Loop Avoidance for BGP Best External

draft-xu-idr-best-external-loop-avoidance-00

Xiaohu Xu
(xuxh@huawei.com)
Erwei Guo
(guoerwei@huawei.com)

IETF 78, Maastricht

7/26/10
Problem Statement

Best External Route

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next-hop</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/8</td>
<td>AS#2</td>
<td>Primary</td>
</tr>
<tr>
<td>10/8</td>
<td>R2</td>
<td>Backup</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next-hop</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/8</td>
<td>R1</td>
<td>Primary</td>
</tr>
<tr>
<td>10/8</td>
<td>AS#3</td>
<td>Backup</td>
</tr>
</tbody>
</table>

Prefix | Next-hop | Status |
--------|----------|--------|
10/8    | R1       | Primary|
10/8    | R2       | Backup |

-> 10.1.1.1.1
Problem Statement

Prefix | Next-hop | Status
--- | --- | ---
10/8 | AS#2 | Invalid
10/8 | R2 | Primary

Prefix | Next-hop | Status
--- | --- | ---
10/8 | R1 | Primary
10/8 | AS#3 | Backup

Prefix | Next-hop | Status
--- | --- | ---
10/8 | R1 | Primary
10/8 | R2 | Backup

7/26/10
Solution Space

• By using IGP within AS#1 for fast flooding of AS#2's ASBR next-hop down event (Disabling BGP next-hop-self), together with the PIC (Prefix Independent Convergence) core technology, the transient loop can be almost eliminated.

• Concerns with the above approach:
  – Is PIC a default capability for all deployed BGP routers?
  – BGP next-hop-self is desired by some carriers in some cases.
  – BGP next-hop-self MUST be enabled in BGP/MPLS VPN [RFC4364] and Softwire Mesh [RFC5565] cases.

• Hence, we need to find out an alternative to the above solution for transient loop.
Proposed Solution

- For normal BGP or Softwire Mesh cases, the best external route SHOULD be processed according to the approaches defined in Carrying Label Information in BGP-4 [RFC3107].
### Diagram Description:

#### Prefixes and Next-hops:
- **10/8**: Next-hop AS#2 Status: Invalid Out-Label: ---
- **10/8**: Next-hop R2 Status: Primary In-Label: L1

#### In-Label and Prefixes:
- **10/8**: Next-hop R1 Status: Primary In-Label: L1
- **10/8**: Next-hop AS#3 Status: Backup

#### Fast Local Repair:
- **IP=10.1.1.1**
- **L=L1**

#### Best External Route:
- **AS#2**
- **AS#3**

---

7/26/10
Proposed Solution (con’t)

• For BGP/MPLS VPN case where one label per VRF mode is used for assigning labels for best routes, the best external route SHOULD be assigned with a unique label distinct from the label for the associated VRF.
  – In either one label per attachment circuit mode or one label per best external route mode.
Best External Route

PE2 assigns a MPLS label for the best external route distinct from that for the best route (one label per VRF mode).

Fast Local Repair

In-L  Prefix  Next-hop  Status  Out-L
---      10/8     PE1     Primary  L1
---      10/8     PE2     Backup   L2
Next Steps

• Solicit more comments from the WG.
• Ask for WG adoption.