ISIS Protocol Extensions for Boundary Node Discovery (BND)
draft-dhody-pce-bn-discovery-isis-05

Abstract

   There are various circumstances where it is highly desirable to be
   able to dynamically and automatically discover a set of Boundary
   Nodes (BN) along with their domain information. For that purpose,
   this document defines extensions to the Intermediate System to
   Intermediate System(IS-IS) routing protocol for the advertisement of
   Boundary Node (BN)Discovery information within an IS-IS area or
   within the entire IS-IS routing domain.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   This document defines extensions to IS-IS [ISO 10589] to allow a
   boundary node in an IS-IS routing domain to advertise its location,
   along with domain information.

   Generic capability advertisement mechanisms for IS-IS are defined in
   [RFC4971]. These allow a router to advertise its capabilities within
   an IS-IS area or an entire IS-IS routing domain. This document
   leverages this generic capability advertisement mechanism to fully
   satisfy the dynamic BN discovery.

   This document defines a new sub-TLV (named the BN Discovery (BND))to
   be carried within the IS-IS Router Capability TLV ([RFC4971]).

   The BN information advertised is detailed in Section 3. Protocol
   extensions and procedures are defined in Section 9 and Section 10.

   A detailed description about the need for auto discovery of Boundary
   Nodes (BN) and thier domains is also provided in this document.

   The IS-IS extensions defined in this document allow for BN discovery
   within an IS-IS routing domain. Boundary Node can be an ABR or ASBR.

   This document defines a set of sub-TLVs that are nested within each
   other. When the degree of nesting TLVs is 2 (a TLV is carried within
   another TLV) the TLV carried within a TLV is called a sub-TLV.
   Strictly speaking, when the degree of nesting is 3, a sub-sub-TLV is
   carried within a sub-TLV that is itself carried within a TLV. For
   the sake of terminology simplicity, a TLV carried within another TLV
   is called a sub-TLV regardless of the degree of nesting.

1.1. 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].

2. Terminology

   The following terminology is used in this document.

   ABR: IS-IS Area Border Router. Routers used to connect two IGP
      areas.

   AS: Autonomous System.

   ASBR: Autonomous System Border Router. Router used to connect
      together ASes of the same or different service providers via one
      or more inter-AS links

   BN: A boundary node is either an ABR in the context of inter-area
      Traffic Engineering or an ASBR in the context of inter-AS Traffic
      Engineering.

   BND: Boundary Node Discovery

   BRPC: Backward Recursive Path Computation

   Domain: Any collection of network elements within a common sphere of
      address management or path computational responsibility. Examples
      of domains include Interior Gateway Protocol (IGP) areas and
      Autonomous Systems (ASs).

   H-PCE: Hierarchical PCE.

   IGP: Interior Gateway Protocol. Either of the two routing
      protocols, Open Shortest Path First (OSPF) or Intermediate System
      to Intermediate System (IS-IS).

   LSA: Link State Advertisement.

   IS-IS: Intermediate System to Intermediate System.

   PCE: Path Computation Element. An entity (component, application,
      or network node) that is capable of computing a network path or
      route based on a network graph and applying computational
      constraints.

   TLV: Type-Length-Variable data encoding.

3. Overview

3.1. Boundary Node (BN) Discovery Information

   The BN discovery information is composed of:

   o The BN location: an IPv4 and/or IPv6 address that is used to reach
      the BN. It is RECOMMENDED to use an address that is always
      reachable from all connected domains;

   o The set of two or more Domain(s) into which the BN has
      connectivity;

Changes in BN discovery information may occur as a result of BN
configuration update or domain status change.

3.2. Flooding Scope

The flooding scope for BN information advertised through IS-IS can be
a single L1 area, an L1 area and the L2 sub-domain, or the entire
IS-IS routing domain.

4. Applications

BRPC procedure as defined in [RFC5441], requires Path Computation
Element (PCE) [RFC4655] to be aware of the BNs for the inter-domain
path computation. This information would be either statically
configured at PCE or learned via some mechanism, as listed in
Section 5.

In case of static configuration, as shown in the Figure 1, incase of
ISIS Backbone area(L2), configuration of BNs at PCE5 is extensive.
BRPC procedure guarantees a best path only if BNs are selected
correctly, any change in BNs at run time may lead to sub-optimal
path. Also Administrator need to configure ABR / ASBR ID in such a
way that it is reachable from all the domains, BND Tlv can take care
of this automatically.

```
   +-------------------------+        +-------------------------+
   |              +------+|        |              +------+|
   |   Area1(L1)   | PCE1 ||        |   Area2(L1)   | PCE2 ||
   |              +--+---+|        |              +--+---+|
   |                 |   |        |                 |   |
   |   +-------------------+----------+-----------------+   |
   |   | +------+    +------+|        |+------+   +------+ |   |
   |   | | BN1  |    | BN2  ||        || BN3  |   | BN4  | |   |
   |   | +--+---+    +--+---+|        |+--+---+   +---+--+ |   |
   +----+----|-----------+----+        +---+----------+----+-----+
        |    |           |   |            |           |   |
        |   +-----------+-------+   +------+----------+   |
        |   |           |   |   |           |
        |   |           |   |   |           |
        |   |           ++----++           |
        |   Backbone(L2)    | PCE5 |           |
        |   |           ++----++           |
        |   |           |   |   |           |
        |   |           |   |   |           |
        |   +-----------+-------+   +------+----------+   |
        |   |           |   |           |   |   |
        |   |           |   |           |   |   |
   +----+----+-----------+----+        +---+----------+-----+-----+
   |   | +--+---+    +--+---+|        |+--+---+   +--+---+ |     |
   |   | | BN5  |    | BN6  ||        || BN7  |   | BN8  | |     |
   |   | +------+    +------+|        |+------+   +------+ |     |
   |   +-------------------+----------+-----------------+     |
   |                 |        |                 |
   |              +------+|        |              +------+|
   |   Area3(L1)   | PCE3 ||        |   Area4(L1)   | PCE4 ||
   |              +--+---+|        |              +--+---+|
   +-------------------------+        +-------------------------+
```

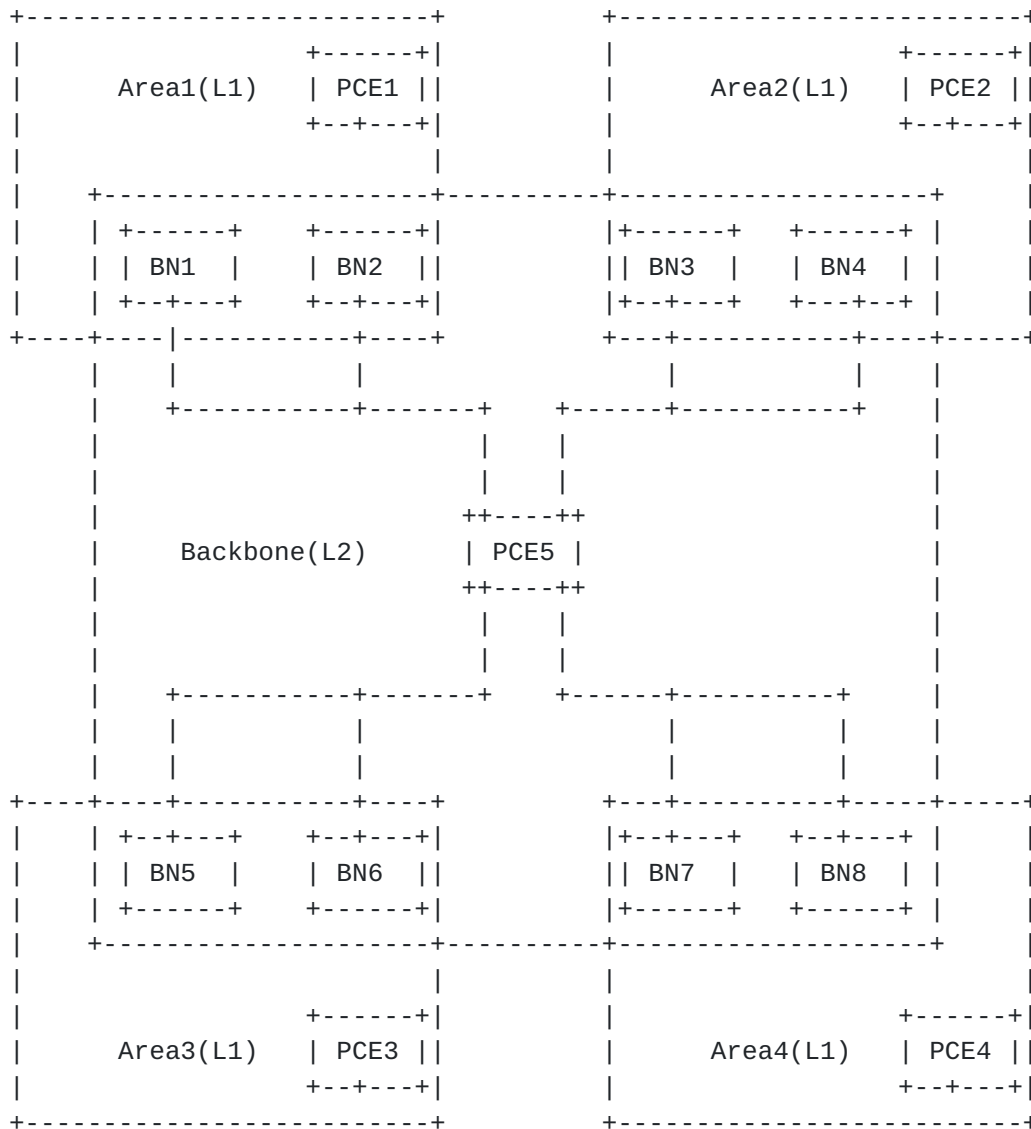   Figure 1: ISIS Area Topology

   The problems with existing mechanism to discover Boundary nodes are
   listed in the next section.

5. Existing Mechanisms

5.1. ISIS

   o As specified in [RFC1142] ABR/ASBR can be identified but not their
      domain information. As stated in Section 7, Selection of correct
      BN is based on domain and thus it is ineffective.

o Selection of ABR/ASBR based on ISIS Database is not a good idea,
   first it requires PCE to look into ISIS Database thus adding to
   coupling, second it MAY require BGP routes to be redistributed
   into ISIS which is also not a good network design principle.

5.2. Inter-AS TE Link

o [RFC5316] specifies how to advertise TE properties of inter-AS
   links; through which ASBR and remote AS can be discovered, but ABR
   and their domain information cannot be discovered via above RFC.

o AS is made up of multiple Area, there maybe a need to clearly
   identify a BN by combination of both AS number and Area-id. Refer
   Section 3.5 of [DOMAIN-SEQ].

5.3. ISIS Area Topology

o To uniquely identify an ISIS L1 area, an unique area Id MUST be
   assigned. There is no other way to learn this information. BND
   can be used to advertise this information.

5.4. HPCE

o Section 4 of [H-PCE] specifies each child PCE should know the
   identity of the domains that neighbor its own domain and
   advertises the same to the parent PCE. No method exist to find
   the neighbor domain which need to be carried in NEIG-PCE-DOMAIN
   Sub-TLV and BN discovery along with neighbor domain information
   can help in generating NEIG-PCE-DOMAIN Sub-TLV.

6. Static Configurations

A simple solution would be to configure BNs [ABR and ASBR] at PCE(s)
along with their domain information. As this information is fairly
static this could work in simple situations. But as PCE are being
used in bigger and multiple domains, any sort of static
configurations would put extra effort on the system administrator.
Selection of correct BNs is the core of the BRPC procedure, we feel
this information should be dynamically learned and maintained.

7. Importance of Domain Information along with BNs

There are methods to learn BNs dynamically from IGP, but the
knowledge of neighboring-domains is not possible to obtain. Without
this the correct BN based on the domain-path can't be selected.
[RFC5441] mentions:

"Note that PCE(i) only considers the entry BNs of domain(i), i.e.,

   only the BNs that provide connectivity from domain(i-1). In other
   words, the set BN-en(k,i) is only made of those BNs that provide
   connectivity from domain (i-1) to domain(i). "

   This selection of correct BNs providing connectivity between correct
   domains cannot be made by the information obtained from IGP. Without
   the correct selection we would not be following [RFC5441].

8. Relationship to Domain-Sequence

   [DOMAIN-SEQ] provides a standard representation of Domain Sequence in
   all deployment scenarios. The Domain Information carried in the BN-
   DOMAIN sub-tlv is same as the sub-objects inside the domain sequence.

9. The IS-IS BND Sub-TLV

   The IS-IS BND sub-TLV contains a non-ordered set of sub-TLVs.

   The format of the IS-IS BND sub-TLV and its sub-TLVs is identical to
   the TLV format used by the Traffic Engineering Extensions to IS-IS
   [RFC5305]. That is, the TLV is comprised of 1 octet for the type, 1
   octet specifying the TLV length, and a value field. The Length field
   defines the length of the value portion in octets.

   The IS-IS BND sub-TLV has the following format:

      TYPE: To be assigned by IANA (suggested value 6)
      LENGTH: Variable
      VALUE: Set of sub-TLVs

   Two sub-TLVs are defined:

        Sub-TLV type   Length   Name
             1   variable   BN-ADDRESS sub-TLV
             2   variable   BN-DOMAIN sub-TLV

   The BN-ADDRESS and BN-DOMAIN sub-TLVs MUST always be present within
   the BND sub-TLV.

   Any unrecognized sub-TLV MUST be silently ignored.

   The BND sub-TLV is carried within an IS-IS CAPABILITY TLV defined in
   [RFC4971].

   The following sub-sections describe the sub-TLVs.

9.1. BN-ADDRESS Sub-TLV

   The BN-ADDRESS sub-TLV specifies an IP address that can be used to
   reach the BN. It is RECOMMENDED to make use of an address that is
   always reachable, provided the BN is alive and reachable.

   The BN-ADDRESS sub-TLV is mandatory; it MUST be present within the
   BND sub-TLV. It MAY appear twice, when the BN has both an IPv4 and
   IPv6 address. It MUST NOT appear more than once for the same address
   type. If it appears more than once for the same address type, only
   the first occurrence is processed and any others MUST be ignored.

   The BN-ADDRESS sub-TLV has the following format:

   TYPE: 1
      LENGTH: 5 for an IPv4 address or 17 for an IPv6 address.
      VALUE: This comprises one octet indicating the address-type and 4
              or 16 octets encoding the IPv4 or IPv6 address to be used
              to reach the BN.

   Address-type:
                   1   IPv4
                   2   IPv6

9.2. BN-DOMAIN Sub-TLV

   The BN-DOMAIN sub-TLV specifies a BN-Domain (area and/or AS) where
   the BN has topology connectivity.

   The BN-DOMAIN sub-TLV is mandatory; it MUST be present within the BND
   TLV.

   A BND sub-TLV MUST include two or more BN-DOMAIN sub-TLVs as the BN
   has connectivity into multiple BN-Domains.

   The BN-DOMAIN sub-TLV has the following format:

   TYPE: 2
      LENGTH: Variable
      VALUE: This is composed of one octet indicating the domain-type
              (area ID or AS Number) and a variable length IS-IS area ID
              or a 32-bit AS number, identifying a BN-Domain where the
              BN has visibility and can compute paths.

   Two domain types are defined:

                   1   Area ID
                   2   AS Number

The Area ID is the area address as defined in [ISO 10589].

When the AS number is coded in two octets, the AS Number field MUST
have its first two octets set to 0.

10. Elements of Procedure

The BND sub-TLV is advertised within an IS-IS Router Capability TLV
defined in [RFC4971]. As such, elements of procedures are inherited
from those defined in [RFC4971].

The flooding scope is controlled by the S flag in the IS-IS Router
Capability TLV (see [RFC4971]). When the scope of the BND sub-TLV is
area local, it MUST be carried within an IS-IS Router Capability TLV
having the S bit cleared. When the scope of the BND sub-TLV is the
entire IS-IS routing domain, it MUST be carried within an IS-IS
Router Capability TLV having the S bit set.

Note that an L1L2 node may include a BND TLV in a Router Capability
TLV with the S bit cleared in both in its L1 and L2 LSPs. This
allows the flooding scope to be restricted to the L1 area and the L2
sub-domain.

When the BN function is deactivated, the IS-IS speaker MUST originate
a new IS-IS LSP that no longer includes the corresponding BND TLV.

The BN address (i.e., the address indicated within the BN-ADDRESS
sub-TLV) SHOULD be reachable via some prefixes advertised by IS-IS.

The BND sub-TLV information regarding a specific BN is only
considered current and useable when the router advertising this
information is itself reachable via IS-IS calculated paths at the
level of the LSP in which the BND sub-TLV appears.

A change in the state of a BN (activate, deactivate, domain change)
MUST result in a corresponding change in the BND sub-TLV information
advertised by an IS-IS router (inserted, removed, updated) in its
LSP. The way BNs determine the information they advertise, and how
that information is made available to IS-IS, is out of the scope of
this document. Some information may be configured and other
information may be automatically determined by ISIS.

A change in information in the BND sub-TLV MUST NOT trigger any SPF
computation at a receiving router.

11. Backward Compatibility

   The BND TLV defined in this document does not introduce any
   interoperability issues.

   An IS-IS router not supporting the BND sub-TLV will just silently
   ignore the sub-TLV as specified in [RFC4971].

12. Impact on Network

   The routers acting as BNs will originate LSP with BND Tlv; As there
   are only few BNs exist in the network, the performance impact in
   flooding is very less.

13. IANA Considerations

   IANA has defined a registry for the sub-TLVs carried in the IS-IS
   Router Capability TLV defined in [RFC4971]. IANA has assigned a new
   sub-TLV codepoint for the BND sub-TLV carried within the Router
   Capability TLV.

   Value   Sub-TLV   References
   -----   --------   ----------
   To be   BND sub-TLV   (this document)
   assigned
   by IANA

14. Security Considerations

   This document defines IS-IS extensions for BN discovery within an
   administrative domain. Hence the security of the BN discovery relies
   on the security of IS-IS.

   Mechanisms defined to ensure authenticity and integrity of IS-IS LSPs
   [RFC5304] and their TLVs, can be used to secure the BND sub-TLV as
   well.

   IS-IS provides no encryption mechanism for protecting the privacy of
   LSPs and, in particular, the privacy of the BN discovery information.

15. Manageability Considerations

   TBD

16. Acknowledgments

We would like to thank Quintin Zhao, Daniel King, Adrian Ferral, Suresh babu, Pradeep Shastry, Saravana Kumar and Srinivasan for their useful comments and suggestions.

17. References

17.1. Normative References

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

[RFC4971]   Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.

[RFC5304]   Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, October 2008.

[RFC5305]   Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

17.2. Informative References

[RFC1142]   Oran, D., "OSI IS-IS Intra-domain Routing Protocol", RFC 1142, February 1990.

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

[RFC5316]   Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, December 2008.

[RFC5441]   Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, April 2009.

[ISO 10589]   ISO, "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/ IEC 10589:2002.

   [H-PCE]   King, D. and A. Farrel, "The Application of the Path
                  Computation Element Architecture to the Determination
                  of a Sequence of Domains in MPLS and GMPLS.
                  (draft-ietf-pce-hierarchy-fwk-04)", June 2012.

   [DOMAIN-SEQ]   Dhody, D., Palle, U., and R. Casellas, "Standard
                  Representation Of Domain Sequence
                  (draft-ietf-pce-pcep-domain-sequence-01)", July 2012.

Authors' Addresses

   Dhruv Dhody
   Huawei Technologies India Pvt Ltd
   Leela Palace
   Bangalore, Karnataka 560008
   INDIA

   EMail: dhruv.dhody@huawei.com

   Udayasree Palle
   Huawei Technologies India Pvt Ltd
   Leela Palace
   Bangalore, Karnataka 560008
   INDIA

   EMail: udayasree.palle@huawei.com