BGP Usage for SDWAN Overlay Networks

draft-dunbar-bess-bgp-sdwan-usage-06

Linda Dunbar & Jim Guichard (Futurewei)
Ali Sajassi (Cisco)
John Drake (Juniper)
Basil Najem (Bell Canada)
Ayan Barnerjee & Dave Carrel (Cisco)

April 2020

Key Items Discussed in the Draft

3. Use Case Scenario Description and Requirements
    - 3.1. Requirements
        - 3.1.1. Supporting Multiple SDWAN Segmentations
        - 3.1.2. Client Service Requirement
        - 3.1.3. Application Flow Based Segmentation
        - 3.1.4. SDWAN Node Provisioning
    - 3.2. Scenarios #1: Homogeneous WAN
    - 3.3. Scenario #2: SDWAN WAN ports to VPN’s PEs and to Internet
    - 3.4. Scenario #3: SDWAN WAN ports to MPLS VPN and the Internet

4. BGP Walk Through
    - 4.1. BGP Walk Through for Homogeneous SDWAN
    - 4.2. BGP Walk Through for Application Flow Based Segmentation
    - 4.3. Client Service Provisioning Model
    - 4.4. WAN Ports Provisioning Model
    - 4.5. Why BGP as Control Plane for SDWAN?

5. SDWAN Traffic Forwarding Walk Through
    - 5.1. SDWAN Network Startup Procedures
    - 5.2. Packet Walk-Through for Scenario #1
    - 5.3. Packet Walk-Through for Scenario #2
    - 5.3.1. SDWAN node WAN Ports Properties Registration
    - 5.3.2. Controller Facilitated IPsec SA & NAT management
    - 5.4. Packet Walk-Through for Scenario #3

key characteristics of “SDWAN” networks

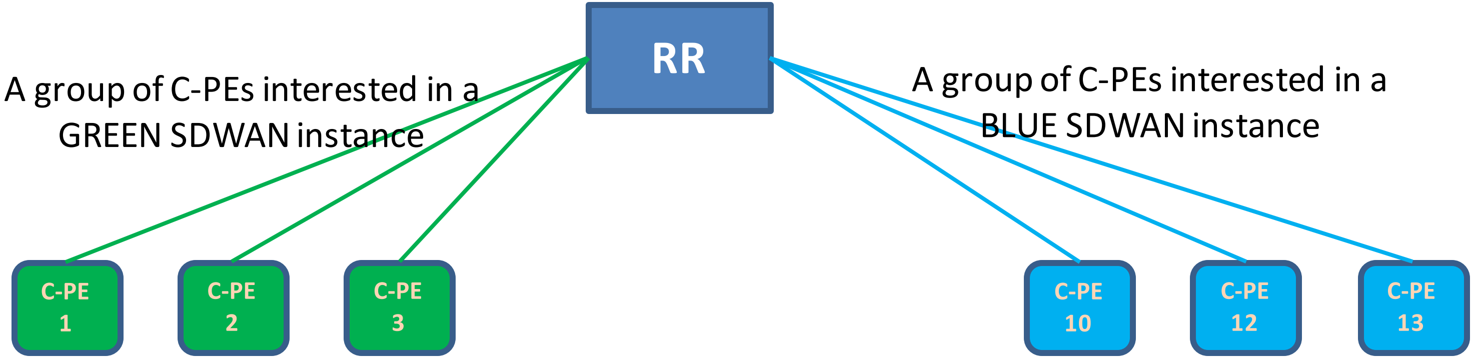
- Augment of transport:
  - ❑ utilizing overlay paths over different underlay networks.
  - ❑ Among the multiple parallel overlay paths between any two SDWAN edges, some are private networks over which traffic can traverse with or without encryption, others require encryption, e.g. over untrusted public networks.
- Enable direct Internet access from remote sites, instead hauling all traffic to Corporate HQ for centralized policy control.
- Some traffic are routed based on application IDs instead of based on destination IP addresses.

Possible approaches to Support Multiple SDWAN Instances (segmentations)

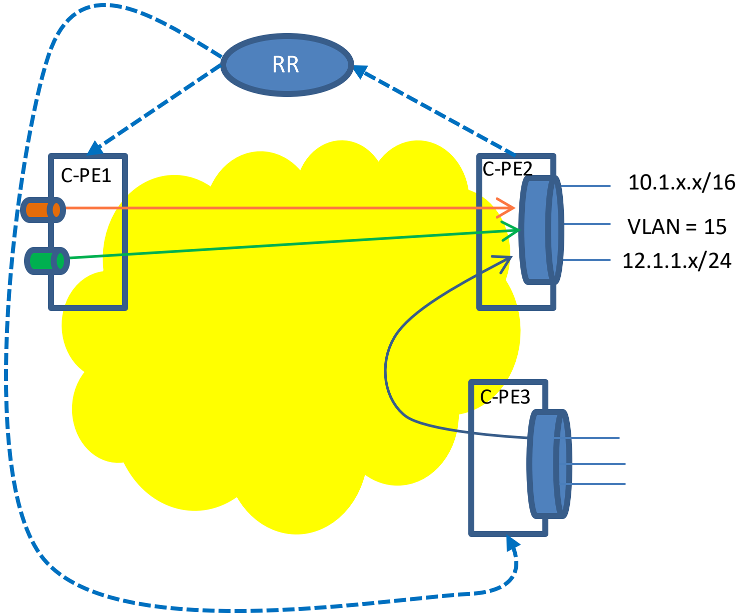
- Create a SDWAN Target ID in the BGP Extended Community to represent different SDWAN Segmentations
  - Same as Route Target, just use a different name to differentiate from VPN If a CPE supports traditional VPN with multiple VRFs, and supports multiple SDWAN Segmentations (instances).
- When the SDWAN Target ID is used,
  - Use the similar approach as VPN Label carried by NLRI Path Attribute [RFC8277] to identify routes belonging to different SDWAN Segmentations.
  - The MPLS VPN SAFI 128 & Route Distinguisher can be used for routes belonging to different SDWAN instances.

Constrained Propagation of Clients routes/info

- Using RFC 4684 to constrain the distribution of BGP UPDATE to only a subset of SDWAN edges
- Using manually provisioned policies on RR to constrain the propagation of BGP UPDATE


RR
A group of C-PEs interested in a GREEN SDWAN instance
A group of C-PEs interested in a BLUE SDWAN instance
C-PE 1
C-PE 2
C-PE 3
C-PE 10
C-PE 12
C-PE 13

BGP Walk Through for Homogeneous SD-WAN multiple routes aggregated in one IPsec



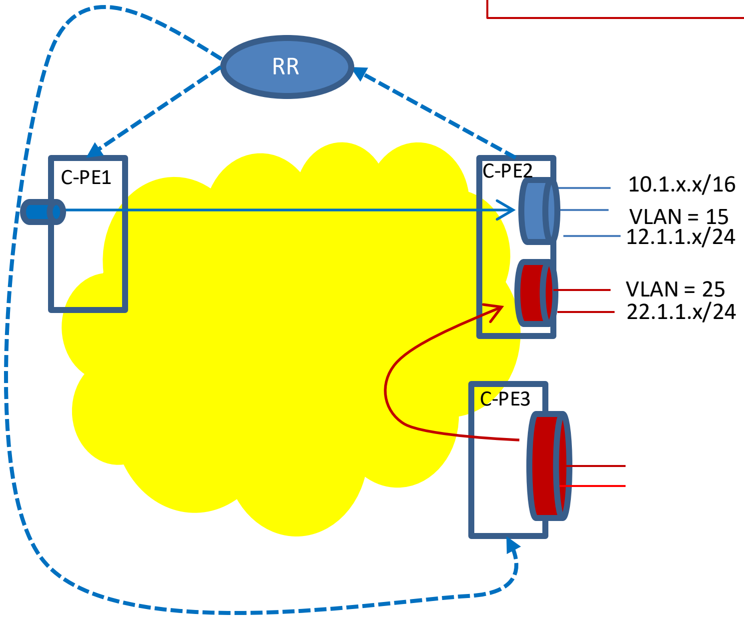
One BGP UPDATE Message from C-PE2 to RR:

- multiple routes encoded in the MP-NLRI Path Attribute
  - 10.1.x.x/16
  - VLAN #15
  - 12.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
  - IPsec attributes for all possible remote nodes, or
  - IPsec attributes for specific remote nodes, or
  - IPsec attributes for specific remote subnets

  ....

BGP Walk Through for Homogeneous SD-WAN Client Routes with different Topologies & Policies

RR sends different UPDATE messages to different edges to reflect different topologies



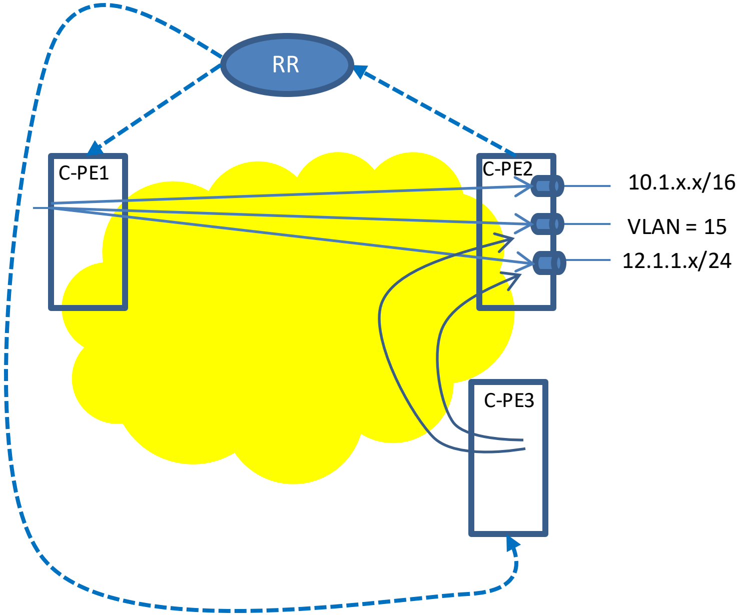
One BGP UPDATE Message to C-PE1 to indicate the BLUE tunnel:

- multiple routes encoded in the MP-NLRI Path Attribute
  - 10.1.x.x/16
  - VLAN #15
  - 12.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
  - IPsec attributes for C-PE1 to C-PE2

One BGP UPDATE Message to C-PE3 to indicate the RED tunnel:

- multiple routes encoded in the MP-NLRI Path Attribute
  - VLAN #25
  - 22.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
  - IPsec attributes for C-PE3 to C-PE2

BGP Walk Through for Homogeneous SD-WAN per Route Encryption (Fine-Grained)

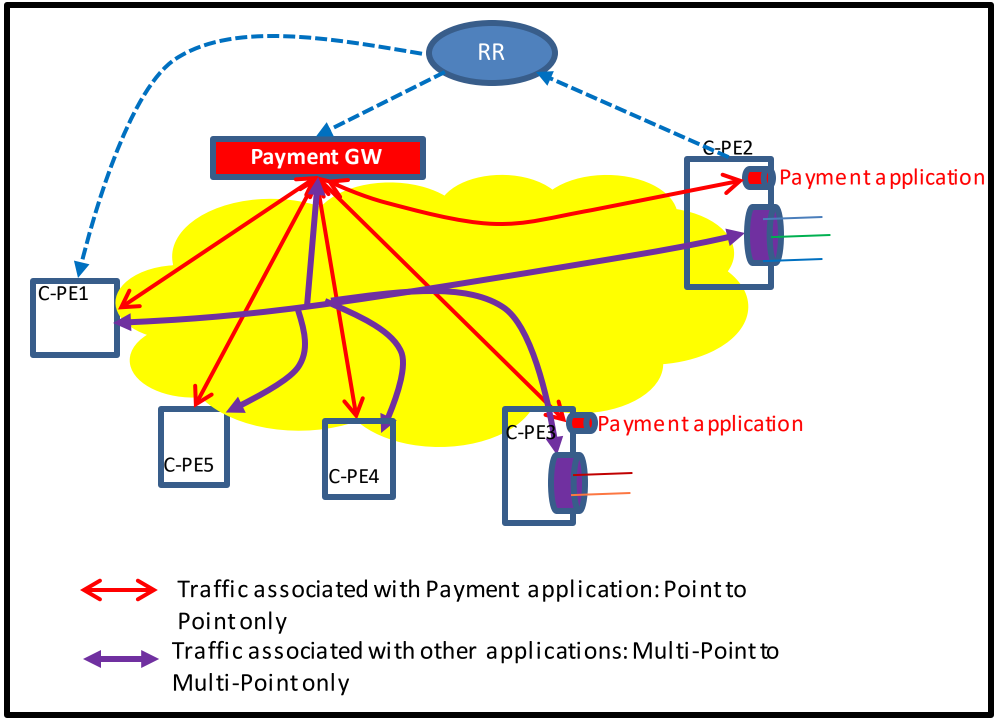


Three Separate BGP UPDATE messages from C-PE2 to RR:
UPDATE 1:

- MP-NLRI Path Attribute
  - 10.1.x.x/16 encoded
  - VLAN #15
  - 12.1.1.x/24

- Tunnel-Encap
  - IPsec SA attributes for any nodes to establish IPsec tunnel C-PE-2 for the routes encoded in the MP-NLRI Path Attribute

Applications Based Segmentation in SDWAN



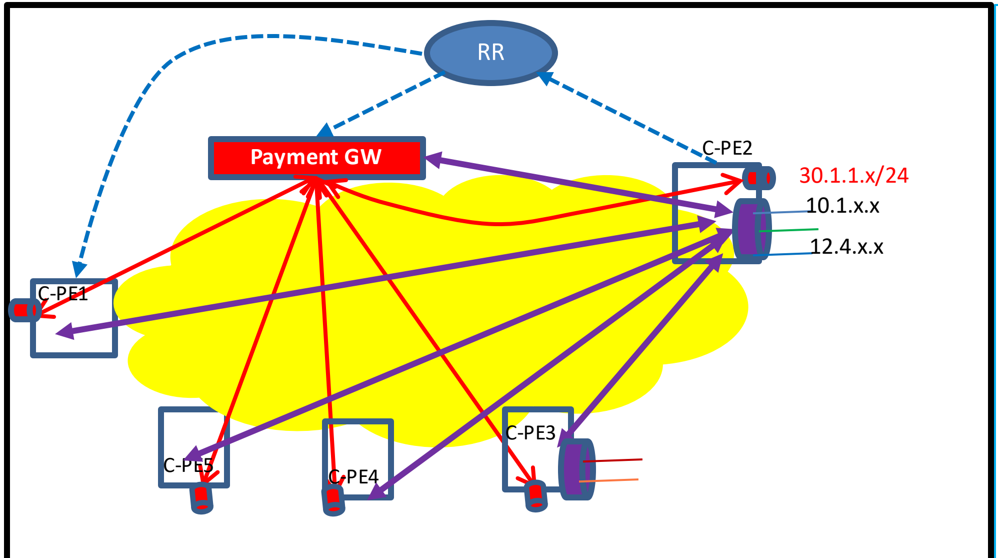
↔ Traffic associated with Payment application: Point to Point only

↔ Traffic associated with other applications: Multi-Point to Multi-Point only

Characteristics:

- RED route (payment applications) can only be propagated to "Payment GW"
- Purple routes needs to be propagated to all other nodes

BGP Walk Through for Applications Based Segmentation in SDWAN



🟥 ↔ Traffic associated with Payment application: Point to Point only

🟪 ↔ Traffic associated with other applications: Multi-Point to Multi-Point only

Assume Payment Application has different IP address than other segments, e.g. 30.1.1.x/24 for Payment application

BGP UPDATE #1 from C-PE2 to RR for the RED P2P topology (only propagated to Payment GW node:

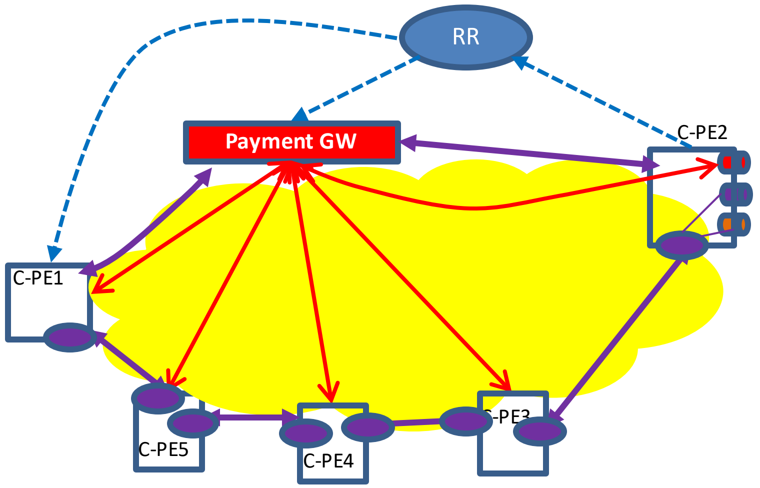
- MP-NLRI Path Attribute:
    - 30.1.1.x/24
- Tunnel Encap Path Attribute
    - IPsec Attributes for PaymentGW ->C-PE2

BGP UPDATE #2 from C-PE2 to RR for the routes to be reached by Purple:

- MP-NLRI Path Attribute:
    - 10.1.x.x
    - 12.4.x.x

- TunnelEncap Path Attribute:
    - Any node to C-PE2

Purple Topology Supported by Multiple IPsec tunnels stitched together



↔ IPsec tunnel as transport that aggregate traffic to different destinations. Each edge route the traffic just like VRF.

BGP UPDATE from C-PE2 to RR for the IPsec C-PE3 -> C-PE2 segment:

- Port Addr encoded in MP-NLRI Path Attribute
- Port based IPsec tunnels encoded in the Tunnel Encap Path Attribute
  - IPsec SA-P-G32 (C-PE3>C-PE2)

BGP UPDATE from C-PE3 to RR for the IPsec C-PE2 -> C-PE3 segment:

- Port Addr encoded in MP-NLRI Path Attribute
- Port based IPsec tunnel encoded in the Tunnel Encap Path Attribute
  - IPsec SA-P-G23 (C-PE2>C-PE3)

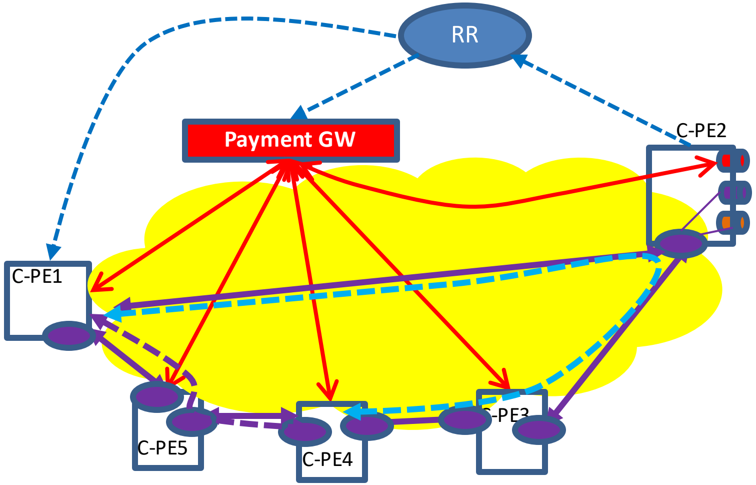
BGP UPDATE from C-PE3 to RR for the IPsec C-PE4 -> C-PE3 segment:

- Port Addr encoded in MP-NLRI Path Attribute
- Port based IPsec tunnel encoded in the Tunnel Encap Path Attribute
  - IPsec SA-P-G43 (C-PE4>C-PE3)

BGP UPDATE from C-PE4 to RR for the IPsec C-PE3 -> C-PE4 segment

BGP UPDATE from C-PE4 to RR for the IPsec C-PE5 -> C-PE4 segment

Analysis of Multiple IPsec tunnels stitched together



IPsec tunnel as transport that aggregate traffic to different destinations. Each edge route the traffic just like VRF.

Pros

- Each C-PEs has much smaller number of IPsec Tunnels to manage
- Less BGP UPDATE messages for keys refreshment and management
- Deterministic route to each destination.

Cons:

- Some C-PEs need to forward packets to another IPsec tunnel,
- Those C-PEs need to process those BGP UPDATE to build the forwarding table.
- There could be multiple ways for packets from A->B.
- TE needs to be enforced by the Controller

Next Step

- **WG Adoption.**
- **Why**
  - Give a clear picture on how BGP is used to scale SDWAN to the industry