Network Working Group   N. So
Internet Draft   A. Malis
Intended Status: Informational   D. McDysan
Expires: September 2010   Verizon
                                                       L. Yong
                                                        Huawei
                                                     F. Jounay
                                                France Telecom
                                                     Y. Kamite
                                                           NTT
                                                March 5, 2010

Framework for MPLS Over Composite Link
draft-so-yong-rtgwg-cl-framework-01

Status of this Memo

Copyright Notice

include Simplified BSD License text as described in Section 4.e of

Abstract

   This document describes a framework for managing aggregated traffic
   over a composite link. A composite link consists of a group of non-
   homogenous links that have the same forward adjacency and can be
   considered as a single TE link or an IP link used for MPLS. The
   document primarily focuses on MPLS data plane and control plane
   traffic over a composite link and composite link advertisement within
   IGP. MPLS traffic can be set up by either RSVP-TE or LDP signaling.
   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

   This document describes a framework of Composite Link in IP/MPLS
   network. Single link and link bundle [RFC4201] have been widely used

in today's IP/MPLS networks. A link bundle 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 Requirement]. This
document states composite link framework in the context of MPLS
network with MPLS control plane.

A composite link is a single logical link 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. To achieve the better component link utilization and
avoid component link congestion, the document describes some new
aspects on the traffic flow assignment to component links. The
document also describes the advertisement of a composite link as a
routing interface in MPLS control plane and the addition of component
links to a composite link.

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. Acronyms

BW: Bandwidth

ECMP: Equal Cost Multi-Path

FRR: Fast Re-Route

LAG: Link Aggregation Group

LDP: Label Distribution Protocol

LSA: Link State Advertisements

LSP: Label Switched Path

MPLS: Multi-Protocol Label Switching

OAM: Operation, Administration, and Management

PDU: Protocol Data Unit

PE: Provider Edge device

RSVP: ReSource reserVation Protocol

RTD: Real Time Delay

TE: Traffic Engineering

VRF: Virtual Routing and Forwarding

2.2. Terminology

Composite Link: a group of component links, which can be considered as a single MPLS TE link or as a single IP link used for MPLS.

Component Link: a physical link (e.g., Lambda, Ethernet PHY, SONET/ SDH, OTN, etc.) with packet transport capability, or a logical link (e.g., MPLS LSP, Ethernet VLAN, MPLS-TP LSP, etc.)

Connection: An aggregation of traffic flows which are treated together as a single unit by Composite Link Interior Functions for the purpose of routing onto a specific component link and measuring traffic volume.

Exterior Functions: These are performed by an MPLS router that makes a composite link useable by the network via control protocols, or by an MPLS router that interacts with other routers to dynamically configure a component link within a composite link. These functions are those that interact via routing and/or signaling protocols with other routers in the same layer network or other layer networks.

Interior Functions: Actions performed by the MPLS routers directly connected by a composite link. This includes the determination of the component link on which a traffic flow is placed. This is local implementation.

Traffic Flow: A set of packets with a common identifier and characteristics that is used by Composite link interior functions to place the set of packets on the same component link. Identifiers can be an MPLS label stack or any combination of IP addresses and protocol types for routing, signaling, and management packets.

Virtual Interface: Composite link is advertised as an interface in IGP

3. Composite Link Framework

3.1. Composite Link Architecture

[ITU-T G.800] defines Composite Link as: multiple parallel links between the same subnetworks can be bundled together into a single

composite link. Each component link of the composite link is
independent in the sense that each component link is supported by a
separate server layer trail. The composite link conveys communication

information using different server layer trails thus the sequence of
symbols crossing this link may not be preserved. This definition
implies that a composite link appears as a single logical link
between two nodes and the component links in a composite link can
have the same or different properties such as latency and capacity.
Composite link framework 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, or can be a logical
link configured on a LSP [LSP Hierarchy]. Even if the transport
technology implementing the component links is identical, the
characteristics (e.g., bandwidth, latency) of the component links may
differ.

Composite Link can apply to MPLS network. As shown in Figure 1, the
composite link may carry LSP traffic flows and control plane 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 between two connected routers 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.

Composite link relies on its component links to carry the traffic. An
important framework concept is the connection shown in Figure 1. The
connection only exists within the composite link and is controlled by
the composite link function. At the transmitting end (R1 in Figure
1), composite link maps a set of traffic flows including control
plane packets into one connection and routes the connection on a
specific component link. This method enables connection based routing
over component links and traffic volume measurement on a per
connection basis. As a result, it guarantees the ordering per traffic
flow. One component link may carry one or more connections. At the
receiving end (R2 in Figure 1), composite link receives traffic flows
from its component links and sends them to MPLS forwarding engine
like a regular link. Note that a special case of this model is where
a single flow is mapped to a single connection.

```
                Management Plane
            Configuration and Measurement <------------+
                        ^                              |
                        |                              |
                        |                              |
                        |                              |
              +-------+-+                        +-+-------+
              |       | |                        | |      |
        CP Packets   V |                        | V   CP Packets
              |  V  +-+-+     Component Link 1    +-+-+  ^  |
              |  |  | |   |=========================|   | |  |
              |  +--|-|*|****** connections ********|*|-|--+  |
             ~~|~~>~~|~|  |=========================|  |~|~~>~~|~~
       LSP   ~~|~~>~~|~|  |     Component Link 2    |  |~|~~>~~|~~ LSP
       Traffic~|~~>~~|~|  |=========================|  |~|~~>~~|Traffic
       Flows ~~|~~>~~|~|*|****** connections ********|*|~|~~>~~|~ Flows
             ~~|~~>~~|~|  |=========================|  |~|~~>~~|~~
             ~~|~~>~~|~|  |     Component Link 3    |  |~|~~>~~|~~
             ~~|~~>~~|~|  |=========================|  |~|~~>~~|~~
              |     | |*|****** connections ********|*|  |     |
              |     | | |=========================|   |     |
              |     +---+                        +---+     |
              +--------+                        +--------+
                   !    !                          !    !
                   !    !<---- Component Links ---->!    !
                   !<------- Composite Link ---------->!
```
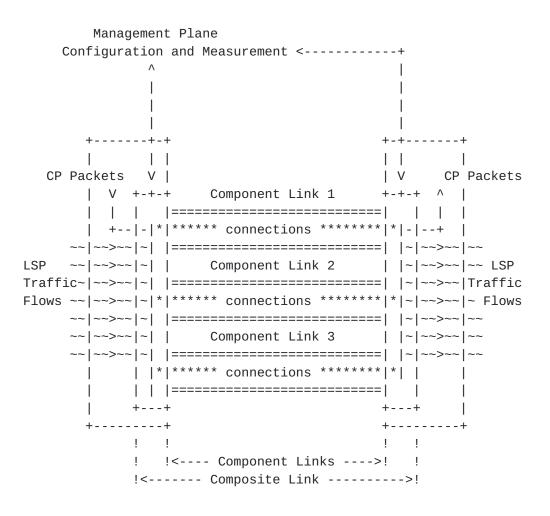
Figure 1 Composite Link Architecture Model

4.

4.1. Interior Functions

The interior functions are implemented within the interior of MPLS
routers connected via a composite link. This includes local
determination of the component link on which a traffic flow is
placed, i.e. the flow to the interface mapping. Although this
assignment is local, management configuration for some aspects of
these interior functions is important to achieve operational
consistency.

As described in section 3.1, the Interior Functions map traffic flows
to a connection and route connections to component links based on
traffic flow parameters, connection parameters, and component link
parameters. The connection parameters may be calculated from the
parameters of the flows mapped to the connection or may derive from

the measurement of the traffic volume on the connection plus
configured connection parameters via the management plane.

A component link in a composite link may fail independently. The
interior functions are able to detect component link failure and re-
assign impacted flows to other active component links. When a
composite link can't recover some impacted flows, it notifies control
plane about the flows. 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.
This action ensures the remaining traffic is transported properly.

The interior functions provide component link fault notification and
composite link fault notification upon the failure of a component
link or a composite link. Component link fault notification is sent
to the management plane. Composite link fault notification is sent to
the control plane and management plane. Composite link allows
operator to trace which component link a LSP is assigned to.

4.1.1. Functions specific to LSP flows with TE information

When a composite link is deployed in a MPLS-TE network, LSPs are
signaled by RSVP-TE protocol. The protocol message contains LSP
parameters like BW, setting/holding priority, latency, etc. The
composite link will take these parameters into account when assigning
LSP to a connection and a connection to a component link.

4.1.2. Functions specific to LSP flows without TE information

When a composite link is deployed in a non-TE MPLS network, LSPs are
signaled by LDP. LDP message does not provide LSP TE information.
These LSPs can be mapped to local configured connections with minimum
bandwidth, maximum bandwidth, preemption priority, and holding
priority parameters. The interior function measures the traffic
volume on the connection and derives the BW parameter for the
connection. The interior functions use these parameters to assign the
connection to a component link.

When the connection bandwidth exceeds the component link capacity,
the interior functions are able to reassign the traffic flows to
other connections that are over different component links.

4.1.3. Functions for LSP flows with and without TE information

When a composite link carries LSPs that are signaled by LDP or RSVP-
TE, the interior functions separates them into different connections
that have independent parameters. If there is a shortage of capacity
in a composite link, interior functions perform preemption, rerouting
or termination of flows based on the traffic flow priority.

[4.2](). Exterior Functions

The Exterior Functions have the aspects that are applicable exterior
to the connected MPLS routers. These exterior functions apply to
routing and signaling protocols. Protocol solutions are for further
study.

   4.2.1. Composite Link Advertisement and Component Link Setup

In IGP, a composite Link is advertised as a single virtual routable
interface between two connected routers, which forms routing and
forwarding adjacency between the connected routers. A composite link
must be advertised as a link in IGP.[RFC2328] The interface
parameters for the composite link can be pre-configured by operator
or be derived from its component links. Composite link latency may be
advertised for other routers to perform route computation.

A composite link may contain the set of component links. A component
link may be configured by operator or signaled by the control plane.
In both cases, it is necessary to convey component link parameters to
the composite link.

When a component link is supported by lower layer network as
described in section 4.2, 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.

Composite link advertisement and component link setup requirements
are specified in [CL Requirement]

   4.2.2. Functions specific to LSP flows with TE information

When RSVP-TE [RFC3209] signals a LSP over a composite link, the
composite link takes the signaled LSP parameters as the flow
parameters and sends the PATH message to the downstream router of the
composite link. The downstream router selects the LSP label that is
used over the composite link and sends RESV message with the label
info. to the upstream LSR. The forward engine at the upstream LSR
will push the label into the LSP packets. Composite link interior
functions map the label to a connection that is assigned to a
component link.

When a composite link is advertised as a TE link in MPLS-TE network
[RFC3630], composite link capacity is advertised. It is possible to
advertise multiple capacities and latencies to reflect individual
component link property. [CL requirement]

   4.2.3. Functions specific to LSP flows without TE information

When LDP [RFC5036] is used to signal LSPs over a composite link, a
LDP session has to be established between two LSRs connected via a

composite link. When signaling a LSP for a FEC over the composite
link, the upstream LSR sends label request message to the downstream
LSR, the downstream LSR selects a label for the FEC and sends the
label mapping message back to the upstream LSR.

4.2.4. Functions to LSP flows with and without TE information

When there is a shortage of capacity in a composite link which
supports LSP flows with and without TE information, interior
functions are able to perform preemption, rerouting or termination of
flows to ensure transport quality, the exterior functions are able to
inform LSR control plan preempted flows. Thus the control plan
informs the LSP head-end if it is singled by RSVP-TE or prunes label
distribution for the LSP that is singled by LDP.

5. Applicability

A composite link may be deployed between two Ps, P and PE, and two
PEs. Two routers connected via a composite link forms forwarding
adjacency and routing adjacency. It can be used in MPLS-TE only, MPLS
non-TE only, or combined environment.

In MPLS-TE application, the composite link is advertised as a link in
IGP and a TE link in IGP TE extension that uses either OSPF-TE or IS-
IS-TE. RSVP-TE signaling protocol is used to establish a LSP. TE
tunnel and private line are examples of applications.

In MPLS non-TE application, the composite link is advertised as a
link within the IGP that uses either OSPF or IS-IS as routing
protocol. In this case, LDP signaling protocol is used to establish a
LSP. L3VPN is the most popular application for this case.

In the hybrid application, the composite link is advertised as a link
and a TE link within IGP by using different types of LSA and both
RSVP-TE and LDP signaling for LSP establishment, as described in
section 3.

5.1. Single-Layer Scenarios

The component links of Figure 1 applies to at least the scenarios
illustrated in Figure 2. The component links may be physical or
logical, or supported by different technologies. In the first
scenario, a set of physical links connect adjacent (P) routers
(R1/R2).

In the second scenario, a set of logical links connect adjacent (P or
PE) routers over other equipment (i.e., R3/R4) that may implement
RSVP-TE signaled MPLS tunnels which may be in the same IGP as R1/R2.
In this case, R3/R4 provides a TE-LSP segment of TE-LSP from R1 and

R2. In another case, R3/R4 are in different IGP from R1/R2, R3 and R4
provide connectivity for R1 and R2. The connectivity provides a

component link in the composite link of R1/R2. R3 and P4 may
implement MPLS-TP.

```
   +-------+                    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 Types

5.2. Multi-Layer Scenario

   In the third scenario in Figure 2, GMPLS lower layer LSPs (e.g.,
   Fiber, Wavelength, TDM) as determined by a lower layer network in a
   multi-layer network deployment as illustrated by R5/R6. In this case,
   R5 and R6 would usually not be part of the same IGP as R1/R2 and may
   have a static interface, or may have a signaling but not a routing
   association with R1 and R2. Note that in scenarios 2 and 3 when the
   intermediate routers are not part of the same IGP as R1/R2 (i.e., can
   be viewed as operating at a lower layer) that the characteristics of
   these links (e.g., latency) may change dynamically, and there is an
   operational desire to handle this type of situation in a more
   automated fashion than is currently possible with existing protocols.
   Note that this problem currently occurs with a single lower-layer
   link in existing networks and it would be desirable for the solution
   to handle the case of a single lower-layer component link as well.
   Note that the interfaces at R1 and R2 are associated with these
   different component links can be configured with IP addresses or use
   unnumbered links as an interior, local function since the individual
   component links are not advertised as the virtual interface.

6. Security Considerations

The solution is a local function on the router to support traffic
engineering management over multiple parallel links. It does not

introduce a security risk for control plane and data plane.

The solution could change the frequency of routing update messages and therefore could change routing convergence time. The solution MUST provide controls to dampen the frequency of such changes so as to not destabilize routing protocols.

7. IANA Considerations

IANA actions to provide solutions are for further study.

8. References

8.1. Normative References

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

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

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

[RFC3630] Katz, D., "Traffic Engineering (TE) Extension to OSPF Version 2", RFC 3630, September 2003.

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

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

8.2. Informative References

[Entropy Label] Kompella, K. and S. Amante, "The Use of Entropy Labels in MPLS Forwarding", November 2008, Work in Progress

[LSP Hierarchy] Shiomoto, K. and A. Farrel, "Procedures for Dynamically   Signaled Hierarchical Label Switched Paths", November 2008, Work in Progress

[Soft Preemption] Meyer, M. and J. Vasseur, "MPLS Traffic Engineering Soft Preemption", February 2009, Work in Progress

[CL Requirement] So, N. and Yong, L, "Requirements for MPLS Over Composite Link", Oct. 2009, Work in Progress

[9](#). Acknowledgments

   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

      Frederic Jounay
      France Telecom
      2, avenue Pierre-Marzin
      22307 Lannion Cedex,
      FRANCE
      Email: frederic.jounay@orange-ftgroup.com

      Yuji Kamite
      NTT Communications Corporation
      Granpark Tower
      3-4-1 Shibaura, Minato-ku
      Tokyo 108-8118
      Japan
      Email: y.kamite@ntt.com