Distributed Bump-in-the-wire Use Case

Author(s): Yubao Wang
Problem Statement (1)

- Bump-in-the-wire Use case
  - Defined in RFC9136 section 4.3
  - These NVEs are all capable of IP-VRF forwarding
  - but the communication between H3 and SN2 still have to pass through the DGWs

- Distributed Inter-Subnet Forwarding
  - The communication between H3 and H4 needn't pass through the DGWs
  - Extends this feature to SN1 and SN2
Problem Statement (1)

- NVEs are capable of IP-VRF forwarding
  - Add IP-VRF instance on NVEs first
  - IP prefixes installed by policy like RFC9136 Bump-in-the-wire
  - or IP prefixes installed by PE-CE routing protocols

- Distributed Inter-Subnet Forwarding
  - Between Bump-in-the-wire subnets (e.g. SN2) and ordinary subnets (e.g. H3)
  - Even between two Bump-in-the-wire instances (e.g. SN1 and SN2)
Solution: Supplementary BD

- **Supplementary BD**
  - RT1_NVE2_10 is advertised for BD-10, but the SBD's export RT (eRT) are carried along with it too.
  - So RT1_NVE3_10 will be imported into SBD or BD-10, not into IP-VRF

- **Distributed Bump-in-the-wire**
  - RT-5 Advertisement is unchanged from RFC9136 Bump-in-the-wire
  - RT-5 Importing is unchanged from RFC9136 Bump-in-the-wire
  - DGW1 and NVE8 follows the behavior of DGW1 of RFC9136 Bump-in-the-wire

Notes: A non-upgraded DGW1 of RFC9136 Bump-in-the-wire use case can take the place of NVE8, whose SBD can be considered to be the BD-10 of that kind of DGW1.
Problem Statement (2)

- RT-5 Route Resolution
  - RT-5's ESI must be resolved in the exact BD which is required
    - RT5_SN1's ESI can't be resolved in BD-20
    - RT5_SN2's ESI can't be resolved in BD-10
  - The required BD should be found out before the route resolution
    - thus BD-10's export RT (eRT) should be carried in RT5_SN1
    - thus BD-20's export RT (eRT) should be carried in RT5_SN2

- RT-1 per EVI routes are imported into BD-10/BD-20

Notes 1: These RT5 routes needn't be recursively resolved in the IP-VRF routing table, because that its L3 out interface (the IRB interface of BD-10/BD-20) can be directly determined by the route-targets carried in itself.

Notes 2: If the ESI of RT5_SN1 is recursively resolved in the IP-VRF routing table, the RT1_NVE3_20 may be mistaken for RT5_SN1's best match. Even worse, RT1_NVE3_10 is not advertised for IP-VRF per RFC9136.
Solution: ACI-Specific Ethernet Auto-Discovery (EAD)

- RT-1 per EVI confliction
  - for BD-20: RT1_NVE2_20 from NVE2
  - for BD-10: RT1_NVE3_10 from NVE3
  - for SBD: RT1s_NVE2_20 and RT1s_NVE3_10

- SBD using ACI-Specific Ether A-D
  - RT1s_NVE2_10, RT1s_NVE3_10 (ET-ID=AC1)
  - RT1s_NVE2_20, RT1s_NVE3_20 (ET-ID=AC2)
  - BD-10 or BD-20 needn't use such A-D mode

- RT-5 Extensions
  - Supplementary Overlay Index (SOI) Extended Community
  - Imported into IP-VRF through SBD
  - RT-5's ET-ID should still be zero
  - Using <ESI, SOI> to select RT-1_per EVI routes in SBD

### Table: ACs for BDx Route Name

<table>
<thead>
<tr>
<th>ACs</th>
<th>for BDx</th>
<th>Route Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC2a</td>
<td>BD-20</td>
<td>RT1_NVE2_20 (two groups of RTs)</td>
</tr>
<tr>
<td>AC2a</td>
<td>SBD</td>
<td>RT1s_NVE2_20</td>
</tr>
<tr>
<td>AC1a</td>
<td>BD-10</td>
<td>RT1_NVE2_10 (two groups of RTs)</td>
</tr>
<tr>
<td>AC1a</td>
<td>SBD</td>
<td>RT1s_NVE2_10</td>
</tr>
</tbody>
</table>

- NVE2 of Non-ACI-Specific A-D mode
- NVE2 of ACI-Specific A-D mode

- Advertised for Bump-in-the-wire instance1
- Conflicting in SBD
**Solution:** Supplementary Overlay Index (SOI) Extended Community

**Figure 5: Supplementary Overlay Index Extended Community**

<table>
<thead>
<tr>
<th>Type=0x06</th>
<th>Sub-Type=TBD</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VLAN2</th>
<th>VLAN1</th>
<th>MBZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FFF</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>I</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 1: VLAN-based SOIs**

<table>
<thead>
<tr>
<th>No.</th>
<th>AC Type</th>
<th>Type</th>
<th>VLAN2</th>
<th>VLAN1</th>
<th>MBZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>untag</td>
<td>type 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>default</td>
<td>type 0</td>
<td>0</td>
<td>FFF</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>dot1q</td>
<td>type 0</td>
<td>0</td>
<td>E</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>QinQ</td>
<td>type 0</td>
<td>E</td>
<td>I</td>
<td>0</td>
</tr>
</tbody>
</table>

**Notes1:**
- E : That field is the External VLAN of the AC.
- I : That field is the Internal VLAN of the AC.

**Notes2:** We can assume that the Ethernet Tag ID (ET-ID) of SBD's IMET routes can still be zero, just like the VLAN-based service interface.

---

**IP-VRF's Routing Table**

<table>
<thead>
<tr>
<th>RT-5 entry</th>
<th>ESI</th>
<th>SOI</th>
<th>L3 Out if</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT5_SN1</td>
<td>ESI23</td>
<td>10</td>
<td>SBD-IRB</td>
</tr>
<tr>
<td>RT5_SN2</td>
<td>ESI23</td>
<td>20</td>
<td>SBD-IRB</td>
</tr>
</tbody>
</table>

**ESI**

**ET-ID**

**Label**

**RT-1 entry**

<table>
<thead>
<tr>
<th>ESI23</th>
<th>10</th>
<th>BD-10 (NVE2)</th>
<th>RT1s_NVE2_10</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESI23</td>
<td>10</td>
<td>BD-10 (NVE3)</td>
<td>RT1s_NVE3_10</td>
</tr>
<tr>
<td>ESI23</td>
<td>20</td>
<td>BD-20 (NVE2)</td>
<td>RT1s_NVE2_20</td>
</tr>
<tr>
<td>ESI23</td>
<td>20</td>
<td>BD-20 (NVE3)</td>
<td>RT1s_NVE3_20</td>
</tr>
</tbody>
</table>

**SBD's Routing Table**