MPLS Working Group
Internet Draft
Updates: 5036 (if approved)
Intended status: Standards Track
Expires: June 8, 2014

Rajiv Asati
Cisco

Vishwas Manral
Hewlett-Packard, Inc.

Rajiv Papneja
Huawei

Carlos Pignataro
Cisco

December 8, 2013

Updates to LDP for IPv6
draft-ietf-mpls-ldp-ipv6-10

Status of this Memo

Copyright Notice

Abstract

The Label Distribution Protocol (LDP) specification defines procedures to exchange label bindings over either IPv4, or IPv6 or both networks. This document corrects and clarifies the LDP behavior when IPv6 network is used (with or without IPv4). This document updates RFC 5036.

Table of Contents

1. Introduction

The LDP [RFC5036] specification defines procedures and messages for exchanging FEC-label bindings over either IPv4 or IPv6 or both (e.g. dual-stack) networks.

However, RFC5036 specification has the following deficiencies in regards to IPv6 usage:

1) LSP Mapping: No rule defined for mapping a particular packet to a particular LSP that has an Address Prefix FEC element containing IPv6 address of the egress router

2) LDP Identifier: No details specific to IPv6 usage

3) LDP Discovery: No details for using a particular IPv6 destination (multicast) address or the source address (with or without IPv4 co-existence)

4) LDP Session establishment: No rule for handling both IPv4 and IPv6 transport address optional objects in a Hello message, and subsequently two IPv4 and IPv6 transport connections

5) LDP Label Distribution: No rule for advertising IPv4 or/and IPv6 FEC-label bindings over an LDP session, and denying the co-existence of IPv4 and IPv6 FEC Elements in the same FEC TLV

6) Next Hop Address & LDP Identifier: No rule for accommodating the usage of duplicate link-local IPv6 addresses

7) LDP TTL Security: No rule for built-in Generalized TTL Security Mechanism (GTSM) in LDP

This document addresses the above deficiencies by specifying the desired behavior/rules/details for using LDP in IPv6 enabled networks (IPv6-only or Dual-stack networks).

Note that this document updates RFC5036.

1.1. Scope

1.1.1. Topology Scenarios

The following scenarios in which the LSRs may be inter-connected via one or more dual-stack interfaces (figure 1), or one or more single-stack interfaces (figure 2 and figure 3) are addressed by this document:

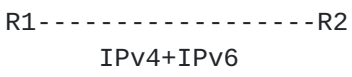

Figure 1 LSRs connected via a Dual-stack Interface

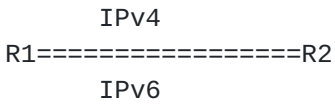

Figure 2 LSRs connected via two single-stack Interfaces

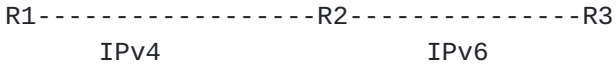

Figure 3 LSRs connected via a single-stack Interface

Note that the topology scenario illustrated in figure 1 also covers the case of a single-stack interface (IPv4, say) being converted to a dual-stacked interface by enabling IPv6 routing as well as IPv6 LDP, even though the IPv4 LDP session may already be established between the LSRs.

Note that the topology scenario illustrated in figure 2 also covers the case of two routers getting connected via an additional single-stack interface (IPv6 routing and IPv6 LDP), even though the IPv4 LDP session may already be established between the LSRs over the existing interface(s).

1.1.2. LDP TTL Security

LDP TTL Security mechanism specified by this document applies only to single-hop LDP peering sessions, but not to multi-hop LDP peering sessions, in line with Section 5.5 of [RFC5082] that describes Generalized TTL Security Mechanism (GTSM).

As a consequence, any LDP feature that relies on multi-hop LDP peering session would not work with GTSM and will warrant (statically or dynamically) disabling GTSM. Please see section 8.

2. Specification 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 [RFC2119].

Abbreviations:

LDP - Label Distribution Protocol

LDPv4 - LDP for enabling IPv4 MPLS forwarding

LDPv6 - LDP for enabling IPv6 MPLS forwarding

LDPoIPv4 - LDP over IPv4 transport session

LDPoIPv6 - LDP over IPv6 transport session

FEC - Forwarding Equivalence Class

TLV - Type Length Value

LSR - Label Switch Router

LSP - Label Switched Path

LSPv4 - IPv4-signaled Label Switched Path [RFC4798]

LSPv6 - IPv6-signaled Label Switched Path [RFC4798]

AFI - Address Family Identifier

3. LSP Mapping

Section 2.1 of [RFC5036] specifies the procedure for mapping a particular packet to a particular LSP using three rules. Quoting the 3rd rule from RFC5036:

> "If it is known that a packet must traverse a particular egress router, and there is an LSP that has an Address Prefix FEC element that is a /32 address of that router, then the packet is mapped to that LSP."

Suffice to say, this rule is correct for IPv4, but not for IPv6, since an IPv6 router may not have any /32 address.

This document proposes to modify this rule by also including a /128 address (for IPv6). In fact, it should be reasonable to just say IPv4 or IPv6 address instead of /32 or /128 addresses as shown below in the updated rule:

> "If it is known that a packet must traverse a particular egress router, and there is an LSP that has an Address Prefix FEC element that is an IPv4 or IPv6 address of that router, then the packet is mapped to that LSP."

4. LDP Identifiers

Section 2.2.2 of [RFC5036] specifies formulating at least one LDP Identifier, however, it doesn't provide any consideration in case of IPv6 (with or without dual-stacking).

The first four octets of the LDP identifier, the 32-bit LSR Id (e.g. (i.e. LDP Router Id), identify the LSR and is a globally unique value within the MPLS network. This is regardless of the address family used for the LDP session. Hence, this document preserves the usage of 32-bit (unsigned non-zero integer) LSR Id on an IPv6 only LSR (note that BGP has also mandated using 32-bit BGP Router ID on an IPv6 only Router [RFC6286]).

> Please note that 32-bit LSR Id value would not map to any IPv4-address in an IPv6 only LSR (i.e., single stack), nor would there be an expectation of it being DNS-resolvable. In IPv4 deployments,

the LSR Id is typically derived from an IPv4 address, generally assigned to a loopback interface. In IPv6 only deployments, this 32-bit LSR Id must be derived by some other means that guarantees global uniqueness within the MPLS network, similar to that of BGP Identifier [RFC6286].

This document qualifies the first sentence of last paragraph of Section 2.5.2 of [RFC5036] to be per address family and therefore updates that sentence to the following: "For a given address family, an LSR MUST advertise the same transport address in all Hellos that advertise the same label space." This rightly enables the per-platform label space to be shared between IPv4 and IPv6.

In summary, this document not only allows the usage of a common LDP identifier i.e. same LSR-Id (aka LDP Router-Id), but also the common Label space id for both IPv4 and IPv6 on a dual-stack LSR.

This document reserves 0.0.0.0 as the LSR-Id, and prohibits its usage.

5. Peer Discovery

5.1. Basic Discovery Mechanism

Section 2.4.1 of [RFC5036] defines the Basic Discovery mechanism for directly connected LSRs. Following this mechanism, LSRs periodically sends LDP Link Hellos destined to "all routers on this subnet" group multicast IP address.

Interesting enough, per the IPv6 addressing architecture [RFC4291], IPv6 has three "all routers on this subnet" multicast addresses:

FF01:0:0:0:0:0:0:2 = Interface-local scope

FF02:0:0:0:0:0:0:2 = Link-local scope

FF05:0:0:0:0:0:0:2 = Site-local scope

[RFC5036] does not specify which particular IPv6 'all routers on this subnet' group multicast IP address should be used by LDP Link Hellos.

This document specifies the usage of link-local scope e.g. FF02:0:0:0:0:0:0:2 as the destination multicast IP address in IPv6 LDP Link Hellos. An LDP Hello packet received on any of the other

destination addresses must be dropped. Additionally, the link-local IPv6 address MUST be used as the source IP address in IPv6 LDP Link Hellos.

Also, the LDP Link Hello packets must have their IPv6 Hop Limit set to 255, and be checked for the same upon receipt before any further processing, as specified in Generalized TTL Security Mechanism (GTSM)[RFC5082]. The built-in inclusion of GTSM automatically protects IPv6 LDP from off-link attacks.

More importantly, if an interface is a dual-stack LDP interface (e.g. enabled with both IPv6 and IPv4 LDP), then the LSR must periodically send both IPv6 and IPv4 LDP Link Hellos (using the same LDP Identifier per section 4) on that interface and be able to receive them. This facilitates discovery of IPv6-only, IPv4-only and dual-stack peers on the interface's subnet.

> An implementation should prefer sending IPv6 LDP link Hellos before sending IPv4 LDP Link Hellos on a dual-stack interface, if possible.

Lastly, the IPv6 and IPv4 LDP Link Hellos must carry the same LDP identifier (assuming per-platform label space usage).

5.2. Extended Discovery Mechanism

Suffice to say, the extended discovery mechanism (defined in section 2.4.2 of [RFC5036]) doesn't require any additional IPv6 specific consideration, since the targeted LDP Hellos are sent to a pre-configured (unicast) destination IPv6 address.

The link-local IP addresses MUST NOT be used as the source or destination IPv6 addresses in extended discovery.

6. LDP Session Establishment and Maintenance

Section 2.5.1 of [RFC5036] defines a two-step process for LDP session establishment, once the peer discovery has completed (LDP Hellos have been exchanged):

1. Transport connection establishment
2. Session initialization

The forthcoming sub-sections discuss the LDP consideration for IPv6 and/or dual-stacking in the context of session establishment and maintenance.

6.1. Transport connection establishment

Section 2.5.2 of [RFC5036] specifies the use of an optional transport address object (TLV) in LDP Link Hello message to convey the transport (IP) address, however, it does not specify the behavior of LDP if both IPv4 and IPv6 transport address objects (TLV) are sent in a Hello message or separate Hello messages. More importantly, it does not specify whether both IPv4 and IPv6 transport connections should be allowed, if there were Hello adjacencies for both IPv4 and IPv6 whether over a single interface or multiple interfaces.

This document specifies that:

1. An LSR MUST NOT send a Hello containing both IPv4 and IPv6 transport address optional objects. In other words, there MUST be at most one optional Transport Address object in a Hello message. An LSR MUST include only the transport address whose address family is the same as that of the IP packet carrying Hello.

2. An LSR SHOULD accept the Hello message that contains both IPv4 and IPv6 transport address optional objects, but MUST use only the transport address whose address family is the same as that of the IP packet carrying Hello. An LSR SHOULD accept only the first transport object for a given Address family in the received Hello message, and ignore the rest, if the LSR receives more than one transport object.

3. An LSR MUST send separate Hellos (each containing either IPv4 or IPv6 transport address optional object) for each IP address family, if LDP was enabled for both IP address families.

4. An LSR MUST use a global unicast IPv6 address in IPv6 transport address optional object of outgoing targeted hellos, and check for the same in incoming targeted hellos (i.e. MUST discard the hello, if it failed the check).

5. An LSR MUST prefer using global unicast IPv6 address for an LDP session with a remote LSR, if it had to choose between global unicast IPv6 address and unique-local or link-local IPv6

address (pertaining to the same LDP Identifier) for the transport connection.

6. An LSR SHOULD NOT create (or honor the request for creating) a TCP connection for a new LDP session with a remote LSR, if they already have an LDP session (for the same LDP Identifier) established over whatever IP version transport.

   This means that only one transport connection is established, regardless of one or two Hello adjacencies (one for IPv4 and another for IPv6) are created & maintained over a single interface (scenario 1 in section 1.1) or multiple interfaces (scenario 2 in section 1.1) between two LSRs.

7. An LSR SHOULD prefer the LDP/TCP connection over IPv6 for a new LDP session with a remote LSR, if it has both IPv4 and IPv6 hello adjacencies for the same (peer) LDP Identifier (over a dual-stack interface, or two or more single-stack IPv4 and IPv6 interfaces). This applies to the section 2.5.2 of RFC5036.

   Each LSR, assuming an active role for whichever address family(s), should enforce this LDP/TCP connection over IPv6 preference for a time-period (default value is 15 seconds), after which LDP/TCP connection over IPv4 should be attempted. This enforcement is independent of whether the LSR is assuming the active role for IPv4. This timer is started upon receiving the first hello from the neighbor.

8. An LSR SHOULD prefer the LDP/TCP connection over IPv6 for a new LDP session with a remote LSR, if they attempted two TCP connections using different transport address families (IPv4 and IPv6) simultaneously.

An implementation may provide an option to favor one AFI (IPv4, say) over another AFI (IPv6, say) for the TCP transport connection, so as to use the favored IP version for the LDP session, and force deterministic active/passive roles.

6.2. Maintaining Hello Adjacencies

In line with the section 2.5.5 of RFC5036, this draft describes that if an LSR has a dual-stack interface, which is enabled with both

IPv6 and IPv4 LDP, then the LSR must periodically send and receive both IPv6 and IPv4 LDP Link Hellos.

This ensures successful LDP discovery and subsequent peering using the appropriate (address family) transport on a multi-access or broadcast interface (even if there are IPv6-only, IPv4-only and dual-stack LSRs connected to that interface).

This document allows an LSR to maintain Rx-side Link Hello adjacency for only one address family that has been used for the establishment of the LDP session.

6.3. Maintaining LDP Sessions

Two LSRs maintain a single LDP session between them (i.e. not tear down an existing session), as described in section 6.1, whether

- they are connected via a dual-stack LDP enabled interface or via two single-stack LDP enabled interfaces;
- a single-stack interface is converted to a dual-stack interface (e.g. figure 1) on either LSR;
- an additional single-stack or dual-stack interface is added or removed between two LSRs (e.g. figure 2).

Needless to say that the procedures defined in section 6.1 should result in preferring LDPoIPv6 session only after the loss of an existing LDP session (because of link failure, node failure, reboot etc.).

If the last hello adjacency for a given address family goes down (e.g. due to dual-stack interfaces being converted into a single-stack interfaces on one LSR etc.), and that address family is the same as the one used in the transport connection, then the transport connection (LDP session) SHOULD be reset. Otherwise, the LDP session should stay intact.

If the LDP session is torn down for whatever reason (LDP disabled for the corresponding transport, hello adjacency expiry etc.), then the LSRs should initiate establishing a new LDP session as per the procedures described in section 6.1 of this document along with RFC5036.

7. Label Distribution

An LSR MUST NOT allocate and advertise FEC-Label bindings for link-local IPv6 address, and ignore such bindings, if ever received. An LSR MUST treat the IPv4-mapped IPv6 address, defined in section 2.5.5.2 of [RFC4291], the same as that of a global IPv6 address and not mix it with the 'corresponding' IPv4 address.

Additionally, to ensure backward compatibility (and interoperability with IPv4-only LDP implementations) in light of section 3.4.1.1 of RFC5036, as rationalized in the Appendix A.1, this document specifies that -

1. An LSR MUST NOT send a label mapping message with a FEC TLV containing FEC Elements of different address family. In other words, a FEC TLV in the label mapping message MUST contain the FEC Elements belonging to the same address family.
2. An LSR MUST NOT send an Address message (or Address Withdraw message) with an Address List TLV containing IP addresses of different address family. In other words, an Address List TLV in the Address (or Address Withdraw) message MUST contain the addresses belonging to the same address family.

An LSR MAY constrain the advertisement of FEC-label bindings for a particular address family by negotiating the IP Capability for a given AFI, as specified in [IPPWCap] document.

8. LDP Identifiers and Next Hop Addresses

RFC5036 section 2.7 specifies the logic for mapping the IP routing next-hop (of a given FEC) to an LDP peer so as to find the correct label entry for that FEC. The logic involves using the IP routing next-hop address as an index into the (peer Address) database (which is populated by the Address message containing mapping between each peer's local addresses and its LDP Identifier) to determine the LDP peer.

However, this logic is insufficient to deal with duplicate IPv6 (link-local) next-hop addresses used by two or more peers. The reason is that all interior IPv6 routing protocols (can) use link-local IPv6 addresses as the IP routing next-hops, and 'IPv6 Addressing Architecture [RFC4291]' allows a link-local IPv6 address to be used on more than one links.

Hence, this logic is extended by this specification to involve not only the IP routing next-hop address, but also the IP routing next-hop interface to uniquely determine the LDP peer(s). The next-hop address-based LDP peer mapping is to be done through LDP peer address database (populated by Address messages received from the LDP peers), whereas next-hop interface-based LDP peer mapping is to be done through LDP hello adjacency/interface database (populated by hello messages from the LDP peers).

This extension solves the problem of two or more peers using the same link-local IPv6 address (in other words, duplicate addresses) as the IP routing next-hops.

Lastly, for better scale and optimization, an LSR may advertise only the link-local IPv6 addresses in the Address message, assuming that the peer uses only the link-local IPv6 addresses as static and/or dynamic IP routing next-hops.

9. LDP TTL Security

This document recommends enabling Generalized TTL Security Mechanism (GTSM) for LDP, as specified in [RFC6720], for the LDP/TCP transport connection over IPv6 (i.e. LDPoIPv6).

[RFC6720] allows for the implementation to statically (configuration) and/or dynamically override the default behavior (enable/disable GTSM) on a per-peer basis. Suffice to say that such an option could be set on either LSR (since GTSM negotiation would ultimately disable GTSM between LSR and its peer(s)).

The GTSM inclusion is intended to automatically protect IPv6 LDP peering session from off-link attacks.

10. IANA Considerations

None.

11. Security Considerations

The extensions defined in this document only clarify the behavior of LDP, they do not define any new protocol procedures. Hence, this document does not add any new security issues to LDP.

While the security issues relevant for the [RFC5036] are relevant for this document as well, this document reduces the chances of off-link attacks when using IPv6 transport connection by including the use of GTSM procedures [RFC5082].

Moreover, this document allows the use of IPsec [RFC4301] for IPv6 protection, hence, LDP can benefit from the additional security as specified in [RFC4835] as well as [RFC5920].

12. Acknowledgments

We acknowledge the authors of [RFC5036], since the text in this document is borrowed from [RFC5036].

Thanks to Bob Thomas for providing critical feedback to improve this document early on. Thanks to Eric Rosen, Lizhong Jin, Bin Mo, Mach Chen, Shane Amante, Pranjal Dutta, Mustapha Aissaoui, Mark Tinka, Tom Petch and Kishore Tiruveedhula for reviewing this document. The authors also acknowledge the help of Manoj Dutta and Vividh Siddha.

Also, thanks to Andre Pelletier (who brought up the issue about active/passive determination, and helped us craft the appropriate solutions).

This document was prepared using 2-Word-v2.0.template.dot.

13. Additional Contributors

The following individuals contributed to this document:

Kamran Raza
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, ON K2K-3E8, Canada
Email: skraza@cisco.com

Nagendra Kumar
Cisco Systems, Inc.
SEZ Unit, Cessna Business Park,
Bangalore, KT, India
Email: naikumar@cisco.com

Andre Pelletier
Email: apelleti@cisco.com

14. References

14.1. Normative References

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

[RFC4291] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 4291, February 2006.

[RFC5036] Andersson, L., Minei, I., and Thomas, B., "LDP Specification", RFC 5036, October 2007.

[RFC5082] Pignataro, C., Gill, V., Heasley, J., Meyer, D., and Savola, P., "The Generalized TTL Security Mechanism (GTSM)", RFC 5082, October 2007.

14.2. Informative References

[RFC4301] Kent, S. and K. Seo, "Security Architecture and Internet Protocol", RFC 4301, December 2005.

[RFC4835] Manral, V., "Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)", RFC 4835, April 2007.

[RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

[RFC4798] De Clercq, et al., "Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)", RFC 4798, February 2007.

[IPPWCap] Raza, K., "LDP IP and PW Capability", draft-ietf-mpls-ldp-ip-pw-capability, June 2011.

15. Appendix

15.1. A.1

It is naive to assume that RFC5036 compliant implementations have supported IPv6 address family (IPv6 FEC processing, in particular) in label advertisement all along. And if that assumption turned out to be not true, then section 3.4.1.1 of RFC5036 would cause LSRs to abort processing the entire label mapping message and generate an error.

This would result in LDPv6 to be somewhat undeployable in existing production networks.

The change proposed in section 7 of this document provides a good safety net and makes LDPv6 incrementally deployable without making any such assumption on the routers' support for IPv6 FEC processing in current production networks.

Author's Addresses

Vishwas Manral
Hewlet-Packard, Inc.
19111 Pruneridge Ave., Cupertino, CA, 95014
Phone: 408-447-1497
Email: vishwas.manral@hp.com

Rajiv Papneja
Huawei Technologies
2330 Central Expressway
Santa Clara, CA 95050
Phone: +1 571 926 8593
EMail: rajiv.papneja@huawei.com

Rajiv Asati
Cisco Systems, Inc.
7025 Kit Creek Road
Research Triangle Park, NC 27709-4987
Email: rajiva@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7200 Kit Creek Road
Research Triangle Park, NC 27709-4987
Email: cpignata@cisco.com