Preferred Path Route Graph Structure
draft-ce-lsr-ppr-graph-04

Abstract

   This document defines a graph structure for the Preferred Path Route
   (PPR) for IS-IS, OSPFv2 and OSPFv3 protocols. This structure helps
   further scale of the PPR and reduce domain level global entries
   needed in some data planes.

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 [RFC2119],
   RFC8174 [RFC8174] when, and only when they appear in all capitals, as
   shown here.

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 https://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 April 1, 2021.

Table of Contents

1. Introduction

   Preferred Path Routing (PPR) is a routing protocol mechanism
   concerned with the creation of a routing path as specified in the
   PPR-Path objects. These can be signaled via appropriate IGPs (IS-IS,
   OSPFv2, OSPFv3) and indicate the path for a data plane identifier
   (PPR-ID). With this, all PPR capable nodes along that path establish
   forwarding state for the PPR-ID and any packet destined to the PPR-ID
   would use that path instead of the IGP computed shortest path to the
   destination.

   PPR-Paths and relevant IGP extensions are defined in
   [I-D.chunduri-lsr-isis-preferred-path-routing] and
   [I-D.chunduri-lsr-ospf-preferred-path-routing]. In these IGP

extensions, PPR-Paths are described as a path structure, which is an ordered linear list of Path Description Elements (PDEs) starting with a sender PDE followed by zero or more transit PDE and finishing with the destination PED. PDEs can indicate the node, a link to the node and services on a node.

A separate PPR-ID is required for every possible PPR-Path, even if one is just a subset of another path with the same destination. To provide PPR-Paths from N possible source nodes to one destination node, N PPR-IDs are therefore necessary. To create full-mesh connectivity via PPR-Paths between N nodes, N^2 PPR-Paths and N^2 PPR-IDs would be needed. Even if PPR-Paths would only be used for a subset of connections, such as for high-value traffic in larger networks, this scale behavior is less than ideal.

To allow scalability, in-terms of number of PPR-IDs needed on the destination nodes, number of forwarding entries needed on the nodes in the paths (for overlapping paths), and to minimize the amount of PPR information needed in the control plane, this document introduces a PPR-Tree structure in Section 2.

The terminology in this document uses the more generic term of PPR Graphs instead of PPR Trees because it is extensible.

1.1. Acronyms

MPLS   - Multi Protocol Label Switching

MSD   - Maximum SID Depth

PDE   - Path Description Element

PPG   - Preferred Path Graph

PPR   - Preferred Path Routing/Route

PPR-ID   - Preferred Path Route Identifier, a data plane identifier

SID   - Segment Identifier

SPF   - Shortest Path First

SR-MPLS   - Segment Routing with MPLS data plane

SRH   - Segment Routing Header - IPv6 routing Extension header

SRv6   - Segment Routing with Ipv6 data plane with SRH

      TE   - Traffic Engineering

2. PPR Graph TLVs

2.1. IS-IS TLVs

   This section describes the encoding of IS-IS PPR Tree TLV. This TLV
   can be seen as having 4 logical section viz., encoding of the PPR-
   Prefix (IS-IS Prefix), encoding of PPG-ID, encoding of path
   description with an ordered PDE (Path Description Element) Sub-TLVs,
   belonging to one or more Branch-IDs and a set of optional PPR
   attribute Sub-TLVs, which can be used to describe PPR Graph common
   parameters. Multiple instances of this TLV MAY be advertised in IS-
   IS LSPs with different PPG-ID Type and with corresponding Branch-ID/
   PDE Sub-TLVS. The PPR Graph TLV has Type TBD (suggested value xxx),
   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    |  Graph-Type   | Graph-Flags   |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |          PPR-Prefix Sub-TLV (variable size)                  |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |L| Frag-ID    |                 PPG-ID                        |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     //     Branch-ID Sub-TLV and PPR-PDE Sub-TLVs (variable)       //
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |               PPR-Attribute Sub-TLVs(variable)               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
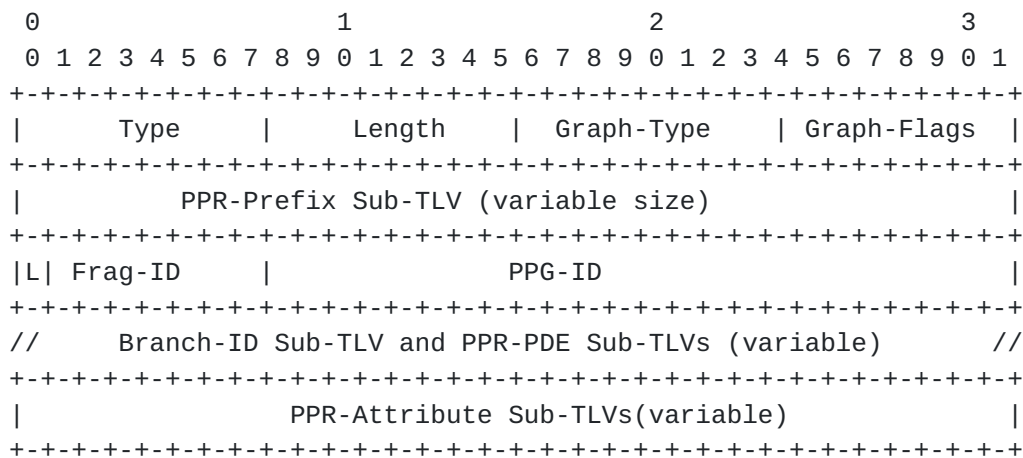
                     Figure 1: PPR Tree TLV Format

   o Type - TBD (IANA) from IS-IS top level TLV registry.

   o Length - Total length of the value field in bytes (variable).

   o Graph-Type - 1 Octet value (0-255, IANA Registry TBD). Value 0
      defines a PPR Tree structure (this document). PPR-Paths can also
      be encoded as PPR-Trees with a single branch.

   o Graph-Flags - 1 Octet flags for this TLV are described below.

   o Frag-ID - 1 Octet TLV Fragment-ID, with 7-bit Identifier value
      (0-127). L bit MUST be set if a graph has only one fragment or if
      it is the last Fragment of the graph. PPG-ID value for all
      fragments MUST be the same.

o PPG-ID - 3 byte Preferred Path Graph Identifier. Originator of
   the graph MUST ensure uniqueness across the domain.

o Branch-ID Sub-TLV is defined in Section 2.1.1. This represents
   the branch-id of the structure followed by PDE Sub-TLVs in that
   branch. Different branches of the graph can be in different
   fragments of this TLV. However, a complete set of PDE Sub-TLVs
   MUST be specified in one TLV fragment.

o PPR-PDE Sub-TLV defined in
   [I-D.chunduri-lsr-isis-preferred-path-routing]. Additional
   information in the PPR-PDE Sub-TLV is described in Section 2.1.2.

o PPR-Attribute Sub-TLVs defined in
   [I-D.chunduri-lsr-isis-preferred-path-routing] are applicable
   here.

PPR-Flags field of PPR TLV has the following flag bits defined.
These flags, at this point mostly related to applicability of this
TLV in an L1 area or entire IS-IS domain or from where the PPR-Prefix
is being originated:

     PPR Graph-Flags Format

         0 1 2 3 4 5 6 7
        +-+-+-+-+-+-+-+-+
        |S|D| Reserved |
        +-+-+-+-+-+-+-+-+

1. S - If set, the PPR Graph TLV MUST be flooded across the entire
    routing domain. If the S flag is not set, the PPR Graph TLV MUST
    NOT be leaked between IS-IS levels. This bit MUST NOT be altered
    during the TLV leaking

2. D - when the PPR Graph TLV is leaked from IS-IS level-2 to level-
    1, the D bit MUST be set. Otherwise, this bit MUST be clear.
    PPR TLVs with the D bit set MUST NOT be leaked from level-1 to
    level-2. This is to prevent TLV looping across levels.

3. Reserved - reserved bits for future use. Reserved bits MUST be
    reset on transmission and ignored on receive.

2.1.1. Branch-ID Sub-TLV

Branch-ID Sub-TLVs represent the branch of the graph described. This
is a new Sub-TLV type (IANA TBD) in PPR TLV
[I-D.chunduri-lsr-isis-preferred-path-routing]. Type TBD (Suggested

Value - IANA TBD), with a length of 1 byte, and Value is the branch
identification number in the range of 0 to 255.

2.1.2. PPR PDE Sub-TLV

PPR PDE Sub-TLV is defined in
[I-D.chunduri-lsr-isis-preferred-path-routing]. This document
extends the same with the following:

1. PPR-PDE Flags (Bit position 2), S: Source Bit. Indicates the PPR
    head-end and MUST be set if this PDE corresponds to the same.

2. PPR-ID Sub-Sub-TLV: Type, length and value fields would be same
    as PPR-ID Sub-TLV defined in
    [I-D.chunduri-lsr-isis-preferred-path-routing]. This Sub-Sub-TLV
    MUST be present only when 'D' flag is set in the PPR-PDE Flags
    field.

PPR-PDE Flags field is defined in PPR-PDE Sub-TLV
[I-D.chunduri-lsr-isis-preferred-path-routing].

2.2. OSPF TLVs

2.2.1. OSPFv2 TLVs

TBD.

2.2.2. OSPFv3 TLVs

TBD.

3. Encoding and Processing details

[I-D.chunduri-lsr-isis-preferred-path-routing] describes how a PPR
path can be established. This document builds on the same base
concept but expands the same with a graph structure as defined in
Section 2. The key new encoding element here over prior PPR Paths is
the existence of multiple Branches in the PPR Graph description.

Each Branch-ID sub-TLV is followed by ordered sequence of PDEs. A
PPR Graph can be constructed from one or more PPR Branches. Branches
are stitched together by using the same PDE in two branches. To
simplify parsing of branches, only the last PDE of a branch can be
stitched to another branch. In result, any PDE can only be a non-
last PDE in one Branch but last PDE in more than one branch. A PPG-
ID field is defined in this document. This MUST be unique in the
domain and represents the graph structure as whole.

A complete Graph may not fit into maximum allowable size of the IS-IS
TLV. To overcome this a 7 bit Frag-ID field is defined (Section 2).
With this, a single PPR Graph is represented via one or more
fragmented PPR Graph TLVs all having the same PPG-ID. Each Fragment
carries the PPG-ID as well as a numeric Frag-ID from 0 to (N-1), when
N fragments are needed to describe the PPR Graph (where N>1). In
this case Fragment (N-1) MUST set the L bit to indicate it is the
last fragment. The optional PPR Attribute Sub-TLVs which describe
the Graph overall MUST be included in the last fragment only.

3.1. S And D bits in PDEs

In PPR Paths as defined in
[I-D.chunduri-lsr-isis-preferred-path-routing], currently only a
simple linear path structure for a destination node is possible.
However, with a bit on path element source and a bit for destination
(refer section) - same path ID/PPR-ID can be used to represent
multiple paths if some of the nodes are also sources and terminating
on the same destination node.

    1. A Linear Path structure:
       PDE1 --> PDE2 --> PDE3 --> PDE4 --> PDE5
       [First PDE always Source and last PDE is always Destination]

    2. A PPR Graph with S and D bits:
       PDE1(with-S-bit-set)-->PDE2-->PDE3(with-S-bit-set)..
       ..-->PDE4(with-D-bit-set)-->PDE5(with-D-bit-Set)

            ==> PDE1 --> PDE2 --> PDE3 --> PDE4
            ==> PDE1 --> PDE2 --> PDE3 --> PDE4 --> PDE5
            ==> PDE3 --> PDE4
            ==> PDE3 --> PDE4 --> PDE5

               Figure 2: PPR Graph with S and D bits

In the above Figure 2 example, in (1) a linear path list of 5 nodes
are described where PDE1 is the source/ingress-point and PDE5 is the
destination/egress point of the path. In (2), the path can be
defined in this document, where some PDEs can have S(ource) and/or
D(estination) bit or both can be set. Here, PDE1 and PDE3 have the
Source bit set, PDE4 and PDE5 the Destination bit set. This Branch
structure is equivalent to the set of 4 PPR-PDE lists as shown:
PDE1->PDE5, PDE1->PDE45, PDE3->PDE4, PDE3->PDE5. This reduces the
amount of information that needs to be sent across the IGP and that
needs to be processed by each node.

   If the bits and branch structure were not used, the 4 PPR PDE lists
   would have required each a unique PPR-ID (and the resulting
   forwarding entries created), but the Branch requires only 2 PPR-IDs:
   one for both paths terminating in PDE4, and one for both paths
   terminating in PDE5.

3.2. Graph processing procedure example

```
         Brach0     Branch1       Branch2

         PDE1       PDE12(S-bit) PDE6
           \           \            /
           PDE2      PDE11      PDE7
             \          \       /
             PDE3     PDE10 PDE8 (S-bit)
               \          \  /
               PDE4      PDE9
                 \      /
                  \    /
                  PDE5
                 (D-bit)
```
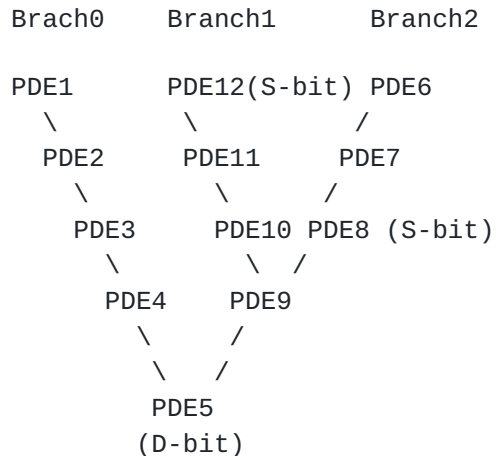
                    Figure 3: PPR Graph (Tree) Example

   With a PPR Tree structure both flooding optimization and reduction in
   the number of SIDs needed at the destination can be achieved. To do
   this encoding as specified in Section 2 (a) Every PDE-ID can be non-
   last-PDE in at most one Branch. It can be last-PDE in one or more
   Branches (ex: PDE9). (b) Branches form a tree by joining nodes with
   same PDE-ID (PDE9 and PDE5 in the above example). Leafs of the tree
   must be S(ources), e.g.: PDE1, PDE12, PDE6. Root of the tree must be
   the only D(estination) of the tree (e.g.: PDE5).

   How to build forwarding entry (referring to the Figure 3 above):

   1. If PPR-ID in PDE of PPR Graph is indicating this node (example:
       PDE5): This node is D(estination) of this tree. Forwarding state
       is built for this PPR-Tree like for PPR-Path, no changes.

   2. If PPR-ID is NOT indicating this node, then this node MAY be
       source (PDE12, PDE8) or midpoint (PDE9, neither source nor
       destination):

   a. Node sequentially examines all branches until it finds a PDE with
       its own PDE-ID. It then establishes a forwarding entry for the
       PPR-ID indicated in the PPR header with the next-hop being the
       next PDE in the current branch.

   b. This nodes PDE may be the last PDE in a Branch, for example PDE9
       in Branch1. In this case, the node ignores this branch because
       it cannot build a complete forwarding entry from it. Instead, it
       will build the forwarding entry from another branch, e.g.: Node
       with PDE9 will build forwarding entry for destination PDE5 when
       it examines Branch2 because there it will have a next hop PDE5.
       After forwarding entry is built, node can stop examining rest of
       Branch or further Branches.

   c. If node does not find its own PDE in any branch it is not on the
       graph and ignores this PPR-Graph.

4. Acknowledgements

   Thanks to Yingzhen Qu and Richard Li for multiple discussions on this
   topic.

5. IANA Considerations

5.1. IS-IS IANA

   This document requests the following new TLV in IANA IS-IS TLV code-
   point registry.

```
       TLV #   Name
       -----   --------------
       TBD     PPR Graph TLV
```

   This document requests IANA to create a new Sub-TLV registry for PPR
   TLV Section 2 with the following initial entries (suggested values):

```
   Sub-TLV #   Sub-TLV Name
   ---------   ----------------------------------------------------------

    TBD         Branch-ID (Section 2)
```

5.2. OSPFv2 IANA

5.3. OSPFv3 IANA

5.4. IGP Parameter IANA

   This document requests additional IANA registries in an IANA managed
   registry "Interior Gateway Protocol (IGP) Parameters" for various PPR
   TLV parameters. The registration procedure is based on the "Expert
   Review" as defined in [RFC8126]. The suggested registry names are:

o "Graph-Type" - Types are an unsigned 8 bit numbers. Values are as
   defined in Section 2 of this document.

o "Graph-Flags" - 1 Octet. Bits as described in Section 2 of this
   document.

6. Security Considerations

Security concerns for IS-IS are addressed in [RFC5304] and [RFC5310].
Further security analysis for IS-IS protocol is done in [RFC7645]
with detailed analysis of various security threats and why [RFC5304]
should not be used in the deployments.

OSPF security extensions are described in [RFC2328] and [RFC7684] and
these apply to the extensions specified in this document. While OSPF
is under a single administrative domain, there can be deployments
where potential attackers have access to one or more networks in the
OSPF routing domain. In these deployments, stronger authentication
mechanisms such as those specified in [RFC7474] SHOULD be used.

Advertisement of the additional information defined in this document
introduces no new security concerns in IS-IS or OSPF protocols.

7. References

7.1. Normative References

[I-D.chunduri-lsr-isis-preferred-path-routing]
          Chunduri, U., Li, R., White, R., Tantsura, J., Contreras,
          L., and Y. Qu, "Preferred Path Routing (PPR) in IS-IS",
          draft-chunduri-lsr-isis-preferred-path-routing-05 (work in
          progress), March 2020.

[I-D.chunduri-lsr-ospf-preferred-path-routing]
          Chunduri, U., Qu, Y., White, R., Tantsura, J., and L.
          Contreras, "Preferred Path Routing (PPR) in OSPF", draft-
          chunduri-lsr-ospf-preferred-path-routing-04 (work in
          progress), March 2020.

[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>.

7.2. Informative References

[RFC2328]   Moy, J., "OSPF Version 2", STD 54, RFC 2328,
            DOI 10.17487/RFC2328, April 1998,
            <https://www.rfc-editor.org/info/rfc2328>.

[RFC5304]   Li, T. and R. Atkinson, "IS-IS Cryptographic
            Authentication", RFC 5304, DOI 10.17487/RFC5304, October
            2008, <https://www.rfc-editor.org/info/rfc5304>.

[RFC5305]   Li, T. and H. Smit, "IS-IS Extensions for Traffic
            Engineering", RFC 5305, DOI 10.17487/RFC5305, October
            2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC5310]   Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R.,
            and M. Fanto, "IS-IS Generic Cryptographic
            Authentication", RFC 5310, DOI 10.17487/RFC5310, February
            2009, <https://www.rfc-editor.org/info/rfc5310>.

[RFC7474]   Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed.,
            "Security Extension for OSPFv2 When Using Manual Key
            Management", RFC 7474, DOI 10.17487/RFC7474, April 2015,
            <https://www.rfc-editor.org/info/rfc7474>.

[RFC7645]   Chunduri, U., Tian, A., and W. Lu, "The Keying and
            Authentication for Routing Protocol (KARP) IS-IS Security
            Analysis", RFC 7645, DOI 10.17487/RFC7645, September 2015,
            <https://www.rfc-editor.org/info/rfc7645>.

[RFC7684]   Psenak, P., Gredler, H., Shakir, R., Henderickx, W.,
            Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute
            Advertisement", RFC 7684, DOI 10.17487/RFC7684, November
            2015, <https://www.rfc-editor.org/info/rfc7684>.

[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>.

Authors' Addresses

Uma Chunduri
Futurewei
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: umac.ietf@gmail.com

Toerless Eckert
Futurewei
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: tte+ietf@cs.fau.de