NEXTHOP_PATH_RECORD ATTRIBUTE for BGP
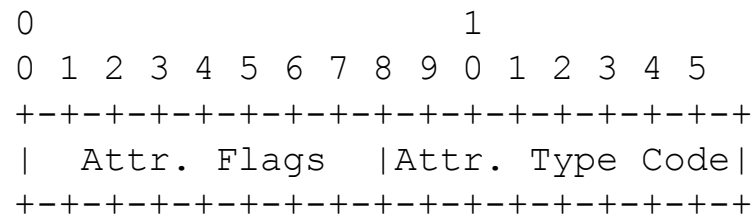
draft-zhang-idr-nexthop-path-record-00

Zhenbin Li, Li Zhang, Susan Hares
Huawei Technologies

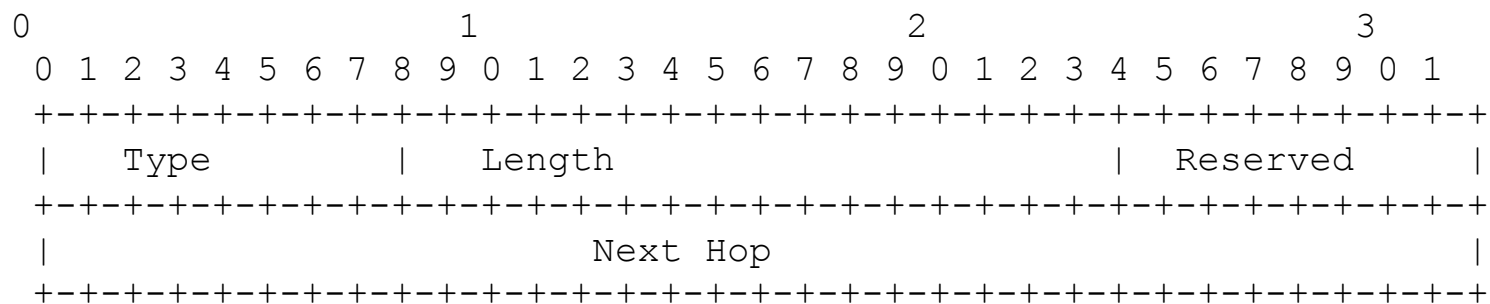IETF 90, Toronto, Canada

NEXTHOP_PATH_RECORD ATTRIBUTE

- NEXTHOP_PATH_RECORD ATTRIBUTE

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Attr. Flags   |Attr. Type Code|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Attr.Flags should be **optional transitive**
Attr.Type Code should be allocated by IANA

- Next hop path segment

```
 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               | Reserved      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Next Hop                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Next Hop is the route next hop address

BGP NEXTHOP_PATH_RECORD ATTRIBUTE Description

- NEXTHOP_PATH_RECORD ATTRIBUTE
  - Optional transitive BGP Path Attribute
  - Records Sequence of next hop path segments

- Operation: Path Record Config on and next-hop-self
  - If originate, add next_hop to the next_hop_segment
  - If passing append next_hop to next_hop_segment

- Operation: Path Record Config on and no next-hop-self
  - Don't originate
  - Pass but do not modify

Deployment Considerations

- Customized Best Path Selection
  - The next_hop_record information gathered on an IBGP or EBP route could be used by off-line decision processing to select paths, and re-inserted as policy to affect the decision making via I2RS

Next Steps

- Get feedback on the NEXTHOP_PATH_RECORD ATTRIBUTE extension and application

- Coordinate with other similar drafts to record next-hop path information through BGP extensions.

- The procedure for next hop path segment usage for IPv6 or other extensions will be discussed later

Questions?

Backup Slides

Online – but not presented

BGP NEXTHOP_PATH_RECORD ATTRIBUTE Description

- NEXTHOP_PATH_RECORD ATTRIBUTE
  - Optional transitive BGP Path Attribute
  - Records Sequence of next hop path segments

- Operation
  - For BGP route selection, it can reduce the route policy complexity
  - It is to get the service path in transport network which can be used for network operation and maintenance
  - will record the distribution path of a route in Seamless MPLS network by the list of next hop path segment

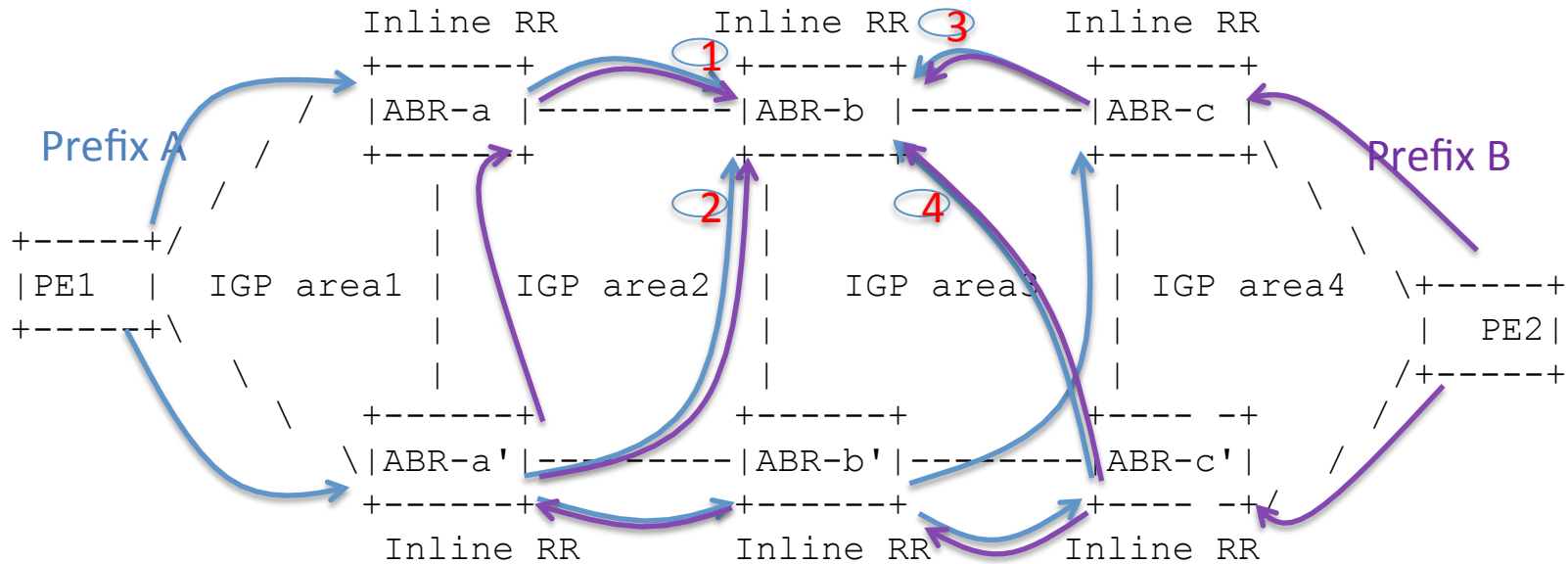# NEXTHOP_PATH_RECORD ATTIBUTE Process in BGP(1)

- Creating and modification process
    1. If the route is originated in this BGP speaker
        - If the attribute is supported, the TLV SHOULD be originated including the BGP speaker's own next hop address in a next hop path segment
        - If the attribute is not supported, the route will be distributed without NEXTHOP_PATH_RECORD ATTRIBUTE
    2. if the route is received from one BGP speaker's UPDATE message
        - If the NEXTHOP_PATH_RECORD ATTRIBUTE is NULL and the local BGP speaker support this attribute, when the route is propagated with NHS, the TLV SHOULD be originated including the BGP speaker's own next hop address in a next hop path segment

NEXTHOP_PATH_RECORD ATTIBUTE Process in BGP(2)

- If the attribute is non-NULL and the local BGP speaker support it, when the route is propagated with NHS, the BGP speaker MUST appends its own next hop address as the last one of the next hop path segments.

- If the attribute is NULL and the local BGP speaker support it, when the route is propagated without NHS, the BGP speaker MUST NOT originate the NEXTHOP_PATH_RECORD ATTRIBUTE.

- If the attribute is non-NULL and the local BGP speaker support it, when the route is propagated without NHS, the BGP speaker MUST NOT change the next hop path sequence.

- If the BGP speaker does not support NEXTHOP_PATH_RECORD ATTRIBUTE, it SHOULD keep the NEXTHOP_PATH_RECORD ATTRIBUTE unchanged.

Deployment Considerations: Seamless MPLS PE-RR

- ## Use in Seamless MPLS case with PE-RR

```
              Inline RR        Inline RR  3      Inline RR
              +------+     1  +------+   3      +------+
          /   |ABR-a |--------|ABR-b |--------|ABR-c |
Prefix A   /  +------+        +------+        +------+\     Prefix B
         /       |         2    |        4       |      \
+-----+/         |              |              |         \
|PE1  |  IGP area1 |  IGP area2 |  IGP area3 | IGP area4 \+-----+
+-----+\          |              |              |          | PE2|
       \          |              |              |          /+-----+
        \     +------+        +------+        +---- -+     /
         \|ABR-a'|----------|ABR-b'|--------|ABR-c'|   /
          +------+        +------+        +---- -+/
          Inline RR        Inline RR      Inline RR
```

- Inline RR will reflect the route with next hop self (NHS)
- For prefix A, ABR-b should select optimal route with next hop of ABR-a or ABR-a'; while for prefix B, ABR-b should select optimal route with next hop of ABR-c or ABR-c'
- To achieve this result, a complex route policy should be predesigned and configured for every peer every prefix

The NEXTHOP_PATH_RECORD ATTRIBUTE can optionally collect the pathway information of the route, which may aid in monitoring paths in this complex path or in offline processing that reduces complex policy.

Draft Updates

- Change the draft name from *draft-zhang-idr-nexthop-path-attr* to *draft-zhang-idr-nexthop-path-record*.

- The BGP path decision process based on the NEXT_HOP ATTIBUTE is removed. NEXT_HOP ATTRIBUTE changes to NEXTHOP_PATH_RECORD ATTRIBUTE.

- Refine drafts based on the collected comments.