IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: November 22, 2013

H. Gredler, Ed.
Juniper Networks, Inc.
S. Amante
Level 3 Communications, Inc.
T. Scholl
Amazon
L. Jalil
Verizon
May 21, 2013

Advertising MPLS labels in IS-IS
draft-gredler-isis-label-advertisement-03

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 a 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 [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 label bindings using 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].

Table of Contents

1. Introduction 4
2. Motivation, Rationale and Applicability 4
2.1. Issue: Bi-directionality semantics 5
2.2. Issue: IP path semantics 5
2.3. Issue: Lack of 'path' notion 5
2.4. Motivation 6
3. MPLS label TLV 6
3.1. Flags 7
3.2. subTLV support 7
3.3. IPv4 Prefix ERO subTLV 7
3.4. IPv6 Prefix ERO subTLV 8
3.5. Unnumbered Interface ID ERO subTLV 9
3.6. IPv4 Prefix Bypass ERO subTLV 10
3.7. IPv6 Prefix Bypass ERO subTLV 10
3.8. Unnumbered Interface ID Bypass ERO subTLV 11
3.9. Prefix ERO and Prefix Bypass ERO subTLV path semantics . . 12
3.10. All Router Block subTLV 12
3.11. All Router ID IPv4 Map subTLV 14
3.12. All Router ID IPv6 Map subTLV 15
4. Advertising Label Examples 15
4.1. Sample Topology 15
4.1.1. Transport IP addresses and Router-IDs 16
4.1.2. Link IP addresses 16
4.2. One-hop LSP to an adjacent Router 17
4.3. One-hop LSP to an adjacent Router using a specific link . 17
4.4. Advertisement of Fast Re-Route LSP for One-Hop LSP 17
4.5. Advertisement of an RSVP LSP 18
4.6. Advertisement of an LDP LSP 18
4.7. Interarea advertisement of diverse paths 18
4.8. Advertisement of SPT labels using 'All Router Block' TLV 19
4.9. Expansion of an 'All Router Block' subTLV 20
5. Inter Area Protocol Procedures 21
5.1. Applicability 21
5.2. Data plane operations 21
5.3. Control plane operations 21
5.3.1. MPLS Label operations 21
5.3.2. MPLS Label Block operations 22
6. Acknowledgements 22
7. IANA Considerations 22
8. Security Considerations 23
9. References 23
9.1. Normative References 23
9.2. Informative References 24
Authors' Addresses 24

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 practical 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 binds the label and the router that acts as a consumer of this binding is the existing remedy for this dilemma.

This document describes an 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:

- o Bi-directionality semantics
- o IP path semantics
- o 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 e.g. one-hop MPLS LSP in a given link 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 MPLS LSP, one is inheriting the semantics of advertising an IP path. Consider router A has got existing MPLS LSPs to its entire one-hop neighborhood and is re-advertising those MPLS 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 MPLS 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 bindings in IS-IS.

The Protocol extensions described in this document are equally applicable to IPv4 and IPv6 carried over MPLS. Furthermore the proposed use of distributing MPLS Labels using IGP prototocols adheres to the architectural principles laid out in [RFC3031].

3. MPLS label TLV

The MPLS Label TLV may be originated by any Traffic Engineering [RFC5305] capable router in an IS-IS domain. The router may advertise a single label binding or a block of label bindings. For single label binding advertisement a router needs to provide at least a single 'nexthop style' anchor. The protocol supports more than one 'nexthop style' anchor to be attached to a Label binding, which results into a simple path description language. In analogy to RSVP the terminology for this is called an 'Explicit Route Object' (ERO). Since ERO style path notation allows to anchor label bindings to to both link and node IP addresses any label switched path, can be described. Furthermore also Label Bindings from other protocols can get easily re-advertised.

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) allocation and ordering of TLV space inside particular IS-IS LSP fragment is significant and needs to be tracked.

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                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: MPLS TLV format

- 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      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: IPv4 Prefix ERO subTLV format

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)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: IPv6 Prefix ERO subTLV format

3.5. Unnumbered Interface ID ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID ERO subTLV (Type 9) describes a path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

- o (TLV 134, 32-Bit format) [RFC5305]

- o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

- o For 32-Bit Router-ID width the subTLV length is set to 8 octets.

- o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

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    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                   Router ID (32 or 128 bits)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Interface ID (32 bits)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Unnumbered Interface ID ERO subTLV format

3.6. IPv4 Prefix Bypass ERO subTLV

The IPv4 Bypass ERO subTLV (Type 3) describes a Bypass LSP 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)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: IPv4 Prefix Bypass ERO subTLV format

3.7. IPv6 Prefix Bypass ERO subTLV

The IPv6 ERO subTLV (Type 4) describes a Bypass LSP 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)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: IPv6 Prefix Bypass ERO subTLV format

3.8. Unnumbered Interface ID Bypass ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID Bypass ERO subTLV (Type 10) describes a Bypass LSP path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

o (TLV 134, 32-Bit format) [RFC5305]

o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

o For 32-Bit Router-ID width the subTLV length is set to 8 octets.

o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

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    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                   Router ID (32 or 128 bits)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Interface ID (32 bits)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: Unnumbered Interface ID Bypass ERO subTLV format

3.9. Prefix ERO and Prefix Bypass ERO subTLV path semantics

All 'Prefix ERO' and 'Prefix Bypass 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. The same ordering applies for the Bypass ERO labels. An implementation SHOULD first encode all primary path EROs followed by the bypass EROs.

3.10. All Router Block subTLV

The 'All Router Block' subTLV (Type 6) denominates the label block size of an MPLS Label advertisement and its semantics to connect to all routers in a given IS-IS domain using a local assigned [RFC3031] label range. Note that the actual mapping of a router within the label range is done using the subTLVs described in Section 3.11 and

Section 3.12. Since generation of an 'All Router ID IPv4 Map' or 'All Router ID IPv6 Map' subTLV is a local policy decision, it might be the case that connectivity is provided not to 'All' but rather a subset of 'All' routers. Keeping policy decisions aside, for simplicity reasons, assume that All Routers in a domain do generate either the 'All Router ID IPv4 Map' or 'All Router ID IPv6 Map' subTLVs and therefore all routers desire construction of a Label switched path from every source router in the network. The basic concept of using label blocks to provide connectivity to a set of routers has been borrowed from [RFC4761] which allows to advertise labels from multiple end-points using a single control-plane message. The difference to [RFC4761] is that rather than advertising where a particular packet came from (=source semantics), destination semantics (where a particular packet will be going to) is advertised.

Along with each label block a router advertises one for more 'IDs'. The 'ID' must be unique within a given domain. The 'ID' serves as ordinal to determine the actual label value inside the set of all advertised label ranges of a given router. A receiving router uses the ordinal to determine the actual label value in order to construct forwarding state to a particular destination router. The 'ID' is separately advertised using the subTLVs described in Section 3.11 and Section 3.12.

The ability to advertise more than one label block eases operational procedures for increasing the number of supported routers within a domain. For example consider a given domain has got support for <M> routers and runs out of ID space. It simply advertises one more label block to cover additional ordinals outside the range of the first label block. An example of label-block expansion is described in more detail in Section 4.9

```
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             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Block Size            | Algo  |      Topology-ID      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: All Router Block subTLV format

The 'Block Size' value contains the size of the label advertisement. The 'value determines the amount of reachable router endpoints within a given Label block. It MUST contain a value greater or equal than two. Note that the label base is inferred from the Label Value in the carrying MPLS Label TLV. For example if a router wants to advertise a label range of 5000-5099 then it would need to generate a

MPLS Label TLV with a Label value of 5000 and a Block Size of 100.

The 'Algo' value denominates the path computation algorithm in order to calculate the forwarding topology. The basic SPF algorithm has an assigned 'Algo' code point of zero. The purpose of the 'Algo' field is to extend the notion of Label Block Signaling to arbitrary algorithms like for example 'MRT' ([I-D.ietf-rtgwg-mrt-frr-architecture]. Advertised Label Blocks with an unknown, unsupported or non-configured algorithm MUST be silently ignored.

The 'Reserved' bits are for future use. They should be zero on transmission and ignored on receipt.

The 'Topology-ID' field contains the Multi Topology ID ([RFC5120]) for which the advertised Label Block does apply. The basic IPv4 unicast Topology has an assigned 'Topology-ID' code point of zero. The basic IPv6 unicast Topology has an assigned 'Topology=ID' code point of 2. Advertised Label Blocks with an unknown, unsupported or non-configured Topology-ID MUST be silently ignored.

A MPLS Label TLV containing the 'All Router Block' subTLV MUST only contain the 'All Router IPv4 Map' subTLV (Section 3.11) or the 'All Router IPv6 Map' subTLV (Section 3.12).

3.11. All Router ID IPv4 Map subTLV

The 'All Router ID IPv4 Map' TLV (Type 7) maps an 'ID' to a given stable transport IPv4 address. Its purpose is to associate a given transport IPv4 IP address to the ordinal inside a label range as described in Section 3.10.

A router MAY advertise more than one 'ID' to 'IPv4 address' mapping pair, in case it has more than one stable transport IPv4 address.

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     IPv4 Address (4 octets)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               ID              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: All Router ID IPv4 Map subTLV format

The 'IPv4 address' contains stable IPv4 transport address of a given

router.

The 'ID' contains the ordinal value of an advertising router inside the set of all advertised label blocks of a given router.

3.12. All Router ID IPv6 Map subTLV

The 'All Router ID IPv6 Map' TLV (Type 8) maps an 'ID' to a given stable transport IPv6 address. Its purpose is to associate a given transport IPv6 IP address to the ordinal inside a label range as described in Section 3.10.

A router MAY advertise more than one 'ID' to 'IPv6 address' mapping pair, in case it has more than one stable transport IPv6 address.

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (16 octets)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IPv6 Address (continued)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               ID              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: All Router ID IPv6 Map subTLV format

The 'IPv6 address' contains the stable IPv6 transport address of a given router.

The 'ID' contains the ordinal value of an advertising router inside the set of all advertised label blocks of a given router.

4. Advertising Label Examples

4.1. Sample Topology

The following topology (Figure 11) and IP addresses shall be used throughout the Label advertisement examples.

```
                         AS1                           :    AS 2
                                                       :
                          :                            :
              Level 1     :     Level 2                :
                          :                            :
                          :                            :
   +-----+             +-----+-IP3--1-IP4--+-----+     :
   | R1  +-IP1--1-IP2--+ R2  |             | R3  |     :
   +--+--+             +--+--+-IP5--3-IP6--+--+--+-IP15-IP16-
      |                   |                   |        :     \
     IP9                 IP7                 IP13      :      \
      |                   |                   |        :    +--+--+
      1                   1                   1        :    |  R7 |
      |                   |                   |        :    +--+--+
    IP10                 IP8                 IP14      :      /
      |                   |                   |        :     /
   +--+--+             +--+--+             +--+--+-IP17-IP18-
   |  R4 +-IP19-2-IP20-+ R5  |-IP11-2-IP12-| R6  |     :
   +-----+             +-----+             +-----+     :
                          :                            :
                          :                            :
                          :                            :
```

                        Figure 11: Sample Topology

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: 10.0.0.9, 10.0.0.10

- 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

The IGP link metrics are displayed in the middle of the link. All of them are assumed to be bi-directional.

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 Fast Re-Route LSP for One-Hop LSP

R2 may advertise a one-hop LSP from R2 to R3, along with a Link Protection Bypass for the directly adjacent links between those two nodes. The Link Protection Bypass would use the path: {R2, R5, R6, R3}. R2 would encode both the primary LSP and Link Protection Bypass LSP as follows:

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

IPv4 Prefix ERO subTLV: 192.168.1.3/32, Strict

IPv4 Prefix Bypass ERO subTLV: 192.168.1.5/32, Strict

IPv4 Prefix Bypass ERO subTLV: 192.168.1.6/32, Strict

IPv4 Prefix Bypass ERO subTLV: 192.168.1.3/32, Strict

4.5. 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

4.6. 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

4.7. 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.

4.8. Advertisement of SPT labels using 'All Router Block' TLV

All routers within a given area MUST advertise their Label Blocks along with an 'ID'.

If R2 would advertise a label block <N1> with a size of 10, declaring SPT label forwarding support to all routers within a given domain, it would encode as follows:

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

All Router Block subTLV: Block Size 10, Algo 0, Topology 0

All Router ID IPv4 Map subTLV: ID 2, 192.168.1.2

If R3 would advertise a label block <N2> with a size of 10, declaring SPT label forwarding support to all routers within a given domain, it would encode as follows:

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

All Router Block subTLV: Block Size 10, Algo 0, Topology 0

All Router ID IPv4 Map subTLV: ID 3, 192.168.1.3

If R5 would advertise a label block <N3> with a size of 10, declaring SPT label forwarding support to all routers within a given domain, it would encode as follows:

```
   TLV 149, MPLS Label <N3>, Flags {}:

      All Router Block subTLV: Block Size 10, Algo 0, Topology 0

      All Router ID IPv4 Map subTLV: ID 5, 192.168.1.5
```

If R6 would advertise a label block <N4> with a size of 10, declaring SPT label forwarding support to all routers within a given domain, it would encode as follows:

```
   TLV 149, MPLS Label <N4>, Flags {}:

      All Router Block subTLV: Block Size 10, Algo 0, Topology 0

      All Router ID IPv4 Map subTLV: ID 6, 192.168.1.6
```

Consider now R2 constructing a SPT label for R6. R2s SPT to R6 is {R2, IP4, R3, R6}. R2 first determines if its downstream router (R3) has advertised a label-block. Since R3 has advertised a label block 'N2' and it has received R6 'ID' of 6 it will be picking the 6th label value inside the advertised range of its downstream neighbor. Specifically R2 MUST be program a MPLS SWAP for its own label range Label(N1+6) to Label(N2+6), NH 10.0.0.4 into its MPLS transit RIB. Furthermore R2 MAY program a MPLS PUSH operation for IP 192.168.1.6 to Label (N2+6), NH 10.0.0.4 into its IPv4 tunnel RIB.

Next walk down to R3, which is the next router on the SPT tree towards R6. R3s SPT to R6 is {R3, R6}. R3 determines if its downstream router (R6) has advertised a label-block. Since R6 has advertised a label block 'N4' and it has received R6 'ID' of 6 it will be picking the 6th label value inside the advertised range of its downstream neighbor. Since R3 is the penultimate router to R6 it MUST program a MPLS POP for its own label range Label(N2+6) NH 10.0.0.14 into its MPLS transit RIB. Furthermore R3 MAY program a MPLS NOP for IP 192.168.1.6, NH 10.0.0.14 into its IPv4 tunnel RIB.

4.9. Expansion of an 'All Router Block' subTLV

All routers within a given area MUST advertise their Label Blocks along with an 'ID'. Now assume that the initial label block size assignment is too small to support all routers which generate an ordinal within an IGP domain. Consider the seven routers in Figure 11, and assume that R7 advertises a new ID '15' using an 'All Router ID Map' subTLV. ID '15' is outside of the range of '10' as

per the previous example in Section 4.8. Now all the routers in an IGP domain need to advertise one more label block in order to map the ID '15' to an actual label value.

All routers would advertise in addition to their label block <N> with a size of 10, a second label block <N2> with a size sufficient enough that the new ordinal can get covered. In this example the same block size 10 is used also for the second label block. For example router R2 would advertise the following label bindings.

```
   TLV 149: MPLS Label <N1>, Flags {}:

      All Router Block subTLV: Block Size 10, Algo 0, Topology 0

      All Router ID IPv4 Map subTLV: ID 2, 192.168.1.2

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

      All Router Block subTLV: Block Size 10, Algo 0, Topology 0
```

Now the upstream router can map the new ID of R7 to an actual label value, as ID '15' corresponds to the 5th label inside the second Label block.

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

5.3.1. MPLS Label operations

If local policy dictates that a given L1L2 router re-advertises a MPLS LSPs into another Level then it MUST prepend its "Traffic-Engineering-ID" as a loose hop in the Prefix ERO subTLV list. If the

LSP is propagated from a higher Level to a lower Level then the 'Down' bit MUST be set.

5.3.2. MPLS Label Block operations

If local policy dictates that a given L1L2 router advertises its 'All Router Block' into another Level, then it also MUST re-advertise all known 'ID' ordinals (again gated by policy) to the respective other Level. Without knowledge of all 'ID's in the network no router is able to construct SPT label switched paths. If a Label Block and its ID mappings are propagated from a higher Level to a lower Level then the 'Down' bit MUST be set.

6. Acknowledgements

Many thanks to Yakov Rekhter and John Drake for their 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
		Unnumbered Interface ID ERO		9	>=0
		IPv4 Prefix Bypass ERO		3	>=0
		IPv6 Prefix Bypass ERO		4	>=0
		Unnumbered Interface ID Bypass ERO		10	>=0
		All Router Block		6	>=0
		All Router ID IPv4 Map		7	>=0
		All Router ID IPv6 Map		8	>=0

Table 1: IANA allocations

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.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[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.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[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.

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

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

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[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.

[RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, February 2011.

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", draft-gredler-rtgwg-igp-label-advertisement-05 (work in progress), May 2013.

[I-D.ietf-rtgwg-mrt-frr-architecture]
Atlas, A., Kebler, R., Envedi, G., Csaszar, A., Tantsura, J., Konstantynowicz, M., White, R., and M. Shand, "An Architecture for IP/LDP Fast-Reroute Using Maximally Redundant Trees", draft-ietf-rtgwg-mrt-frr-architecture-02 (work in progress), February 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", draft-previdi-filsfils-isis-segment-routing-02 (work in progress), 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