

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 10, 2017

S. Previdi, Ed.
Cisco Systems, Inc.
J. Dong, Ed.
M. Chen
Huawei Technologies
H. Gredler
RtBrick Inc.
J. Tantsura
Individual
January 6, 2017

Distribution of Traffic Engineering (TE) Policies and State using BGP-LS [draft-ietf-idr-te-lsp-distribution-06](#)

Abstract

This document describes a mechanism to collect the Traffic Engineering and Policy information that is locally available in a router and advertise it into BGP-LS updates. Such information can be used by external components for path computation, re-optimization, service placement, network visualization, etc.

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

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 10, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Carrying TE Policy Information in BGP	5
2.1.	TE Policy Information	5
2.2.	TE Policy NLRI	5
2.2.1.	TE Policy Descriptors	6
2.3.	TE Policy State	12
2.3.1.	RSVP Objects	14
2.3.2.	PCE Objects	15
2.3.3.	SR TE Policy Sub-TLVs	15
3.	Operational Considerations	21
4.	IANA Considerations	22
4.1.	BGP-LS NLRI-Types	22
4.2.	BGP-LS Protocol-IDs	22
4.3.	BGP-LS Descriptors TLVs	22
4.4.	BGP-LS LSP-State TLV Protocol Origin	23
5.	Security Considerations	23
6.	Acknowledgements	23
7.	References	23
7.1.	Normative References	24
7.2.	Informative References	25
	Authors' Addresses	26

[1.](#) Introduction

In many network environments, traffic engineering policies are instantiated into various forms:

- o MPLS Traffic Engineering Label Switched Paths (TE-LSPs).
- o IP based tunnels (IP in IP, GRE, etc).

- o Segment Routing Traffic Engineering Policies (SR TE Policy) as defined in [[I-D.previdi-idr-segment-routing-te-policy](#)]
- o Local cross-connect configuration

All this information can be grouped into the same term: TE Policies. In the rest of this document we refer to TE Policies as the set of information related to the various instantiation of policies: MPLS TE LSPs, IP tunnels (IPv4 or IPv6), SR TE Policies, etc.

TE Policies are generally instantiated by the head-end and are based on either local configuration or controller based programming of the node using various protocols and APIs, e.g., PCEP or BGP.

In many network environments, the configuration and state of each TE Policy that is available in the network is required by a controller which allows the network operator to optimize several functions and operations through the use of a controller aware of both topology and state information.

One example of a controller is the stateful Path Computation Element (PCE) [[I-D.ietf-pce-stateful-pce](#)], which could provide benefits in path reoptimization. While some extensions are proposed in Path Computation Element Communication Protocol (PCEP) for the Path Computation Clients (PCCs) to report the LSP states to the PCE, this mechanism may not be applicable in a management-based PCE architecture as specified in [section 5.5 of \[RFC4655\]](#). As illustrated in the figure below, the PCC is not an LSR in the routing domain, thus the head-end nodes of the TE-LSPs may not implement the PCEP protocol. In this case a general mechanism to collect the TE-LSP states from the ingress LERs is needed. This document proposes an TE Policy state collection mechanism complementary to the mechanism defined in [[I-D.ietf-pce-stateful-pce](#)].

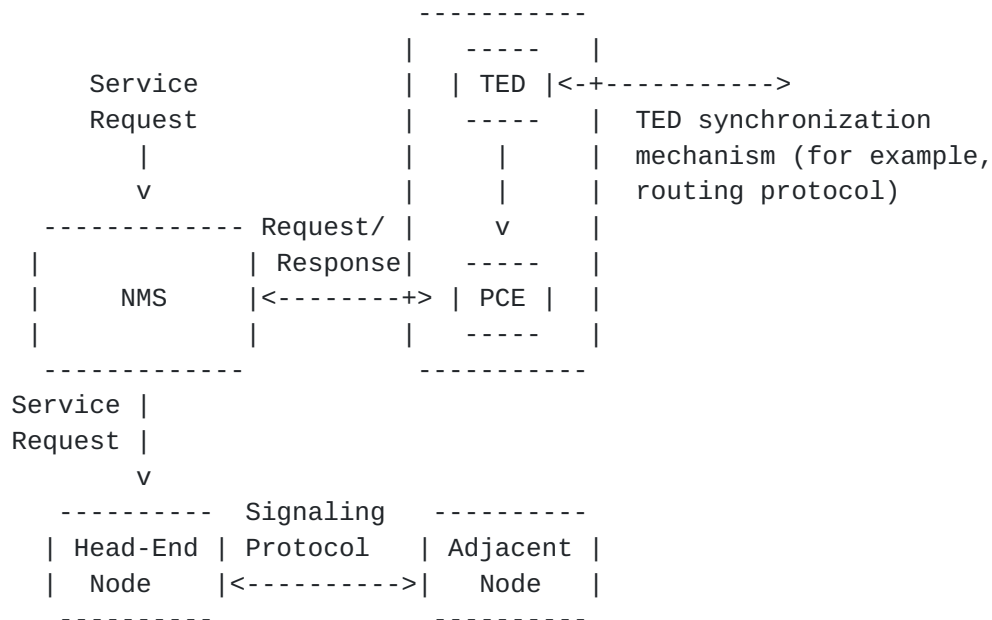


Figure 1. Management-Based PCE Usage

In networks with composite PCE nodes as specified in [section 5.1 of \[RFC4655\]](#), PCE is implemented on several routers in the network, and the PCCs in the network can use the mechanism described in [\[I-D.ietf-pce-stateful-pce\]](#) to report the TE Policy information to the PCE nodes. An external component may also need to collect the TE Policy information from all the PCEs in the network to obtain a global view of the LSP state in the network.

In multi-area or multi-AS scenarios, each area or AS can have a child PCE to collect the TE Policies in its own domain, in addition, a parent PCE needs to collect TE Policy information from multiple child PCEs to obtain a global view of LSPs inside and across the domains involved.

In another network scenario, a centralized controller is used for service placement. Obtaining the TE Policy state information is quite important for making appropriate service placement decisions with the purpose to both meet the application's requirements and utilize network resources efficiently.

The Network Management System (NMS) may need to provide global visibility of the TE Policies in the network as part of the network visualization function.

BGP has been extended to distribute link-state and traffic engineering information to external components [[RFC7752](#)]. Using the same protocol to collect Traffic Engineering and Policy information

is desirable for these external components since this avoids introducing multiple protocols for network information collection. This document describes a mechanism to distribute traffic engineering and policy information (MPLS, IPv4 and IPv6) to external components using BGP-LS.

2. Carrying TE Policy Information in BGP

2.1. TE Policy Information

TE Policy information is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. The "Link-State NLRI" defined in [RFC7752] is extended to carry the TE Policy information. BGP speakers that wish to exchange TE Policy information MUST use the BGP Multiprotocol Extensions Capability Code (1) to advertise the corresponding (AFI, SAFI) pair, as specified in [RFC4760].

The format of "Link-State NLRI" is defined in [RFC7752]. A new "NLRI Type" is defined for TE Policy Information as following:

- o NLRI Type: TE Policy NLRI (suggested codepoint value 5, to be assigned by IANA).

[RFC7752] defines the BGP-LS NLRI as follows:

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	2	3	4	5	6	7	8	9																																								
+																+																+																+																															
																NLRI Type																																Total NLRI Length																															
+																+																+																+																															
//																Link-State NLRI (variable)																																//																															
+																+																+																+																															

This document defines a new NLRI-Type and its format: the TE Policy NLRI defined in the following section.

2.2. TE Policy NLRI

The TE Policy NLRI (NLRI Type 5. Suggested value, to be assigned by IANA) is shown in the following figure:


```

      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
+---+---+---+---+---+---+
| Protocol-ID |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Identifier                                     |
|                                     (64 bits)                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     TE Policy Descriptors (variable)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Protocol-ID field specifies the signaling protocol of the TE Policy. The following Protocol-IDs are defined (suggested values, to be assigned by IANA) and apply to the TE Policy NLRI:

```

+-----+-----+
| Protocol-ID | NLRI information source protocol |
+-----+-----+
|      8      |      RSVP-TE                               |
|      9      |      Segment Routing                       |
+-----+-----+

```

- o "Identifier" is an 8 octet value as defined in [\[RFC7752\]](#).
- o Following TE Policy Descriptors are defined:

```

+-----+-----+
| Codepoint | Descriptor TLV |
+-----+-----+
| 267      | Tunnel ID      |
| 268      | LSP ID         |
| 269      | IPv4/6 Tunnel Head-end address |
| 270      | IPv4/6 Tunnel Tail-end address |
| 271      | SR TE Policy   |
| 272      | Local Cross Connect |
+-----+-----+

```

[2.2.1.](#) TE Policy Descriptors

This sections defines the TE Policy Descriptors TLVs.

[2.2.1.1.](#) Tunnel Identifier (Tunnel ID)

The Tunnel Identifier TLV contains the Tunnel ID defined in [\[RFC3209\]](#) 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               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               Tunnel ID           |
+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 267)
- o Length: 2 octets.
- o Tunnel ID: 2 octets as defined in [\[RFC3209\]](#).

2.2.1.2. LSP Identifier (LSP ID)

The LSP Identifier TLV contains the LSP ID defined in [\[RFC3209\]](#) 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               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               LSP ID             |
+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 268)
- o Length: 2 octets.
- o LSP ID: 2 octets as defined in [\[RFC3209\]](#).

2.2.1.3. IPv4/IPv6 Tunnel Head-End Address

The IPv4/IPv6 Tunnel Head-End Address TLV contains the Tunnel Head-End Address defined in [\[RFC3209\]](#) and has 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//          IPv4/IPv6 Tunnel Head-End Address (variable)          //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 269)
- o Length: 4 or 16 octets.

When the IPv4/IPv6 Tunnel Head-end Address TLV contains an IPv4 address, its length is 4 (octets).

When the IPv4/IPv6 Tunnel Head-end Address TLV contains an IPv6 address, its length is 16 (octets).

2.2.1.4. IPv4/IPv6 Tunnel Tail-End Address

The IPv4/IPv6 Tunnel Tail-End Address TLV contains the Tunnel Tail-End Address defined in [[RFC3209](#)] and has 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//          IPv4/IPv6 Tunnel Tail-End Address (variable)          //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 270)
- o Length: 4 or 16 octets.

When the IPv4/IPv6 Tunnel Tail-end Address TLV contains an IPv4 address, its length is 4 (octets).

When the IPv4/IPv6 Tunnel Tail-end Address TLV contains an IPv6 address, its length is 16 (octets).

2.2.1.5. SR TE Policy TLV

The SR TE Policy TLV identifies a SR TE Policy as defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#) 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                 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
|               Distinguisher (4 octets)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
|               Policy Color (4 octets)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
|               Endpoint (4 or 16 octets)                             |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 271)
- o Length: 12 octets.
- o Distinguisher, Policy Color and Endpoint are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.2.1.6. MPLS Cross Connect

The MPLS Cross Connect TLV identifies a local MPLS state in the form of incoming label and interface followed by an outgoing label and interface. Outgoing interface may appear multiple times (for multicast states).

The Local Cross Connect TLV 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Incoming label (4 octets)          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Outgoing label (4 octets)         |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     Sub-TLVs (variable)                //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 271)
- o Length: variable.
- o Incoming and Outgoing labels: 4 octets each.
- o Sub-TLVs: following Sub-TLVs are defined:
 - * Interface Sub-TLV
 - * Forwarding Equivalent Class (FEC)

The MPLS Cross Connect TLV:

MUST have an incoming label.

MUST have an outgoing label.

MAY contain an Interface Sub-TLV having the I-flag set.

MUST contain at least one Interface Sub-TLV having the I-flag unset.

MAY contain multiple Interface Sub-TLV having the I-flag unset.
This is the case of a multicast MPLS cross connect.

MAY contain a FEC Sub-TLV.

[2.2.1.6.1](#). MPLS Cross Connect Sub-TLVs

2.2.1.6.1.1. Interface Sub-TLV

The Interface sub-TLV is optional and contains the identifier of the interface (incoming or outgoing) in the form of an IPv4 address or an IPv6 address.

The Interface sub-TLV 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                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|  Flags                      |  Interface Address (4 or 16 octets)  //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 1)
- o Length: 5 or 17.
- o Flags: 1 octet of flags defined as follows:

```

      0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
| I |           |
+---+---+---+---+---+---+

```

where:

- * I-Flag is the Interface flag. When set, the Interface Sub-TLV describes an incoming interface. If the I-flag is not set, then the Interface Sub-TLV describes an outgoing interface.

- o Interface address: a 4 octet IPv4 address or a 16 octet IPv6 address.

2.2.1.6.1.2. Forwarding Equivalent Class (FEC) Sub-TLV

The FEC sub-TLV is optional and contains the FEC associated to the incoming label.

The FEC sub-TLV 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          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|  Flags      | Masklength | Prefix (variable)      //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//              Prefix (variable)                  //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: To be assigned by IANA (suggested value: 2)
- o Length: variable.
- o Flags: 1 octet of flags defined as follows:

```

0 1 2 3 4 5 6 7
+---+---+---+---+
|4|          |
+---+---+---+---+

```

where:

- * 4-Flag is the IPv4 flag. When set, the FEC Sub-TLV describes an IPv4 FEC. If the 4-flag is not set, then the FEC Sub-TLV describes an IPv6 FEC.

- o Mask Length: 1 octet of prefix length.
- o Prefix: an IPv4 or IPv6 prefix whose mask length is given by the "Mask Length" field.

2.3. TE Policy State

A new TLV called "TE Policy State TLV" (codepoint to be assigned by IANA), is used to describe the characteristics of the TE Policy, which is carried in the optional non-transitive BGP Attribute "LINK_STATE Attribute" defined in [RFC7752]. These TE Policy characteristics include the characteristics and attributes of the policy, its dataplane, explicit path, Quality of Service (QoS) parameters, route information, the protection mechanisms, etc.


```

      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                 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
//               TE Policy State Information (variable)               //
|                                     |                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

TE Policy State TLV

Type: Suggested value 1158 (to be assigned by IANA)

TE Policy State Information: Consists of a set of objects as defined in [RFC3209], [RFC3473], [RFC5440] and [I-D.previdi-idr-segment-routing-te-policy]. Rather than replicating all these objects in this document, the semantics and encodings of the objects are reused. These objects are carried in the "TE Policy State Information" with 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Protocol-Origin|   Reserved   |               Length                 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
//               Protocol specific TE Policy objects               //
|                                     |                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

TE Policy State Information

The Protocol-Origin field identifies the protocol from which the contained object originated. This allows for objects defined in different protocols to be collected while avoiding the possible code collisions among these protocols. Three Protocol-Origins are defined in this document (suggested values, to be assigned by IANA)

```

+-----+-----+
| Protocol | LSP Object |
| Origin  | Origin    |
+-----+-----+
|    1    | RSVP-TE   |
|    2    | PCE       |
|    3    | BGP SR TE Policy |
+-----+-----+

```


The 8-bit Reserved field SHOULD be set to 0 on transmission and MUST be ignored on receipt.

The Length field is set to the Length of the value field, which is the total length of the contained object.

The Valued field is a TE Policy object which is defined in the protocol identified by the Protocol-Origin field.

2.3.1. RSVP Objects

RSVP-TE objects are encoded in the "Value" field of the LSP State TLV and consists of MPLS TE LSP objects defined in RSVP-TE [[RFC3209](#)] [[RFC3473](#)]. Rather than replicating all MPLS TE LSP related objects in this document, the semantics and encodings of the MPLS TE LSP objects are re-used. These MPLS TE LSP objects are carried in the LSP State TLV.

When carrying RSVP-TE objects, the "Protocol-Origin" field is set to "RSVP-TE" (suggested value 1, to be assigned by IANA).

The following RSVP-TE Objects are defined:

- o SENDER_TSPEC and FLOW_SPEC [[RFC2205](#)]
- o SESSION_ATTRIBUTE [[RFC3209](#)]
- o EXPLICIT_ROUTE Object (ERO) [[RFC3209](#)]
- o ROUTE_RECORD Object (RRO) [[RFC3209](#)]
- o FAST_REROUTE Object [[RFC4090](#)]
- o DETOUR Object [[RFC4090](#)]
- o EXCLUDE_ROUTE Object (XRO) [[RFC4874](#)]
- o SECONDARY_EXPLICIT_ROUTE Object (SERO) [[RFC4873](#)]
- o SECONDARY_RECORD_ROUTE (SRRO) [[RFC4873](#)]
- o LSP_ATTRIBUTES Object [[RFC5420](#)]
- o LSP_REQUIRED_ATTRIBUTES Object [[RFC5420](#)]
- o PROTECTION Object [[RFC3473](#)][[RFC4872](#)][[RFC4873](#)]
- o ASSOCIATION Object [[RFC4872](#)]

- o PRIMARY_PATH_ROUTE Object [[RFC4872](#)]
- o ADMIN_STATUS Object [[RFC3473](#)]
- o LABEL_REQUEST Object [[RFC3209](#)][RFC3473]

For the MPLS TE LSP Objects listed above, the corresponding sub-objects are also applicable to this mechanism. Note that this list is not exhaustive, other MPLS TE LSP objects which reflect specific characteristics of the MPLS TE LSP can also be carried in the LSP state TLV.

2.3.2. PCE Objects

PCE objects are encoded in the "Value" field of the MPLS TE LSP State TLV and consists of PCE objects defined in [[RFC5440](#)]. Rather than replicating all MPLS TE LSP related objects in this document, the semantics and encodings of the MPLS TE LSP objects are re-used. These MPLS TE LSP objects are carried in the LSP State TLV.

When carrying PCE objects, the "Protocol-Origin" field is set to "PCE" (suggested value 2, to be assigned by IANA).

The following PCE Objects are defined:

- o METRIC Object [[RFC5440](#)]
- o BANDWIDTH Object [[RFC5440](#)]

For the MPLS TE LSP Objects listed above, the corresponding sub-objects are also applicable to this mechanism. Note that this list is not exhaustive, other MPLS TE LSP objects which reflect specific characteristics of the MPLS TE LSP can also be carried in the LSP state TLV.

2.3.3. SR TE Policy Sub-TLVs

Segment Routing Traffic Engineering Policy (SR TE Policy) as described in [[I-D.previdi-idr-segment-routing-te-policy](#)] makes use of the Tunnel Encapsulation Attribute defined in [[I-D.ietf-idr-tunnel-encaps](#)] and defines following sub-TLVs:

- o Preference
- o Binding SID
- o Weight

- o Segment List
- o Segment

The equivalent sub-TLVs are defined hereafter and carried in the TE Policy State TLV. When carrying SR TE Policy objects, the "Protocol-Origin" field is set to "BGP SR TE Policy" (suggested value 3, to be assigned by IANA).

2.3.3.1. Preference Object

The Preference sub-TLV 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 |   Flags   | RESERVED |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Preference (4 octets)
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

All fields, including type and length, are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.2. SR TE Binding SID Sub-TLV

The Binding SID sub-TLV 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 |   Flags   | RESERVED |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Binding SID (variable, optional)
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

All fields, including type and length, are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

[I-D.previdi-idr-segment-routing-te-policy] specifies the Binding SID sub-TLV which carries an indication of which value to allocate as Binding SID to the SR TE Policy. In the context of the BGP-LS extensions defined in this document, the Binding SID sub-TLV to the receiver of the , the Binding SID TLThe Binding SID sub-TLV contains the Binding SID the originator of the BGP-LS update has allocated to the corresponding SR TE Policy.

In the context of BGP-LS, the Binding SID sub-TLV defined in this document, contains the effective value of the Binding SID that the router allocated to the SR TE Policy. The router is the SR TE Policy receiver (as described in [\[I-D.previdi-idr-segment-routing-te-policy\]](#)) and it is also the originator of the corresponding BGP-LS update with the extensions defined in this document.

2.3.3.3. Weight Sub-TLV

The Weight sub-TLV 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      |      Flags      |  RESERVED  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                                           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

All fields, including type and length, are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.4. Segment List Sub-TLV

The Segment List object contains sub-TLVs (which in fact are sub-sub-TLVs) and has 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      |  RESERVED  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     sub-TLVs                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- o All fields, including type and length, are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).
- o Length is the total length (not including the Type and Length fields) of the sub-TLVs encoded within the Segment List sub-TLV.
- o sub-objects:
 - * An optional single Weight sub-TLV.
 - * One or more Segment sub-TLVs.

The Segment List sub-TLV is mandatory.

Multiple occurrences of the Segment List sub-TLV MAY appear in the SR TE Policy.

2.3.3.5. Segment Sub-TLV

The Segment sub-TLV describes a single segment in a segment list (i.e.: a single element of the explicit path). Multiple Segment sub-TLVs constitute an explicit path of the SR TE Policy.

[I-D.previdi-idr-segment-routing-te-policy] defines 8 different types of Segment Sub-TLVs:

- Type 1: SID only, in the form of MPLS Label
- Type 2: SID only, in the form of IPv6 address
- Type 3: IPv4 Node Address with optional SID
- Type 4: IPv6 Node Address with optional SID
- Type 5: IPv4 Address + index with optional SID
- Type 6: IPv4 Local and Remote addresses with optional SID
- Type 7: IPv6 Address + index with optional SID
- Type 8: IPv6 Local and Remote addresses with optional SID

2.3.3.5.1. Type 1: SID only, in the form of MPLS Label

The Type-1 Segment Sub-TLV encodes a single SID in the form of an MPLS label. The format is as follows:

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 Flags RESERVED																																							
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+									
Label TC S TTL																																							
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+									

Type, Length and values are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.2. Type 2: SID only, in the form of IPv6 address

The Type-2 Segment Sub-TLV encodes a single SID in the form of an IPv6 address. The format is as follows:


```

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   |   Flags   |   RESERVED   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                               IPv6 SID (16 octets)                               //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in
[\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.3. Type 3: IPv4 Node Address with optional SID

The Type-3 Segment Sub-TLV encodes an IPv4 node address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```

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   |   Flags   |   RESERVED   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               IPv4 Node Address (4 octets)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                               SID (optional, 4 or 16 octets)                               //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in
[\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.4. Type 4: IPv6 Node Address with optional SID

The Type-4 Segment Sub-TLV encodes an IPv6 node address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```

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   |   Flags   |   RESERVED   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                               IPv6 Node Address (16 octets)                               //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                               SID (optional, 4 or 16 octets)                               //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in
[\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.5. Type 5: IPv4 Address + index with optional SID

The Type-5 Segment Sub-TLV encodes an IPv4 node address, an interface index and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```

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 |   Flags   | RESERVED |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.6. Type 6: IPv4 Local and Remote addresses with optional SID

The Type-6 Segment Sub-TLV encodes an IPv4 node address, an adjacency local address, an adjacency remote address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```

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 |   Flags   | RESERVED |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in [\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.7. Type 7: IPv6 Address + index with optional SID

The Type-7 Segment Sub-TLV encodes an IPv6 node address, an interface index and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:


```

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   |   Flags   |   RESERVED   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     IfIndex (4 octets)                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     IPv6 Node Address (16 octets)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     SID (optional, 4 or 16 octets)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in
[\[I-D.previdi-idr-segment-routing-te-policy\]](#).

2.3.3.5.8. Type 8: IPv6 Local and Remote addresses with optional SID

The Type-8 Segment Sub-TLV encodes an IPv6 node address, an adjacency local address, an adjacency remote address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```

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   |   Flags   |   RESERVED   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     Local IPv6 Address (16 octets)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     Remote IPv6 Address (16 octets)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
//                                     SID (4 or 16 octets)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type, Length and values are defined in
[\[I-D.previdi-idr-segment-routing-te-policy\]](#).

3. Operational Considerations

The Existing BGP operational procedures apply to this document. No new operation procedures are defined in this document. The operational considerations as specified in [\[RFC7752\]](#) apply to this document.

In general, it is assumed that the TE Policy head-end nodes are responsible for the distribution of TE Policy state information, while other nodes, e.g. the nodes in the path of a policy, MAY report the TE Policy information (if available) when needed. For example,

the border routers in the inter-domain case will also distribute LSP state information since the ingress node may not have the complete information for the end-to-end path.

4. IANA Considerations

This document requires new IANA assigned codepoints.

4.1. BGP-LS NLRI-Types

IANA maintains a registry called "Border Gateway Protocol - Link State (BGP-LS) Parameters" with a sub-registry called "BGP-LS NLRI-Types".

The following codepoints is suggested (to be assigned by IANA):

Type	NLRI Type	Reference
5	TE Policy NLRI type	this document

4.2. BGP-LS Protocol-IDs

IANA maintains a registry called "Border Gateway Protocol - Link State (BGP-LS) Parameters" with a sub-registry called "BGP-LS Protocol-IDs".

The following Protocol-ID codepoints are suggested (to be assigned by IANA):

Protocol-ID	NLRI information source protocol	Reference
8	RSVP-TE	this document
9	Segment Routing	this document

4.3. BGP-LS Descriptors TLVs

IANA maintains a registry called "Border Gateway Protocol - Link State (BGP-LS) Parameters" with a sub-registry called "Node Anchor, Link Descriptor and Link Attribute TLVs".

The following TLV codepoints are suggested (to be assigned by IANA):

TLV Code Point	Description	Value defined in
1158	TE Policy State TLV	this document
267	Tunnel ID TLV	this document
268	LSP ID TLV	this document
269	IPv4/6 Tunnel Head-end address TLV	this document
270	IPv4/6 Tunnel Tail-end address TLV	this document
271	SR TE Policy Identifier TLV	this document

4.4. BGP-LS LSP-State TLV Protocol Origin

This document requests IANA to maintain a new sub-registry under "Border Gateway Protocol - Link State (BGP-LS) Parameters". The new registry is called "Protocol Origin" and contains the codepoints allocated to the "Protocol Origin" field defined in [Section 2.3](#). The registry contains the following codepoints (suggested values, to be assigned by IANA):

Protocol Origin	Description
1	RSVP-TE
2	PCE
3	BGP SR TE Policy

5. Security Considerations

Procedures and protocol extensions defined in this document do not affect the BGP security model. See [[RFC6952](#)] for details.

6. Acknowledgements

The authors would like to thank Dhruv Dhody, Mohammed Abdul Aziz Khalid, Lou Berger, Acee Lindem, Siva Sivabalan, Arjun Sreekantiah, Dhanendra Jain and Clarence Filsfils for their review and valuable comments.

7. References

7.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", [RFC 2205](#), DOI 10.17487/RFC2205, September 1997, <<http://www.rfc-editor.org/info/rfc2205>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#), DOI 10.17487/RFC3473, January 2003, <<http://www.rfc-editor.org/info/rfc3473>>.
- [RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", [RFC 4090](#), DOI 10.17487/RFC4090, May 2005, <<http://www.rfc-editor.org/info/rfc4090>>.
- [RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", [RFC 4760](#), DOI 10.17487/RFC4760, January 2007, <<http://www.rfc-editor.org/info/rfc4760>>.
- [RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", [RFC 4872](#), DOI 10.17487/RFC4872, May 2007, <<http://www.rfc-editor.org/info/rfc4872>>.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", [RFC 4873](#), DOI 10.17487/RFC4873, May 2007, <<http://www.rfc-editor.org/info/rfc4873>>.
- [RFC4874] Lee, CY., Farrel, A., and S. De Cnodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", [RFC 4874](#), DOI 10.17487/RFC4874, April 2007, <<http://www.rfc-editor.org/info/rfc4874>>.

- [RFC5420] Farrel, A., Ed., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", [RFC 5420](#), DOI 10.17487/RFC5420, February 2009, <<http://www.rfc-editor.org/info/rfc5420>>.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", [RFC 5440](#), DOI 10.17487/RFC5440, March 2009, <<http://www.rfc-editor.org/info/rfc5440>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", [RFC 7752](#), DOI 10.17487/RFC7752, March 2016, <<http://www.rfc-editor.org/info/rfc7752>>.

7.2. Informative References

- [I-D.ietf-idr-tunnel-encaps]
Rosen, E., Patel, K., and G. Velde, "The BGP Tunnel Encapsulation Attribute", [draft-ietf-idr-tunnel-encaps-03](#) (work in progress), November 2016.
- [I-D.ietf-pce-stateful-pce]
Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", [draft-ietf-pce-stateful-pce-18](#) (work in progress), December 2016.
- [I-D.previdi-idr-segment-routing-te-policy]
Previdi, S., Filsfils, C., Sreekantiah, A., Sivabalan, S., Mattes, P., Rosen, E., and S. Lin, "Advertising Segment Routing Traffic Engineering Policies in BGP", [draft-previdi-idr-segment-routing-te-policy-03](#) (work in progress), December 2016.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", [RFC 4655](#), DOI 10.17487/RFC4655, August 2006, <<http://www.rfc-editor.org/info/rfc4655>>.
- [RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", [RFC 6952](#), DOI 10.17487/RFC6952, May 2013, <<http://www.rfc-editor.org/info/rfc6952>>.

Authors' Addresses

Stefano Previdi (editor)
Cisco Systems, Inc.
Via Del Serafico, 200
Rome 00142
Italy

Email: sprevidi@cisco.com

Jie Dong (editor)
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Mach(Guoyi) Chen
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: mach.chen@huawei.com

Hannes Gredler
RtBrick Inc.

Email: hannes@rtbrick.com

Jeff Tantsura
Individual

Email: jefftant@gmail.com

