Technical Disclosure Commons

Defensive Publications Series

July 2020

PIM FLOODING MECHANISM AND SOURCE DISCOVERY (PFM-SD) EXTENSION TO AVOID FLOOD BETWEEN MULTI HOME PEER

Mankamana Mishra

Ali Sajassi

Ijsbrand Wijnands

Jayashree Subramanian

Anuj Budhiraja

PIM FLOODING MECHANISM AND SOURCE DISCOVERY (PFM-SD) EXTENSION TO AVOID FLOOD BETWEEN MULTI HOME PEER

AUTHORS:

Mankamana Mishra
Ali Sajassi
Ijsbrand Wijnands
Jayashree Subramanian
Anuj Budhiraja

ABSTRACT

Techniques are provided to support an extension to PFM-SD that avoids multicast traffic flooding across multi-home provide edge nodes, and maintains a faster convergence capability provided by multi-homing. These techniques allow a last hop router to create two trees, and provides a framework to ensure that Ethernet Segment failure has minimum traffic close for a receiver. In addition, these techniques involve a mechanism to avoid traffic flood over a core network between peers.

DETAILED DESCRIPTION

Ethernet Virtual Private Networking (EVPN) is being used to provide access redundancy. Multicast source and multicast receivers both reside behind a multi-homed segment. Though originally EVPN was designed and defined in context of a Layer-2 network, it has been proposed to use EVPN as a redundancy service provider for Layer 2 and Layer 3 access redundancy.

There are challenges with multicast sources behind an all active multi-homed segment.

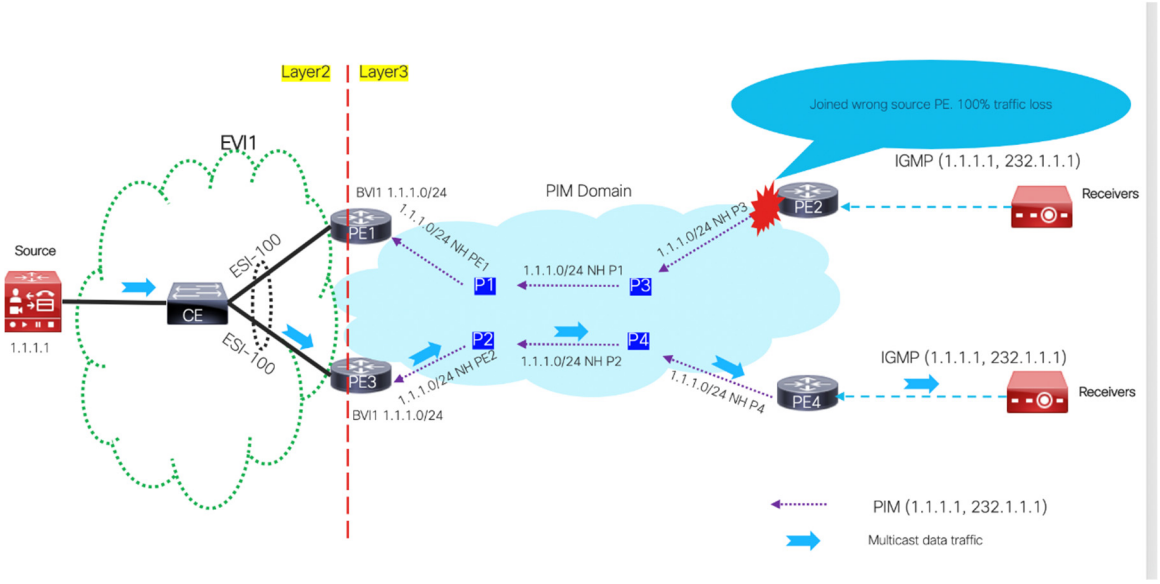

Figure 1

Consider the topology shown in Figure 1.

Multicast Source: Multicast source is sitting behind all active multi-homing segment.
Multicast Receivers: Multicast receivers are behind PIM domain in Layer 3 network.
Initial data flow: Multicast source (1.1.1.1) starts originating multicast traffic for group (232.1.1.1). Once the flow reaches a customer edge (CE) device, it does hashing and picks one of the link to send it to first hop router (FHR). The CE device picks provider edge node 3 (PE3) as the FHR and multicast traffic is sent to PE3.

Control plane operation: For simplicity Source-Specific Multicast (SSM) case is described here.

Receiver behind PE4

1. Receiver behind PE4 sends Internet Group Management Protocol (IGMP) join for (1.1.1.1, 232.1.1.1) to nearest multicast router. PE4 traditionally looks at Interior Gateway Protocol (IGP) to find the next hop to reach 1.1.1.0/24 prefix.
2. IGP provides P4 as next hop. PE4 originates PIM join (1.1.1.1, 232.1.1.1) to P4.
3. Same procedure continues at each hop till join reaches PE3.
4. PE3 forwards multicast traffic to newly built tree. Receiver behind PE4 starts getting multicast traffic.

Receiver behind PE2

1. Receiver behind PE2 sends IGMP join for (1.1.1.1, 232.1.1.1) to PE2. PE2 looks for IGP next hop to reach 1.1.1.0/24 and IGP points to P3. PE2 originates Protocol Independent Multicast (PIM) join towards P3.
2. P3 does exact same procedure and sends join to P1.
3. Finally, PIM join reaches PE1 (following IGP next hop to source prefix).
4. At this point of time, PE1 does not have any multicast traffic so there is no traffic flow along this multicast tree.

End result: Receiver behind PE2 does not get any multicast traffic.

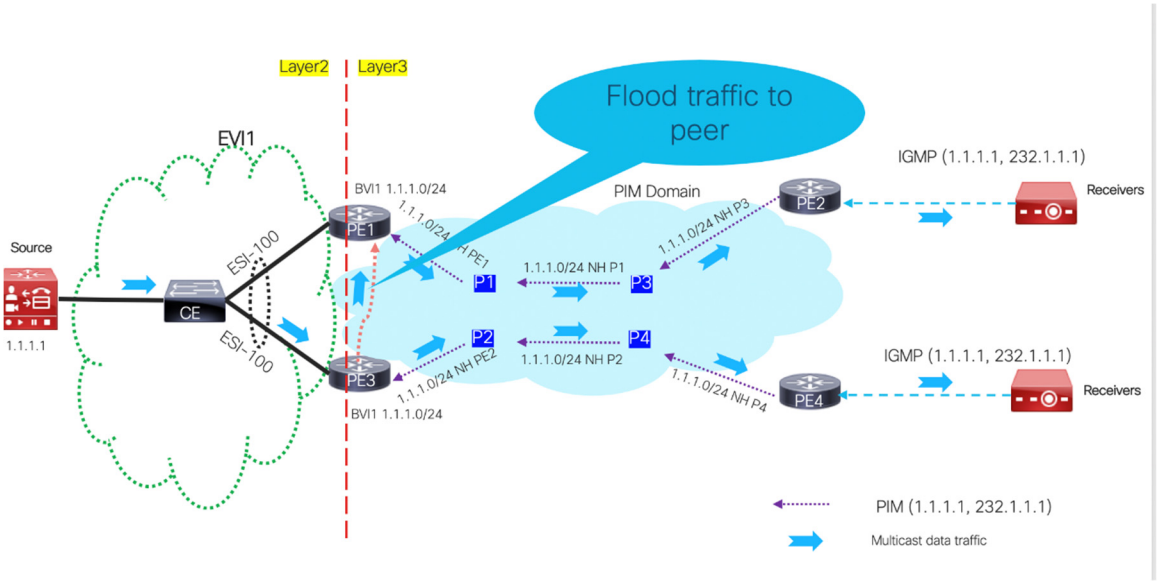

Figure 2

One of the brute force methods to solve this problem is to flood multicast traffic to all of the peers who are participating in all active multi-homing. This is shown in Figure 2 above. In this case, PE1 and PE3 both would have multicast traffic. Depending on which ingress router gets the multicast traffic, that ingress router would forward it to the multicast tree.

This solution is easy to implement and is valid for small scale implementations. However, bandwidth overutilization may occur from flooding of the traffic via the core network. Multicast traffic potentially could travel twice over the core network. Also, if multicast scale gets high, this would use up all bandwidth. Video traffic usually has high bandwidth flow. Therefore, this solution would not scale well for video traffic.

A scalable solution needs to be able to be protected behind all active multi-homing segments, be able to provide fast convergence in case of access or FHR failure, and void any extra flooding in the network.

To this end, an extension to PIM Flooding Mechanism Source Discovery (PFM-FM) is provided to achieve multicast source protection behind all active multi-homing segments, faster convergence in case of failure and avoidance of any extra flooding of multicast traffic in the core network.

(ES, Prefix) Announcement Using PFM

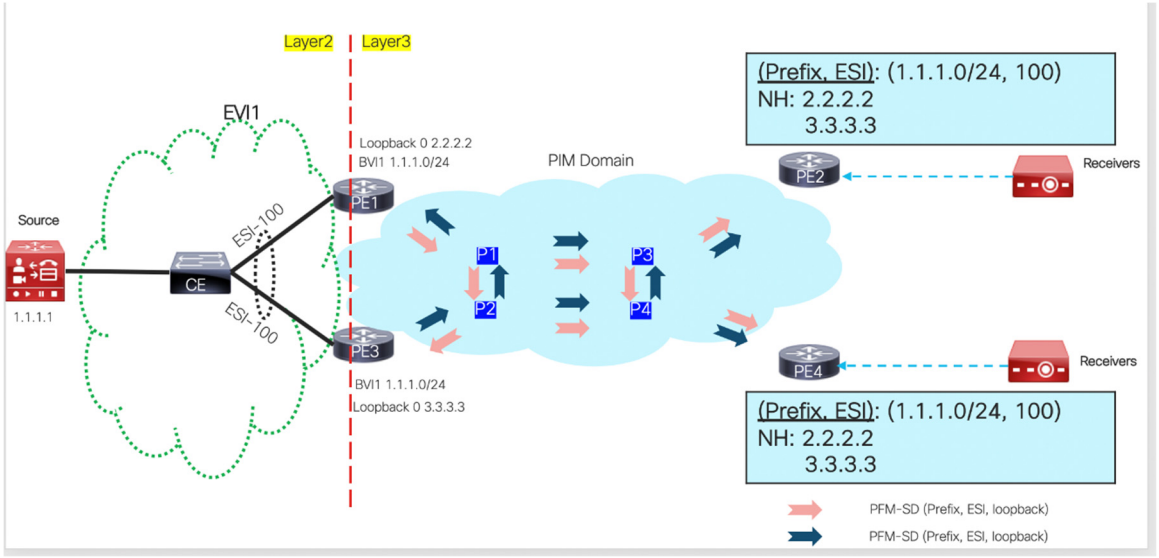

Figure 3

Reference is made to Figure 3 above. If a multicast source is being protected by an all active multi-homing segment, as soon as an Ethernet Segment (ES) comes up in given Bridge Domain, a new PFM message is originated in the PIM domain which carries that (ES, Prefix) mapping. It also carries a unique router ID that would be used to identify originator of (ES, Prefix) mapping.

In the example of Figure 3, when ES 100 comes up, PE1 originates (100, 1.1.1.0/24) Originator: 2.2.2.2 and PE3 originates (100, 1.1.1.0/24) Originator: 3.3.3.3 message. Once this message reaches PE2 and PE4, they maintain a mapping of next hop per (ES, Prefix).

Source-Specific Multicast (SSM)

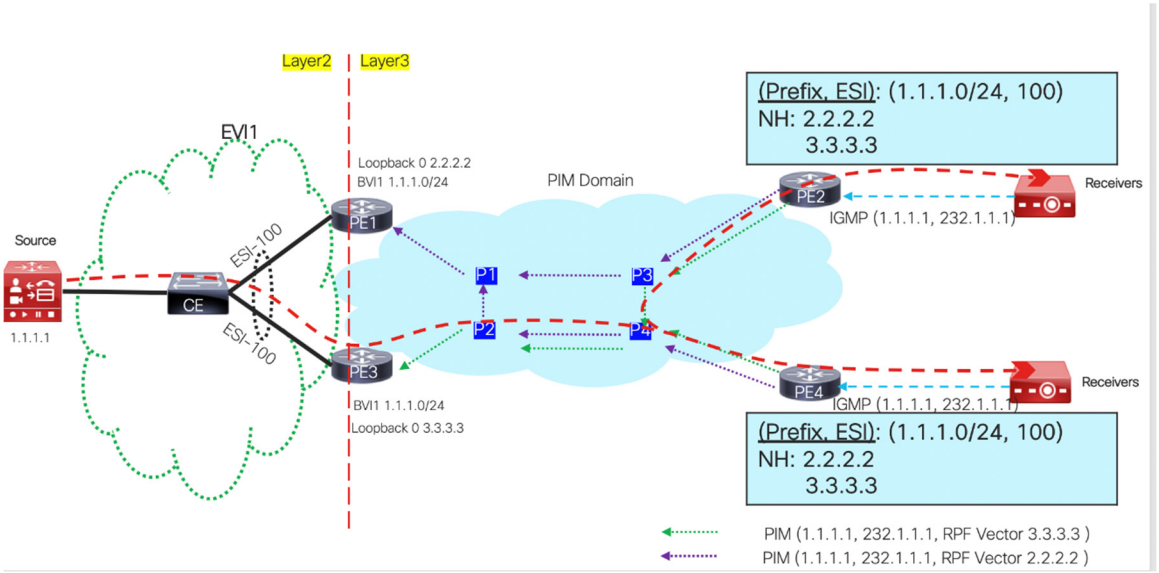

Figure 4

Control plane:

In case of SSM, when the Last Hop Router receives the (S,G) join, it would look at an internal database if there are multiple next hops stored for this prefix. If so, then it would originates two PIM targeted join requests to both of the next hops. There would be two multicast trees built towards both of the ingress PEs which are part of redundancy group.

Data Plane:

Since there was two trees built towards both of ingress routers, whichever ingress router gets the multicast traffic would forward it towards the receiver.

Reverse Path Forwarding (RPF) Check:

While sending a join request, it would not be possible to determine which tree the multicast traffic came from. An implementation can pick one of the trees as accepting while listening on both of the trees. Depending on which the tree traffic comes from, RPF can be updated accordingly. At any given point of time only one of the trees would be in the accepting state.

Node Failure / AC failure in SSM case:

Once the traffic starts reaching the redundancy peer and the multicast tree has been setup, traffic starts flowing right away. The last hop router now would stop getting multicast traffic from the initial RPF and secondary RPF starts getting multicast traffic.

Any-Source Multicast (ASM)

In the case of SSM, the multicast source are known in advance, whereas for ASM once multicast flow starts from source, whichever first hop router gets the multicast traffic would start PFM-SD with the procedure defined in PFM-SD-RFC . Once the last hop router receives source information, the rest of the procedure follows as for SSM. Since (ES, Prefix) mapping has already been advertised by the first message, even in the case of ASM, there would be a tree created towards both of the peers.

In summary, techniques are provided to support an extension to PFM-SD that avoids multicast traffic flooding across multi-home provide edge nodes, and maintains a faster convergence capability provided by multi-homing. These techniques allow a last hop router to create two trees, and provides a framework to ensure that Ethernet Segment failure has minimum traffic close for a receiver. In addition, these techniques involve a mechanism to avoid traffic flood over a core network between peers.