draft-ietf-bess-rfc7432bis-10.txt

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

IETF 120, June 2024

Vancouver

Changes Relative to Rev07

1. Last update was provided in IETF 116 (March 2023)
2. All the changes since then are mainly clarifications in nature. But it is good to review them with the WG.
3. Changes along with their description are covered in next few pages.

Cisco Confidential

Change-1: Multihoming Redundancy Mode

- RFC7432bis defines two-bit field in the flag bits of ESI Label Extended Community instead of single-bit for RFC 7432
- The value for this two-bit field are assigned such that to endure backward compatibility with RFC 7432 (i.e., 0=All-Active, 1=Single-Active)
- There were some places in RFC7432bis that referred to the single-bit flag instead of two-bit field
- Now the entire document consistently refers to this two-bit field – i.e.,

Cisco Confidential

Change-2:
Section 6: Port-based Service Interface and Untagged User Data Traffic

- RFC7432 and RFC7432bis doesn’t explicitly talk about untagged user data traffic for port-based service interface
- Added couple of sentences to each of the sections 6.2.1 and 6.3.1 to clarify the handling of untagged user data traffic
- **Section 6.2.1:** Furthermore, untagged user data traffic is mapped to the same bridge table as all other tagged user data traffic (i.e., VLANs).
- **Section 6.3.1:** Furthermore, untagged user data traffic is mapped to its own bridge table under the same MAC-VRF associated with the EVI -- i.e., the MAC-VRF for a PE consists of one additional bridge table for untagged user data traffic.

Cisco Confidential

Change-3:
Originating Router’s IP Address

- Added the following paragraph to sections 8.1.1 and 11.1:
- “The Originating Router's IP Address field value MUST be set to an IP address of the PE (e.g., this address may be the PE's loopback address). The IP Address Length field is in bits. The Originating Router's IP address does not need to be a routable address and its purpose is to identify the originator of that EVPN route uniquely. It can be either IPv4 or IPv6 address independent of the BGP next hop address type for that NLRI and it must remain the same for all EVPN routes advertised by that PE (across all EVIs).”

Cisco Confidential

Change-4:
Section 8.6: Signaling Primary & Backup DF Elected PEs

- Modified section 8.6 to make it concise
- From: “Once the Primary and Backup-DF Elected PEs for a given <ES, EVI> are determined, the multi-homed PEs for that ES are able to set the P and B bits in the L2-Attr Extended Community to reflect the advertising PE's role for that EVI. Dual-homed PEs for that ES MAY advertise the P and B bits to reflect the advertising PE's role for that EVI, and muli-homed PEs of 3 or more peers for that ES SHOULD each advertise the P and B bits for that ES to prevent higher convergence time.”
- To: ”Once the Primary and Backup DF Elected PEs for a given <ES, EVI> are determined, the multi-homed PEs for that ES will each advertise an Ethernet A-D per EVI route for that EVI and each will include an L2-Attr extended community with the P and B bits set to reflect the advertising PE's role for that EVI.”

Cisco Confidential

Change-5:
MAC Address Aliasing in section 10.1

- Removed the paragraph corresponding to MAC address aliasing as it was not being used
- “If all PEs of a given EVPN instance act as a default gateway for that EVPN instance, and the same default gateway MAC address is used across all gateway devices, then no such advertisement is needed. However, if each default gateway uses a different MAC address, then each default gateway needs to be aware of other gateways' MAC addresses and thus the need for such an advertisement. This is called MAC address aliasing, since a single default gateway can be represented by multiple MAC addresses.”

Cisco Confidential

Change-6:
Section 11.2 – P-tunnel Identificaion

- The following text was added as clarification for IMET route to cover inter-AS option-B
- “A PE receiving an Inclusive Multicast Ethernet Tag route (with ingress replication as P-tunnel) SHOULD use the Next Hop field of the MP_REACH_NLRI attribute when resolving the route to an LSP.”

Cisco Confidential

Change-7:
Section 15.2 - sticky MAC

- Changed from:
- “If a PE receives such advertisements and later learns the same MAC address(es) via local learning, then the PE MUST alert the operator.”
- To:
- ”If a PE receives such advertisements, then it MAY program to drop any received frames with that MAC SA over its local ACs. When a PE later learns the same MAC address(es) via local learning for remote PEs or via a different ES for the advertising PE, then the PE MUST alert the operator and MAY drop the received frames.
- [RFC9135] describes symmetric and asymmetric IRB operation where an access-facing IRB interface is associcated with each subnet (i.e., VLAN). Each of these IRB interfaces is configured with a MAC address (typically Anycast) and an Anycast IP address. The MAC address associated with an IRB interface should be considered as sticky MAC address and be programmed as such for local ACs. If this MAC address is not Anycast, then it is advertised with both Gateway Extended EC and MAC Mobility EC with static flag set; however, if it is Anycast, then no EVPN MAC/IP route advertisement is needed.

Cisco Confidential

Change-8:
Section 15.3 – Loop protection

- Changed From:
- “When the PE programs M as a Black-Hole MAC in the Bridge Table, M is no longer associated to the backdoor Attachment Circuit (AC), but to a Black-Hole destination.”
- To:
- “When the PE programs M as a "Drop" MAC, then it drops any received frames with MAC-SA that is the same as "Drop" MAC (e.g., duplicate MAC). The PE MAY keep a counter for such discarded frames for each duplicate (dropped) MAC address or an aggregate counter for all duplicate (dropped) MAC addresses. The PE MAY program M as a "Drop" MAC on its local ACs if it receives from remote PE(s) a MAC/IP route update for M with the sticky-bit set (in the MAC Mobility extended community).”

Cisco Confidential

Change-9:
Section 18 – Frame Ordering

- Changed From:
- “If a network uses entropy labels per [RFC6790], then the control word SHOULD NOT be used.”
- To:
- “If a network (inclusive of all PE and P nodes) uses entropy labels per [RFC6790] for ECMP load balancing, then the control word may not be used.”

Cisco Confidential

Next Step

- Addressed all comments by Mike McBridge for early review by RTGDIR
- Addressed all comments by Stewart Bryant for early review by GENART
- Document is ready for wider IESG review

Cisco Confidential