)
                                                    ADVA Optical Networking

Expires: August 12, 2014   February 12, 2014

Mutually Exclusive Link Group (MELG)
draft-beeram-ccamp-melg-03

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 August 12, 2014.

Copyright Notice

Abstract

This document introduces the concept of MELG ("Mutually Exclusive Link Group") and discusses its usage in the context of mutually exclusive Virtual TE Links.

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].

Table of Contents

1. Introduction

A Virtual TE Link (as defined in [RFC6001]) advertised into a Client Network Domain represents a potentiality to setup an LSP in the Server Network Domain to support the advertised TE link. The Virtual TE Link gets advertised like any other TE link and follows the same rules that are defined for the advertising, processing and use of regular TE links [RFC4202]. However, "mutual exclusivity" is one attribute that is specific to Virtual TE links. This document discusses the different types of mutual exclusivity (Static vs Dynamic) that come into play and explains the need to advertise this

information into the Client TEDB. It then goes onto introduce a new TE construct (MELG) to carry static mutual exclusivity information.

2. Virtual TE Link - Semantics

A Virtual TE Link (as per existing definitions) represents the potentiality to setup a server layer LSP, but there are currently no strict guidelines imposed on how the underlying server layer LSP would need to get set up. The characteristics of the underlying server-path are not necessarily pinned down until the Virtual TE Link gets actually committed. This means that some important characteristics of the Virtual TE Link like shared-risk and delay (and mutual exclusivity information) may not be known until the corresponding server layer LSP is set up. This makes resource planning (for example - pre-configuring network failure recovery schemes) in a multi-layer network that includes Virtual TE Links a very hard problem.

This document uses a slightly enhanced view of a Virtual TE Link. In the context of this document, the Virtual TE Link (even when it is uncommitted) is always aware of the key characteristics of the underlying server-path. The creation and maintenance of this Virtual TE Link is strictly driven by policy. Policy not only determines which Virtual TE Link to create (What termination points?), but it may also constrain how the corresponding underlying server layer LSP (What path?) needs to get set up. The basic idea behind this "enhanced view" is that it makes the "Virtual TE Link" get as close as it can to representing a "Real TE Link".

Also, as per this document, a Virtual TE Link remains a Virtual TE Link through-out its life-time (until it gets deleted by the user/policy). It may get committed (underlying server LSP gets set up) and uncommitted (underlying server LSP gets deleted) from time to time, but it never really loses it "Virtual" property.

3. Mutually Exclusive Virtual TE Links

Mutual Exclusivity comes into play when multiple Virtual TE Links are dependent on the usage of the same underlying server resource. Since not all of these Virtual TE Links can get committed at the same time, they are deemed to be mutually exclusive.

The existence of this "mutual exclusivity" property would need to be advertised into the Client TE Domain. This is of relevance to Client Path Computation engines; especially those that are capable of doing concurrent computations. If this information is absent, there exists the risk of the Computation engine yielding erroneous concurrent path computation results where only a subset of the computed paths get successfully provisioned.

The "Mutual Exclusivity" property of a Virtual TE Link can be either static or dynamic in nature.

3.1. Static vs Dynamic

Static Mutual Exclusivity: This type of mutual exclusivity exists permanently within a given network configuration. It comes into play when two or more Virtual TE Links depend on the usage of the same non-shareable underlying server network domain resource. This resource gets used up in its entirety by a single Virtual TE Link when committed. Such resources exist only in the WDM layer.

Dynamic Mutual Exclusivity: This type of mutual exclusivity exists temporarily within a given network configuration. It comes into play when two or more Virtual TE Links depend on the usage of the same shareable underlying server network domain resource. Mutual Exclusivity exists when the amount of the server resource that is available for sharing is limited; it ceases to exist when sufficient amount of the resource is available for accommodating all corresponding Virtual TE Links. Such resources exist in all layers.

Because of their inherent difference, the advertisement paradigm of the TE construct required to carry static mutual exclusivity information is quite different from that of the TE construct required to carry dynamic mutual exclusivity information. Static mutual exclusivity Information can get advertised per TE-Link using a simple sub-TLV construct. There wouldn't be any scaling issues with this approach because of the static nature of the information that gets advertised. On the contrary, advertising dynamic mutual exclusivity information per TE-Link poses serious scaling concerns and hence requires a different type of construct/paradigm.

This document introduces a new TE construct for carrying static mutual exclusivity information. The mechanisms to address dynamic mutual exclusivity are discussed in a separate document [SRcLG].

4. Static Mutual Exclusivity

Consider the network topology depicted in Figure 1a. This is a typical packet optical transport deployment scenario where the WDM layer network domain serves as a Server Network Domain providing transport connectivity to the packet layer network Domain (Client Network Domain).

```
                              |
                              | +---+            /-\
                              | |   | Router    (   ) WDM
                              | +---+ Node       \-/  node
                              |________________________________

 +---+        /-\          /-\           /-\          +---+
 | R1|-------( A )--------( C )---------( E )---------| R3|
 +---+        \-/          \-/           \-/          +---+
                          /   \         /   \
                         /     \       /     \
                        /       \     /       \
                       /         \   /         \
                      /           \ /           \
     +---+          /-\           /-\           /-\          +---+
     | R2|---------( B )---------( D )---------( F )---------| R4|
     +---+          \-/           \-/           \-/          +---+
```

Figure 1a: Sample topology

```
     -------------                        |  [ ] Client TE Node
     | Client TE |                        |  +++ Client TE Link
     | DataBase  |                        |_____________________
     -------------
        [R1] ++++++++ [A]                  [E] +++++++++ [R3]


        [R2] ++++++++ [B]                  [F] +++++++++ [R4]
```

Figure 1b: Client TE Database

Nodes R1, R2, R3 and R4 are IP routers that are connected to an Optical WDM transport network. A, B, C, D, E and F are WDM nodes that constitute the Server Network Domain. The border nodes (A, B, E and F) operate in both the server and client domains. Figure 1b depicts how the Client Network Domain TE topology looks like when there are no Client TE Links provisioned across the optical domain.

```
                             | *****  B-F WDM Path
                             | @@@@@  B-E WDM Path
                             |________________________________

 +---+        /-\          /-\ @@@@@@@@@ /-\          +---+
 | R1|-------( A )--------( C )---------( E )---------| R3|
 +---+        \-/         @\-/           \-/          +---+
                         @/   \         /   \
                        @/     \       /     \
                       @/       \     /       \
                      @/         \   /         \
                     @/           \ /           \
     +---+          /-\ ********* /-\ ********* /-\          +---+
     | R2|---------( B )---------( D )---------( F )---------| R4|
     +---+          \-/           \-/           \-/          +---+
```

Figure 2a: Mutually Exclusive potential WDM paths

```
      ------------   |  TE-Links B-F and B-E are mutually exclusive;
      | Client-TE|   |  They depend on the usage of the same
      | Database |   |  underlying non-shareable server resource
      ------------   |_____________________________________________

        [R1] ++++++++ [A]                      [E] +++++++++ [R3]
                                          +++++
                                      ++++
                                  ++++
                              ++++
                          ++++
        [R2] ++++++++ [B] ++++++++++++++++++++ [F] +++++++++ [R4]
```

Figure 2b: Client TE Database - Mutually Exclusive Virtual TE Links

Now consider augmenting the Client TE topology by creating a couple of Virtual TE Links across the optical domain. The potential paths in the WDM network catering to these two virtual TE links are as shown in Fig 2a and the corresponding augmented Client TE topology is as illustrated in Fig 2b.

In this particular example, the potential paths in the WDM layer network supporting the Virtual TE Links require the usage of the same source transponder (on "Node B"). Because the Virtual TE Links depend on the same uncommitted network resource, only one of them could get activated at any given time. In other words they are mutually exclusive. This scenario is encountered when the potential paths depend on any common physical resource (e.g. transponder, regenerator, wavelength converter, etc.) that could be used by only one Server Network Domain LSP at a time.

This document proposes the use of "Mutually Exclusive Link Group (MELG)" for catering to this scenario.

5. Mutually Exclusive Link Group

The Mutually Exclusive Link Group (MELG) construct defined in this document has 2 purposes

- To indicate via a separate network unique number (MELG ID) an element or a situation that makes the advertised Virtual TE Link belong to one or more Mutually Exclusive Link Groups. Path computing element will be able to decide on whether two or more Virtual TE Links are mutually exclusive or not by finding an overlap of advertised MELGs (similar to deciding on whether two or more TE links share fate or not by finding common SRLGs)

- To indicate whether the advertised Virtual TE Link is committed or not at the moment of the advertising. Such information is important for a path computation element: Committing new Virtual TE links (vs. re-using already committed ones) has a consequence of allocating more server layer resources and disabling other Virtual TE Links that have common MELGs with newly committed Virtual TE Links; Committing a new Virtual TE Link also means a longer setup time for the Client LSP and higher risk of setup-failure.

6. Protocol Extensions

6.1. OSPF

The MELG is a sub-TLV of the top level TE Link TLV. It may occur at most once within the Link TLV. The format of the MELGs sub-TLV is defined as follows:

Name: MELG
Type: TBD
Length: Variable

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Sub-TLV Type      |             Sub-TLV Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    VTE-Flags (16 bits)     |U |  Number of MELGs (16 bits)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MELGID1 (64 bits)                             |
|                 MELGID2 (64 bits)                             |
|                ........................                       |
|                 MELGIDn (64 bits)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Number of MELGs: number of MELGS advertised for the Virtual TE Link;

VTE-Flags: Virtual TE Link specific flags;

MELGID1,MELGID2,...,MELGIDn: 64-bit network domain unique numbers associated with each of the advertised MELGs

Currently defined Virtual TE Link specific flags are:

U bit (bit 1): Uncommitted - if set, the Virtual TE Link is uncommitted at the time of the advertising (i.e. the server layer network LSP is not set up); if cleared, the Virtual TE Link is committed (i.e. the server layer LSP is fully provisioned and functioning). All other bits of the "VTE-Flags" field are reserved for future use and MUST be cleared.

Note: A Virtual TE Link advertisement MAY include MELGs sub-TLV with zero MELGs for the purpose of communicating to the TE domain whether the Virtual TE Link is currently committed or not.

6.2. ISIS

The MELG TLV (of type TBD) contains a data structure consisting of:

```
6        octets of System ID
1        octet of Pseudonode Number
1        octet Flag
4        octets of IPv4 interface address or 4 octets of a Link
         Local Identifier
4        octets of IPv4 neighbor address or 4 octets of a Link
         Remote Identifier
2        octets MELG-Flags
2        octets - Number of MELGs
variable List of MELG values, where each element in the list
         has 8 octets
```

The following illustrates encoding of the value field of the MELG TLV.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           System ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        System ID (cont.)      |Pseudonode num |    Flags      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-++-+-+-+-+-+-+-+-+-+-+-+-++-+-+-+-+-+
|          Ipv4 interface address/Link Local Identifier         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Ipv4 neighbor address/Link Remote Identifier         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    VTE-Flags (16 bits)     |U |  Number of MELGs (16 bits)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MELGID1 (64 bits)                             |
|                 MELGID2 (64 bits)                             |
|                ........................                       |
|                 MELGIDn (64 bits)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The neighbor is identified by its System ID (6 octets), plus one octet to indicate the pseudonode number if the neighbor is on a LAN interface.

The least significant bit of the Flag octet indicates whether the interface is numbered (set to 1) or unnumbered (set to 0). All other bits are reserved and should be set to 0.

The length of the TLV is 20 + 8 * (number of MELG values).

The semantics of "VTE-Flags", "Number of MELGs" and "MELGID Values" are the same as the ones defined under OSPF extensions.

The MELG TLV MAY occur more than once within the IS-IS Link State Protocol Data Units.

7. Security Considerations

TBD

8. IANA Considerations

8.1. OSPF

IANA is requested to allocate a new sub-TLV type for MELG (as defined in Section 6.1) under the top-level TE Link TLV.

8.2. ISIS

IANA is requested to allocate a new IS-IS TLV type for MELG (as defined in Section 6.2).

9. Normative References

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

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

[RFC6001] D.Papadimitriou, M.Vigoureax, K.Shiomoto, D.Brungard and JL. Le Roux, "GMPLS Protocol Extensions for Multi-Layer and Multi-Region Networks", RFC 6001, October 2010.

[SRcLG] Beeram, V., "Shared Resource Link Group", draft-beeram-ccamp-srclg, February 2014

10. Acknowledgments

Chris Bowers [cbowers@juniper.net]

Authors' Addresses

Vishnu Pavan Beeram
Juniper Networks
Email: vbeeram@juniper.net

Igor Bryskin
ADVA Optical Networking
Email: ibryskin@advaoptical.com

John Drake
Juniper Networks
Email: jdrake@juniper.net

Gert Grammel
Juniper Networks
Email: ggrammel@juniper.net

Wes Doonan
Email: wddlists@gmail.com

Manuel Paul
Deutsche Telekom
Email: Manuel.Paul@telekom.de

Ruediger Kunze
Deutsche Telekom
Email: Ruediger.Kunze@telekom.de

Oscar Gonzalez de Dios
Telefonica
Email: ogondio@tid.es

Cyril Margaria
Juniper Networks
Email: cmargaria@juniper.net

Friedrich Armbruster
Coriant GmbH

Email: friedrich.armbruster@coriant.com

Daniele Ceccarelli
Ericsson
Email: daniele.ceccarelli@ericsson.com

Fatai Zhang
Huawei Technologies
Email: zhangfatai@huawei.com