PCE Working Group   S. Sivabalan
Internet-Draft   C. Filsfils
Intended status: Standards Track   S. Previdi
Expires: April 16, 2017   Cisco Systems, Inc.
                                                          J. Tantsura
                                                             Ericsson
                                                          J. Hardwick
                                                 Metaswitch Networks
                                                           M. Nanduri
                                               Microsoft Corporation
                                                    October 13, 2016

        Carrying Binding Label/Segment-ID in PCE-based Networks.
            draft-sivabalan-pce-binding-label-sid-02.txt

Abstract

   It is possible to associate a binding label to RSVP-TE signaled
   Traffic Engineering Label Switching Path or binding Segment-ID (SID)
   to Segment Routed Traffic Engineering path. Such a binding label/SID
   can be used by an upstream node for steering traffic into the
   appropriate TE path to enforce TE policies. This document proposes
   an approach for reporting binding label/SID to Path Computation
   Element (PCE) for supporting PCE-based Traffic Engineering policies.

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 [RFC2119].

Table of Contents

1. Introduction

   A PCE can compute Traffic Engineering paths (TE paths) through a
   network that are subject to various constraints. Currently, TE paths
   are either set up using the RSVP-TE signaling protocol or Segment
   Routed (SR). We refer to such paths as RSVP-TE paths and SR-TE paths
   respectively in this document.

   Similar to assigning label to a Forwarding Equivalence Class (FEC)
   via Label Distribution Protocol (LDP), a binding label can be
   assigned to a RSVP-TE LSP. If the topmost label of an incoming
   packet is the binding label, the packet is steered onto the RSVP-TE
   LSP. As such, any upstream node can use binding labels to steer the
   packets that it originates to appropriate TE LSPs to enforce TE
   policy. Similarly, a binding SID (see
   [I-D.ietf-isis-segment-routing-extensions] and
   [I-D.ietf-ospf-segment-routing-extensions]) can be used to enforce TE
   policy with SR-TE path. Note that if an SR-TE path is represented as
   a forwarding-adjacency, then the corresponding adjacency SID can be

used as the binding SID. In such case, the path is advertised using
the routing protocols as described in [RFC5440]. The binding SID
provides an alternate mechanism without additional overhead on
routing protocols.

[RFC5440] describes the Path Computation Element Protocol (PCEP) for
communication between a Path Computation Client (PCC) and a PCE or
between a pair of PCEs.[I-D.ietf-pce-stateful-pce] specifies
extension to PCEP that allows a PCC to delegate its LSPs to a PCE.
The PCE can then update the state of LSPs delegated to it.
[I-D.ietf-pce-pce-initiated-lsp] specifies a mechanism allowing a PCE
to dynamically instantiate an LSP on a PCC by sending the path and
characteristics of the LSP. The PCEP extension to setup and maintain
SR-TE paths is specified in [I-D.ietf-pce-segment-routing].

Binding label/SID has local significance to the ingress node of the
corresponding TE path. When a stateful PCE is deployed for setting
up TE paths, it may be desirable to report the binding label or SID
to the PCE for the purpose of enforcing end-to-end TE policy. A
sample Data Center (DC) use-case is illustrated in the following
diagram. In the MPLS DC network, an SR LSP (without traffic
engineering) is established using a prefix SID advertised by BGP (see
[I-D.keyupate-idr-bgp-prefix-sid]). In IP/MPLS WAN, an SR-TE LSP is
setup using the PCE. The list of SIDs of the SR-TE LSP is {A, B, C,
D}. The gateway node 1 (which is the PCC) allocates a binding SID X
and reports it to the PCE. In order for the access node to steer the
traffic over the SR-TE LSP, the PCE passes the SID stack {Y, X} where
Y is the prefix SID of the gateway node-1 to the access node. In the
absence of the binding SID X, the PCE should pass the SID stack {Y,
A, B, C, D} to the access node. This example also illustrates the
additional benefit of using the binding SID to reduce the number of
SIDs imposed on the access nodes with a limited forwarding capacity.

```
          SID stack
          {Y, X}                    +-----+
   _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _| PCE |
   |                                +-----+
   |                                  ^
   |                                  | Binding
   |            .------.              | SID (X)        .------.
   |          (        )             |              (          )
   V       .--(          )--.        |        .--(            )--.
+------+  (                  ) +-------+  (                      ) +-------+
|Access|_(  MPLS DC Network  )_|Gateway|_(     IP/MPLS WAN     )_|Gateway|
| Node | (  ==============>  ) |Node-1 | ( ================> ) |Node-2 |
+------+  (      SR path     ) +-------+  (     SR-TE path     ) +-------+
           '--(          )--'   Prefix     '--(          )--'
              (        )        SID of        (        )
               '-----'         Node-1          '-----'
                                is Y             SIDs for SR-TE LSP:
                                                 {A, B, C, D}
```

Figure 1: A sample Use-case of Binding SID

   It may be possible for a PCE to request a PCC to allocate a specific
   binding label/SID by sending an update message. If the PCC can
   successfully allocate the specified binding value, it reports the
   binding value to the PCE. Otherwise, the PCC sends an error message
   to the PCE indicating the cause of the failure.

   In this document, we introduce a new OPTIONAL TLV that a PCC can use
   in order to report the binding label/SID associated with a TE LSP, or
   a PCE to request a PCC to allocate a binding label/SID. This TLV is
   intended for TE LSPs established using RSVP-TE, SR, or any other
   future method. Also, in the case of SR-TE LSPs, the TLV can carry an
   MPLS label (for SR-TE path with MPLS data-plane) or a binding SID
   (e.g., IPv6 address for SR-TE paths with IPv6 data-plane). However,
   use of this TLV for carrying non-MPLS binding SID will be described
   in separate document(s). Binding value means either MPLS label or
   SID throughout this document.

2. Terminology

   The following terminologies are used in this document:

   LER: Label Edge Router.
   LSP: Label Switched Path.
   LSR: Label Switching Router.
   PCC: Path Computation Client.

PCE: Path Computation Element
PCEP: Path Computation Element Protocol.
SID: Segment Identifier.
SR: Segment Routing.
TLV: Type, Length, and Value.

3. Path Binding TLV

The new optional TLV is called "TE-PATH-BINDING TLV" whose format is
shown in the diagram below is defined to carry binding label or SID
for a TE path. This TLV is associated with the LSP object specified
in ([I-D.ietf-pce-stateful-pce]). The type of this TLV is to be
allocated by IANA.

```
    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            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |       Binding Type (BT)       |         Binding Value         |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   ~           Binding Value (continued) (variable length)        ~
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
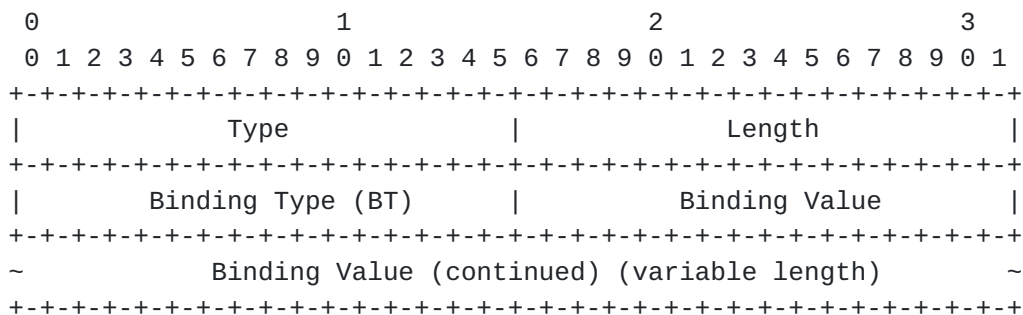
Figure 2: TE-PATH-BINDING TLV

TE-PATH-BINDING TLV is a generic TLV such that it is able to carry
MPLS label binding as well as other types of future bindings (e.g.,
IPv6 SR path). It is formatted according to the rules specified in
[RFC5440]. The two octet Binding Type (BT) field identifies the type
of binding included in the TLV. This document specifies the
following BT values:

o BT = 0: The binding value is an MPLS label carried in the format
   specified in [RFC5462] where only the label value is valid, and
   other fields (TC, S, and TTL)fields MUST be considered invalid.
   The Length MUST be set to 6.
o BT = 1: Similar to the case where BT is 0 except that all the
   fields on the MPLS label entry are set on transmission. However,
   the receiver MAY choose to override TC, S, and TTL values
   according its local policy.

4. Operation

The binding value is allocated by PCC and reported to PCE via PCRpt
message. If a PCE does not recognize the TE-PATH-BINDING TLV, it
MUST ignore the TLV in accordance with ([RFC5440]). If a PCE

recognizes the TLV but does not support the TLV, it MUST send PCErr
with Error-Type = 2 (Capability not supported).

If a TE-PATH-BINDING TLV is absent in PCRpt message, PCE MUST assume
that the corresponding LSP does not have any binding. If there are
more than one TE-PATH-BINDING TLVs, only the first TLV MUST be
processed and the rest MUST be silently ignored. If PCE recognizes
an invalid binding value (e.g., label value from the reserved label
space when MPLS label binding is used), it MUST send the PCErr
message with Error-Type = 10 ("Reception of an invalid object") and
Error Value = TBD ("Bad label value") as specified in
[I-D.ietf-pce-segment-routing].

If a PCE requires a PCC to allocate a specific binding value, it may
do so by sending a PCUpd message containing a TE-PATH-BINDING TLV.
If the value can be successfully allocated, the PCC reports the
binding value to the PCE. If the PCC considers the binding value
specified by the PCE invalid, it MUST send a PCErr message with
Error-Type = TBD ("Binding SID failure") and Error Value = TBD
("Invalid SID"). If the binding value is valid, but the PCC is
unable to allocate the binding value, it MUST send a PCErr message
with Error-Type = TBD ("Binding label/SID failure") and Error Value =
TBD ("Unable to allocate the specified label/SID").

If a PCC receives TE-PATH-BINDING TLV in any message other than
PCUpd, it MUST close the corresponding PCEP session with the reason
"Reception of a malformed PCEP message" according ([RFC5440]).
Similarly, if a PCE receives a TE-PATH-BINDING TLV in any message
other than a PCRpt or if the TE-PATH-BINDING TLV is associated with
any object other than LSP object, the PCE MUST close the
corresponding PCEP session with the reason "Reception of a malformed
PCEP message" according ([RFC5440]).

If a PCC wishes to withdraw or modify a previously reported binding
value, it MUST send a PCRpt message without any TE-PATH-BINDING TLV
or with the TE-PATH-BINDING TLV containing the new binding value
respectively.

If a PCE wishes to modify a previously requested binding value, it
MUST send a PCUpd message with TE-PATH-BINDING TLV containing the new
binding value. Absense of TE-PATH-BINDING TLV in PCUpd message means
that the PCE does not specify a binding value in which case the
binding value allocation is governed by the PCC's local policy.

If a PCC receives a valid binding value from a PCE which is different
than the current binding value, it MUST try to allocate the new
value. If the new binding value is successfully allocated, the PCC
MUST report the new value to the PCE. Otherwise, it MUST send a

PCErr message with Error-Type = TBD ("Binding label/SID failure") and
Error Value = TBD ("Unable to allocate the specified label/SID").

5. Security Considerations

No additional security measure is required.

6. IANA Considerations

6.1. TE-PATH-BINDING TLV

IANA is requested to allocate a new TLV type (recommended value is
31)for TE-PATH-BINDING TLV specified in this document.

This document requests that a registry is created to manage the value
of the Binding Type field in the TE-PATH-BINDING TLV.

Value	Description	Reference
0	MPLS Label	This document

6.2. PCEP Error Type and Value

IANA is requested to allocate code-points in the PCEP-ERROR Object
Error Types and Values registry for the following new error-values:

```
Error-Type    Meaning
----------    -------
TBD (recommended 22)  Binding SID failure:

              Error-value = TBD (recommended 1):  Invalid SID
              Error-value = TBD (recommended 2):  Unable to allocate
                                                  binding SID
```

7. Acknowledgements

We like to thank Milos Fabian for his valuable comments.

8. Normative References

[I-D.ietf-isis-segment-routing-extensions]
          Previdi, S., Filsfils, C., Bashandy, A., Gredler, H.,
          Litkowski, S., Decraene, B., and j. jefftant@gmail.com,
          "IS-IS Extensions for Segment Routing", draft-ietf-isis-
          segment-routing-extensions-08 (work in progress), October
          2016.

   [I-D.ietf-ospf-segment-routing-extensions]
              Psenak, P., Previdi, S., Filsfils, C., Gredler, H.,
              Shakir, R., Henderickx, W., and J. Tantsura, "OSPF
              Extensions for Segment Routing", draft-ietf-ospf-segment-
              routing-extensions-09 (work in progress), July 2016.

   [I-D.ietf-pce-pce-initiated-lsp]
              Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP
              Extensions for PCE-initiated LSP Setup in a Stateful PCE
              Model", draft-ietf-pce-pce-initiated-lsp-07 (work in
              progress), July 2016.

   [I-D.ietf-pce-segment-routing]
              Sivabalan, S., Medved, J., Filsfils, C., Crabbe, E.,
              Raszuk, R., Lopez, V., Tantsura, J., Henderickx, W., and
              J. Hardwick, "PCEP Extensions for Segment Routing", draft-
              ietf-pce-segment-routing-08 (work in progress), October
              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-16 (work in progress), September 2016.

   [I-D.keyupate-idr-bgp-prefix-sid]
              Patel, K., Previdi, S., Filsfils, C., Sreekantiah, A.,
              Ray, S., and H. Gredler, "Segment Routing Prefix SID
              extensions for BGP", draft-keyupate-idr-bgp-prefix-sid-05
              (work in progress), July 2015.

   [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>.

   [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP)
              Hierarchy with Generalized Multi-Protocol Label Switching
              (GMPLS) Traffic Engineering (TE)", RFC 4206,
              DOI 10.17487/RFC4206, October 2005,
              <http://www.rfc-editor.org/info/rfc4206>.

   [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>.

   [RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching
              (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic
              Class" Field", RFC 5462, DOI 10.17487/RFC5462, February
              2009, <http://www.rfc-editor.org/info/rfc5462>.

Authors' Addresses

   Siva Sivabalan
   Cisco Systems, Inc.
   2000 Innovation Drive
   Kanata, Ontario K2K 3E8
   Canada

   Email: msiva@cisco.com

   Clarence Filsfils
   Cisco Systems, Inc.
   Pegasus Parc
   De kleetlaan 6a, DIEGEM BRABANT 1831
   BELGIUM

   Email: cfilsfil@cisco.com

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

   Email: sprevidi@cisco.com

   Jeff Tantsura
   Ericsson
   300 Holger Way
   San Jose, CA 95134
   USA

   Email: jeff.tantsura@ericsson.com

Jonathan Hardwick
Metaswitch Networks
100 Church Street
Enfield, Middlesex
UK

Email: Jonathan.Hardwick@metaswitch.com

Mohan Nanduri
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
USA

Email: mnanduri@microsoft.com