draft-ietf-bess-rfc7432bis-13.txt

A. Sajassi (Cisco), LA. Burdet (Cisco), J. Drake (Independent), J. Rabadan (Nokia)

IETF 123, July 2025

Madrid

Changes in This Rev.

1. Further description of EVPN VLAN-bundle model
2. Added texts to each EVPN BGP route to indicate what constitutes a BGP error
3. Added a new section (7.14) for BGP Error Handling
   - Special thanks to our AD, Ketan, for encouraging us to write this section and to for helping us to gather the materials for it
4. Clarified time sequence example in section 9.2.2 (Route Resolution)

Cisco Confidential

1. VLAN-bundle Model

- **From**: In VLAN-aware bundle mode, there is one EVI with multiple BDs where each BD is represented by a VLAN. Furthermore, there are multiple BTs in a single MAC-VRF.
- **To**: In VLAN-bundle model, which can be considered as analogous to SVL mode in 802.1Q, there are multiple VLANs per EVI; however, all these VLANs map to the same BT per PE. Furthermore, there is one BT per MAC-VRF per PE. Since EVPN does not perform any VLAN specific operation in this model (e.g., VLAN pruning as done in IEEE 802.1Q), from EVPN perspective multiple VLANs look like a single VLAN/BD for VLAN-bundle model. That is why Ethernet Tag ID in all EVPN routes is set to zero for the EVPN-bundle model.

Cisco Confidential

3. BGP Error Handling

- The error handling actions as described in [RFC7606] are applicable for handling of BGP update messages for BGP EVPN SAFI.
- As indicated in [RFC7606], the goal is to minimize the disruption of a session reset or 'AFI/SAFI disable' to the extent possible.
- BGP Error Handling is covered for both:
  - NLRI Processing
  - Attribute Processing

Cisco Confidential

NLRI Processing

- NLRI length must be parsed and validate to be consistent with the total length of the enclosing MP_REACH or MP_UNREACH attribute. If invalid, "Session Reset" behavior MUST be applied.
- For unrecognized Route Types, the remainder of the NLRI (following the Length field) MUST be read (up to 'Length' octets) and MUST be ignored (i.e., discarded).

Cisco Confidential

NLRI Processing – Cont.

- For recognized routes, the following procedure needs to be done:
  - If the value in the Length field is inconsistent with the minimum required or maximum possible length for the specific Route Type, then "Session Reset" behavior MUST be applied.
  - The remainder of the NLRI (following the Length field) MUST be read (up to 'Length' octets) and processed as per the specific Route Type. If a syntactic error is encountered in reading any of the Key fields, then "Session Reset" behavior MUST be applied.
  - If a non-syntatic error is encountered in reading any of the Key fields, then "Treat-it-as-withdraw" behavior MUST be applied.
  - If an error is encountered in reading any of the non-Key fields, then "Treat-it-as-withdraw" behavior MUST be applied.

Cisco Confidential

Attribute Processing

- As a default behavior, if an attribute is not relevant for processing of a particular NLRI Type, then that attribute SHOULD be ignored.
- As a default behavior, if an error is encountered in an attribute that affects the forwarding of EVPN traffic, "Treat-it-as-withdraw" behavior SHOULD be applied to all the NLRIs that are encoded in the MP_REACH attribute of the Update message.
- As a default behavior, if an extended community type or instance, within the received EXTENDED COMMUNITIES attribute, is not relevant for processing of a particular NLRI Type, then that extended community instance SHOULD be ignored.

Cisco Confidential

Attribute Processing – Cont.

- If multiple instances of ESI Label extended community (EC) type are encountered, then all but the first one MUST be ignored.
- If multiple instances of ES Import Route Target EC type are encountered, then all all but the first one MUST be ignored.
- If multiple instances of MAC Mobility EC type are encountered, then all but the first one MUST be ignored.
- If multiple instances of Default Gateway EC type are encountered, then all but the first one MUST be ignored.
- If multiple instances of Layer2 Attributes EC type are encountered, then all but the first one MUST be

Cisco Confidential

Clarified time sequence example in section 9.2.2

T4:

- From nothing in Rev 11
- To (rev 12): “If after T3, PE1 and PE2 both advertise a MAC route for M1, but PE1 withdraws its Ethernet A-D per EVI route, then PE3 forwards traffic destined to M1 to PE2 only.”
- To (rev 13): “If after T3, PE1 and PE2 both advertise a MAC route for M1, but PE1 withdraws its Ethernet A-D per EVI route, then PE3 may still forward traffic destined to M1 to both PE1 and PE2. This is because the corresponding ES on PE1 might have failed and PE1 has withdrawn the associated routes but the withdrawal of Ethernet A-D per EVI has arrived ahead of Ethernet A-D per ES route.”

Cisco Confidential

Clarified time sequence example in section 9.2.2 - Cont.

T4’:

- From nothing in Rev 11
- To (rev 12): “If after T3, PE1 and PE2 both advertise a MAC route for M1, but PE1 withdraws its Ethernet A-D per EVI route and PE2 withdraws its MAC route for M1, then PE3 forwards traffic destined to M1 to PE2 only.”
- To (rev 13): “If after T3, PE1 and PE2 both advertise a MAC route for M1, but PE1 withdraws its Ethernet A-D per EVI route and PE2 withdraws its MAC route for M1, then PE3 forwards traffic destined to M1 to both PE1 and PE2 similar to T4. If PE3 later receives a withdrawal for Ethernet A-D per ES route from PE1, then it treats traffic to M1 as unknown unicast.”

Cisco Confidential

Next Steps

- WGLC

Cisco Confidential