           Using LDP Multipoint Extensions on Targeted LDP Sessions

                  draft-napierala-mpls-targeted-mldp-02.txt

Abstract

   Label Distribution Protocol (LDP) can be used to set up
   Point-to-Multipoint (P2MP) and Multipoint-to-Multipoint (MP2MP) Label
   Switched Paths. The existing specification for this functionality
   assumes that a pair of LDP neighbors are directly connected.
   However, the LDP base specification allows for the case where a pair
   of LDP neighbors are not directly connected; the LDP session between
   such a pair of neighbors is known as a "Targeted LDP" session. This
   document specifies the use of the LDP P2MP/MP2MP extensions over a
   Targeted LDP session.

Table of Contents

1. Introduction

1.1. Specification of requirements

   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].

1.2. Targeted mLDP

   The Label Distribution Protocol (LDP) extensions for setting up
   Point-to-MultiPoint (P2MP) Label Switched Paths (LSPs) and
   Multipoint-to-Multipoint (MP2MP) LSPs are specified in [mLDP]. This
   set of extensions is generally known as "Multipoint LDP" (mLDP).

   A pair of Label Switched Routers (LSRs) that are the endpoints of an
   LDP session are considered to be "LDP neighbors". When a pair of LDP
   neighbors are "directly connected" (e.g., they are connected by a
   layer 2 medium, or are otherwise considered to be neighbors by the a
   network's interior routing protocol), the LDP session is said to be a
   "directly connected" LDP session. When the pair of LDP neighbors are
   not directly connected, the session between them is said to be a
   "Targeted" LDP session.

   The base specification for mLDP does not explicitly cover the case
   where the LDP multipoint extensions are used over a targeted LDP
   session. This document provides that specification.

   We will use the term "Multipoint" to mean "either P2MP or MP2MP".

1.3. Targeted mLDP and the Upstream LSR

1.3.1. Selecting the Upstream LSR

   In mLDP, a multipoint LSP (MP-LSP) has a unique identifier that is an
   ordered pair of the form <root, opaque value>. The first element of
   the ordered pair is the IP address of the MP-LSP's "root node". The
   second element of the ordered pair is an identifier that is unique in
   the context of the root node.

   If LSR D is setting up the MP-LSP <R, X>, D must determine the
   "upstream LSR" for <R, X>. In [mLDP], the upstream LSR for <R, X>,
   U, is defined to be the "next hop" on D's path to R, and "next hop"
   is tacitly assumed to mean "IGP next hop". It is thus assumed that
   there is a direct LDP session between D and U. In this
   specification, we extend the notion of "upstream LSR" to cover the

following cases:

  - U is the "BGP next hop" on D's path to R, where U and D are not
    IGP neighbors, and where there is a Targeted LDP session between
    U and D. In this case, we allow D to select U as the "upstream
    LSR" for <R,X>.

  - If the "next hop interface" on D's path to R is an RSVP-TE P2P
    tunnel whose remote endpoint is U, and if there is known to be an
    RSVP-TE P2P tunnel from U to D, and if there is a Targeted LDP
    session between U and D, then we allow D to select U as the
    "upstream LSR" for <R,X>. This is useful when D and U are part
    of a network area that is fully meshed via RSVP-TE P2P tunnels.

The particular method used to select an "upstream LSR" is determined
by the SP. Other methods than the ones above MAY be used.

1.3.2. Sending data from U to D

By using Targeted mLDP, we can construct an MP-LSP <R,X> containing
an LSR U, where U has one or more downstream LSR neighbors (D1, ...,
Dn) to which it is not directly connected. In order for a data
packet to travel along this MP-LSP, U must have some way of
transmitting the packet to D1, ..., Dn. We will cover two methods of
transmission:

  - Unicast Replication.

    In this method, U creates n copies of the packet, and unicasts
    each copy to exactly one of D1, ..., Dn.

  - Multicast tunneling.

    In this method, U becomes the root node of a multicast tunnel,
    with D1, ..., Dn as leaf nodes. When a packet traveling along
    the MP-LSP <R,X> arrives at U, U transmits it through the
    multicast tunnel, and as a result it arrives at D1, ..., Dn.

    When this method is used, it may be desirable to carry traffic of
    multiple MP-LSPs through a single multicast tunnel. We specify
    procedures that allow for the proper demultiplexing of the MP-
    LSPs at the leaf nodes of the multicast tunnel. We do not assume
    that all the leaf nodes of the tunnel are on all the MP-LSPs
    traveling through the tunnel; thus some of the tunnel leaf nodes
    may need to discard some of the packets received through the
    tunnel. For example, suppose MP-LSP <R1,X1> contains node U with
    downstream LSRs D1 and D2, while MP-LSP <R2,X2> contains node U

        with downstream LSRs D2 and D3. Suppose also that there is a
        multicast tunnel with U as root and with D1, D2, and D3 as leaf
        nodes. U can aggregate both MP-LSPs in this one tunnel.
        However, D1 will have to discard packets that are traveling on
        <R2,X1>, while D3 will have to discard packets that are traveling
        on <R1,X2>.

1.4. Applicability of Targeted mLDP

   When LSR D is setting up MP-LSP <R,X>, it MUST NOT use targeted mLDP
   unless D can select the "upstream LSR" for <R,X> using one of the
   procedures discussed in section 1.3.1.

   Whether D uses Targeted mLDP when this condition holds is determined
   by provisioning, or by other methods that are outside the scope of
   this specification.

   When Targeted mLDP is used, the choice between unicast replication
   and multicast tunneling is determined by provisioning, or by other
   methods that are outside the scope of this specification.

1.5. LDP Capabilities

   Per [mLDP], any LSR that needs to set up an MP-LSP must support the
   procedures of [LDP-CAP], and in particular must send and receive the
   P2MP Capability and/or the MP2MP Capability. This specification does
   not define any new capabilities; the advertisement of the P2MP and/or
   MP2MP Capabilities on a Targeted LDP session means that the
   advertising LSR is capable of following the procedures of this
   document.

   Some of the procedures of this document require the use of upstream-
   assigned labels [LDP-UP]. In order to use upstream-assigned labels
   as part of Targeted mLDP, an LSR must advertise the LDP Upstream-
   Assigned Label Capability [LDP-UP] on the Targeted LDP session.

2. Targeted mLDP with Unicast Replication

   When unicast replication is used, the mLDP procedures are exactly the
   same as described in [mLDP], with the following exception. If LSR D
   is setting up MP-LSP <R,X>, its "upstream LSR" is selected according
   to the procedures of section 1.3.1, and is not necessarily the "IGP
   next hop" on D's path to R.

   Suppose that LSRs D1 and D2 are both setting up the P2MP MP-LSP

<R,X>, and that LSR U is the upstream LSR on each of their paths to
R. D1 and D2 each binds a label to <R,X>, and each uses a label
mapping message to inform U of the label binding. Suppose D1 has
assigned label L1 to <R,X> and D2 has assigned label L2 to <R,X>.
(Note that L1 and L2 could have the same value or different values;
D1 and D2 do not coordinate their label assignments.) When U has a
packet to transmit on the MP-LSP <R,X>, it makes a copy of the
packet, pushes on label L1, and unicasts the resulting packet to D1.
It also makes a second copy of the packet, pushes on label L2, and
then unicasts the resulting packet to D2.

This procedure also works when the MP-LSP <R,X> is a MP2MP LSP.
Suppose that in addition to labels L1 and L2 described above, U has
assigned label L3 for <R,X> traffic received from D1, and label L4
for <R,X> traffic received from D2. When U processes a packet with
label L3 at the top of its label stack, it knows the packet is from
D1, so U sends a unicast copy of the packet to D2, after swapping L3
for L2. U does not send a copy back to D1.

Note that all labels used in this procedure are downstream-assigned
labels.

The method of unicast is a local matter, outside the scope of this
specification. The only requirement is that D1 will receive the copy
of the packet carrying label L1, and that D1 will process the packet
by looking up label L1. (And similarly, D2 must receive the copy of
the packet carrying label L2, and must process the packet by looking
up label L2.)

Note that if the method of unicast is MPLS, U will need to push
another label on each copy of the packet before transmitting it.
This label needs to ensure that delivery of the packet to the
appropriate LSR, D1 or D2. Use of penultimate-hop popping for that
label is perfectly legitimate.

[3](3). Targeted mLDP with Multicast Tunneling

Suppose that LSRs D1 and D2 are both setting up MP-LSP <R,X>, and
that LSR U is the upstream LSR on each of their paths to R. Since
multicast tunneling is being used, when U has a packet to send on
this MP-LSP, it does not necessarily send two copies, one to D1 and
one to D2. It may send only one copy of the packet, which will get
replicated somewhere downstream in the multicast tunnel. Therefore,
the label that gets bound to the MP-LSP must be an upstream-assigned
label, assigned by U. This requires a change from the procedures of
[mLDP]. D1 and D2 do not send label mapping messages to U; instead
they send label request messages to U, asking U to assign a label to

the MP-LSP <R,X>. U responds with a label mapping message containing
an upstream-assigned label, L (using the procedures specified in
[LDP-UP]). As part of the same label mapping message, U also sends
an Interface TLV (as specified in [LDP-UP]) identifying the multicast
tunnel in which data on the MP-LSP will be carried. When U transmits
a packet on this tunnel, it first pushes on the upstream-assigned
label L, and then pushes on the label that corresponds to the
multicast tunnel.

If the numerical value L of the upstream-assigned label is the value
3, defined in [LDP] and [RFC3032] as "Implicit NULL", then the
specified multicast tunnel will carry only the specified MP-LSP.
That is, aggregation of multiple MP-LSPs into a single multicast
tunnel is not being done. In this case, no upstream-assigned label
is pushed onto a packet that is transmitted through the multicast
tunnel.

Various types of multicast tunnel may be used. The choice of tunnel
type is determined by provisioning, or by some other method that is
outside the scope of this document. [LDP-UP] specifies encodings
allowing U to identify an mLDP MP-LSP, and RSVP-TE P2MP LSP, as well
as other types of multicast tunnel.

This document does not specify procedures for tunneling one or more
MP2MP LSPs through P2MP tunnels. While it is possible to do this, it
is highly RECOMMENDED that MP2MP LSPs be tunneled through MP2MP LSPs
(unless, of course, unicast replication is being used).

If the multicast tunnel is an mLDP MP-LSP or an RSVP-TE P2MP LSP,
when U transmits a packet on the MP-LSP <R,X>, the upstream-assigned
label L will be the second label in the label stack. Penultimate-hop
popping MUST NOT be done, because the top label provides the context
in which the second label is to be interpreted. See [RFC5331].

When LSR U uses these procedures to inform LSR D that a particular
MP-LSP is being carried in a particular multicast tunnel, U and D
MUST take appropriate steps to ensure that packets U sends into this
tunnel will be received by D. The exact steps to take depend on the
tunnel type. As long as U is D's upstream LSR for any MP-LSP that
has been assigned to this tunnel, D must remain joined to the tunnel.

Note that U MAY assign the same multicast tunnel for multiple
different MP-LSPs. However, U MUST assign a distinct upstream-
assigned label to each MP-LSP. This allows the packets traveling
through the tunnel to be demultiplexed into the proper MP-LSPs.

If U has an MP-LSP <R1,X1> with downstream LSRs D1 and D2, and an MP-
LSP <R2,X2> with downstream LSRs D2 and D3, U may assign both MP-LSPs

to the same multicast tunnel. In this case, D3 will receive packets
traveling on <R1,X1>. However, the upstream-assigned label carried
by those packets will not be recognized by D3, hence D3 will discard
those packets. Similarly, D1 will discard the <R2,X2> packets.

This document does not specify any rules for deciding whether to
aggregate two or more MP-LSPs into a single multicast tunnel. Such
rules are outside the scope of this document.

Except for the procedures explicitly detailed in this document, the
procedures of [mLDP] and [LDP-UP] apply unchanged.

4. IANA Considerations

This document has no considerations for IANA.

5. Security Considerations

This document raises no new security considerations beyond those
discussed in [LDP], [LDP-UP], and [RFC5331].

6. Acknowledgments

The authors wish to think Lizhong Jin for his comments.

7. Authors' Addresses

Maria Napierala
AT&T Labs
200 Laurel Avenue, Middletown, NJ 07748
E-mail: mnapierala@att.com

Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA, 01719
E-mail: erosen@cisco.com

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a Diegem 1831
Belgium
E-mail: ice@cisco.com

8. Normative References

[LDP] Loa Andersson, Ina Minei, Bob Thomas, editors, "LDP Specification", RFC 5036, October 2007

[LDP-CAP] Bob Thomas, Kamran Raza, Shivani Aggarwal, Rahul Aggarwal, Jean-Louis Le Roux, "LDP Capabilities", RFC 5561, July 2009

[mLDP] Ina Minei, Kireeti Kompella, IJsbrand Wijnands, Bob Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", draft-ietf-mpls-ldp-p2mp-15.txt, August 2011

[LDP-UP] Rahul Aggarwal, Jean-Louis Le Roux, "MPLS Upstream Label Assignment for LDP", draft-ietf-mpls-ldp-upstream-10.txt, February 2011

[RFC2119] "Key words for use in RFCs to Indicate Requirement Levels.", Bradner, March 1997

[RFC3032] Eric Rosen, et. al., "MPLS Label Stack Encoding", RFC 3032, January 2001

[RFC5331] Rahul Aggarwal, Yakov Rekhter, Eric Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, August 2009