December 2019

# INTERWORKING BETWEEN LEGACY AND NEXT-GENERATION MULTICAST VIRTUAL PRIVATE NETWORK (MVPN) TRANSPORTS

Arvind Venkateswaran

Rishabh Parekh

Ijsbrand Wijnands

Jayashree Subramanian

# INTERWORKING BETWEEN LEGACY AND NEXT-GENERATION MULTICAST VIRTUAL PRIVATE NETWORK (MVPN) TRANSPORTS

AUTHORS:
Arvind Venkateswaran
Rishabh Parekh
Ijsbrand Wijnands
Jayashree Subramanian

## ABSTRACT

Proposed herein is a technique that allows next generation (next-gen) Provider Edge (PE) devices with minimal investment in legacy transports to interwork with legacy multicast Virtual Private Network (MVPN) PE devices.

## DETAILED DESCRIPTION

Customers and service providers are migrating from legacy to next-gen MVPN transports like Multiprotocol Label Switching (MPLS), Bit-Indexed Explicit Replication (BIER), and Tree Segment Identifier (Tree-SID). One option for migration is to have next-gen devices provide legacy and next-gen encapsulation and decapsulation. However, some customers want to deploy products that do not expend resources for supporting legacy transport but still need to interwork with legacy devices in the network.

This proposal provides a technique that allows devices that only support next-gen transports to interwork with legacy devices in a network. Consider a network as shown below in Figure 1 in which there may be some devices support only legacy transport and some devices support only next-gen transport. Further, consider that there may be devices within the network that support both legacy and next-gen transports and perform the role of stitching the legacy transport with the next-gen transport. For purposes of the present disclosure, these devices may be referred to as translation devices.

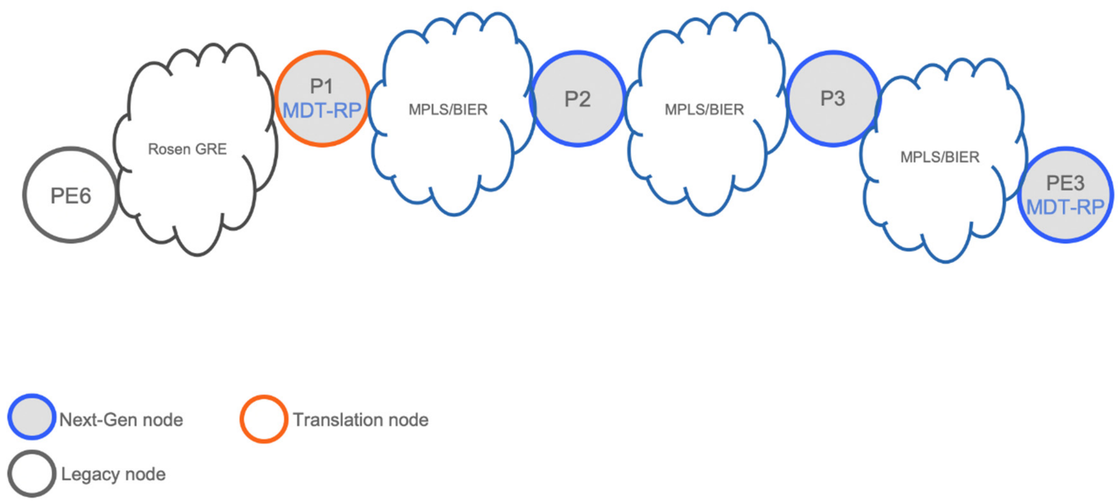


*Figure 1*

For Figure 1, consider that: Provider Edge (PE) device PE6 is a pure legacy device; P2, P3, and PE3 (which may be any combination of Area Border Routers (ABRs), Autonomous System Border Routers (ASBRs), and/or PEs) are next-gen (NG) devices; and P1 (ABR/ASBR/PE) is a translation device that supports legacy and next-gen core-transports. Note that there could be intermediate (P) routers in each of the segment clouds above that participate in legacy or next-gen transports.

A fundamental concept for the technique of this proposal is for translation device(s) to maintain a one-to-one mapping between an Internet Protocol (IP)-multicast flow and a context that is encoded in a next-gen transport. For example, consider an implementation in which legacy devices may utilize Protocol-Independent Multicast (PIM) procedures to build control plane trees, while next-gen devices may utilize Border Gateway Protocol (BGP) MVPN procedures to setup next-gen transport. In this example, the translation devices will utilize both PIM and BGP MVPN procedures to build and maintain the one-to-one mapping. In the data plane, the translation devices can utilize this mapping to translate from legacy to next-gen transport and vice versa.

Figures 2 and 3, below, illustrate features associated with the control plane and data plane flow for the proposed technique using ingress replication as the core-transport.

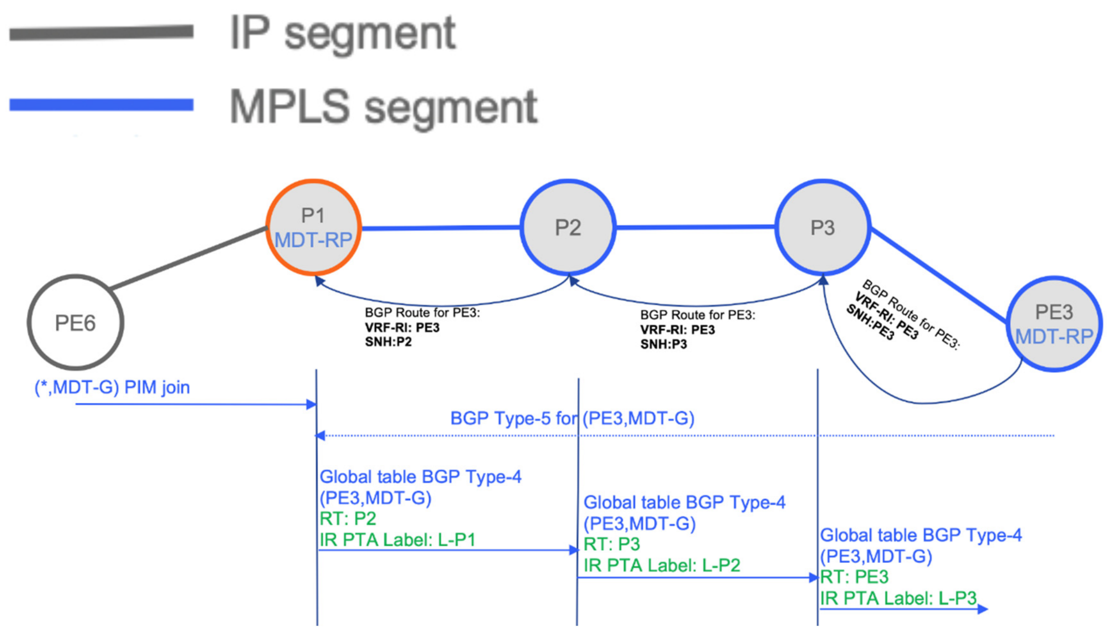


*Figure 2*

For the control plane, this technique may involve the announcement of unicast prefixes (of the translation and the edge devices) using BGP, as described in Internet Engineering Task Force (IETF) Request For Comments (RFC) 7524. For Figure 2, P1 may be the rendezvous point (RP) for the Rosen Generic Routing Encapsulation (GRE) Multicast Distribution Tree (MDT) group address. During operation, legacy PE node PE6 may send a (*,MDT-G) join to P1. NG PE, PE3, may announce itself as a source participating in the Rosen GRE group in which PE3 may announce a BGP source-active message for the announcement. Once the source is discovered, the translation device (P1, in this example) builds the core-tree hop-by-hop towards the edge NG device using BGP Leaf-AD (auto-discovery) routes (P1->P2->P3->PE3, in the above example).

For data plane communications, as shown in Figure 3, PE3 encapsulates customer payload with MPLS ingress-replication encapsulation with a label corresponding to the Rosen GRE (S,G) (source, group) announced for the customer flow. As described above the translation device, P1, has a one-to-one mapping of the Rosen GRE (S,G) state to the ingress-replication label. Upon receiving the encapsulated customer payload, P1 can strip the ingress-replication label and re-impose the Rosen GRE encapsulation. PE6 operating as a legacy PE can strip the Rosen GRE header and forward out the customer payload.

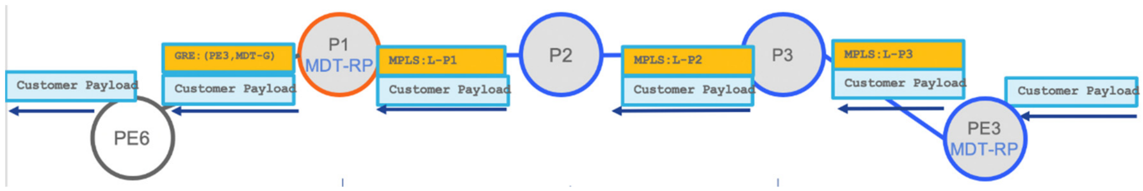


*Figure 3*

Figure 4, below, illustrates an example implementation similar to the example discussed above with the source and the receiver positions swapped.

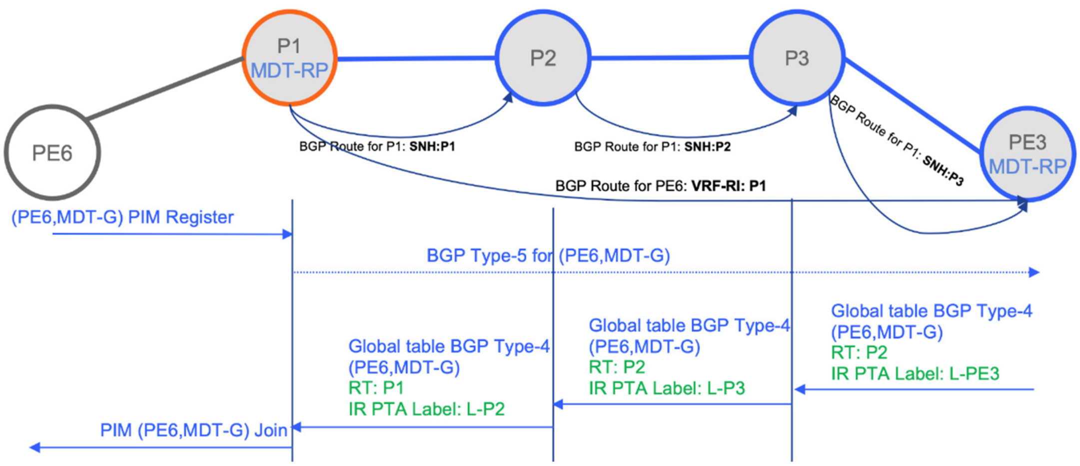


*Figure 4*

For the control plane with regard to Figure 4, P1 is the RP for the Rosen GRE MDT group address. As shown in Figure 4, legacy PE node PE6 sends a PIM register to P1. Upon receiving the PIM register message, P1 triggers a BGP Source-active message. Once the source is discovered, NG PE PE3 builds the core-tree hop-by-hop towards the P1 using BGP Leaf-AD routes (PE3->P3->P2->P1, in this example).

For data plane communications for this example, as shown below in Figure 5, PE6 encapsulates the customer payload with Rosen GRE encapsulation. The translation device, P1, has a one-to-one mapping of the Rosen GRE (S,G) state to the ingress-replication label. Upon receiving the encapsulated customer payload, P1 will strip the Rosen GRE header and re-impose the ingress replication MPLS encapsulation (L-P2). PE3 can decapsulate the MPLS packet and forward the customer payload.

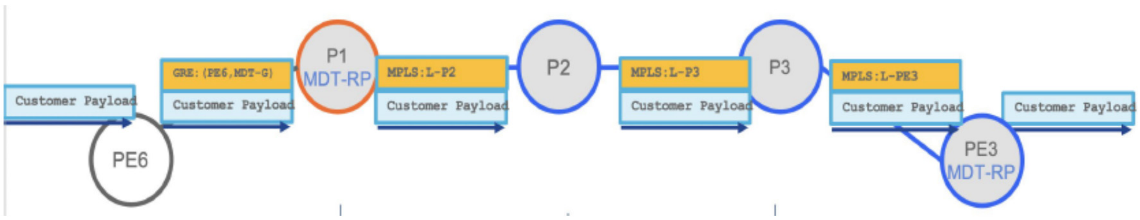


*Figure 5*

Although the above examples/illustrations use ingress replication as the core-transport, the technique of this proposal can be realized using procedures for other NG core-transports, as described below, in which the control plane procedures may differ slightly from those described above.

For example, when a NG node receives a BGP Leaf-AD route from a downstream node, as shown in the illustrations above, it can originate a BGP Type-3 Selective P-Multicast Service Interface) S-PMSI for the (S,G) with a PMSI tunnel attribute that carries the core-context. In other examples, this can be a Forwarding Equivalent Class (FEC) in a multicast Label Distribution Protocol (mLDP) implementation, a tunnel-identifier a Resource Reservation Protocol-Traffic Engineering (RSVPTE) implementation, a Tree-Identifier in a Tree-SID implementation, or an upstream assigned context in a BIER implementation. A downstream router can join the core-tree with the context learned from the BGP Type-3 S-PMSI route and can build a one-to-one mapping between this core-context and the Rosen GRE (S,G) state.

Thus, the technique proposed herein allows next-gen mVPN core-transport protocols to interwork with legacy transports using a translation node in which different MVPN transports, specifically legacy IP to next-gen MPLS or BIER, can be stitched together via the translation node. Additionally, the translation node can translate control-plane message between different MVPN transport setup protocols (e.g., PIM <-> BGP). A one-to-one mapping translation table can be maintained for different MVPN transports to enable stitching. For each lookup in the translation table, a specific lookup key can be created (e.g., MPLS, Multicast IP version 4 (IPv4), Multicast IP version 6 (IPv6), BIER, etc.) in the translation table. Thus, the technique of this proposal solves problems with migration for which there may currently be no other solutions.

While existing solutions for MVPN transport migration typically involve a new PE router to also support legacy transport (e.g., dual-encap on both legacy and next-gen transport), the technique proposed herein allows a next-gen PE to support only next-gen MVPN transport but still interoperate with other PEs that support only the legacy transport. A translation device can stitches (instead of encapsulation) next-gen mVPN transports to legacy transports. Hence, the technique does not require investment in legacy transports on a new platform in order to interwork with legacy devices in a customer network, thereby reducing the cost of the newer devices.

Accordingly, the technique proposed herein allows next-gen MVPN core-transport protocols to be stitched with legacy transports using a translation node. The technique does not involve configuration changes on existing legacy devices. Rather, the technique allows customers to plug next-gen devices into existing networks and provide interworking with the legacy devices. Additionally, the technique does not involve software or hardware upgrades on existing legacy devices. Further, the technique does not involve protocol extensions for next-gen nodes to communicate with the translation node. Still further, the technique does not involve investment in legacy transports to interwork with legacy devices in a customer network.

In summary, this proposal provides a technique that allows next-gen PE devices with minimal investment in legacy transports to interwork with legacy MVPN PE devices.