IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: October 05, 2013

H. Gredler, Ed.
Juniper Networks, Inc.
S. Amante
Level 3 Communications, Inc.
T. Scholl
Amazon
L. Jalil
Verizon

April 5, 2013

# Advertising MPLS labels in IS-IS
## draft-gredler-isis-label-advertisement-00

Abstract

Historically MPLS label distribution was driven by protocols like LDP, RSVP and LBGP. All of those protocols are session oriented.  In order to obtain label binding for a given destination FEC from a given router one needs first to establish an LDP/RSVP/LBGP session with that router.

Advertising MPLS labels in IGPs advertisement [I-D.gredler-rtgwg-igp-label-advertisement] describes several use cases where utilizing the flooding machinery of link-state protocols for MPLS label distribution allows to obtain the binding without requiring to establish an LDP/RSVP/LBGP session with that router.

This document describes the protocol extension to distribute MPLS labels by the IS-IS protocol.

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].

Status of this Memo

Copyright Notice



Table of Contents

## 1. Introduction

MPLS label allocations are predominantly distributed by using the LDP [RFC5036], RSVP [RFC5151] or labeled BGP [RFC3107] protocol. All of those protocols have in common that they are session oriented, which means that in order to obtain label binding for a given destination FEC from a given router one needs first to establish a direct control plane (LDP/RSVP/LBGP) session with that router.

There are a couple of use cases [I-D.gredler-rtgwg-igp-label-advertisement] where the consumer of a MPLS label binding may not be adjacent to the router that performs the binding. Bringing up an explicit session using the existing label distribution protocols between the non-adjacent router that bind the label and the router that acts as a consumer of this binding is the existing remedy for this dilemma.

This document describes a single IS-IS protocol extension which allows routers to advertise MPLS label bindings within and beyond an IGP domain, and controlling inter-area distribution.

## 2. Motivation, Rationale and Applicability

One possible way of distributing MPLS labels using IS-IS has been described in Segment Routing [I-D.previdi-filsfils-isis-segment-routing]. The authors propose to re-use the IS-Reach TLVs (22, 23, 222) and Extended IP Prefix TLVs (135, 236) for carrying the label information. While retrofitting existing protocol machinery for new purposes is generally a good thing, Segment Routing [I-D.previdi-filsfils-isis-segment-routing] falls short of addressing some use-cases defined in [I-D.gredler-rtgwg-igp-label-advertisement].

The dominant issue around re-using IS-Reach TLVs and the extended IP Prefix TLVs is that both family of TLVs have existing protocol semantics, which might not be well suitable to advertising MPLS label switched paths in a generic fashion. These are specifically:

- Bi-directionality semantics

- IP path semantics

- Lack of 'path' notion

### 2.1. Issue: Bi-directionality semantics

'Bi-directionality semantics', affects the complexity around advertisement of unidirectional LSPs. Label advertisement of per-link labels or 'Adj-SIDs' [I-D.previdi-filsfils-isis-segment-routing] is done using IS-reach TLVs. Usually implementations need to have an adjacency in 'Up' state prior to advertising this adjacency as IS-reach TLV in its Link State PDUs (LSPs). In order to advertise a per-link LSP an implementation first needs to have an adjacency, which only transitions to 'Up' state after passing the 3-way check. This implies bi-directionality. If an implementation wants to advertise per-link LSPs to e.g. outside the IGP domain then it would need to fake-up an adjacency. Changing existing IGP Adjacency code to support such cases defeats the purpose of re-using existing functionality as there is not much common functionality to be shared.

### 2.2. Issue: IP path semantics

LSPs pointing to a Node are advertised as 'Node-SIDs' [I-D.previdi-filsfils-isis-segment-routing] using the family of extended IP Reach TLVs. That means that in order to advertise a LSP, one is inheriting the semantics of advertising an IP path. Consider router A has got existing LSPs to its entire one-hop neighborhood and is re-advertising those LSPs using IP reachability semantics. Now we have two exact matching IP advertisements. One from the owning router (router B) which advertises its stable transport loopback address and another one from router A re-advertising a LSP path to router B. Existing routing software may get confused now as the 'stable transport' address shows up from multiple places in the network and more worse the IP forwarding path for control-plane protocols may get mingled with the MPLS data plane.

### 2.3. Issue: Lack of 'path' notion

Both IS-Reach TLVs and IP Prefix Reachability TLVs have a limited semantics describing MPLS label-switched paths in the sense of a 'path'. Both encoding formats allow to specify a pointer to some specific router, but not to describe a MPLS label switched path containing all of its path segments. [I-D.previdi-filsfils-isis-segment-routing] allows to define 'Forwarding Adjacencies' as per [RFC4206]. The way to describe a path of a given forwarding adjacency is to carry a list of "Segment IDs". That implies that nodes which do not yet participate in 'Segment routing' or are outside of a 'Segment routing' domain can not be expressed using those path semantics.

A protocol for advertising MPLS label switched paths, should be generic enough to express paths sourced by existing MPLS LSPs, such that ingress routers can flexibly combine them according to application needs.

### 2.4. Motivation

IGP advertisement of MPLS label switched paths requires a new set of protocol semantics (path paradigm), which hardly can be expressed using the existing IS-IS protocol. This document describes IS-IS protocol extensions which allows generic advertisement of MPLS label switched paths in IS-IS.

## 3. MPLS label TLV

The MPLS Label TLV may be originated by any Traffic Engineering [RFC5305] capable router in an IS-IS domain. A router may advertise MPLS labels along with so called 'ERO' path segments describing the label switched path. Since ERO style path notation allows to express pointers to link and node IP addresses any label switched path, sourced by any protocol can be described.

Due to the limited size of subTLV space (See [RFC5311] section 4.5 for details), The MPLS Label TLV has cumulative rather than canceling semantics. If a router originates more than one MPLS Label TLV with the same Label value, then the subTLVs of the second, third, etc. TLV are accumulated. Since some subTLVs represent an ordered set (e.g. ERO subTLVs) and in order to not complicate the cross-fragment tracking logic, all subTLVs for a given label value MUST be present in a single IS-IS fragment.

The MPLS Label TLV has type 149 and has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type      |     Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|U|R|R|R|            MPLS Label                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o  4 bits of flags, consisting of:

   *  1 bit of up/down information (U bit)

   *  3 bits are reserved for future use

o  20 bits of MPLS label information

o  0-252 octets of sub-TLVs, where each sub-TLV consists of a sequence of:

   *  1 octet of sub-TLV type

   *  1 octet of length of the value field of the sub-TLV

*  0-250 octets of value

### 3.1. Flags

Flags

Up/Down Bit: A router may flood MPLS label information across level boundaries.  In order to prevent flooding loops, a router will Set the Up/Down (U-Bit) when propagating from Level 2 down to Level 1. This is done as per the procedures for IP Prefixes lined out in [RFC5302].

## 3.2.  subTLV support

An originating router MAY want to attach one or more subTLVs to the MPLS label TLV.  SubTLVs presence is inferred from the length of the MPLS Label TLV.  If the MPLS Label TLV Length field is > 3 octets then one or more subTLVs may be present.

## 3.3.  IPv4 Prefix ERO subTLV

The IPv4 ERO subTLV (Type 1) describes a path segment using IPv4 Prefix style of encoding.  Its appearance and semantics have been borrowed from Section 4.3.3.2 [RFC3209].

The 'Prefix Length' field contains the length of the prefix in bits. Only the most significant octets of the prefix are encoded.  I.e.  1 octet for prefix length 1 up to 8, 2 octets for prefix length 9 to 16, 3 octets for prefix length 17 up to 24 and 4 octets for prefix length 25 up to 32, etc.

The 'L' bit in the subTLV is a one-bit attribute.  If the L bit is set, then the value of the attribute is 'loose.'  Otherwise, the value of the attribute is 'strict.'

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    | Prefix Length |  IPv4 Prefix  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv4 Prefix  (continued, variable-length      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

## 3.4.  IPv6 Prefix ERO subTLV

The IPv6 ERO subTLV (Type 2) describes a path segment using IPv6 Prefix style of encoding.  Its appearance and semantics have been borrowed from Section 4.3.3.3 [RFC3209].

The 'Prefix Length' field contains the length of the prefix in bits. Only the most significant octets of the prefix are encoded.  I.e.  1 octet for prefix length 1 up to 8, 2 octets for prefix length 9 to

16, 3 octets for prefix length 17 up to 24 and 4 octets for prefix length 25 up to 32, ...., 16 octets for prefix length 113 up to 128.

The 'L' bit in the subTLV is a one-bit attribute.  If the L bit is set, then the value of the attribute is 'loose.'  Otherwise, the value of the attribute is 'strict.'

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |L|    Type     |     Length    | Prefix Length | IPv6 Prefix   |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | IPv6 Prefix (continued)                                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | IPv6 Prefix (continued)                                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | IPv6 Prefix (continued)                                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | IPv6 Prefix (continued, variable length)      |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

### 3.5. Prefix ERO subTLV path semantics

All 'Prefix ERO' information represents an ordered set which describes the segments of a label-switched path.  The last Prefix ERO subTLV describes the segment closest to the egress point of the LSP. Contrary the first Prefix ERO subTLV describes the first segment of a label switched path.  If a router extends or stitches a label switched path it MUST prepend the new segments path information to the Prefix ERO list.

## 4. Advertising Label Examples

### 4.1. Sample Topology

The following topology (Figure 4) and IP addresses shall be used throughout the Label advertisement examples.

```
                       AS1                         :    AS 2
                                                   :
                        :                          :
            Level 1     :     Level 2              :
                        :                          :
   +-----+           +-----+-IP3--IP4--+-----+     :
   | R1  +-IP1--IP2--+ R2  |           | R3  |     :
   +--+--+           +--+--+-IP5--IP6--+--+--+-IP15-IP16-
      |                 |                 |        :     \
     IP3               IP7               IP13      :   +--+--+
      |                 |                 |        :   |  R7 |
     IP4               IP8               IP14      :   +--+--+
      |                 |                 |        :     /
   +--+--+           +--+--+           +--+--+-IP17-IP18-
   |  R4 +-IP19-IP20-+ R5  |-IP11-IP12-| R6  |     :
   +-----+           +-----+           +-----+     :
                        :                          :
                        :                          :
                        :                          :
```

### 4.1.1. Transport IP addresses and router-IDs

- o R1: 192.168.1.1
- o R2: 192.168.1.2
- o R3: 192.168.1.3
- o R4: 192.168.1.4
- o R5: 192.168.1.5
- o R6: 192.168.1.6
- o R7: 192.168.1.7

### 4.1.2. Link IP addresses

- o R1 to R2 link: 10.0.0.1, 10.0.0.2
- o R1 to R4 link: 10.0.0.3, 10.0.0.4
- o R2 to R3 link #1: 10.0.0.3, 10.0.0.4
- o R2 to R3 link #2: 10.0.0.5, 10.0.0.6
- o R2 to R5 link: 10.0.0.7, 10.0.0.8

o  R3 to R6 link: 10.0.0.13, 10.0.0.14

o  R3 to R7 link: 10.0.0.15, 10.0.0.16

o  R4 to R5 link: 10.0.0.19, 10.0.0.20

o  R5 to R6 link: 10.0.0.11, 10.0.0.12

o  R6 to R7 link: 10.0.0.17, 10.0.0.18

### 4.2.   One-hop LSP to an adjacent Router

If R1 would advertise a label <N> bound to a one-hop LSP from R1 to R2 it would encode as follows:

TLV 149: MPLS label N, Flags {}:

IPv4 Prefix ERO subTLV: 192.168.1.2/32, Strict

### 4.3.   One-hop LSP to an adjacent Router using a specific link

If R2 would advertise a label <N>bound to a one-hop LSP from R2 to R3, using the link #2 it would encode as follows

TLV 149: MPLS label <N>, Flags {}:

IPv4 Prefix ERO subTLV: 10.0.0.6/32, Strict

### 4.4.   Advertisement of an RSVP LSP

Consider a RSVP LSP name "R2-to-R6" traversing (R2 to R3 using link #1, R6):

If R2 would advertise a label <N> bound to the RSVP LSP named 'R2-to-R6', it would encode as follows

TLV 149: MPLS label <N>, Flags {}:

IPv4 Prefix ERO subTLV: 10.0.0.4/32, Strict

IPv4 Prefix ERO subTLV: 192.168.1.6/32, Strict

### 4.5.   Advertisement of an LDP LSP

Consider R2 that creates a LDP label binding for FEC 172.16.0.0./12 using label <N>.

If R2 would re-advertise this binding in IS-IS it would encode as follows

TLV 149: MPLS label <N>, Flags {}:

IPv4 Prefix ERO subTLV: 172.16.0.0/12, Loose

### 4.6. Interarea advertisement of diverse paths

Consider two R2->R6 paths: {R2, R3, R6} and {R2, R5, R6}

Consider two R5->R3 paths: {R5, R2, R3} and {R5, R6, R3}

R2 encodes its two paths to R6 as follows:

TLV 149: MPLS label <N1>, Flags {}:

IPv4 Prefix ERO subTLV: 192.168.1.3, Strict

IPv4 Prefix ERO subTLV: 192.168.1.6, Strict

TLV 149: MPLS label <N2>, Flags {}:

IPv4 Prefix ERO subTLV: 192.168.1.5, Strict

IPv4 Prefix ERO subTLV: 192.168.1.6, Strict

R5 encodes its two paths to R3 as follows:

TLV 149: MPLS label <N1>, Flags {}:

IPv4 Prefix ERO subTLV: 192.168.1.2, Strict

IPv4 Prefix ERO subTLV: 192.168.1.3, Strict

TLV 149: MPLS label <N2>, Flags {}:

IPv4 Prefix ERO subTLV: 192.168.1.6, Strict

IPv4 Prefix ERO subTLV: 192.168.1.3, Strict

A receiving L1 router does see now all 4 paths and may decide to load-balance across all or a subset of them.

# 5. Inter Area Protocol Procedures

## 5.1. Applicability

Propagation of a MPLS LSP across a level boundary is a local policy decision.

## 5.2. Data plane operations

If local policy dictates that a given L1L2 router needs to re-advertise a MPLS LSPs from one Level to another then it MUST allocate a new label and program its label forwarding table to connect the new label to the path in the respective other level.  Depending on how to reach the re-advertised LSP, this is typically done using a MPLS 'SWAP' or 'SWAP/PUSH' data plane operation.

## 5.3. Control plane operations

If local policy dictates that a given L1L2 router needs to re-advertise a MPLS LSPs from one Level to another then it must prepend its "Traffic-Engineering-ID" as a loose hop in the Prefix ERO subTLV list.

## 6. Acknowledgements

Many thanks to Yakov Rekhter for his useful comments.

## 7. IANA Considerations

This documents request allocation for the following TLVs and subTLVs.

| PDU | TLV | subTLV | Type | subType | #Occurence |
|---|---|---|---|---|---|
| LSP | MPLS Label | | 149 | | >=0 |
| | | IPv4 Prefix ERO | | 1 | >=0 |
| | | IPv6 Prefix ERO | | 2 | >=0 |

The MPLS Label TLV requires a new sub-registry. Type value 149 has been assigned, with a starting sub-TLV value of 1, range from 1-127, and managed by Expert Review.

## 8. Security Considerations

This document does not introduce any change in terms of IS-IS security. It simply proposes to flood MPLS label information via the IGP. All existing procedures to ensure message integrity do apply here.

## 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.

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, May 2001.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V. and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

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

[RFC5151] Farrel, A., Ayyangar, A. and JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering -- Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, February 2008.

[RFC5302] Li, T., Smit, H. and T. Przygienda, "Domain-Wide Prefix Distribution with Two-Level IS-IS", RFC 5302, October 2008.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

[RFC5311] McPherson, D., Ginsberg, L., Previdi, S. and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, February 2009.

### 9.2. Informative References

[I-D.gredler-rtgwg-igp-label-advertisement]
Gredler, H., Amante, S., Scholl, T. and L. Jalil, "Advertising MPLS labels in IGPs", Internet-Draft draft-gredler-rtgwg-igp-label-advertisement-03, April 2013.

[I-D.previdi-filsfils-isis-segment-routing]
Previdi, S., Filsfils, C., Bashandy, A., Horneffer, M., Decraene, B., Litkowski, S., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W. and J. Tantsura, "Segment Routing with IS-IS Routing Protocol", Internet-Draft draft-previdi-filsfils-isis-segment-routing-02, March 2013.

Authors' Addresses

Hannes Gredler, editor
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Shane Amante
Level 3 Communications, Inc.
1025 Eldorado Blvd
Broomfield, CO 80021
US

Email: shane@level3.net

Tom Scholl
Amazon
Seattle, WN
US

Email: tscholl@amazon.com


Luay Jalil
Verizon
1201 E Arapaho Rd.
Richardson, TX 75081
US

Email: luay.jalil@verizon.com