BESS Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: November 7, 2021

G. Mishra
Verizon Inc.
M. Mishra
Cisco Systems
J. Tantsura
L. Wang
Juniper Networks, Inc.
Q. Yang
Arista Networks
A. Simpson
Nokia
S. Chen
Huawei Technologies
May 6, 2021

Deployment Guidelines for Edge Peering IPv4-NLRI with IPv6-NH
draft-ietf-bess-deployment-guide-ipv4nlri-ipv6nh-00

Abstract

As Enterprises and Service Providers upgrade their brown field or green field MPLS/SR core to an IPv6 transport, Multiprotocol BGP (MP-BGP)now plays an important role in the transition of their Provider (P) core network as well as Provider Edge (PE) Edge network from IPv4 to IPv6. Operators must be able to continue to support IPv4 customers when both the Core and Edge networks are IPv6-Only.

This document details an important External BGP (eBGP) PE-CE Edge IPv6-Only peering design that leverages the MP-BGP capability exchange by using IPv6 peering as pure transport, allowing both IPv4 Network Layer Reachability Information (NLRI) and IPv6 Network Layer Reachability Information (NLRI)to be carried over the same (Border Gateway Protocol) BGP TCP session. The design change provides the same Dual Stacking functionality that exists today with separate IPv4 and IPv6 BGP sessions as we have today. With this design change from a control plane perspective a single IPv6 is required for both IPv4 and IPv6 routing updates and from a data plane forwarindg perspective an IPv6 address need only be configured on the PE and CE interface for both IPv4 and IPv6 packet forwarding.

This document provides a much needed solution for Internet Exchange Point (IXP) that are facing IPv4 address depletion at large peering points. With this design, IXP can now deploy PE-CE IPv6-Only eBGP Edge peering design to eliminate IPv4 provisioning at the Edge. This core and edge IPv6-Only peering design paradigm change can apply to any eBGP peering, public internet or private, which can be either Core networks, Data Center networks, Access networks or can be any

eBGP peering scenario. This document provides interoperability test cases for the IPv6-Only peering design as well as test results between five major vendors stakeholders in the routing and switching indusrty, Cisco, Juniper, Arista, Nokia and Huawei. With the test results provided for the IPv6-Only Edge peering design, the goal is that all other vendors around the world that have not been tested will begin to adopt and implement this new Best Current Practice for eBGP IPv6-Only Edge peering.

As this issue with IXP IPv4 address depletion is a critical issue around the world, it is imperative for an immediate solution that can be implemented quickly. This Best Current Practice IPv6-only eBGP peering design specification will help proliferate IPv6-Only deployments at the eBGP Edge network peering points to starting immediately at a minimum with operators around the world using Cisco, Juniper, Arista, Nokia and Huawei. As other vendors start to implement this Best Current Practice, the IXP IPv4 address depletion gap will eventually be eliminated.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

As Enterprises and Service Providers upgrade their brown field or

green field MPLS/SR core to an IPv6 transport such as MPLS LDPv6, SR-MPLSv6 or SRv6, Multiprotocol BGP (MP-BGP) now plays an important role in the transition of the Provider (P) core networks and Provider Edge (PE) edge networks from IPv4 to IPv6. Operators have a requirement to support IPv4 customers and must be able to support IPv4 address family and Sub-Address-Family Virtual Private Network (VPN)-IPv4, and Multicast VPN IPv4 customers.

IXP are also facing IPv4 address depletion at their peering points, which are large Layer 2 transit backbones that service providers peer and exchange IPv4 and IPv6 Network Layer Reachability Information (NLRI). Today, these transit exchange points are Dual Stacked. With

this IPv6-only BGP peering design, only IPv6 is configured on the PE-CE interface, the Provider Edge (PE) - Customer Edge (CE), the IPv6 BGP peer is now used to carry IPv4 (Network Layer Reachability Information) NLRI over an IPv6 next hop using IPv6 next hop encoding defined in [RFC8950], while continuing to forward both IPv4 and IPv6 packets. In the framework of this design the PE is no longer Dual Stacked. However in the case of the CE, PE-CE link CE side of the link is no longer Dual Stacked, however all other internal links within the CE domain may or maynot be Dual stacked.

MP-BGP specifies that the set of usable next-hop address families is determined by the Address Family Identifier (AFI) and the Subsequent Address Family Identifier (SAFI). Historically the AFI/SAFI definitions for the IPv4 address family only have provisions for advertising a Next Hop address that belongs to the IPv4 protocol when advertising IPv4 or VPN-IPv4. [RFC8950] specifies the extensions necessary to allow advertising IPv4 NLRI, Virtual Private Network Unicast (VPN-IPv4) NLRI, Multicast Virtual Private Network (MVPN-IPv4) NLRI with a Next Hop address that belongs to the IPv6 protocol. This comprises of an extended next hop encoding MP-REACH BGP capability exchange to allow the address of the Next Hop for IPv4 NLRI, VPN-IPv4 NLRI and MVPN-IPv4 NLRI to also belong to the IPv6 Protocol. [RFC8950] defines the encoding of the Next Hop to determine which of the protocols the address actually belongs to, and a new BGP Capability allowing MP-BGP Peers to discover dynamically whether they can exchange IPv4 NLRI and VPN-IPv4 NLRI with an IPv6 Next Hop.

The current specification for carrying IPv4 NLRI of a given address

family via a Next Hop of a different address family is now defined in [RFC8950], and specifies the extended next hop encoding MP-REACH capability extension necessary to do so. This comprises an extension of the AFI/SAFI definitions to allow the address of the Next Hop for IPv4 NLRI or VPN-IPv4 NLRI to belong to either the IPv4 or the IPv6 protocol, the encoding of the Next Hop information to determine which of the protocols the address belongs to, and a new BGP Capability allowing MP-BGP peers to dynamically discover whether they can exchange IPv4 NLRI and VPN- IPv4 NLRI with an IPv6 Next Hop.

With the new extensions defined in [RFC8950] supporting NLRI and next hop address family mismatch, the BGP peer session can now be treated as a pure TCP transport and carry both IPv4 and IPv6 NLRI at the Provider Edge (PE) - Customer Edge (CE) over a single IPv6 TCP session. This allows for the elimination of dual stack from the PE-CE peering point, and now enable the peering to be IPv6-ONLY. The elimination of IPv4 on the PE-CE peering points translates into OPEX expenditure savings of point-to-point infrastructure links as well as /31 address space savings and administration and network management

of both IPv4 and IPv6 BGP peers. This reduction decreases the number of PE-CE BGP peers by fifty percent, which is a tremendous cost savings for operators.

While the savings exists at the Edge eBGP PE-CE peering, on the core side PE to Route Reflector (RR) peering carrying <AFI/SAFI> IPv4 <1/1>, VPN-IPV4 <1/128>, and Multicasat VPN <1/129>, there is no savings as the Provider (P) Core is IPv6 Only and thus can only have an IPv6 peer and must use [RFC8950] extended next hop encoding to carrying IPv4 NLRI IPV4 <2/1>, VPN-IPV4 <2/128>, and Multicasat VPN <2/129> over an IPv6 next hop.

This document provides a much needed solution for Internet Exchange Point (IXP) that are facing IPv4 address depletion at large peering points. With this design, IXP can now use deploy PE-CE IPv6-Only eBGP Edge peering design to eliminate IPv4 provisioning at the Edge. This core and edge IPv6-Only peering design paradigm change can apply to any eBGP peering, public internet or private, which can be either Core networks, Data Center networks, Access networks or can be any eBGP peering scenario. This document provides interoperability test cases for the IPv6-Only peering design as well as successful test results between five major vendors stakeholders in the routing and

switching indusrty, Cisco, Juniper, Arista, Nokia and Huawei. With the test results provided for the IPv6-Only Edge peering design, the goal is that all other vendors around the world that have not been tested will begin to adopt and implement this new best practice for eBGP IPv6-Only Edge peering.

As this issue with IXP address depletion is a critical issue around the world, it is imperative for an immediate solution that can be implemented quickly. This best practice IPv6-only eBGP peering design specification will help proliferate IPv6-Only deployments at the eBGP Edge network peering points starting immediately at a minimum with operators around the world using Cisco, Juniper, Arista, Nokia and Huawei.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Terminology

Terminolgoy used in defining the IPv6-Only Edge specification.

AFBR: Address Family Border Router Provider Edge (PE).

Edge: PE-CE Edge Network Provider Edge - Customer Edge

Core: P Core Network Provider (P)

4to6 Softwire : IPv4 edge over an IPv6-Only core

6to4 Softwire: IPv6 edge over an IPv4-Only core

E2E: End to End

4. IPv6-Only Edge Peering Architecture

4.1. Problem Statement

This specification addresses a real issue that has been discussed at many operator groups around the world related to IXP major peering points where hundreds of AS's have both IPv4 and IPv6 dual stacked peering. IPv4 address depletion have been a major issue issue for many years now. Operators around the world are clamoring for a solution that can help solve issues related to IPv4 address depletion at these large IXP peering points. With this solution IXPs as well as all infrastructure networks such as Core networks, DC networks, Access networks as well as any PE-CE public or private network can now utilize this IPv6-Only Edge solution and reap the benefits immediately on IPv4 address space saving.

```
                              IXP Problem Statement

                       Dual Stacked            Dual Stacked
                            CE                      PE

                        +-------+ IPv4 BGP Peer +-------+
                        |       |---------------|       |
                        |   CE  | IPv6 BGP Peer |  PE   |
                        |       |---------------|       |
                        +-------+               +-------+
                        IPv4 forwarding           IPv4 forwarding
                        IPv6 forwarding           IPv6 forwarding
```

Figure 1: Problem Statement - IXP Dual Stack Peering


```
                               _________
    Dual Stacked     _____    /         \             Dual Stacked
      PE / CE       /     \__/           \___            PE / CE
  +-----+  +-----+   /                        \       +-------+   +------+
  |     |  |     |  |0====VPN Overlay Tunnel ==0|     |       |   |      |
  |     |  |     |  |                            \    |       |   |      |
  | CE  |--| PE  |--\         IPv6-Only Core     |----|  PE   |---|  CE  |
  |     |  |     |   \0=========Underlay =======0|    |       |   |      |
```

```
  +-----+  +-----+     \                          __/     +-------+   +------+
  IPv4 IPv6 BGP peer    \ IP / MPLS / SR domain /      IPv4 and IPv6 BGP peer
  IPv4 forwarding        \__          __       /          IPv4 forwarding
  IPv6 forwarding           \_______/  \_____/            IPv6 forwarding
```

Figure 2: Problem Statement - E2E Dual Stack Edge

4.2. IPv6-Only PE-CE Design Solution

The IPv6-Only Edge design solution provides a means of E2E single protocol design solution extension of [RFC5565] Softwire Mesh framework from the PE-CE Edge to the Core from ingres so egress through the entire operators domain. This solution eliminates all IPv4 addressing from end to end while still providing the same Dual Stack functionality of IPv4 and IPv6 packet forwarding from a data plane perspective by leveraging the [RFC8950] extended next hop encoding so that IPv4 NLRI can be advertised over a single IPv6 pure transport TCP session. This IPv6-Only E2E architecture eliminates all IPv4 peering and IPv4 addressing E2E from the ingress CE to ingress PE to egress PE to egress CE and all hops along the operator E2E path.

```
                   Solution applicable to
        any Edge peering scenario - IXP, Core, DC, Access, etc

         +--------+                +--------+
         |        |  IPv6 Only     |        |
         |   CE   |----------------|   PE   |
         |        |  IPv6 BGP Peer |        |
         +--------+                +--------+
        IPv4 forwarding             IPv4 forwarding
        IPv6 forwarding             IPv6 forwarding
```

Figure 3: IPv6-Only Solution Applicability

```
      IPv6-Only       _____    /        \                 IPv6-Only
       PE / CE       /     \__/          \___              PE / CE
   +----+  +----+   /                        \        +------+   +-----+
   |    |  |    |  |0====VPN Overlay Tunnel ==0|      |      |   |     |
   |    |  |    |  |                             \    |      |   |     |
   | CE |--| PE |--\         IPv6-Only Core      |----|  PE  |---|  CE |
   |    |  |    |    \0=========Underlay ===== ==0    |      |   |     |
   +----+  +----+     \                        __/    +------+   +-----+
   IPv6 BGP peer        \IP / MPLS / SR domain /        IPv6 BGP peer
   IPv4 forwarding      \__         __        /         IPv4 forwarding
   IPv6 forwarding         \_______/  \_____/           IPv6 forwarding
```

Figure 4: E2E VPN Solution

4.3. IPv6-Only Edge Peering Design

4.3.1. IPv6-Only Edge Peering Packet Walk

The IPv6-Only Edge Peering design utilizes two key E2E Softwire Mesh Framework scenario's, 4to6 softwire and 6to4 softwire. The Softwire mesh framework concept is based on the overlay and underlay MPLS or SR based technology framework, where the underlay is the transport layer and the overlay is a Virtual Private Network (VPN) layer, and is the the tunneled virtualization layer containing the customer payload. The concept of a 6to4 Softwire is based on transmission of IPv6 packets at the edge of the network by tunneling the IPv6 packets over an IPv4-Only Core. The concept of a 4to6 Softwire is also based on transmission of IPv4 packets at the edge of the network by tunneling the IPv4 packets over an IPv6-Only Core.

This document describes End to End (E2E) test scenarios that follow a packet flow from IPv6-Only attachment circuit from ingress PE-CE to egress PE-CE tracing the routing protocol control plane and data plane forwarding of IPv4 packets in a 4to6 softwire or 6to4 softwire within the IPv4-Only or IPv6-Only Core network. In both secneario we are focusing on IPv4 packets and the control plane and data plane forwarding aspects of IPv4 packets from the PE-CE Edge network over an IPv6-Only P (Provider) core network or IPv4-Only P (Provider) core network. With this IPv6-Only Edge peering design, the Softwire Mesh Framework is not extended beyond the Provider Edge (PE) and continues to terminate on the PE router.

4.3.2. 6to4 Softwire IPv4-Only Core packet walk

6to4 softwire where IPv6-Edge eBGP IPv6 peering where IPv4 packets at network Edge traverse a IPv4-Only Core

In the scenario where IPv4 packets originating from a PE-CE edge are tunneled over an MPLS or Segment Routing IPv4 underlay core network, the PE and CE only have an IPv6 address configured on the interface. In this scenario the IPv4 packets that ingress the CE from within the CE AS are over an IPv6-Only interface and are forwarded to an IPv4 NLRI destination prefix learned from the Pure Transport Single IPv6 BGP Peer. In the IPv6-Only Edge peering architecture the PE is IPv6-Only as all PE-CE interfaces are IPv6-Only. However, on the CE, the PE-CE interface is the only interface that is IPv6-Only and all other interfaces may or may not be IPv6-Only. Following the data plane packet flow, IPv4 packets are forwarded from the ingress CE to the IPv6-Only ingress PE where the VPN label imposition push per prefix, per-vrf, per-CE occurs and the labeled packet is forwarded over a 6to4 softwire IPv4-Only core, to the egress PE where the VPN label disposition pop occurs and the native IPv4 packet is forwarded to the egress CE. In the reverse direction IPv4 packets are forwarded from the egress CE to egress PE where the VPN label imposition per prefix, per-vrf, per-CE push occurs and the labeled packet is forwarded back over the 6to4 softwire IPv4-Only core, to the ingress PE where the VPN label disposition pop occurs and the native IPv4 packet is forwarded to the ingress CE. . The functionality of the IPv4 forwarding plane in this scenario is identical from a data plane forwarding perspective to Dual Stack IPv4 forwarding scenario.

```
                      +--------+   +--------+
                      |  IPv4  |   |  IPv4  |
                      | Client |   | Client |
                      | Network|   | Network|
                      +--------+   +--------+
                          |   \     /   |
                          |    \   /    |
                          |     \ /     |
                          |      X      |
                          |     / \     |
                          |    /   \    |
                          |   /     \   |
                      +--------+   +--------+
                      |  AFBR  |   |  AFBR  |
                   +--| IPv4/6 |---| IPv4/6 |--+
                   |  +--------+   +--------+  |
   +--------+      |                           |       +--------+
   |  IPv4  |      |                           |       |  IPv4  |
   | Client |      |                           |       | Client |
   | Network|------|            IPv4           |-------| Network|
   +--------+      |            only           |       +--------+
                   |                           |
                   |  +--------+   +--------+  |
                   +--|  AFBR  |---|  AFBR  |--+
                      | IPv4/6 |   | IPv4/6 |
                      +--------+   +--------+
                        |   \     /   |
                        |    \   /    |
                        |     \ /     |
                        |      X      |
                        |     / \     |
                        |    /   \    |
                        |   /     \   |
                     +--------+   +--------+
                     |  IPv6  |   |  IPv4  |
                     | Client |   | Client |
                     | Network|   | Network|
                     +--------+   +--------+
```

Figure 5: 6to4 Softwire - IPv6 Edge over an IPv4-Only Core

4.3.3. 4to6 Softwire IPv6-Only Core packet walk

4to6 softwire where IPv6-Edge eBGP IPv6 peering where IPv4 packets at network Edge traverse a IPv6-Only Core

In the scenario where IPv4 packets originating from a PE-CE edge are tunneled over an MPLS or Segment Routing IPv4 underlay core network, the PE and CE only have an IPv6 address configured on the interface. In this scenario the IPv4 packets that ingress the CE from within the CE AS are over an IPv6-Only interface and are forwarded to an IPv4 NLRI destination prefix learned from the Pure Transport Single IPv6 BGP Peer. In the IPv6-Only Edge peering architecture the PE is IPv6-Only as all PE-CE interfaces are IPv6-Only. However, on the CE, the PE-CE interface is the only interface that is IPv6-Only and all other interfaces may or may not be IPv6-Only. Following the data plane packet flow, IPv4 packets are forwarded from the ingress CE to the IPv6-Only ingress PE where the VPN label imposition push per prefix, per-vrf, per-CE occurs and the labeled packet is forwarded over a 4to6 softwire IPv6-Only core, to the egress PE where the VPN label disposition pop occurs and the native IPv4 packet is forwarded to the egress CE. In the reverse direction IPv4 packets are forwarded from the egress CE to egress PE where the VPN label imposition per prefix, per-vrf, per-CE push occurs and the labeled packet is forwarded back over the 4to6 softwire IPv6-Only core, to the ingress PE where the VPN label disposition pop occurs and the native IPv4 packet is forwarded to the ingress CE. . The functionality of the IPv4 forwarding plane in this scenario is identical from a data plane forwarding perspective to Dual Stack IPv4 forwarding scenario.

```
                             +--------+   +--------+
                             |  IPv4  |   |  IPv4  |
                             | Client |   | Client |
                             | Network|   | Network|
                             +--------+   +--------+
                                 |   \     /   |
                                 |    \   /    |
                                 |     \ /     |
                                 |      X      |
                                 |     / \     |
                                 |    /   \    |
                                 |   /     \   |
                             +--------+   +--------+
                             |  AFBR  |   |  AFBR  |
                          +--| IPv4/6 |---| IPv4/6 |--+
                          |  +--------+   +--------+  |
          +--------+      |                           |       +--------+
          |  IPv6  |      |                           |       |  IPv6  |
          | Client |      |                           |       | Client |
          | Network|------|           IPv6            |-------| Network|
          +--------+      |           only            |       +--------+
                          |                           |
                          |  +--------+   +--------+  |
                          +--|  AFBR  |---|  AFBR  |--+
                             | IPv4/6 |   | IPv4/6 |
```

```
                             +--------+   +--------+
                               |   \     /   |
                               |    \   /    |
                               |     \ /     |
                               |      X      |
                               |     / \     |
                               |    /   \    |
                               |   /     \   |
                            +--------+   +--------+
                            |  IPv4  |   |  IPv4  |
                            | Client |   | Client |
                            | Network|   | Network|
                            +--------+   +--------+
```

        Figure 6: 4to6 Softwire - IPv4 Edge over an IPv6-Only Core

4.4. RFC5549 and RFC8950 Applicability

4.4.1. IPv6-Only Edge Peering design next-hop encoding

This section describes [RFC8950] next hop encoding updates to [RFC5549] applicability to this specification. IPv6-only eBGP Edge PE-CE peering to carry IPv4 Unicast NLRI <AFI/SAFI> IPv4 <1/1> over an IPv6 next hop BGP capability extended hop encoding IANA capability codepoint value 5 defined is applicable to both [RFC5549] and [RFC8950] as IPv4 Unicast NLRI <AFI/SAFI> IPv4 <1/1> does not change in the RFC updates.

IPv4 packets over an IPv6-Only core 4to6 Softwire E2E packet flow is part of the IPv6-Only design vendor interoperaiblity test cases and in that respect is applicable as [RFC8950] updates [RFC5549] for <AFI/SAFI> VPN-IPV4 <1/128>, and Multicasat VPN <1/129>

4.4.2. RFC8950 updates to RFC5549 applicability

This section describes the [RFC8950] next hop encoding updates to [RFC5549]

In [RFC5549] when AFI/SAFI 1/128 is used, the next-hop address is encoded as an IPv6 address with a length of 16 or 32 bytes. This document modifies how the next-hop address is encoded to accommodate all existing implementations and bring consistency with VPNv4oIPv4 and VPNv6oIPv6. The next-hop address is now encoded as a VPN-IPv6 address with a length of 24 or 48 bytes [RFC8950] (see Sections 3 and 6.2 of this document). This change addresses Erratum ID 5253 (Err5253). As all known and deployed implementations are interoperable today and use the new proposed encoding, the change does not break existing interoperability. Updates to [RFC8950] is applicable to the IPv6-Only PE-CE edge design for the IPv6 next hop encoding E2E test case of IPv4 packets over and IPv6-Only core 4to6 Softwire. In this test case IPv4 Unicast NLRI <AFI/SAFI> IPv4 <1/1> is advertised over the PE to RR core peering 4to6 softwire in <AFI/SAFI> VPN-IPV4 <1/128>. In this test case label allocation mode comes into play which is discussed in section 8.9.

[RFC5549] next hop encoding of MP_REACH_NLRI with:

o NLRI= NLRI as per current AFI/SAFI definition

Advertising with [RFC4760] MP_REACH_NLRI with:

o AFI = 1

o SAFI = 128 or 129

o Length of Next Hop Address = 16 or 32

o NLRI= NLRI as per current AFI/SAFI definition

[RFC8950] next hop encoding of MP_REACH_NLRI with:

o NLRI= NLRI as per current AFI/SAFI definition

Advertising with [RFC4760] MP_REACH_NLRI with:

o AFI = 1

o SAFI = 128 or 129

o Length of Next Hop Address = 24 or 48

o Next Hop Address = VPN-IPv6 address of next hop with an 8-octet RD set to zero (potentially followed by the link-local VPN-IPv6 address of the next hop with an 8-octet RD is set to zero).

o NLRI= NLRI as per current AFI/SAFI definition

5. IPv6-Only Design Edge E2E Test Cases

Proof of conept interoperability testing of the 4 test cases between the 5 vendors Cisco, Juniper, Arista, Nokia and Huawei.

5.1. Test-1 IPv6-Only PE-CE Global Table over IPv4-Only Core

```
                                ________
      IPv6-Only       _____    /        \                 IPv6-Only
       PE / CE       /     \__/          \___              PE / CE
   +----+  +----+   /                        \        +------+   +-----+
   |    |  |    |  |                          |_      |      |   |     |
   |    |  |    |  |                             \    |      |   |     |
   | CE |--| PE |--\         IPv4-Only Core      |----|  PE  |---|  CE |
   |    |  |    |    \0=========Underlay =======0|    |      |   |     |
   +----+  +----+     \                        __/    +------+   +-----+
   IPv6 BGP peer       \     MPLS / SR domain /         IPv6 BGP peer
   IPv4 forwarding      \__         __       /          IPv4 forwarding
   IPv6 forwarding         \_______/  \_____/           IPv6 forwarding
```

Figure 7: Test-1 E2E IPv6-Only PE-CE Global Table 6to4 Softwire

Cisco, Juniper, Arista, Nokia, Huawei Test case Results documented here.

5.2. Test-2 E2E IPv6-Only PE-CE VPN over IPv4-Only Core

```
                                ________
      IPv6-Only       _____    /        \                 IPv6-Only
       PE / CE       /     \__/          \___              PE / CE
   +----+  +----+   /                        \        +------+   +-----+
   |    |  |    |  | 0====VPN Overlay Tunnel ==0|     |      |   |     |
   |    |  |    |  |                             \    |      |   |     |
   | CE |--| PE |--\         IPv4-Only Core      |----|  PE  |---|  CE |
   |    |  |    |    \0=========Underlay =======0|    |      |   |     |
   +----+  +----+     \                        __/    +------+   +-----+
   IPv6 BGP peer       \   MPLS / SR domain   /         IPv6 BGP peer
   IPv4 forwarding      \__         __       /          IPv4 forwarding
   IPv6 forwarding         \_______/  \_____/           IPv6 forwarding
```

       Figure 8: Test-2 E2E IPv6-Only PE-CE Design VPN 6to4 Softwire

   Cisco, Juniper, Arista, Nokia, Huawei Test case Results documented
   here.

5.3. Test-3 IPv6-Only PE-CE Global Table over IPv6-Only Core

```
                                ________
      IPv6-Only       _____    /        \                 IPv6-Only
       PE / CE       /     \__/          \___              PE / CE
   +----+  +----+   /                        \        +------+   +-----+
   |    |  |    |  |                          |_      |      |   |     |
   |    |  |    |  |                             \    |      |   |     |
   | CE |--| PE |--\         IPv6-Only Core      |----|  PE  |---|  CE |
   |    |  |    |    \0=========Underlay =======0|    |      |   |     |
   +----+  +----+     \                        __/    +------+   +-----+
   IPv6 BGP peer       \     MPLS / SR domain /         IPv6 BGP peer
   IPv4 forwarding      \__         __       /          IPv4 forwarding
   IPv6 forwarding         \_______/  \_____/           IPv6 forwarding
```

      Figure 9: Test-3 E2E IPv6-Only PE-CE Global Table 4to6 Softwire

Cisco, Juniper, Arista, Nokia, Huawei Test case Results documented here.

5.4. Test-4 IPv6-Only PE-CE VPN over IPv6-Only Core

```
                                ________
      IPv6-Only       _____    /        \                 IPv6-Only
       PE / CE       /     \__/          \___              PE / CE
   +----+  +----+   /                        \        +------+   +-----+
   |    |  |    |  | 0====VPN Overlay Tunnel ==0|     |      |   |     |
   |    |  |    |  |                             \    |      |   |     |
   | CE |--| PE |--\         IPv6-Only Core      |----|  PE  |---|  CE |
   |    |  |    |    \0=========Underlay =======0|    |      |   |     |
   +----+  +----+     \                        __/    +------+   +-----+
   IPv6 BGP peer       \    MPLS / SR domain  /         IPv6 BGP peer
   IPv4 forwarding      \__         __       /          IPv4 forwarding
   IPv6 forwarding         \_______/  \_____/           IPv6 forwarding
```

Figure 10: Test-4 E2E IPv6-Only PE-CE VPN 4to6 Softwire

Cisco, Juniper, Arista, Nokia, Huawei Test case Results documented here.

5.5. IPv6-Only PE-CE Operational Considerations Testing

```
                       Ping CE to PE when destination prefix is withdrawn
                       Traceroute CE to PE and test all ICMPv4 and ICMPv6 type

                   +-------+                +-------+
                   |       |  IPv6 Only     |       |
                   |   CE  |----------------|  PE   |
                   |       |  IPv6 BGP Peer |       |
                   +-------+                +-------+
                  IPv4 forwarding            IPv4 forwarding
                  IPv6 forwarding            IPv6 forwarding
```

Figure 11: Ping and Trace Test Case

Cisco, Juniper, Arista, Nokia, Huawei Test case Results documented here.

6. Operational Considerations

With a single IPv6 Peer carrying both IPv4 and IPv6 NLRI there are some operational considerations in terms of what changes and what does not change.

What does not change with a single IPv6 transport peer carrying IPv4 NLRI and IPv6 NLRI below:

Routing Policy configuration is still separate for IPv4 and IPv6 configured by capability as previously.

Layer 1, Layer 2 issues such as one-way fiber or fiber cut will impact both IPv4 and IPv6 as previously.

If the interface is in the Admin Down state, the IPv6 peer would go down, and IPv4 NLRI and IPv6 NLRI would be withdrawn as previously.

Changes resulting from a single IPv6 transport peer carrying IPv4 NLRI and IPv6 NLRI below:

Physical interface is no longer dual stacked.

Any change in IPv6 address or DAD state will impact both IPv4 and IPv6 NLRI exchange.

Single BFD session for both IPv4 and IPv6 NLRI fate sharing as the session is now tied to the transport, which now is only IPv6 address family.

Both IPv4 and IPv6 peer now exists under the IPv6 address family configuration.

Fate sharing of IPv4 and IPv6 address family from a logical perspective now carried over a single physical IPv6 peer.

From an operations perspective, prior to elimination of IPv4 peers, an audit is recommended to identify and IPv4 and IPv6 peering incongruencies that may exist and to rectify them. No operational impacts or issues are expected with this change.

With MPLS VPN overlay, per-CE next-hop label allcoation mode where both IPv4 and IPv6 prefixes have the same label in no table lookup pop-n-forward mode should be taken into consideration.

7. IANA Considerations

There are not any IANA considerations.

8. Security Considerations

The extensions defined in this document allow BGP to propagate reachability information about IPv4 prefixes over an MPLS or SR IPv6-Only core network. As such, no new security issues are raised beyond those that already exist in BGP-4 and the use of MP-BGP for IPv6. Both IPv4 and IPv6 peers exist under the IPv6 address family configuration. The security features of BGP and corresponding security policy defined in the ISP domain are applicable. For the inter-AS distribution of IPv6 routes according to case (a) of Section 4 of this document, no new security issues are raised beyond those that already exist in the use of eBGP for IPv6 [RFC2545].

9. Acknowledgments

Thanks to Kaliraj Vairavakkalai, Linda Dunbar, Aijun Wang, Eduardfor Vasilenko, Joel Harlpern, Michael McBride, Ketan Talaulikar for review comments.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2545] Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol

Extensions for IPv6 Inter-Domain Routing", RFC 2545, DOI 10.17487/RFC2545, March 1999, <https://www.rfc-editor.org/info/rfc2545>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

10.2. Informative References

[I-D.ietf-idr-dynamic-cap] Ramachandra, S. and E. Chen, "Dynamic Capability for BGP-4", draft-ietf-idr-dynamic-cap-14 (work in progress), December 2011.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for

IPv6 VPN", RFC 4659, DOI 10.17487/RFC4659, September 2006, <https://www.rfc-editor.org/info/rfc4659>.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, DOI 10.17487/RFC4684, November 2006, <https://www.rfc-editor.org/info/rfc4684>.

[RFC4798] De Clercq, J., Ooms, D., Prevost, S., and F. Le Faucheur, "Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)", RFC 4798, DOI 10.17487/RFC4798, February 2007, <https://www.rfc-editor.org/info/rfc4798>.

[RFC4925] Li, X., Ed., Dawkins, S., Ed., Ward, D., Ed., and A. Durand, Ed., "Softwire Problem Statement", RFC 4925, DOI 10.17487/RFC4925, July 2007, <https://www.rfc-editor.org/info/rfc4925>.

[RFC5549] Le Faucheur, F. and E. Rosen, "Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop", RFC 5549, DOI 10.17487/RFC5549, May 2009, <https://www.rfc-editor.org/info/rfc5549>.

[RFC5565] Wu, J., Cui, Y., Metz, C., and E. Rosen, "Softwire Mesh Framework", RFC 5565, DOI 10.17487/RFC5565, June 2009, <https://www.rfc-editor.org/info/rfc5565>.

[RFC6074] Rosen, E., Davie, B., Radoaca, V., and W. Luo, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074, DOI 10.17487/RFC6074, January 2011, <https://www.rfc-editor.org/info/rfc6074>.

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <https://www.rfc-editor.org/info/rfc6513>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8950] Litkowski, S., Agrawal, S., Ananthamurthy, K., and K. Patel, "Advertising IPv4 Network Layer Reachability Information (NLRI) with an IPv6 Next Hop", RFC 8950, DOI 10.17487/RFC8950, November 2020, <https://www.rfc-editor.org/info/rfc8950>.

Authors' Addresses

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Mankamana Mishra
Cisco Systems
821 Alder Drive,
MILPITAS CALIFORNIA 95035

Email: mankamis@cisco.com

Jeff Tantsura
Juniper Networks, Inc.

Email: jefftant.ietf@gmail.com

Lili Wang
Juniper Networks, Inc.
10 Technology Park Drive,
Westford MA 01886
US

Email: liliw@juniper.net

Qing Yang
Arista Networks

Email: qyang@arista.com

Adam Simpson
Nokia

Email: adam.1.simpson@nokia.com

Shuanglong Chen
Huawei Technologies

Email: chenshuanglong@huawei.com