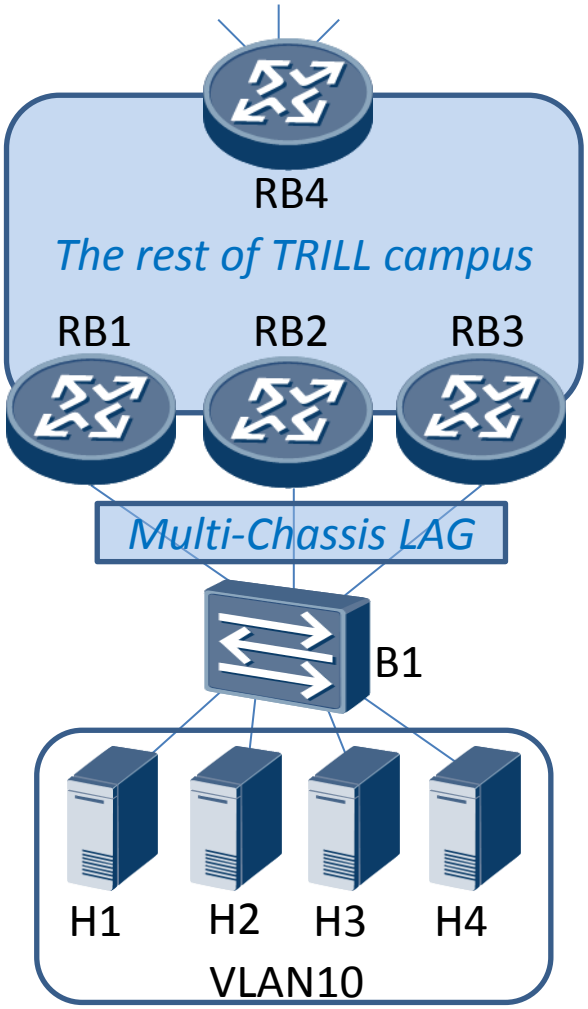
TRILL Active-Active Edge (AAE) Using Multiple MAC Attachments

Mingui, Radia, Hongjun, Mukhtiar, Muhammad

Overview

- AAE can balance the load of a VLAN among members.
- AAE can locally recover a link failure without triggering routing convergence.
- Issue: if RB1,2,3 ingress a packet using their own nicknames, RB4 will observe MAC flip-flop.



Arising potential solutions

- In order to address MAC flip-flopping
  - Solution1 : AAE RBridges use pseudo-nickname instead of their own as the ingress nickname
  - Solution2: AAE Rbridges split the work and assign responsible Rbridges for each MAC (e.g., MAC mod 3). If the receiver is the responsible one, encapsulates the packet and sends it out. Otherwise,
    - Known unicast: sends it using the nickname of its responsible RBridge
    - Multicast: redirects it to its responsible RBridge for encapsulation
  - Solution3: AAE RBridges use their own nicknames, remote RBridges keep multiple locations of each MAC per VLAN
- This draft explores solution 3. It requires no coordination among AAE RBridges. However, it requires changes on remote edge RBridges. So three ways have been proposed for backward compatibility.

No MAC flip-flopping

- The remote RBridge keeps all MAC locations.
- It keeps using the closest one, so no flip-flop.
  - E.g., RB4 keeps using RB1 as the egress RB for H1.

OLD: RB4's MAC table

MAC	VLAN	Egress RB
H1	10	RB1/RB2/RB3

Egress RB of H1 keeps changing

NEW: RB4's MAC table

#	MAC	VLAN	Egress RB
→ 1	H1	10	RB1
2	H1	10	RB2
3	H1	10	RB3

All egress RBs are remembered, while RB4 keeps using a single one.

Regular unicast/multicast ingress

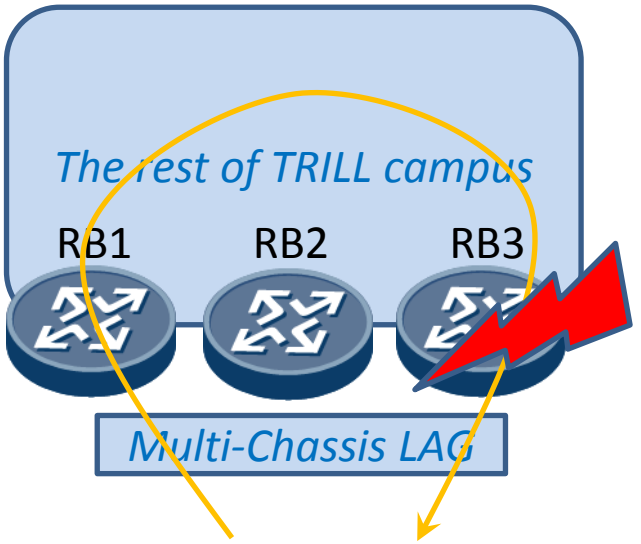
- AAE member RBridges simply ingress unicast and multicast frames per [RFC6325].
- No pseudonode nickname, no coordination

No duplication (single exit point)

- When a multicast is going to be egressed onto the MC-LAG, only one of the AAE members can act as the exit point (otherwise traffic duplication will happen).
  - Static algorithms, e.g., VLAN/MAC mod 3, can be used.
  - Let AAE members listen to LACP and use the hashing method of MC-LAG.

No Echo (Split Horizon)

- Echo of a multicast frame will cause a forwarding loop.
- Split horizon is the answer.
- The key for split horizon
  - ingress nickname+data label



RB3 splits horizon to prevent the forwarding loop.

No black-hole & no triangular forwarding

- If a link in the MC-LAG fails, the corresponding RB sends out MAC withdraw message to flush the MAC remembered by remote RBridges. This prevents black-hole issue.

- Other two solutions resort to tunneling technique to solve the black-hole issue, which requires new silicon. Also, the tunneling technique brings the triangle forwarding issue.

Load Balance Towards the AAE

- The remote RBridge can balance the load towards a MAC among its attachments.
- This is the south-bound flow-level traffic spreading.

Thanks!