Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 10 September 2023

S. Previdi

K. Talaulikar, Ed.
Cisco Systems
J. Dong
Huawei Technologies
H. Gredler
RtBrick Inc.
J. Tantsura
Microsoft
9 March 2023

Advertisement of Traffic Engineering Paths using BGP Link-State
draft-ietf-idr-bgp-ls-te-path-00

Abstract

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

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In many network environments, traffic engineering (TE) paths are instantiated into various forms:

* MPLS Traffic Engineering Label Switched Paths (TE-LSPs).

* Local MPLS cross-connect configuration

All this information can be grouped into the same term: TE Paths. In the rest of this document we refer to TE Paths as the set of information related to the various instantiation of policies: MPLS TE LSPs, Local MPLS cross-connects, etc.

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

In many network environments, the configuration, and state of each TE Path 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) [RFC8231], which could provide benefits in path optimization. While some extensions are proposed in the 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 a TE Path state collection mechanism complementary to the mechanism defined in [RFC8231].

```
                                    -----------
                                   |   -----   |
               Service             |  | TED |<-+------------>
               Request             |   -----   |  TED synchronization
                  |                |     |     |  mechanism (e.g.,
                  v                |     |     |  routing protocol)
            ------------- Request/ |     v     |
           |             | Response|   -----   |
           |     NMS     |<--------+> | PCE |  |
           |             |         |   -----   |
            -------------           -----------
          Service |
          Request |
                  v
             ----------  Signaling   ----------
            | Head-End | Protocol   | Adjacent |
            |  Node    |<---------->|   Node   |
             ----------              ----------
```

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 [RFC8231] to report the TE Path information to the PCE nodes. An external component may also need to collect the TE Path 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 Paths in its domain, in addition, a parent PCE needs to collect TE Path 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 Path state information is quite important for making appropriate service placement decisions with the purpose of both meeting the application's requirements and utilizing network resources efficiently.

The Network Management System (NMS) may need to provide global visibility of the TE Paths 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]. BGP-LS is extended to carry TE Path information via #draft-ietf-idr-bgp-ls-sr-policy# so that the same protocol may be used to also collect Segment

Routing traffic engineering paths information such that external components like controllers can use the same protocol for network information collection. This document specifies similar extensions to BGP-LS for the advertisement of information other TE Paths to external components.

1.1. Requirements Language

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.

2. Carrying TE Policy Information in BGP

The "Link-State NLRI" defined in [RFC7752] is extended to carry the TE Path information. New TLVs carried in the Link_State Attribute defined in [RFC7752] are also defined to carry the attributes of a TE Path in the subsequent sections.

The format of "Link-State NLRI" is defined in [RFC7752] 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            NLRI Type          |     Total NLRI Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                  Link-State NLRI (variable)                 //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Additional "NLRI Types" are defined for TE Path Information as following:

* MPLS-TE LSP NLRI (value TBD)

* MPLS Local Cross-connect NLRI (value TBD)

The common format for these NLRI types is defined in Section 3 below.

3. TE Path NLRI Types

This document defines TE Path NLRI Types with their common format as 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)                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //            Node Descriptor TLV (for the Headend)            //
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //                 TE Path Descriptors (variable)              //
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

* Protocol-ID field specifies the component that owns the TE Path state in the advertising node. The existing protocol-id value of 5 for Static Configuration applies for some of the NLRI types and the "RSVP-TE" Protocol-ID (value 8) is defined for some of the other types in this document.

* "Identifier" is an 8 octet value as defined in [RFC7752].

* "Local Node Descriptor" (TLV 256) as defined in [RFC7752] that describes the headend node.

* "TE Path Descriptors" consists of one or more of the TLVs listed as below for use with the respective NLRI type advertisements as specified in Section 4:

Codepoint	Descriptor TLVs
550	Tunnel ID
551	LSP ID
552	IPv4/6 Tunnel Head-end address
553	IPv4/6 Tunnel Tail-end address
555	Local MPLS Cross Connect

The Local Node Descriptor TLV MUST include the following Node Descriptor TLVs:

* BGP Router-ID (TLV 516) [RFC9086], which contains a valid BGP Identifier of the node originating the TE Path advertisement.

* Autonomous System Number (TLV 512) [RFC7752], which contains the ASN or AS Confederation Identifier (ASN) [RFC5065], if confederations are used, of the node originating the TE Path advertisement.

The Local Node Descriptor TLV SHOULD include at least one of the following Node Descriptor TLVs:

* IPv4 Router-ID of Local Node (TLV 1028) [RFC7752], which contains the IPv4 TE Router-ID of the local node when one is provisioned.

* IPv6 Router-ID of Local Node (TLV 1029) [RFC7752], which contains the IPv6 TE Router-ID of the local node when one is provisioned.

The Local Node Descriptor TLV MAY include the following Node Descriptor TLVs:

* BGP Confederation Member (TLV 517) [RFC9086], which contains the ASN of the confederation member (i.e. Member-AS Number), if BGP confederations are used, of the local node.

* Node Descriptors as defined in [RFC7752].

4. TE Path Descriptors

This section defines the TE Path Descriptors TLVs which are used to describe the TE Path being advertised by using the NLRI types defined in Section 3.

4.1. Tunnel Identifier

The Tunnel Identifier TLV contains the Tunnel ID defined in [RFC3209] and is used with the Protocol-ID set to RSVP-TE to advertise the MPLS-TE LSP NLRI Type. It is a mandatory TE Path Descriptor TLV for MPLS-TE LSP NLRI type. It 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:

* Type: 550

* Length: 2 octets.

* Tunnel ID: 2 octets as defined in [RFC3209].

4.2. LSP Identifier

The LSP Identifier TLV contains the LSP ID defined in [RFC3209] and is used with the Protocol-ID set to RSVP-TE to advertise the MPLS-TE LSP NLRI Type. It is a mandatory TE Path Descriptor TLV for MPLS-TE LSP NLRI type. It 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:

* Type: 551

* Length: 2 octets.

* LSP ID: 2 octets as defined in [RFC3209].

4.3. IPv4/IPv6 Tunnel Head-End Address

The IPv4/IPv6 Tunnel Head-End Address TLV contains the Tunnel Head-End Address defined in [RFC3209] and is used with the Protocol-ID set to RSVP-TE to advertise the MPLS-TE LSP NLRI Type. It is a mandatory TE Path Descriptor TLV for MPLS-TE LSP NLRI type. It 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               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//        IPv4/IPv6 Tunnel Head-End Address (variable)         //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

* Type: 552

* 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).

4.4. IPv4/IPv6 Tunnel Tail-End Address

The IPv4/IPv6 Tunnel Tail-End Address TLV contains the Tunnel Tail-End Address defined in [RFC3209] and is used with the Protocol-ID set to RSVP-TE to advertise the MPLS-TE LSP NLRI Type. It is a mandatory TE Path Descriptor TLV for MPLS-TE LSP NLRI type. It 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//        IPv4/IPv6 Tunnel Tail-End Address (variable)         //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

* Type: 553

* 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).

4.5. Local MPLS Cross Connect

The Local MPLS Cross Connect TLV identifies a local MPLS state in the form of an incoming label and interface followed by an outgoing label and interface. The outgoing interface may appear multiple times (for multicast states). It is used with Protocol ID set to "Static Configuration" value 5 as defined in [RFC7752]. It is a mandatory TE Path Descriptor TLV for MPLS Local Cross-connect NLRI type.

The Local MPLS 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Incoming label (4 octets)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Outgoing label (4 octets)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                           Sub-TLVs (variable)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

* Type: 555

* Length: variable.

* Incoming and Outgoing labels: 4 octets each.

* Sub-TLVs: following Sub-TLVs are defined:

  - Interface Sub-TLV

  - Forwarding Equivalent Class (FEC)

The Local 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 an FEC Sub-TLV.

The following sub-TLVs are defined for the Local MPLS Cross Connect TLV:

Codepoint	Descriptor TLV
556	MPLS Cross Connect Interface
557	MPLS Cross Connect FEC

These are defined in the following sub-sections.

4.5.1. MPLS Cross Connect Interface

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

The MPLS Cross Connect 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     |
+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Local Interface Identifier (4 octets)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//         Interface Address (4 or 16 octets)                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

* Type: 556

* Length: 9 or 21.

* 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.

* Local Interface Identifier: a 4-octet identifier.

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

4.5.2. MPLS Cross Connect FEC

The MPLS Cross Connect FEC sub-TLV is optional and contains the FEC associated with the incoming label.

The MPLS Cross Connect 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:

* Type: 557

* Length: variable.

* 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.

* Mask Length: 1 octet of prefix length.

* Prefix: an IPv4 or IPv6 prefix whose mask length is given by the "Mask Length" field padded to an octet boundary.

5. MPLS-TE Path State TLV

A new TLV called "MPLS-TE Path State TLV", is used to describe the characteristics of the MPLS-TE LSP NLRI type and it is carried in the optional non-transitive BGP Attribute "LINK_STATE Attribute" defined in [RFC7752]. These MPLS-TE LSP characteristics include the characteristics and attributes of the LSP, its dataplane, explicit path, Quality of Service (QoS) parameters, route information, the protection mechanisms, etc.

The MPLS-TE Path State 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Object-origin | Address Family|            RESERVED           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//          MPLS-TE Path State Objects (variable)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

where:
```

Figure 1: MPLS-TE Path State TLV

* Type: 1200

* Length: the total length of the MPLS-TE Path State TLV not including the Type and Length fields.

* Object-origin: identifies the component (or protocol) from which the contained object originated. This allows for objects defined in different components to be collected while avoiding the possible codepoint collisions among these components. The following object-origin codepoints are defined in this document.

Code Point	Object Origin
1	RSVP-TE
2	PCEP
3	Local/Static

* Address Family: describes the address family used to set up the MPLS-TE path. The following address family values are defined in this document:

Code Point	Dataplane
1	MPLS-IPv4
2	MPLS-IPv6

* RESERVED: 16-bit field. SHOULD be set to 0 on transmission and MUST be ignored on receipt.

* TE Path State Objects: Rather than replicating all these objects in this document, the semantics and encodings of the objects as defined in RSVP-TE and PCEP are reused.

The state information is carried in the "MPLS-TE Path State Objects" with the following format as described in the sub-sections below.

5.1. RSVP Objects

RSVP-TE objects are encoded in the "MPLS-TE Path State Objects" field of the MPLS-TE Path 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 MPLS-TE Path State TLV.

When carrying RSVP-TE objects, the "Object-Origin" field is set to "RSVP-TE".

The following RSVP-TE Objects are defined:

* SENDER_TSPEC and FLOW_SPEC [RFC2205]

* SESSION_ATTRIBUTE [RFC3209]

* EXPLICIT_ROUTE Object (ERO) [RFC3209]

* ROUTE_RECORD Object (RRO) [RFC3209]

* FAST_REROUTE Object [RFC4090]

* DETOUR Object [RFC4090]

* EXCLUDE_ROUTE Object (XRO) [RFC4874]

* SECONDARY_EXPLICIT_ROUTE Object (SERO) [RFC4873]

* SECONDARY_RECORD_ROUTE (SRRO) [RFC4873]

* LSP_ATTRIBUTES Object [RFC5420]

* LSP_REQUIRED_ATTRIBUTES Object [RFC5420]

* PROTECTION Object [RFC3473][RFC4872][RFC4873]

* ASSOCIATION Object [RFC4872]

* PRIMARY_PATH_ROUTE Object [RFC4872]

* ADMIN_STATUS Object [RFC3473]

* 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.

5.2. PCEP Objects

PCEP objects are encoded in the "MPLS-TE Path State Objects" field of the MPLS-TE Path State TLV and consist of PCEP 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 MPLS-TE Path State TLV.

When carrying PCEP objects, the "Object-Origin" field is set to "PCEP".

The following PCEP Objects are defined:

* METRIC Object [RFC5440]

* 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 TE Path State TLV.

6. Procedures

The BGP-LS advertisements for the TE Path NLRI types are originated by the headend node for the TE Paths that are instantiated on its local node.

For MPLS TE LSPs signaled via RSVP-TE, the NLRI descriptor TLVs as specified in Section 4.1, Section 4.2, Section 4.3, and Section 4.4 are used. Then the TE LSP state is encoded in the BGP-LS Attribute field as MPLS-TE Path State TLV as described in Section 5. The RSVP-TE objects that reflect the state of the LSP are included as defined in Section 5.1. When the TE LSP is setup with the help of PCEP signaling then another MPLS-TE Path State TLV SHOULD be used to encode the related PCEP objects corresponding to the LSP as defined in Section 5.2.

When a SR Policy [RFC9256] is setup with the help of PCEP signaling [RFC8664] then a MPLS-TE Path State TLV MAY be used to encode the related PCEP objects corresponding to the LSP as defined in Section 5.2 specifically to report information and status that is not covered by the SR Policy State TLVs specified in #draft-ietf-idr-bgp-ls-sr-policy#. In the event of a conflict of information, the receiver MUST prefer the information originated via the SR Policy State TLVs over the PCEP objects reported via the TE Path State TLV.

7. Manageability Considerations

The Existing BGP operational and management procedures apply to this document. No new procedures are defined in this document. The considerations as specified in [RFC7752] apply to this document.

In general, it is assumed that the TE Path head-end nodes are responsible for the advertisement of TE Path state information, while other nodes, e.g. the nodes in the path of a policy, MAY report the TE Path 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.

8. IANA Considerations

This section describes the code point allocation by IANA for this document.

8.1. BGP-LS NLRI-Types

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

The following table lists the code points pending allocation by IANA:

Type	NLRI Type	Reference
TBD	MPLS-TE LSP NLRI	this document
TBD	MPLS Local Cross-connect NLRI	this document

8.2. BGP-LS Protocol-IDs

IANA maintains a registry called "BGP-LS Protocol-IDs" in the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group.

The following Protocol-ID codepoints have been allocated by IANA:

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

8.3. BGP-LS TLVs

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

The following table lists the status of TLV code points that have been allocated by IANA:

Code Point	Description	Value defined in
550	Tunnel ID	this document
551	LSP ID	this document
552	IPv4/6 Tunnel Head-end address	this document
553	IPv4/6 Tunnel Tail-end address	this document
555	MPLS Local Cross Connect	this document
556	MPLS Cross Connect Interface	this document
557	MPLS Cross Connect FEC	this document
1200	MPLS-TE Path State	this document

8.4. BGP-LS TE State Object Origin

This document requests IANA to maintain a new registry under "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group with the allocation policy of "Expert Review" [RFC8126] using the guidelines for Designated Experts as specified in [RFC9029]. The new registry is called "TE State Path Origin" and contains the codepoints allocated to the "Object Origin" field defined in Section 5. The registry contains the following codepoints, with initial values, to be assigned by IANA with the reference set to this document:

Code Point	Object Origin
0	Reserved (not to be used)
1	RSVP-TE
2	PCEP
3	Local/Static
4-250	Unassigned
251-255	Private Use (not to be assigned by IANA)

8.5. BGP-LS TE State Address Family

This document requests IANA to maintain a new registry under "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group with the allocation policy of "Expert Review" [RFC8126] using the guidelines for Designated Experts as specified in [RFC9029]. The new registry is called "TE State Address Family" and contains the codepoints allocated to the "Address Family" field defined in Section 5. The registry contains the following codepoints, with initial values, to be assigned by IANA with the reference set to this document:

Code Point	Address Family
0	Reserved (not to be used)
1	MPLS-IPv4
2	MPLS-IPv6
3-250	Unassigned
251-255	Private Use (not to be assigned by IANA)

9. Security Considerations

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

10. Contributors

The following people have substantially contributed to the editing of this document:

Clarence Filsfils
Cisco Systems
Email: cfilsfil@cisco.com

Mach (Guoyi) Chen
Huawei Technologies
Email: mach.chen@huawei.com

11. Acknowledgements

The authors would like to thank Dhruv Dhody, Mohammed Abdul Aziz Khalid, Lou Berger, Acee Lindem, Siva Sivabalan, Arjun Sreekantiah, Dhanendra Jain, Francois Clad, Zafar Ali, Stephane Litkowski, and Aravind Babu Mahendra Babu for their review and valuable comments.

12. References

12.1. 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>.

[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, <https://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, <https://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, <https://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, <https://www.rfc-editor.org/info/rfc4090>.

[RFC4872] Lang, J.P., 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, <https://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, <https://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, <https://www.rfc-editor.org/info/rfc4874>.

[RFC5420] Farrel, A., Ed., Papadimitriou, D., Vasseur, JP., and A. Ayyangar, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, DOI 10.17487/RFC5420, February 2009, <https://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, <https://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, <https://www.rfc-editor.org/info/rfc7752>.

[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>.

[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>.

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/info/rfc8664>.

[RFC9029] Farrel, A., "Updates to the Allocation Policy for the Border Gateway Protocol - Link State (BGP-LS) Parameters Registries", RFC 9029, DOI 10.17487/RFC9029, June 2021, <https://www.rfc-editor.org/info/rfc9029>.

[RFC9086] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Patel, K., Ray, S., and J. Dong, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing BGP Egress Peer Engineering", RFC 9086, DOI 10.17487/RFC9086, August 2021, <https://www.rfc-editor.org/info/rfc9086>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

12.2. Informative References

[RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, DOI 10.17487/RFC5065, August 2007, <https://www.rfc-editor.org/info/rfc5065>.

[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, <https://www.rfc-editor.org/info/rfc6952>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/info/rfc8231>.

Authors' Addresses

Stefano Previdi
Email: stefano@previdi.net

Ketan Talaulikar (editor)
Cisco Systems
India
Email: ketant.ietf@gmail.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing
100095
China
Email: jie.dong@huawei.com

Hannes Gredler
RtBrick Inc.
Email: hannes@rtbrick.com

Jeff Tantsura
Microsoft
Email: jefftant.ietf@gmail.com