     Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR)
          IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs)
                          with MPLS Data Planes

Abstract

   A Segment Routing (SR) architecture leverages source routing and
   tunneling paradigms and can be directly applied to the use of a
   Multiprotocol Label Switching (MPLS) data plane. A node steers a
   packet through a controlled set of instructions called "segments" by
   prepending the packet with an SR header.

   The segment assignment and forwarding semantic nature of SR raises
   additional considerations for connectivity verification and fault
   isolation for a Label Switched Path (LSP) within an SR architecture.
   This document illustrates the problem and defines extensions to
   perform LSP Ping and Traceroute for Segment Routing IGP-Prefix and
   IGP-Adjacency Segment Identifiers (SIDs) with an MPLS data plane.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8287.

Copyright Notice

Table of Contents

1. Introduction

    "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures"
    [RFC8029] defines a simple and efficient mechanism to detect data-
    plane failures in Label Switched Paths (LSPs) by specifying
    information to be carried in an MPLS "echo request" and "echo reply"
    for the purposes of fault detection and isolation. Mechanisms for
    reliably sending the echo reply are defined. The functionality
    defined in [RFC8029] is modeled after the Ping/Traceroute paradigm
    (ICMP echo request [RFC792]) and is typically referred to as "LSP
    Ping" and "LSP Traceroute". [RFC8029] supports hierarchical and
    stitching LSPs.

    [SR] introduces and describes an SR architecture that leverages the
    source routing and tunneling paradigms. A node steers a packet
    through a controlled set of instructions called "segments" by
    prepending the packet with an SR header. A detailed definition of
    the SR architecture is available in [SR].

    As described in [SR] and [SR-MPLS], the SR architecture can be
    directly applied to an MPLS data plane, the SID will be 20 bits, and
    the SR header is the label stack. Consequently, the mechanics of
    data-plane validation of [RFC8029] can be directly applied to SR
    MPLS.

    Unlike LDP or RSVP, which are the other well-known MPLS control plane
    protocols, the basis of Segment ID assignment in SR architecture is
    not always on a hop-by-hop basis. Depending on the type of Segment
    ID, the assignment can be unique to the node or within a domain.

    This nature of SR raises additional considerations for validation of
    fault detection and isolation in an SR network. This document
    illustrates the problem and describes a mechanism to perform LSP Ping
    and Traceroute for Segment Routing IGP-Prefix and IGP-Adjacency SIDs
    within an MPLS data plane.

1.1. Coexistence of SR-Capable and Non-SR-Capable Node Scenarios

   [INTEROP] describes how SR operates in a network where SR-capable and
   non-SR-capable nodes coexist. In such a network, one or more
   SR-based LSPs and non-SR-based LSPs are stitched together to achieve
   an end-to-end LSP. This is similar to a network where LDP and RSVP
   nodes coexist and the mechanism defined in Section 4.5.2 of [RFC8029]
   is applicable for LSP Ping and Trace.

   Section 8 of this document explains one of the potential gaps that is
   specific to SR-Capable and non-SR-capable node scenarios and explains
   how the existing mechanism defined in [RFC8029] handles it.

2. Requirements Notation

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in
   BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
   capitals, as shown here.

3. Terminology

   This document uses the terminology defined in [SR] and [RFC8029];
   readers are expected to be familiar with those terms.

4. Challenges with Existing Mechanisms

   The following example describes the challenges with using the current
   MPLS Operations, Administration, and Maintenance (OAM) mechanisms on
   an SR network.

4.1. Path Validation in Segment Routing Networks

   [RFC8029] defines the MPLS OAM mechanisms that help with fault
   detection and isolation for an MPLS data-plane path by the use of
   various Target Forwarding Equivalence Class (FEC) Stack sub-TLVs that
   are carried in MPLS echo request packets and used by the responder
   for FEC validation. While it is obvious that new sub-TLVs need to be
   assigned for SR, the unique nature of the SR architecture raises the
   need for additional operational considerations for path validation.
   This section discusses the challenges.

```
                        L1
             +---------+
             |    L2   |
             R3-------R6
            /           \
    R1----R2             R7----R8
            \           /
             \         /
             R4-------R5
```
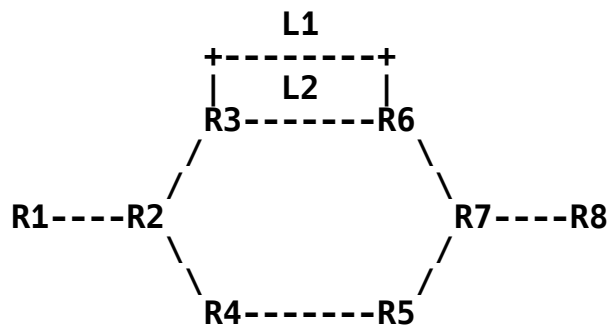
              Figure 1: Segment Routing Network

   The Node Segment IDs for R1, R2, R3, R4, R5, R6, R7, and R8 are 5001,
   5002, 5003, 5004, 5005, 5006, 5007, and 5008, respectively.

      9136 --> Adjacency Segment ID from R3 to R6 over link L1.
      9236 --> Adjacency Segment ID from R3 to R6 over link L2.
      9124 --> Adjacency segment ID from R2 to R4.
      9123 --> Adjacency Segment ID from R2 to R3.

   The forwarding semantic of the Adjacency Segment ID is to pop the
   Segment ID and send the packet to a specific neighbor over a specific
   link. A malfunctioning node may forward packets using the Adjacency
   Segment ID to an incorrect neighbor or over an incorrect link. The
   exposed Segment ID (of an incorrectly forwarded Adjacency Segment ID)
   might still allow such a packet to reach the intended destination,
   even though the intended strict traversal was broken.

   In the topology above, assume that R1 sends traffic with a segment
   stack as {9124, 5008} so that the path taken will be
   R1-R2-R4-R5-R7-R8. If the Adjacency Segment ID 9124 is misprogrammed
   in R2 to send the packet to R1 or R3, the packet may still be
   delivered to R8 (if the nodes are configured with the same SR Global
   Block (SRGB)) [SR] but not via the expected path.

   MPLS traceroute may help with detecting such a deviation in the
   above-mentioned scenario. However, in a different example, it may
   not be helpful, for example, if R3 forwards a packet with Adjacency
   Segment ID 9236 via link L1 (due to misprogramming) when it was
   expected to be forwarded over link L2.

5. Segment ID Sub-TLV

   The format of the following Segment ID sub-TLVs follows the
   philosophy of the Target FEC Stack TLV carrying FECs corresponding to
   each label in the label stack. When operated with the procedures
   defined in [RFC8029], this allows LSP Ping/Traceroute operations to
   function when the Target FEC Stack TLV contains more FECs than
   received label stacks at the responder nodes.

   Three new sub-TLVs are defined for the Target FEC Stack TLV (Type 1),
   the Reverse-Path Target FEC Stack TLV (Type 16), and the Reply Path
   TLV (Type 21).

           Sub-Type   Sub-TLV Name
           --------   ----------------
              34   IPv4 IGP-Prefix Segment ID
              35   IPv6 IGP-Prefix Segment ID
              36   IGP-Adjacency Segment ID

   See Section 9.2 for the registry for the Protocol field specified
   within these sub-TLVs.

5.1. IPv4 IGP-Prefix Segment ID

   The IPv4 IGP-Prefix Segment ID is defined in [SR]. The format is as
   specified below:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                          IPv4 Prefix                          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |Prefix Length  |    Protocol    |           Reserved           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
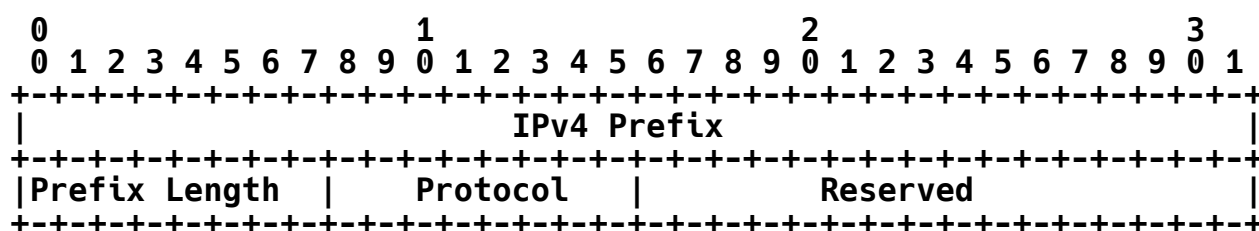
   IPv4 Prefix

      This field carries the IPv4 Prefix to which the Segment ID is
      assigned. In case of an Anycast Segment ID, this field will carry
      the IPv4 Anycast address. If the prefix is shorter than 32 bits,
      trailing bits SHOULD be set to zero.

   Prefix Length

      The Prefix Length field is one octet. It gives the length of the
      prefix in bits (values can be 1-32).

   Protocol

      This field is set to 1, if the responder MUST perform FEC
      validation using OSPF as the IGP protocol. Set to 2, if the
      responder MUST perform Egress FEC validation using the
      Intermediate System to Intermediate System (IS-IS) as the IGP
      protocol. Set to 0, if the responder can use any IGP protocol for
      Egress FEC validation.

   Reserved

      The Reserved field MUST be set to 0 when sent and MUST be ignored
      on receipt.

5.2. IPv6 IGP-Prefix Segment ID

   The IPv6 IGP-Prefix Segment ID is defined in [SR]. The format is as
   specified below:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   |                          IPv6 Prefix                          |
   |                                                               |
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |Prefix Length  |    Protocol    |            Reserved           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   IPv6 Prefix

      This field carries the IPv6 prefix to which the Segment ID is
      assigned. In case of an Anycast Segment ID, this field will carry
      the IPv4 Anycast address. If the prefix is shorter than 128 bits,
      trailing bits SHOULD be set to zero.

   Prefix Length

      The Prefix Length field is one octet, it gives the length of the
      prefix in bits (values can be 1-128).

   Protocol
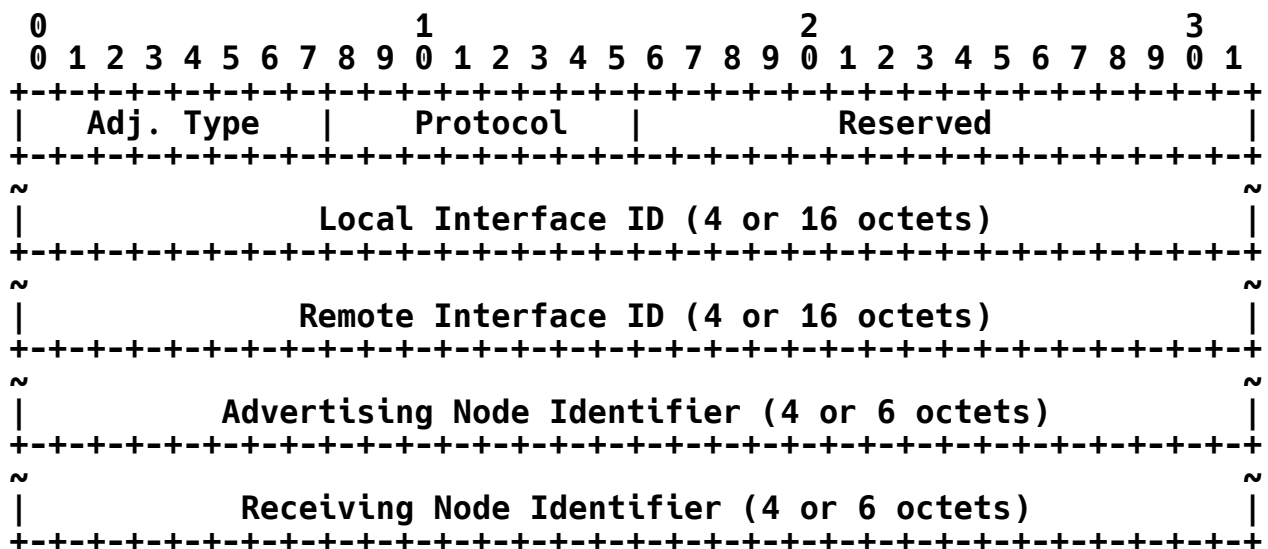
      Set to 1 if the responder MUST perform FEC validation using OSPF
      as the IGP protocol. Set to 2 if the responder MUST perform
      Egress FEC validation using IS-IS as the IGP protocol. Set to 0
      if the responder can use any IGP protocol for Egress FEC
      validation.

   Reserved

      MUST be set to 0 on send and MUST be ignored on receipt.

5.3. IGP-Adjacency Segment ID

   This sub-TLV is applicable for any IGP-Adjacency defined in [SR].
   The format is as specified below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Adj. Type   |    Protocol   |             Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|              Local Interface ID (4 or 16 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|              Remote Interface ID (4 or 16 octets)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|           Advertising Node Identifier (4 or 6 octets)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|            Receiving Node Identifier (4 or 6 octets)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   Adj. Type (Adjacency Type)

      Set to 1 when the Adjacency Segment is a Parallel Adjacency as
      defined in [SR]. Set to 4 when the Adjacency Segment is IPv4
      based and is not a Parallel Adjacency. Set to 6 when the
      Adjacency Segment is IPv6 based and is not a Parallel Adjacency.
      Set to 0 when the Adjacency Segment is over an unnumbered
      interface.

Protocol

   Set to 1 if the responder MUST perform FEC validation using OSPF
   as the IGP protocol. Set to 2 if the responder MUST perform
   Egress FEC validation using IS-IS as the IGP protocol. Set to 0
   if the responder can use any IGP protocol for Egress FEC
   validation.

Reserved

   MUST be set to 0 on send and MUST be ignored on receipt.

Local Interface ID

   An identifier that is assigned by the local Label Switching Router
   (LSR) for a link to which the Adjacency Segment ID is bound. This
   field is set to a local link address (IPv4 or IPv6). For IPv4,
   this field is 4 octets; for IPv6, this field is 16 octets. If
   unnumbered, this field is 4 octets and includes a 32-bit link
   identifier as defined in [RFC4203] and [RFC5307]. If the
   Adjacency Segment ID represents Parallel Adjacencies [SR], this
   field is 4 octets and MUST be set to 4 octets of zeroes.

Remote Interface ID

   An identifier that is assigned by the remote LSR for a link on
   which the Adjacency Segment ID is bound. This field is set to the
   remote (downstream neighbor) link address (IPv4 or IPv6). For
   IPv4, this field is 4 octets; for IPv6, this field is 16 octets.
   If unnumbered, this field is 4 octets and includes a 32-bit link
   identifier as defined in [RFC4203] and [RFC5307]. If the
   Adjacency Segment ID represents Parallel Adjacencies [SR], this
   field is 4 octets and MUST be set to 4 octets of zeroes.

Advertising Node Identifier

   This specifies the Advertising Node Identifier. When the Protocol
   field is set to 1, then this field is 4 octets and carries the
   32-bit OSPF Router ID. If the Protocol field is set to 2, then
   this field is 6 octets and carries the 48-bit IS-IS System ID. If
   the Protocol field is set to 0, then this field is 4 octets and
   MUST be set to zero.

   Receiving Node Identifier

      This specifies the downstream node identifier. When the Protocol
      field is set to 1, then this field is 4 octets and carries the
      32-bit OSPF Router ID. If the Protocol field is set to 2, then
      this field is 6 octets and carries the 48-bit IS-IS System ID. If
      the Protocol field is set to 0, then this field is 4 octets and
      MUST be set to zero.

6. Extension to Downstream Detailed Mapping TLV

   In an echo reply, the Downstream Detailed Mapping TLV [RFC8029] is
   used to report for each interface over which a FEC could be
   forwarded. For a FEC, there are multiple protocols that may be used
   to distribute label mapping. The Protocol field of the Downstream
   Detailed Mapping TLV is used to return the protocol that is used to
   distribute the label carried in the Downstream Label field. The
   following protocols are defined in [RFC8029]:

      Protocol #   Signaling Protocol
      ----------   ------------------
         0   Unknown
         1   Static
         2   BGP
         3   LDP
         4   RSVP-TE

   With SR, OSPF or IS-IS can be used for label distribution. This
   document adds two new protocols as follows:

      Protocol #   Signaling Protocol
      ----------   ------------------
         5   OSPF
         6   IS-IS

   See Section 9.4.

7. Procedures

   This section describes aspects of LSP Ping and Traceroute operations
   that require further considerations beyond [RFC8029].

7.1. FECs in Target FEC Stack TLV

   When LSP echo request packets are generated by an initiator, FECs
   carried in the Target FEC Stack TLV may need to differ to support an
   SR architecture. The following defines the Target FEC Stack TLV
   construction mechanics by an initiator for SR scenarios.

Ping

   The initiator MUST include FEC(s) corresponding to the
   destination segment.

   The initiator MAY include FECs corresponding to some or all of
   the segments imposed in the label stack by the initiator to
   communicate the segments traversed.

Traceroute

   The initiator MUST initially include FECs corresponding to all
   segments imposed in the label stack.

   When a received echo reply contains the FEC Stack Change TLV
   with one or more of the original segments being popped, the
   initiator MAY remove a corresponding FEC(s) from the Target FEC
   Stack TLV in the next (TTL+1) traceroute request, as defined in
   Section 4.6 of [RFC8029].

   When a received echo reply does not contain the FEC Stack
   Change TLV, the initiator MUST NOT attempt to remove any FECs
   from the Target FEC Stack TLV in the next (TTL+1) traceroute
   request.

As defined in [SR-OSPF] and [SR-IS-IS], the Prefix SID can be
advertised as an absolute value, an index, or as a range. In any of
these cases, the initiator MUST derive the Prefix mapped to the
Prefix SID and use it in the IGP-Prefix Segment ID defined in
Sections 5.1 and 5.2. How the responder uses the details in the
SR-FEC sub-TLV to perform the validation is a local implementation
matter.

7.2. FEC Stack Change Sub-TLV

   [RFC8029] defines a FEC Stack Change sub-TLV that a router must
   include when the FEC stack changes.

   The network node that advertised the Node Segment ID is responsible
   for generating a FEC Stack Change sub-TLV with the Post Office
   Protocol (POP) operation type for the Node Segment ID, regardless of
   whether or not Penultimate Hop Popping (PHP) is enabled.

   The network node that is immediately downstream of the node that
   advertised the Adjacency Segment ID is responsible for generating the
   FEC Stack Change sub-TLV for POP operation for the Adjacency Segment
   ID.

7.3. Segment ID POP Operation

   The forwarding semantic of the Node Segment ID with the PHP flag is
   equivalent to usage of Implicit Null in MPLS protocols. The
   Adjacency Segment ID is also similar in a sense that it can be
   thought of as a locally allocated segment that has PHP enabled when
   destined for the next-hop IGP Adjacency Node. Procedures described
   in Section 4.4 of [RFC8029] rely on the Stack-D and Stack-R
   explicitly having the Implicit Null value. Implementations SHOULD
   use the Implicit Null for the Node Segment ID PHP and Adjacency
   Segment ID PHP cases.

7.4. Segment ID Check

   This section modifies the procedure defined in Section 4.4.1 of
   [RFC8029]. Step 4 defined in Section 4.4.1 of [RFC8029] is modified
   as below:

      4. If the label mapping for FEC is Implicit Null, set the
         FEC-status to 2 and proceed to step 4a. Otherwise,
         if the label mapping for FEC is Label-L, proceed to step 4a.
         Otherwise, set the FEC-return-code to 10 ("Mapping for this
         FEC is not the given label at stack-depth"), set the
         FEC-status to 1, and return.

     4a. Segment Routing IGP-Prefix and IGP-Adjacency SID Validation:

       If the Label-stack-depth is 0 and the Target FEC Stack sub-TLV
       at FEC-stack-depth is 34 (IPv4 IGP-Prefix Segment ID), {

          Set the Best-return-code to 10, "Mapping for this FEC is not
          the given label at stack-depth <RSC>" if any below
          conditions fail:

          /* The responder LSR is to check if it is the egress of the
          IPv4 IGP-Prefix Segment ID described in the Target FEC Stack
          sub-TLV, and if the FEC was advertised with the PHP bit
          set.*/

          - Validate that the Node Segment ID is advertised for the
             IPv4 Prefix by IGP Protocol {

             o When the Protocol field in the received IPv4 IGP-
                Prefix Segment ID sub-TLV is 0, use any locally
                enabled IGP protocol.

          o When the Protocol field in the received IPv4 IGP-
             Prefix Segment ID sub-TLV is 1, use OSPF as the IGP
             protocol.

          o When the Protocol field in the received IPv4 IGP-
             Prefix Segment ID sub-TLV is 2, use IS-IS as the IGP
             protocol.

          o When the Protocol field in the received IPv4 IGP-
             Prefix Segment ID sub-TLV is an unrecognized value, it
             MUST be treated as a Protocol value of 0.

        }

     - Validate that the Node Segment ID is advertised with the
       No-PHP flag. {

          o When the Protocol is OSPF, the NP-Flag defined in
             Section 5 of [SR-OSPF] MUST be set to 0.

          o When the Protocol is IS-IS, the P-Flag defined in
             Section 6.1 of [SR-IS-IS] MUST be set to 0.

        }

     If it can be determined that no protocol associated with the
     Interface-I would have advertised the FEC-Type at FEC-stack-
     depth, set the Best-return-code to 12, "Protocol not
     associated with interface at FEC-stack-depth" and return.

     Set FEC-Status to 1 and return.

    }

    Else, if the Label-stack-depth is greater than 0 and the Target
    FEC Stack sub-TLV at FEC-stack-depth is 34 (IPv4 IGP-Prefix
    Segment ID), {

     Set the Best-return-code to 10 if any below conditions fail:

     - Validate that the Node Segment ID is advertised for the
       IPv4 Prefix by the IGP protocol {

          o When the Protocol field in the received IPv4 IGP-
             Prefix Segment ID sub-TLV is 0, use any locally
             enabled IGP protocol.

   o When the Protocol field in the received IPv4 IGP-
      Prefix Segment ID sub-TLV is 1, use OSPF as the IGP
      protocol.

   o When the Protocol field in the received IPv4 IGP-
      Prefix Segment ID sub-TLV is 2, use IS-IS as the IGP
      protocol.

   o When the Protocol field in the received IPv4 IGP-
      Prefix Segment ID sub-TLV is an unrecognized value, it
      MUST be treated as a Protocol value of 0.

   }

   If it can be determined that no protocol associated with
   Interface-I would have advertised the FEC-Type at FEC-stack-
   depth, set the Best-return-code to 12, "Protocol not
   associated with interface at FEC stack-depth" and return.

   Set FEC-Status to 1 and return.

}

Else, if the Label-stack-depth is 0 and the Target FEC sub-TLV
at FEC-stack-depth is 35 (IPv6 IGP-Prefix Segment ID), {

   Set the Best-return-code to 10 if any of the below
   conditions fail:

   /* The LSR needs to check if it is being a tail-end for the
   LSP and have the prefix advertised with the PHP bit set*/

   - Validate that the Node Segment ID is advertised for the
      IPv6 Prefix by the IGP protocol {

      o When the Protocol field in the received IPv6 IGP-
         Prefix Segment ID sub-TLV is 0, use any locally
         enabled IGP protocol.

      o When the Protocol field in the received IPv6 IGP-
         Prefix Segment ID sub-TLV is 1, use OSPF as the IGP
         protocol.

      o When the Protocol field in the received IPv6 IGP-
         Prefix Segment ID sub-TLV is 2, use IS-IS as the IGP
         protocol.

         o When the Protocol field in the received IPv6 IGP-
            Prefix Segment ID sub-TLV is an unrecognized value, it
            MUST be treated as a Protocol value of 0.

         }

      - Validate that the Node Segment ID is advertised with the
         No-PHP flag. {

         o When the Protocol is OSPF, the NP-flag defined in
            Section 5 of [SR-OSPFV3] MUST be set to 0.

         o When the Protocol is IS-IS, the P-Flag defined in
            Section 6.1 of [SR-IS-IS] MUST be set to 0.

         }

      If it can be determined that no protocol associated with
      Interface-I would have advertised the FEC-Type at FEC-stack-
      depth, set the Best-return-code to 12, "Protocol not
      associated with interface at FEC stack-depth" and return.

      Set the FEC-Status to 1 and return.

   }

   Else, if the Label-stack-depth is greater than 0 and the Target
   FEC sub-TLV at FEC-stack-depth is 35 (IPv6 IGP-Prefix Segment
   ID), {

      Set the Best-return-code to 10 if any below conditions fail:

      - Validate that the Node Segment ID is advertised for the
         IPv4 Prefix by the IGP protocol {

         o When the Protocol field in the received IPv6 IGP-
            Prefix Segment ID sub-TLV is 0, use any locally
            enabled IGP protocol.

         o When the Protocol field in the received IPv6 IGP-
            Prefix Segment ID sub-TLV is 1, use OSPF as the IGP
            protocol.

         o When the Protocol field in the received IPv6 IGP-
            Prefix Segment ID sub-TLV is 2, use IS-IS as the IGP
            protocol.

o When the Protocol field in the received IPv6 IGP-
   Prefix Segment ID sub-TLV is an unrecognized value, it
   MUST be treated as a Protocol value of 0.

}

If it can be determined that no protocol associated with
Interface-I would have advertised the FEC-Type at FEC-stack-
depth, set the Best-return-code to 12, "Protocol not
associated with interface at FEC stack-depth" and return.

Set the FEC-Status to 1 and return.

}

Else, if the Target FEC sub-TLV at FEC-stack-depth is 36
(IGP-Adjacency Segment ID), {

Set the Best-return-code to 35 (Section 9.5) if any below
conditions fail:

When the Adj. Type is 1 (Parallel Adjacency):

o Validate that the Receiving Node Identifier is the
   local IGP identifier.

o Validate that the IGP-Adjacency Segment ID is
   advertised by the Advertising Node Identifier of the
   Protocol in the local IGP database {

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 0, use any locally
      enabled IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 1, use OSPF as the
      IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 2, use IS-IS as the
      IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is an unrecognized
      value, it MUST be treated as a Protocol value of 0.

   }

When the Adj. Type is 4 or 6 (IGP Adjacency or LAN
Adjacency):

o Validate that the Remote Interface ID matches the
   local identifier of the interface (Interface-I) on
   which the packet was received.

o Validate that the Receiving Node Identifier is the
   local IGP identifier.

o Validate that the IGP-Adjacency Segment ID is
   advertised by the Advertising Node Identifier of
   Protocol in the local IGP database {

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 0, use any locally
      enabled IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 1, use OSPF as the
      IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is 2, use IS-IS as the
      IGP protocol.

   * When the Protocol field in the received IGP-
      Adjacency Segment ID sub-TLV is an unrecognized
      value, it MUST be treated as a Protocol value of 0.

   }

Set the FEC-Status to 1 and return.

}

7.5. TTL Consideration for Traceroute

   The LSP Traceroute operation can properly traverse every hop of the
   SR network for the Uniform Model as described in [RFC3443]. If one
   or more LSRs employ a Short Pipe Model, as described in [RFC3443],
   then the LSP Traceroute may not be able to properly traverse every
   hop of the SR network due to the absence of TTL copy operation when
   the outer label is popped. The Short Pipe is one of the most
   commonly used models. The following TTL manipulation technique MAY
   be used when the Short Pipe Model is used.

   When tracing an LSP according to the procedures in [RFC8029], the TTL
   is incremented by one in order to trace the path sequentially along
   the LSP. However, when a source-routed LSP has to be traced, there
   are as many TTLs as there are labels in the stack. The LSR that
   initiates the traceroute SHOULD start by setting the TTL to 1 for the
   tunnel in the LSP's label stack it wants to start the tracing from,
   the TTL of all outer labels in the stack to the max value, and the
   TTL of all the inner labels in the stack to zero. Thus, a typical
   start to the traceroute would have a TTL of 1 for the outermost label
   and all the inner labels would have a TTL of 0. If the FEC Stack TLV
   is included, it should contain only those for the inner-stacked
   tunnels. The Return Code/Subcode and FEC Stack Change TLV should be
   used to diagnose the tunnel as described in [RFC8029]. When the
   tracing of a tunnel in the stack is complete, then the next tunnel in
   the stack should be traced. The end of a tunnel can be detected from
   the Return Code when it indicates that the responding LSR is an
   egress for the stack at depth 1. Thus, the traceroute procedures in
   [RFC8029] can be recursively applied to traceroute a source-routed
   LSP.

8. Backward Compatibility with Non-SR Devices

   [INTEROP] describes how SR operates in a network where SR-capable and
   non-SR-capable nodes coexist. In such networks, there may not be any
   FEC mapping in the responder when the initiator is SR-capable, while
   the responder is not (or vice-versa). But this is not different from
   RSVP and LDP interoperation scenarios. When LSP Ping is triggered,
   the responder will set the FEC-return-code to Return 4, "Replying
   router has no mapping for the FEC at stack-depth".

   Similarly, when an SR-capable node assigns Adj-SID for a non-SR-
   capable node, the LSP traceroute may fail as the non-SR-capable node
   is not aware of the "IGP Adjacency Segment ID" sub-TLV and may not
   reply with the FEC Stack Change sub-TLVs. This may result in any
   further downstream nodes replying back with a Return Code of 4,
   "Replying router has no mapping for the FEC at stack-depth".

9. IANA Considerations

9.1. New Target FEC Stack Sub-TLVs

IANA has assigned three new sub-TLVs from the "sub-TLVs for TLV Types
1, 16, and 21" subregistry of the "Multi-Protocol Label Switching
(MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA].

Sub-Type	Sub-TLV Name	Reference
34	IPv4 IGP-Prefix Segment ID	Section 5.1
35	IPv6 IGP-Prefix Segment ID	Section 5.2
36	IGP-Adjacency Segment ID	Section 5.3

9.2. Protocol in the Segment ID Sub-TLV

IANA has created a new "Protocol in the Segment ID sub-TLV" (see
Section 5) registry under the "Multi-Protocol Label Switching (MPLS)
Label Switched Paths (LSPs) Ping Parameters" registry. Code points
in the range of 0-250 will be assigned by Standards Action [RFC8126].
The range of 251-254 is reserved for experimental use and will not be
assigned. The value of 255 is marked "Reserved". The initial
entries into the registry are:

Value	Meaning	Reference
0	Any IGP protocol	This document
1	OSPF	This document
2	IS-IS	This document

9.3. Adjacency Type in the IGP-Adjacency Segment ID

IANA has created a new "Adjacency Type in the IGP-Adjacency Segment
ID" registry (see Section 5.3) under the "Multi-Protocol Label
Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters"
registry. Code points in the range of 0-250 will be assigned by
Standards Action. The range of 251-254 is reserved for experimental
use and will not be assigned. The value of 255 is marked "Reserved".
The initial entries into the registry are:

Value	Meaning
0	Unnumbered Interface Adjacency
1	Parallel Adjacency
4	IPv4, Non-parallel Adjacency
6	IPv6, Non-parallel Adjacency

9.4. Protocol in the Label Stack Sub-TLV of the Downstream Detailed
      Mapping TLV

   IANA has created a new "Protocol in the Label Stack sub-TLV of the
   Downstream Detailed Mapping TLV" registry under the "Multi-Protocol
   Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters"
   registry. Code points in the range of 0-250 will be assigned by
   Standards Action. The range of 251-254 is reserved for experimental
   use and will not be assigned. The value of 255 is marked "Reserved".
   The initial entries into the registry are:

Value	Meaning	Reference
0	Unknown	Section 3.4.1.2 of RFC 8029
1	Static	Section 3.4.1.2 of RFC 8029
2	BGP	Section 3.4.1.2 of RFC 8029
3	LDP	Section 3.4.1.2 of RFC 8029
4	RSVP-TE	Section 3.4.1.2 of RFC 8029
5	OSPF	Section 6 of this document
6	IS-IS	Section 6 of this document
7-250	Unassigned	
251-254	Reserved for Experimental Use	This document
255	Reserved	This document

9.5. Return Code

   IANA has assigned a new Return Code from the "Multi-Protocol Label
   Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" in the
   0-191 (Standards Action) range from the "Return Codes" subregistry.

Value	Meaning	Reference
35	Mapping for this FEC is not associated with the incoming interface	Section 7.4 of this document

10. Security Considerations

   This document defines additional MPLS LSP Ping sub-TLVs and follows
   the mechanisms defined in [RFC8029]. All the security considerations
   defined in [RFC8029] will be applicable for this document and, in
   addition, they do not impose any additional security challenges to be
   considered.

11. References

11.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://www.rfc-editor.org/info/rfc2119>.

   [RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing
              in Multi-Protocol Label Switching (MPLS) Networks",
              RFC 3443, DOI 10.17487/RFC3443, January 2003,
              <https://www.rfc-editor.org/info/rfc3443>.

   [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in
              Support of Generalized Multi-Protocol Label Switching
              (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005,
              <https://www.rfc-editor.org/info/rfc4203>.

   [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions
              in Support of Generalized Multi-Protocol Label Switching
              (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008,
              <https://www.rfc-editor.org/info/rfc5307>.

   [RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N.,
              Aldrin, S., and M. Chen, "Detecting Multiprotocol Label
              Switched (MPLS) Data-Plane Failures", RFC 8029,
              DOI 10.17487/RFC8029, March 2017,
              <https://www.rfc-editor.org/info/rfc8029>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

11.2. Informative References

   [IANA]   IANA, "Multi-Protocol Label Switching (MPLS) Label
              Switched Paths (LSPs) Ping Parameters",
              <http://www.iana.org/assignments/
              mpls-lsp-ping-parameters>.

   [INTEROP] Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., and
              S. Litkowski, "Segment Routing interworking with LDP",
              Work in Progress, draft-ietf-spring-segment-routing-ldp-
              interop-09, September 2017.

   [RFC792]   Postel, J., "Internet Control Message Protocol", STD 5,
               RFC 792, DOI 10.17487/RFC0792, September 1981,
               <https://www.rfc-editor.org/info/rfc792>.

   [RFC8126]   Cotton, M., Leiba, B., and T. Narten, "Guidelines for
               Writing an IANA Considerations Section in RFCs", BCP 26,
               RFC 8126, DOI 10.17487/RFC8126, June 2017,
               <https://www.rfc-editor.org/info/rfc8126>.

   [SR]   Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B.,
               Litkowski, S., and R. Shakir, "Segment Routing
               Architecture", Work in Progress, draft-ietf-spring-
               segment-routing-14, December 2017.

   [SR-IS-IS] Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A.,
               Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura,
               "IS-IS Extensions for Segment Routing", Work in Progress,
               draft-ietf-isis-segment-routing-extensions-15, December
               2017.

   [SR-MPLS]   Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
               Litkowski, S., and R. Shakir, "Segment Routing with MPLS
               data plane", Work in Progress, draft-ietf-spring-segment-
               routing-mpls-11, October 2017.

   [SR-OSPF]   Psenak, P., Previdi, S., Filsfils, C., Gredler, H.,
               Shakir, R., Henderickx, W., and J. Tantsura, "OSPF
               Extensions for Segment Routing", Work in Progress,
               draft-ietf-ospf-segment-routing-extensions-24, December
               2017.

   [SR-OSPFV3]
               Psenak, P., Previdi, S., Filsfils, C., Gredler, H.,
               Shakir, R., Henderickx, W., and J. Tantsura, "OSPFv3
               Extensions for Segment Routing", Work in Progress,
               draft-ietf-ospf-ospfv3-segment-routing-extensions-10,
               September 2017.

Acknowledgements

Contributors

   The following are key contributors to this document:

      Hannes Gredler, RtBrick, Inc.
      Tarek Saad, Cisco Systems, Inc.
      Siva Sivabalan, Cisco Systems, Inc.
      Balaji Rajagopalan, Juniper Networks
      Faisal Iqbal, Cisco Systems, Inc.

Authors' Addresses

   Nagendra Kumar (editor)
   Cisco Systems, Inc.
   7200-12 Kit Creek Road
   Research Triangle Park, NC 27709-4987
   United States of America

   Email: naikumar@cisco.com

   Carlos Pignataro (editor)
   Cisco Systems, Inc.
   7200-11 Kit Creek Road
   Research Triangle Park, NC 27709-4987
   United States of America

   Email: cpignata@cisco.com

   George Swallow
   Southend Technical Center

   Email: swallow.ietf@gmail.com

   Nobo Akiya
   Big Switch Networks

   Email: nobo.akiya.dev@gmail.com

   Sriganesh Kini
   Individual

   Email: sriganeshkini@gmail.com

   Mach(Guoyi) Chen
   Huawei

   Email: mach.chen@huawei.com