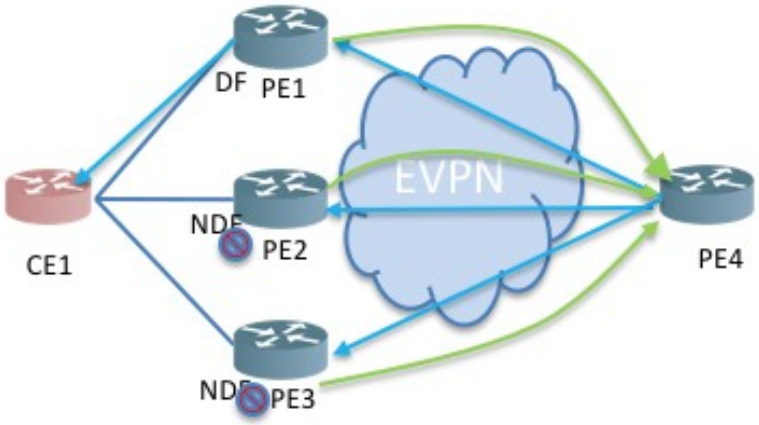
EVPN BUM Optimization

[draft-mohanty-bess-evpn-bum-opt-00]

Problem Statement

With increasing demand for EVPN multi-homing (> 2), it becomes increasingly challenging to control BUM traffic - which gets flooded per EVPN Instance to all multi-homed PEs.



- This is an unnecessary use of bandwidth in the EVPN Core.
- All NDF PE's receive traffic which they drop, and hence non-optimal use of bandwidth and line card processing.
- Remote PE's replicate a copy of the Ethernet Frame to all NDF PE's, which is later dropped. This consumes cycles at remote PE's.

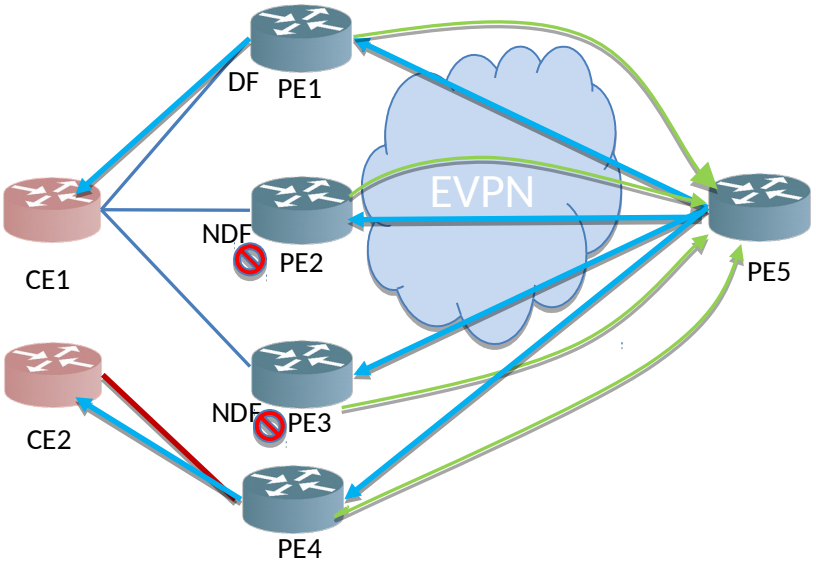
Initial Solution

- PE_k only needs to advertise “Inclusive Multicast Ethernet Tag route” (Type-3 route) for an EVPN Instance, EVI_i **if and only if** EVI_i is configured on **at least one** Ethernet Segment (which also has a presence in another PE_j, i.e Multihomed) and PE_k is the DF for that specific Ethernet Segment.

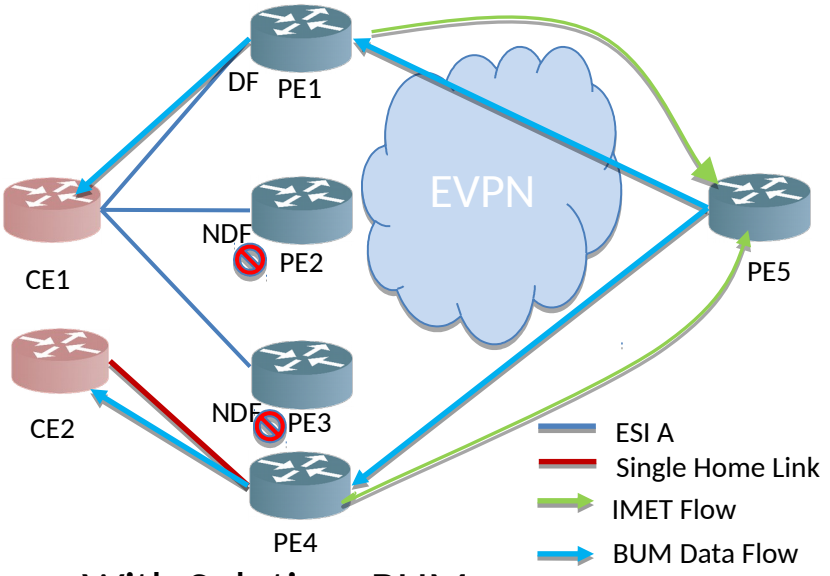
- The Type-3 should always be advertised for a “Single-Homed” Ethernet Segment on an EVI.

- For an EVPN Instance in a PE on the first DF Ethernet-Segment, the IMET should be advertised, whereas on the Last DF to Non-DF transition, it should be withdrawn

Initial Solution



Present Behavior of BUM flow for L2VRF



With Solution, BUM flow for L2VRF

Solution:

- PE2 and PE3 do not advertise IMET, as there is no other Ethernet-Segment for which these PEs are DF for that EVI
- PE1 advertises the IMET as has a Ethernet-Segment for which it is DF for that EVI.
- PE4 advertises IMET, as it has a “Single-Home” Ethernet Segment, for that EVI.
- PE5 thereby will only send BUM traffic to PE1 and PE4.

Caveats; With this approach, on a DF PE(PE1) failure, there will be BUM drop till the IMET from the next elected DF [PE2 or PE3], is received at PE5.

Note; in existing behavior, there is BUM drop based on Type 4 withdraw in the PEs. However with this approach, convergence delay will be MAX[Type 4, Type 3 Propagation delays] after the New DF is elected.

Final Solution

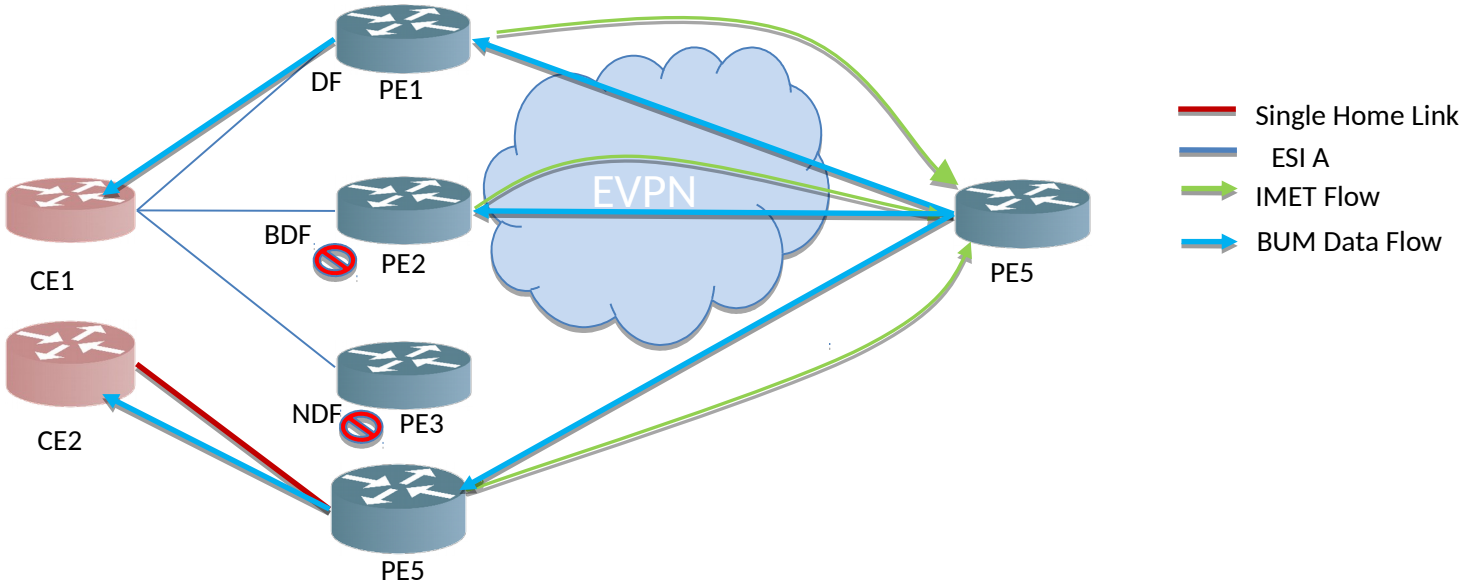
- Where failover performance cannot be compromised and the operator agrees to trade this off against BUM arriving at the “**Backup DF**” as determined by the DF election algorithm amongst the Multi home PEs, then the BDF also sends the Type-3 to attract BUM, in addition to the DF.

- Benefits increase with the number of Multihomed PEs.

- Improves convergence for known multicast streams, using EVPN IGMP[draft-sajassi-bess-evpn-igmp-mld-proxy] for known Multicast. SMET is presently only sent from the DF (Section 4); we are proposing to send SMET from the Backup DF, as well, to preserve convergence.

Final Solution

- Multi-homed PEs can easily compute the Backup DF, based on the DF election mode in operation.
- Proposing that a PE should only advertise Type-3 for an EVI **if and only if** one of the conditions hold:
  - It has an Single Home Ethernet Segment, in the EVI
  - It is DF for at least one Ethernet-Segment, for that EVI
  - It is BDF for at least one Ethernet-Segment, for that EVI



No convergence overhead, as the backup DF always receives the BUM traffic.

Reward / Advantages

- The proposal imparts additional control and limits the EVPN Flooding to the PE that really requires it. This ensures more optimal bandwidth usage of the EVPN core. This is especially true in ingress replication, with no/reduced protocol overhead.
- PEs that does not require BUM traffic are not sent BUM traffic. This increases the utilization at these PEs
- Solution applicable for Single-Active and All-Active MH scenarios.
- Requires no changes to the standardized protocol per se.
- Reduction in the number of EVPN Type-3 route advertisements
- Easy to detect by remote PE's, and to interop between vendors
- Also applicable to EVPN-IGMP to preserve convergence
- Benefit increases with the increase in the number of Multi-home PEs per ESI and the number of EVPN Instances. The benefit is more prominent with scale.