Preferred Path Routing (PPR) in IGPs

draft-chunduri-lsr-isis-preferred-path-routing-01
draft-chunduri-lsr-ospf-preferred-path-routing-01

Uma Chunduri, Richard Li [Huawei USA]

Russ White [LinkedIn]

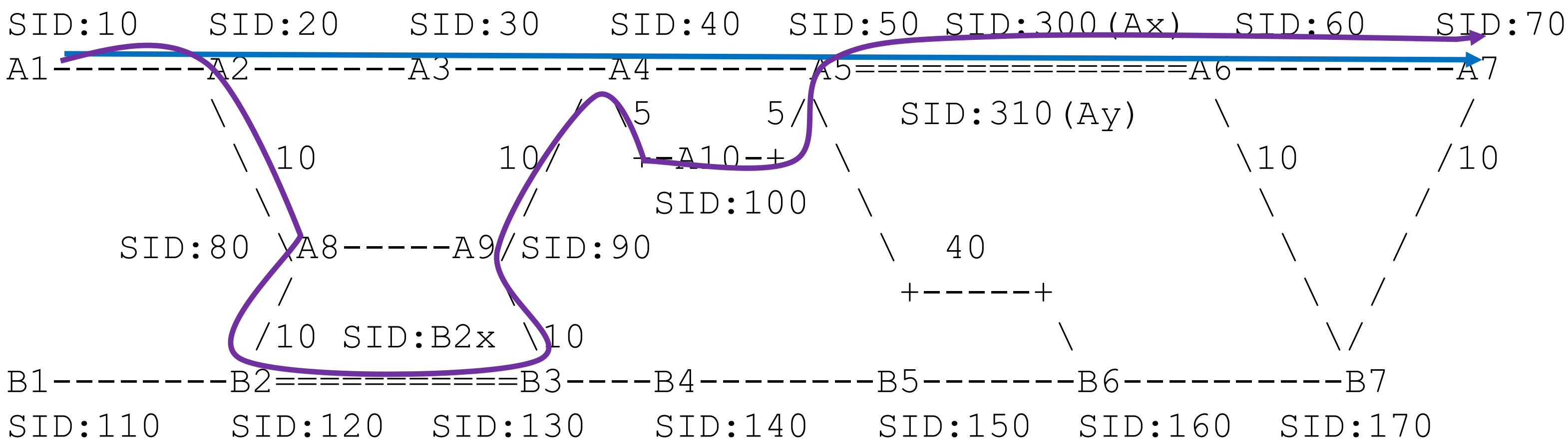
Jeff Tantsura [Nuage Networks]

Luis M. Contreras [Telefonica]

Yingzhen Qu [Huawei USA]

LSR WG, IETF 102, Montreal

SR Network



SR Path:

- All nodes: A1 to A7 and B1 to B7 (In the above Picture uses SR-MPLS)
- Shortest Path from A1 to A7: A2-A3-A4-A5-A6-A7 (→)
- SR-PATH-1: From A1 to A7 - A2-A8-B2-B2x-A9-A10-Ax-A7(→)
- The above SR path uses both node and adjacency SIDS a stack of 8 labels.

- This number can be higher based on MSD Capabilities of intermediate nodes and ELI/EL pair or because of Service Labels
- If this were to use for IPv6 data plane – RFC8200 IPv6 Encap. + SRH with 8 SRv6 SIDs

Deployment Scenarios & Issues

Issues:

- **Hardware capabilities:** Not all nodes can push/read label stack needed.
  - MSD only helps to mitigate, if there is an alternate NSP, which meets the operator requirements

- **Line Rate:** Potential performance issues with increased size of SRH with IPv6 SIDs

- **MTU:** Potential MTU/Fragmentation issues with large SID stack (SR-MPLS, SRH)

- **Header Tax:** NW/Path overhead relative to actual application data, especially for small payload packets (mIOT and uRLLC in 5G or in various fixed scenarios).

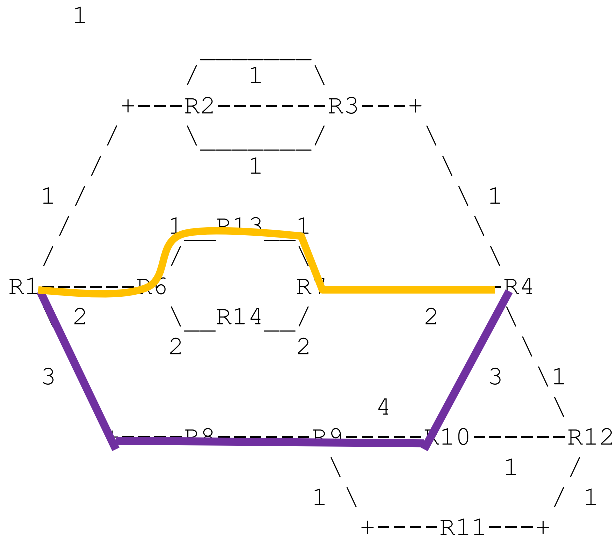
PPR TLV (IS-IS)

```
 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     |  PPR-Flags                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          PPR-Prefix Sub-TLV (variable size)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          PPR-ID Sub-TLV (variable size)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          PPR-PDE Sub-TLVs (variable)                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          PPR-Attribute Sub-TLVs (variable)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- New top level MT aware TLV with control plane prefix
- Data plane identifier (PPR-ID) with data plane type (e.g. MPLS, SRH, IPv6 etc..)
- Path Description Element Sub-TLV (ordered path info of underlying data plane)
- PPR Attribute Sub TLVs –path attributes including but not limited traffic accounting

Computation

- **Strict PPR**: R1-R8-R9-R10-R4 for a prefix from R4
- After R9 receives the path, after SPF computation it check if R9 is itself on the path
- For E.g.: @R9: Without the above PPR, for R4's FEC, R9 would have set the NH to R11,but it changes to R10.
- Same applies to all the nodes in the path (R1, R8 etc..).
- **Loose PPR:** R1-R6-R7-R4 for a reachable prefix of R4 with 'L' bit set.
- Control plane processing is same Strict PPR. Various data plane differences in the document.



In Summary:

- Change the next hop from actual shortest path NH to the NH of the immediate segment as described in the PPR TLV.
- Data-plane: IGP programs the received PPR-ID to the corresponding NH calculated as above (based on the data plane type)
- Existing protections/IP-FRR works as is; per new NH (but more can be done).

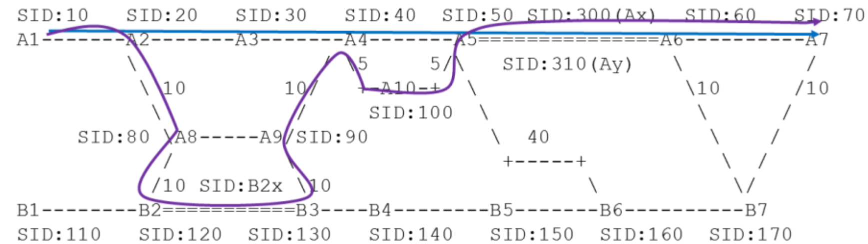
Relation To Binding SID

- Not related
  - Earlier version of SR draft has EROs
- IS-IS Binding SID advertises the SID on behalf of one or *more* nodes in the network

Traffic Accounting

- Traffic Accounting Statistics through PPR with PPR-ID
  - **OPT1:** Offline Provisioning and collecting from the nodes as needed based on the operator
    - PPR-ID representing the PPR simplifies this operation for Operators
  - **OPT2:** Dynamically enable stats for certain PPRs through optional PPR Attribute Sub-TLVs(traffic accounting with fine granularity for some PPRs as needed automatically). No need for additional labels and MSD/RLD compatability issues.

Data plane with PPR-ID (SR-MPLS)

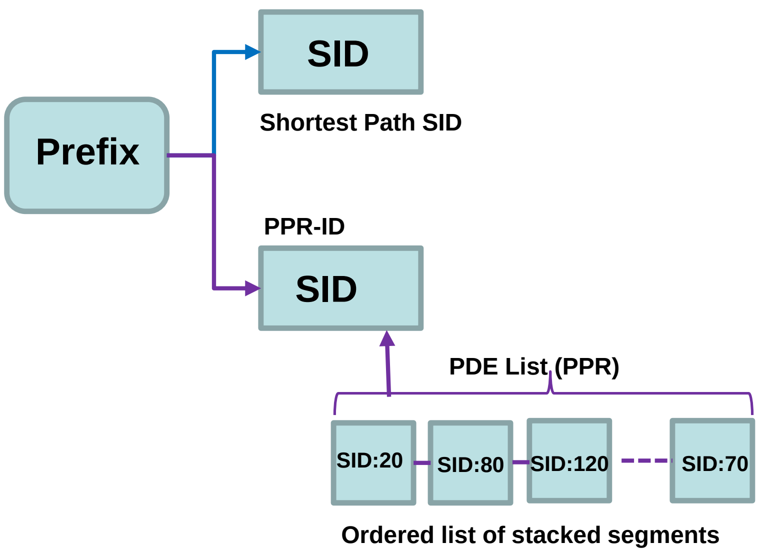


8 Node NSP with RLD = 6

```
+---------------+
|  A2's Label   |
+---------------+
|  A8's Label   |
+---------------+
|  B2's Label   |
+---------------+
|  B3's Label   |
+---------------+
|  ELI Label    |
+---------------+
|  EI Label     |
+---------------+
|  A9's Label   |
+---------------+
|  A5's Label   |
+---------------+
|  Ax's Label   |
+---------------+
|  A7's Label   |
+---------------+
|  ELI Label    |
+---------------+
|  EI Label     |
+---------------+
| Service Label |
+---------------+
|               |
|    Payload    |
|               |
+---------------+
```

```
+---------------+
| PPR-ID/Label  |
+---------------+
| Service Label |
+---------------+
|               |
|    Payload    |
|               |
+---------------+
```

Summary:



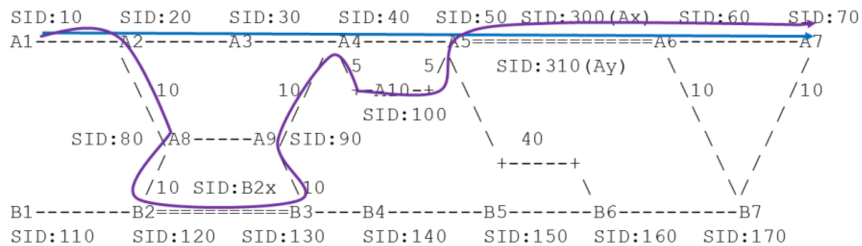
- This can be seen as equivalent of single shortest path SID (LDP equivalent); i.e., PPR is represented with one label as in the DP
- Though Service Label is shown as one, it can be more in most deployments/various scenarios

IETF102

Data plane with PPR-ID (with SRH)

SRH with 8 Segment PPR

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Next Header   |  Hdr Ext Len  | Routing Type  | Segments Left |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Last Entry   |     Flags     |              Tag              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[0] (A2's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[1] (A8's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[3] (B2's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[4] (B3's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[5] (A9's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[6] (A5's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[7] (Ax's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[8] (A7's 128 bits IPv6 address)        |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[9](Service SID 128 bits IPv6 address)  |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                                                             //
//         Optional Type Length Value objects (variable)       //
//                                                             //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```




```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Next Header   |  Hdr Ext Len  | Routing Type  | Segments Left |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Last Entry   |     Flags     |              Tag              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[0] (NSPF-ID 128 bits IPv6 address)     |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|           Segment List[1](Service SID 128 bits IPv6 address)  |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                                                             //
//         Optional Type Length Value objects (variable)       //
//                                                             //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Yang Data Model

- Yang data model for Preferred Path Routing

https://tools.ietf.org/html/draft-qct-lsr-ppr-yang-00

Scaling & Other Aspects

- **Scalability**
  - In a network with N nodes and with O(N^2) total unidirectional paths with (k) multiple such path only small set are preferred paths based on the deployment (for high value traffic Detnet, 5G Slices)
  - However to address the scaling of preferred paths a TREE structure can be used and details are in [I-D.draft-ce-ppr-graph-00].
  - Each PPR Tree uses one label/SID and defines paths from any set of nodes to one destination, thus reduces the number of entries needed from SRGB at each node (more details in the draft).
  - In other word, PPR Tree identifiers are destination identifiers and with this scaling simplifies to linear in N i.e., O(k*N).
- **Support for native IP data planes (IPv4 and IPv6) with only control plane upgrades**
  - Needs respective IP encapsulation with destination IP as PPR-ID
  - Needed for slow migration and backward compatibility, More details in the draft

DMM WG is responding to 3GPP Study item for optimized 5G user plane

- 3GPP Study Item http://www.3gpp.org/ftp/tsg_ct/WG4_protocollars_ex-CN4/TSGCT4_82_Gothenburg/Docs/C4-181380.zip
- 3GPP Scenarios, Requirements, Solution Comparison https://www.ietf.org/id/draft-bogineni-dmm-optimized-mobile-user-plane-01.txt (SRv6, LISP, ILA)

✓ **This work helps most of the proposals to reduce the transport overhead on N9 interface.**

Status & Next Steps

- Concept Presented at IETF101; lot of offline feedback and updates
- Comments?
- Request WG Adoption

Thank you!