Network Working Group   L. Jin
Internet-Draft   F. Jounay
Intended status: Informational   France Telecom
Expires: June 12, 2014   M. Bhatia
                                                    Alcatel-Lucent
                                                           S. Kini
                                                          Ericsson
                                                December 12, 2013

Hub and Spoke Multipoint Label Switched Path Tunnels
draft-ietf-mpls-rsvp-te-hsmp-lsp-01

Abstract

   There are applications that require bi-directional, co-routed and
   guaranteed communication from a root node to several leaf nodes in a
   hub and spoke fashion. To meet such application requirements in a
   Multi-protocol Label Switching (MPLS) network this draft defines a
   Hub and Spoke Multipoint Traffic Engineered Label Switched Path (HSMP
   TE LSP) with resource reservations for guaranteed communication.
   This draft also defines a protocol to setup such LSPs by re-using and
   extending P2MP RSVP-TE.

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 December 05, 2013.

Copyright Notice

Table of Contents

1. Introduction

   There are many applications that require one-to-many bi-directional
   communication. Some of these applications are described in Section 3
   along with their requirements from the network. Making such
   applications work over a MPLS network by using both P2MP and P2P
   constructs results in scalability issues and these are discussed in
   Section 4. This document defines a technique to do one-to-many bi-
   directional communication over an MPLS network that re-uses the
   existing P2MP and P2P constructs in MPLS but combines them in a
   scalable manner. This technique re-uses and extends the current
   traffic-engineered P2MP and P2P constructs and protocols. It is
   described in detail in Section 5.

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. Abbreviations and Terminology

   HSMP LSP - Hub and Spoke Multipoint LSP

3. Applications

  We describe two representative applications that require one-to-many
  bi-directional communication. The first is the 'Time
  Synchronization' application described in 2.1. The second is the
  P2MP pseudowire application and is described in section 2.

3.1. Time Synchronization

  Time Synchronization [IEEE1588] over an MPLS network is being defined
  in [I-D.ietf-tictoc-1588overmpls]. A scalable time-sync architecture
  requires the master to provide time synchronization to a large number
  of slaves. It requires the PTP messages to flow bi-directionally
  between master and slave in a hub and spoke manner. More importantly
  these messages must have the same delay in both directions. This
  requires the underlying network to reserve resources to transport PTP
  messages and also to co-route them in both directions to avoid any
  differences in the delays of the paths in both directions.

3.2. P2MP pseudowire

  A P2MP PW [I-D.ietf-pwe3-p2mp-pw] is required for the VPMS service
  [I-D.ietf-l2vpn-vpms-frmwk-requirements]. In this application the
  root PE requires bi-directional communication with several leaf PEs.
  The underlying MPLS transport should support this type of
  communication for the P2MP PW in a reliable and efficient manner.

4. Scalability issues

  A straightforward method to achieve one-to-many bi-directional
  communication with resource guarantees is to use a P2MP RSVP-TE
  tunnel from the hub PE to the spoke PEs and use P2P RSVP-TE tunnels
  from each spoke PE to the hub PE. The spoke-to-hub P2P tunnels can
  be explicitly routed such that they are co-routed along the reverse
  direction of the P2MP tunnel. In this model, scalability issues
  arise both in the data plane and the control plane as explained below
  in this section.

For the purpose of this discussion, an application with a hub and
spoke bi-directional communication over a tree topology MPLS network
is illustrated in Figure 1. The hub LSR is A and the spoke LSRs are
E, F, G, and H. For communication from the hub to the spokes, a P2MP
LSP can be setup with A as the root and E, F, G and H as leaves. For
communication from the spokes to the hub, a P2P LSP can be setup from
each spoke to the hub.

```
                         A
                         |
                         B
                        / \
                       /   \
                      /     \
                    C         D
                   / \       / \
                  E   F     G   H
```
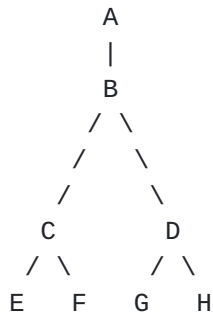
            Figure 1: Hub and spoke LSP over a tree topology

Each LSR along this tree will have to allocate a unique label for
each of the P2P LSPs that go from spoke to hub through it. This
leads to a linear increase in forwarding state at each LSR in
proportion to the number of spoke nodes that are in its sub-tree.
This has poor scaling characteristics in the data plane as the number
of spoke nodes increase.

Each LSR also has to allocate control plane state for each of the P2P
LSPs that go from spoke to hub through it. Each P2P LSP will need a
separate path state block (PSB) and a reservation state block (RSB)
and these will store additional information on signaling attributes.
This state is in addition to the state maintained for the P2MP LSP.
Clearly this state too increases linearly with the number of spoke
nodes that are in its sub-tree. This too has poor scaling
characteristics in the control plane as the number of spoke nodes
increase. Also the number of signaling messages increases linearly
though some of it may be mitigated by using refresh reduction
[RFC2961].

5. Hub and Spoke Multipoint LSP

To solve the issues identified in Section 4 this document defines a
hub and spoke traffic-engineered multipoint LSP (HSMP TE LSP) with
resource reservations. Such an LSP is a combination of an explicitly
routed uni-directional traffic-engineered P2MP LSP from the hub to
the spokes and a co-routed uni-directional MP2P LSP from the spokes
to the hub. The data plane for a HSMP TE LSP is explained in
Section 5.1 and the control plane is explained in Section 5.2

5.1. Data plane

In the direction from hub-to-spoke the data plane processing is the
same as that of a P2MP LSP. In the direction from the spokes to the
hub, each LSR allocates labels for its upstream LSRs. Each LSR
merges the traffic received from multiple upstream LSRs before
forwarding it on the LSP towards the hub. It should be noted that
due to label merging the GAL processing in the direction from spoke
to hub is not defined.

5.2. Control Plane

The signaling protocol to setup a HSMP TE LSP can re-use the
signalling protocol for P2MP RSVP-TE [RFC4875] with some extensions.
The hub and spokes of a HSMP LSP can be modeled the same as the
source and leaves respectively of a P2MP LSP. A source-to-leaf (S2L)
sub-LSP defined in [RFC4875] for the P2MP LSP is used to represent a
hub-to-spoke communication of the HSMP LSP. To signal the bi-
directional co-routed nature of the communication from the hub to the
spoke, the extensions defined in section 3 of [RFC3473] must be used.
Each Path message of a HSMP LSP LSR MUST have a Upstream_Label
object. If a PathErr is received in response with a "Routing problem
/Unacceptable label value" indication then the Acceptable Label Set
(if present) must be examined to allocate a label for the
Upstream_Label object. If an LSR signaling an HSMP LSP receives
PathErr messages with different Acceptable Label Sets from different
neighboring LSRs then it may need to allocate more than one label to
satisfy all the Acceptable Label Sets. The LSR should try to
minimize the number of unique labels allocated for a HSMP LSP in such
a case.

Pruning and grafting for a HSMP LSP follow the same procedures as for
a P2MP LSP. During re-merge in addition to the procedures in section
18.1.1 of [RFC4875] the ingress or transit LSR that creates the
branch would also be a re-merge LSR for the traffic from the spokes
towards the hub. Also the re-merge node for the traffic from hub to
spoke would be a branching node for traffic from the spokes to the
hub. The LSR that is branching the traffic from the spokes to the
hub would duplicate the traffic whereas the LSR that is re-merging
the traffic should forward traffic only from one the incoming
interfaces.

The HSMP LSP also requires bandwidth allocation that is asymmetric
between the hub-to-spoke and the spoke-to-hub direction. At the same
time it requires the spokes to be able to request different amounts
of bandwidth towards the hub. The protocol extensions defined in
[RFC6387] are used for asymmetric bandwidth allocation between the
hub-to-spoke and spoke-to-hub directions. The UPSTREAM_TSPEC,

UPSTREAM_ADSPEC and UPSTREAM_FLOWSPEC objects are also used in a HSMP
LSP. However in case of a HSMP LSP an intermediate LSR can receive
different UPSTREAM_TSPECs in the Resv messages from neighboring LSRs.
Combining these Tspecs and generating an appropriate
UPSTREAM_FLOWSPEC towards each spoke is still under discussion.
Also, processing a received UPSTREAM_FLOWSPEC and generating
appropriate UPSTREAM FLOWSPECs in the hub to spoke direction is also
under discussion.

6. Acknowledgements

We would like to thank Dimitri Papadimitriou, Yuji Kamite, Sebastien
Jobert for their comments and feedback on the document.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

The same security considerations apply as for the RSVP-TE P2MP LSP
[RFC4875] specification.

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.

[RFC3473]   Berger, L., "Generalized Multi-Protocol Label Switching
            (GMPLS) Signaling Resource ReserVation Protocol-Traffic
            Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC4875]   Aggarwal, R., Papadimitriou, D., and S. Yasukawa,
            "Extensions to Resource Reservation Protocol - Traffic
            Engineering (RSVP-TE) for Point-to-Multipoint TE Label
            Switched Paths (LSPs)", RFC 4875, May 2007.

[RFC6387]   Takacs, A., Berger, L., Caviglia, D., Fedyk, D., and J.
            Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label
            Switched Paths (LSPs)", RFC 6387, September 2011.

9.2. Informative References

   [I-D.ietf-l2vpn-vpms-frmwk-requirements]
              Kamite, Y., JOUNAY, F., Niven-Jenkins, B., Brungard, D.,
              and L. Jin, "Framework and Requirements for Virtual
              Private Multicast Service (VPMS)", draft-ietf-l2vpn-vpms-
              frmwk-requirements-05 (work in progress), October 2012.

   [I-D.ietf-pwe3-p2mp-pw]
              Sivabalan, S., Boutros, S., and L. Martini, "Signaling
              Root-Initiated Point-to-Multipoint Pseudowire using LDP",
              draft-ietf-pwe3-p2mp-pw-04 (work in progress), March 2012.

   [I-D.ietf-tictoc-1588overmpls]
              Davari, S., Oren, A., Bhatia, M., Roberts, P., and L.
              Montini, "Transporting Timing messages over MPLS
              Networks", draft-ietf-tictoc-1588overmpls-05 (work in
              progress), June 2013.

   [IEEE1588]
              IEEE 1588-2008, "IEEE Standard for a Precision Clock
              Synchronization Protocol for Networked Measurement and
              Control Systems ", 2008.

   [RFC2961] Berger, L., Gan, D., Swallow, G., Pan, P., Tommasi, F.,
              and S. Molendini, "RSVP Refresh Overhead Reduction
              Extensions", RFC 2961, April 2001.

   [RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute
              Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May
              2005.

Authors' Addresses

   Lizhong Jin

   Email: lizho.jin@gmail.com

   Frederic Jounay
   France Telecom

   Email: frederic.Jounay@orange.ch

Manav Bhatia
Alcatel-Lucent

Email: manav.bhatia@alcatel-lucent.com

Sriganesh Kini
Ericsson

Email: sriganesh.kini@ericsson.com