Network Working Group
Internet Draft
Intended Status: Informational
Expires: December 2011

N. So
A. Malis
D. McDysan
Verizon
L. Yong
Huawei
C. Villamizar
Infinera
T. Li
Cisco

June 9, 2011

Composite Link Framework in Multi Protocol Label Switching (MPLS)
draft-so-yong-rtgwg-cl-framework-04

Status of this Memo

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

This Internet-Draft is submitted to IETF 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 This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.

This Internet-Draft is submitted to IETF 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or 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 December 9, 2011.

Copyright Notice

Copyright (c) 2011 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.

Abstract

This document specifies a composite link framework in MPLS network. A composite link consists of a group of homogenous or non-homogenous links that have the same forward adjacency and can be considered as a single TE link or an IP link in routing. A composite link relies on its component links to carry the traffic over the composite link. The document specifies composite link framework. Applicability is described for a single pair of MPLS-capable nodes, a sequence of MPLS-capable nodes, or a set of layer networks connecting MPLS-capable nodes.

Table of Contents

1. Introduction

Composite link functional requirements are specified in [CL-REQ]. This document specifies a framework of Composite Link in MPLS network to meet the requirements. Single link and LAG based link bundle have been widely used in today's MPLS networks. A link bundle [RFC4201] bundles a group of homogeneous links as a TE link to make routing approach more scalable. A composite link allows bundling non-homogenous links together as a single logical link. The motivations for using a composite link are descried in the document [CL-REQ]. This document describes composite link framework in the context of MPLS network with MPLS control plane.

A composite link is a single logical link in MPLS network that contains multiple parallel component links between two routers. Unlike a link bundle [RFC4201], the component links in a composite link can have different properties such as cost or capacity. A composite link can transport aggregated traffic as other physical links from the network perspective and use its component links to carry the traffic internally.

Specific protocol solutions are outside the scope of this document.

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].

2.1. Terminology

Composite Link: A composite link is a logical link composed of a set of parallel point-to-point component links, where all links in the set share the same endpoints. A composite link may itself be a component of another composite link, but only a strict hierarchy of links is allowed.

Component Link: A point-to-point physical or logical link that preserves ordering in the steady state. A component link may have transient out of order events, but such events must not exceed the network's specific NPO. Examples of a physical link are: Lambda, Ethernet PHY, and OTN. Examples of a logical link are: MPLS LSP, Ethernet VLAN, and MPLS-TP LSP.

Flow: A sequence of packets that must be transferred in order on one component link.

Flow identification: The label stack and other information that uniquely identifies a flow. Other information in flow identification may include an IP header, PW control word, Ethernet MAC address, etc. Note that an LSP may contain one or more Flows or an LSP may be equivalent to a Flow. Flow identification is used to locally select a component link, or a path through the network toward the destination.

Network Performance Objective (NPO): Numerical values for performance measures, principally availability, latency, and delay variation. See Appendix A for more details.

3. Composite Link Framework

A Composite Link in the context of MPLS network is a set of parallel links between two routers that form a single logical link within the network. Composite link model is illustrated in Figure 1, where a composite link is configured between routers R1 and R2. The composite link has three component links. Individual component links in a composite link may be supported by different transport technologies such as wavelength, Ethernet VLAN. Even if the transport technology implementing the component links is identical, the characteristics (e.g., bandwidth, latency) of the component links may differ.

As shown in Figure 1, the composite link may carry LSP traffic flows and control plane packets that appear as IP packets. A LSP may be established over the link by either RSVP-TE or LDP signaling protocols. All component links in a composite link have the same forwarding adjacency. The composite link forms one routing interface at the composite link end points for MPLS control plane. In other words, two routers connected via a composite link have forwarding adjacency and routing adjacency. Each component link only has significance to the composite link, i.e. it does not appear as a link in the control plane.


```
            Management Plane
        Configuration and Measurement <------------+
                   ^                                |
                   |                                |
       +-------+-+                        +-+-------+
       |       | |                        | |       |
  CP Packets   V |                        | V     CP Packets
       |  V    | |     Component Link 1   | |    ^  |
```

```
             |  |    |=|===========================|=|    |  |
             |  +----| |     Component Link 2      | |----+  |
             |       |=|===========================|=|       |
   Aggregated LSPs   | |                           | |       |
            ~|~~~~~~>| |     Component Link 3      | |~~~~>~~|~~
             |       |=|===========================|=|       |
             |       | |                           | |       |
             | LSR     |                           |    LSR  |
             +---------+                           +---------+
                     !                               !
                     !                               !
                     !<------ Composite Link ------->!
```

Figure 1 Composite Link Architecture Model

A component link in a composite link may be constructed in different ways.[CL-REQ] Figure 2 shows three common ways that may be deployed in a network.

```
      +-------+                 1. Physical Link            +-------+
      |     |-|----------------------------------------------|-|     |
      |     | |                                              | |     |
      |     | |     +------+                     +------+    | |     |
      |     | |     | MPLS |    2. Logical Link  | MPLS |    | |     |
      |     |.|.... |......|.....................|......|....|.|     |
      |     | |-----|  R3  |---------------------|  R4  |----| |     |
      |     | |     +------+                     +------+    | |     |
      |     | |                                              | |     |
      |     | |                                              | |     |
      |     | |     +------+                     +------+    | |     |
      |     | |     |GMPLS |    3. Logical Link  |GMPLS |    | |     |
      |     |.|. ...|......|.....................|......|....|.|     |
      |     | |-----|  R5  |---------------------|  R6  |----| |     |
      |       |     +------+                     +------+    |       |
      | R1    |                                              |    R2 |
      +-------+                                              +-------+
            |<------------- Composite Link ------------------->|
```

Figure 2 Illustration of Component Link Variances

As shown, the first component link is configured with direct physical media wire. The second component link is a TE tunnel that traverses R3 and R4. Both R3 and R4 are the nodes in the MPLS. The third component link is formed by lower layer network that has GMPLS enabled. In this case, R5 and R6 are not the nodes controlled by the MPLS but provide the connectivity for the component link. Note: if two unidirectional LSPs are used to construct a component link, they MUST be co-routed.

Composite link forms one logical link between connected routers and is used to carry aggregated traffic.[CL-REQ] Composite link relies on its component links to carry the traffic over the composite link. This means that a composite link maps incoming traffic into component links. The router (R1 in Figure 1) of composite link ingress maps a set of traffic flows including control plane packets to a specific component link. The router (R2 in Figure 1) of composite link egress receives the packets from its component links and sends them to MPLS forwarding engine like a regular link. The traffic from R2 to R1 is distributed by the router R2.

Traffic mapping to component links may be done by control plane, management plane, or data plane.[CL-REQ] The objectives are to keep the individual flow packets in sequence and meet its QoS criteria, do not overload any component link, and be able to perform local recovery when one of component link fails.[CL-REQ] Operator may have other objectives such as placing a bi-directional flow or LSP on the same component link in both direction, load balance over component links, composite link energy saving, and etc. A flow may be a LSP, or sub-LSP [MLSP], PW, a flow within PW [FAT-PW], entropy flow in LSP [ENTROPY].

4. Composite Link in Control Plane

A composite Link is advertised as a single logical interface between two connected routers, which forms routing and forwarding adjacency between the routers in IGP. The interface parameters for the composite link can be pre-configured by operator or be derived from its component links. Composite link advertisement requirements are specified in [CL-REQ].

In IGP-TE, a composite link is advertised as a single TE link between two connected routers. This is similar to a link bundle [RFC4201]. Link bundle applies to a set of homogenous component links. Composite link allows homogenous and non-homogenous component links. In order for route computing engine to calculate a proper path for a LSP, it is necessary for composite link to advertise the summarized available BW as well as the maximum BW for single LSP. If a composite link contains some non-homogeneous component links, the composite link also need to advertise the summarized BW and the maximum BW for single LSP per each homogeneous component link group. The protocol extension for composite link advertisement will be described in the separated draft.

Current IGP Link announcement protocol does not give such capability. The protocol only allows advertising one cost based TE metrics. RFC4201 only advertises a largest available BW for a link bundle that includes a set of homogenous component links. Signaling protocol does not allow signaling an aggregated LSP where its BW may be larger than the capacity of any component links. [ENTROPY]

A composite link may contain the set of component links. A component link may be configured by operator or signaled by the control plane. If two unidirectional LSPs are used to construct a component link, they MUST be co-routed. In both cases, it is necessary to convey component link parameters to the composite link.[CL-REQ] In an often case, a component link is configured to carry traffic under normal operation; under some situation, operator may want to configure a component link as a link for local recovery purpose, in which the composite link should not count the component link TE parameters into the composite link TE parameters for the advertisement. However, the composite link may send the traffic over the component link when a component link failure occurs. The protocol extension for singling component link parameter is for further study. This capability is not supported in current protocol because a link bundle contains a set of homogenous links.

When a component link is supported by lower layer network (third component link in figure 2), the control plane that the composite link resides is able to interoperate with the GMPLS or MPLS-TP control plane that lower layer network uses for component link addition and deletion.[CL-REQ]

It is possible for operator to configure one or multiple interface (s) over a composite link.

Both LDP [RFC5036] and RSVP-TE [RFC3209] can be used to signal a LSP over a composite link. The router of composite link ingress MUST place the LSP on the component link that meets the LSP criteria indicated in the signaling message. Since the composite link brings some unique characteristics, some signaling protocol extensions are expected to facilitate the composite link to place LSP to a proper component link. Several cases may be considered as below. Signaling extension for configuring such LSP is for further study.

A composite link may contain non-homogeneous component links. The route computing engine may select one group of component links for a LSP, it is necessary for signaling protocol to be able to indicate which group of component links for signaled LSP. Composite link MUST place the LSP to the component link whose performance such as delay or jitter is equal or better than the required.

Since composite link capacity is aggregated capacity and is larger than individual component link capacity, it is possible to signal a LSP whose BW is larger than individual component link capacity.[CL-REQ] Assumption is such LSP carrying an aggregated traffic.

When a bi-directional LSP request is signaled over a composite link, if the request indicates that the LSP must be placed on the same component link, the routers of the composite link MUST place the LSP traffic in both directions on a same component link.

Individual component link may fail independently. Upon component link failure, a composite link needs to perform local recovery, i.e. use some available capacity in other component link to carry impacted traffic. Composite link may use LDP or RSVP-TE signaling protocol to facilitate the recovery process between two routers that composite resides. To avoid looped crankback, during the local recovery process, the composite link should advertise its available BW as zero; after finishing the local recovery, it should update its proper available BW in IGP.

5. Composite Link in Data Plane

The traffic over a composite link is distributed over individual component links. Traffic dissemination may be determined by control plane, management plane, or data plane, and may be changed due to component link status change.[CL-REQ] The distribution function is local to the routers in which a composite link belongs to and is not specified here. However, if a bi-directional LSP is required to be placed on the same component link in both directions, the routers at both composite link end points need incorporation in determining the component link for the LSP. The protocol extension of that is for further study.

When performing traffic placement, a composite link does not differentiate multicast traffic vs. unicast traffic.

A component link in a composite link may fail independently. The routers at a composite link MUST maintain each component link status. Two routers may use the control plane to sync up a component link state. When a component link fails, the routers of a composite link MUST re-assign impacted flows to other active component links in minimal disruptive manner. This is local function and do not incorporate with LSP head-end routers. When a composite link is not able to transport all flows, it preempts some flows based upon local management configuration and informs the control plane on these preempted flows. The composite link MUST support soft preemption [RFC5712]. This action ensures the remaining traffic is transported properly. Note: as mentioned in section 4, in order to prevent lopped crankback symptom, when composite link performs local recovery process, it should advertise its available BW as zero; when the local process completes, the composite link should update its link state with the proper available BW.

The composite link functions provide component link fault notification and composite link fault notification. Component link fault notification MUST be sent to the management plane. Composite

link fault notification MUST be sent to management plane and distribute via link state message in IGP.

Operator may want to perform an optimization function such as load balance or energy saving over a composite link, which may conduct some traffic moving from one component link to another. The process MUST support locally and gracefully traffic movement process among component links. The protocol that facilitates this process between two composite link end points is for further study.

6. Composite Link in Management Plane

Management Plane MUST keep tracking a composite link and its individual composite link status and configuration. Management Plane MUST be able to make any component link in a composite link active and de-activate in order to facilitate operation maintenance task. The routers of a composite link resides MUST perform the redistribution of the traffic flows on a de-activated link to other component links based on the traffic flow TE criteria.

Management Plane MUST be able to configure a LSP over a composite link and be able to select a component link for the LSP.

Management Plane MUST be able to trace which component link a LSP is assigned to and monitor individual component link and composite link performance.

Management Plane MUST be able to ping individual component link within a composite link.

Management Plane should build the proper commands to allow operator execute an optimization process.

7. Security Considerations

For further study.

8. IANA Considerations

IANA actions to provide solutions are for further study.

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.

[RFC3209] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," December 2001

[RFC4201] Kompella, K., "Link Bundle in MPLS Traffic Engineering", RFC 4201, March 2005.

[RFC5036] Andersson, L., "LDP Specification", RFC 5036 , October 2007.

[RFC5712] Meyer, M., "MPLS Traffic Engineering Soft Preemption", January 2010.

9.2. Informative References

[CL-REQ] Villamizar, C. and McDysan, D, "Requirements for MPLS Over Composite Link", Oct. 2010, Work in Progress

[ENTROPY] Kompella, K. and S. Amante, "The Use of Entropy Labels in MPLS Forwarding", draft-ietf-mpls-entropy-label-00.txt, May 2011, Work in Progress

[FAT-PW] Bryan, S., et. Al, "Flow Aware Transport of Pseudowire over an MPLS PSN", draft-ietf-pwe3-fat-pw-06, May. 2011, Work in progress

[MLSP] Kompella, K. "Multi-path Label Switched Paths Signaled Using RSVP-TE", draft-kompella-mpls-rsvp-ecmp-00.txt, July 2010, Work in Progress

10. Acknowledgments

Authors would like to thank Adrian Farrel, Fred Jounay, Yuji Kamite for his extensive comments and suggestions, Ron Bonica, Nabil Bitar, Eric Gray, Lou Berger, and Kireeti Kompella for their reviews and great suggestions.

Authors' Addresses

So Ning
Verizon
2400 N. Glem Ave.,
Richerdson, TX 75082
Phone: +1 972-729-7905
Email: ning.so@verizonbusiness.com

Andrew Malis
Verizon
117 West St.
Waltham, MA 02451
Phone: +1 781-466-2362
Email: andrew.g.malis@verizon.com

Dave McDysan
Verizon
22001 Loudoun County PKWY
Ashburn, VA 20147
Email: dave.mcdysan@verizon.com

Lucy Yong
Huawei USA
1700 Alma Dr. Suite 500
Plano, TX 75075
Phone: +1 469-229-5387
Email: lucyyong@huawei.com

Curtis Villamizar
Infinera
Email: cvillamizar@infinera.com

Tony Li
Cisco Systems
Email: tony.li@tony.li