Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 16, 2014

J. Dong
M. Chen
Huawei Technologies
H. Gredler
Juniper Networks, Inc.
S. Previdi
Cisco Systems, Inc.

July 15, 2013

Distribution of MPLS Traffic Engineering (TE) LSP State using BGP
draft-dong-idr-te-lsp-distribution-03

Abstract

This document describes a mechanism to collect the Traffic Engineering (TE) LSP information using BGP. Such information can be used by external components for path reoptimization, service placement and network visualization.

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

Table of Contents

1. Introduction

In some network environments, the states of established Multi-Protocol Label Switching (MPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) in the network are required by some components external to the network domain. Usually this information is directly maintained by the ingress Label Edge Routers (LERs) of the MPLS TE LSPs.

One example of using the LSP information is 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-LSP may not implement the PCEP protocol. In this case some general mechanism to collect the TE-LSP states from the ingress LERs is needed. This document

proposes an LSP state collection mechanism complementary to the mechanism defined in [I-D.ietf-pce-stateful-pce].

```
                                    -----------
                                   |   -----   |
               Service             |  | TED |<-+----------->
               Request             |   -----   |  TED synchronization
                  |                |     |     |  mechanism (for example,
                  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], the 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 LSP information to the PCE nodes. An external component may further need to collect the LSP information from all the PCEs in the network to get a global view of the LSP states in the network.

In some networks, a centralized controller is used for service placement. Obtaining the TE LSP state information is quite important for making appropriate service placement decisions with the purpose of both meeting the application's requirements and utilizing the network resource efficiently.

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

BGP has been extended to distribute link-state and traffic engineering information and share with some external components [I-D.ietf-idr-ls-distribution]. Using the same protocol to collect other network layer information would be desired by the external components, which avoids introducing multiple protocols for network

information collection. This document describes a mechanism to distribute the TE LSP information to external components using BGP.

2. Carrying LSP State Information in BGP

2.1. LSP Identifier Information

The TE LSP Identifier information is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. The "Link State NLRI" defined in [I-D.ietf-idr-ls-distribution] is extended to carry the TE LSP Identifier information. BGP speakers that wish to exchange TE LSP 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 [I-D.ietf-idr-ls-distribution]. Two new "NLRI Type" are defined for TE LSP Identifier Information as following:

o NLRI Type = 5: IPv4 TE LSP NLRI

o NLRI-Type = 6: IPv6 TE LSP NLRI

The IPv4 TE LSP NLRI (NLRI Type = 5) 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)                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    IPv4 Tunnel Sender Address                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Tunnel ID        |             LSP ID            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   IPv4 Tunnel End-point Address               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2. IPv4 TE LSP NLRI

The IPv6 TE LSP NLRI (NLRI Type = 6) 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)                          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   +                                                               +
   |                  IPv6 Tunnel Sender Address                   |
   +                          (16 octets)                          +
   |                                                               |
   +                                                               +
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |          Tunnel ID            |             LSP ID            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   +                                                               +
   |                 IPv6 Tunnel End-point Address                 |
   +                          (16 octets)                          +
   |                                                               |
   +                                                               +
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3. IPv6 TE LSP NLRI

For IPv4 TE LSP NLRI and IPv6 TE LSP NLRI, the Protocol-ID field is set to 6, which indicates that the NLRI information has been sourced by RSVP-TE.

The Identifier field is used to discriminate between instances with different LSP technology - e.g. one identifier can identify the instance for packet path, and another one is to identify the instance of optical path.

The other fields in the IPv4 TE LSP NLRI and IPv6 TE LSP NLRI are the same as specified in [RFC3209].

2.2. LSP State Information

The LSP State TLV is used to describe the characteristics of the TE LSPs, which is carried in the optional non-transitive BGP Attribute "LINK_STATE Attribute" defined in [I-D.ietf-idr-ls-distribution].

The "Value" field of the LSP State TLV corresponds to the format and semantics of a set of objects defined in [RFC3209], [RFC3473] and [RFC5440] for TE LSPs. Rather than replicating all RSVP-TE related objects in this document the semantics and encodings of existing RSVP-TE objects are re-used. Hence all RSVP-TE LSP objects are regarded as sub-TLVs. The LSP State TLV SHOULD only be used with IPv4/IPv6 TE LSP NLRI.

```
 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 LSP Objects (variable)                   ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4. LSP State TLV

Currently the TE LSP Objects that can be carried in the LSP State TLV include:

- o LSP Attributes (LSPA) Object [RFC5440]
- o Explicit Route Object (ERO) [RFC3209]
- o Record Route Object (RRO) [RFC3209]
- o BANDWIDTH Object [RFC5440]
- o METRIC Object [RFC5440]
- o Protection Object [RFC3473]
- o Admin_Status Object [RFC3473]

Other TE LSP objects may also be carried in LSP state TLV, which is for further study.

3. IANA Considerations

IANA needs to assign one new TLV type for "LSP State TLV" from the TLV registry of Link_State Attribute.

IANA needs to assign one Protocol-ID for 'RSVP-TE' from the BGP-TE/LS registry of Protocol-IDs.

4. Security Considerations

TBD

5. References

5.1. Normative References

[I-D.ietf-idr-ls-distribution]
Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", draft-ietf-idr-ls-distribution-03 (work in progress), May 2013.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

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

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

5.2. Informative References

[I-D.ietf-pce-stateful-pce]
Crabbe, E., Medved, J., Minei, I., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-05 (work in progress), July 2013.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Building, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

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

Email: mach.chen@huawei.com

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

Email: hannes@juniper.net

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

Email: sprevidi@cisco.com