BGP Usage for SDWAN Overlay Networks

draft-dunbar-bess-bgp-sdwan-usage-08

Linda Dunbar & Jim Guichard (Futurewei)
Ali Sajassi (Cisco)
John Drake (Juniper)
Basil Najem (Bell Canada)
Ayan Barnerjee & Dave Carrel (Cisco)
July 2020

Key Items Discussed in the Draft

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.

➢ The Application Routing can also be based on specific performance criteria (e.g. packets delay, packet loos, jitter) to provide better application performance by choosing the right underlay that meets or exceeds the specified criteria.

To Support Multiple SDWAN Instances (segmentations)

- **Control Plane Messages:**
  - 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).
  - Overlay Topology:
    - Overlay connectivity is established based on SDWAN Route Target
- **Data Plane:**
  - Public network: SDWAN ID carried by the VNID (virtual network ID) or GRE keys carried by the inner encapsulation within IPsec ESP Tunnel.

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
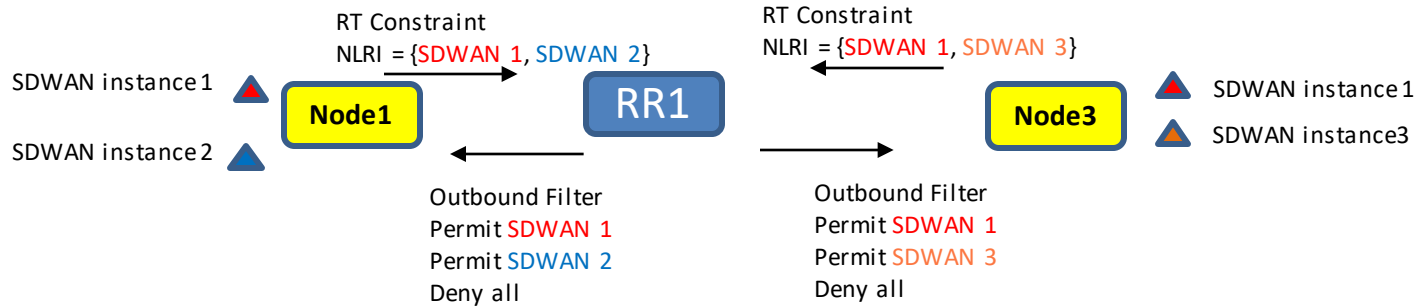
BGP Walk Through for Homogeneous SD-WAN multiple routes aggregated in one IPsec
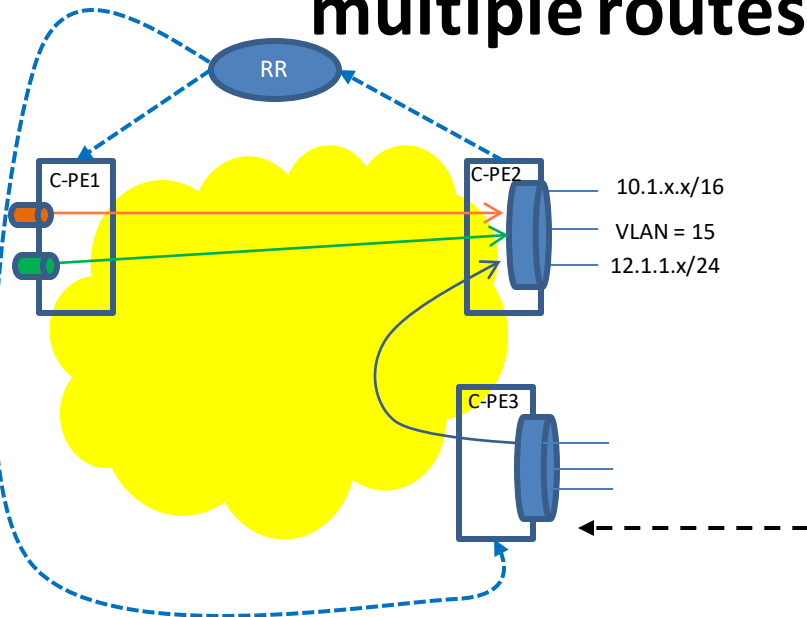
One BGP UPDATE Message from C-PE2 to RR:
- multiple routes encoded in the MP-NLRI Path Attribute
  - 10.1.x.x/16
  - VLAN #15
  - 12.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
  - IPsec attributes for all possible remote nodes, or
  - IPsec attributes for specific remote nodes, or
  - IPsec attributes for specific remote subnets
  ....

SDWAN Instance 1
Import RT 1:1
Export RT 1:1
SDWAN Instance 2
Import RT 1:2
Export RT 1:2
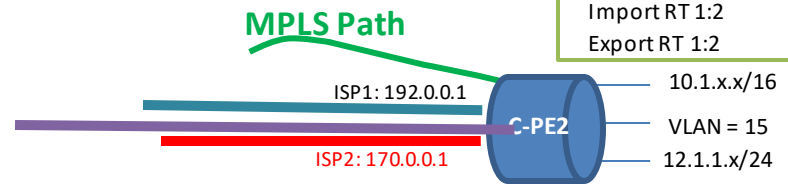
BGP Walk Through for Homogeneous SD-WAN Client Routes with different Topologies & Policies

RR sends different UPDATE messages to different edges to reflect different topologies
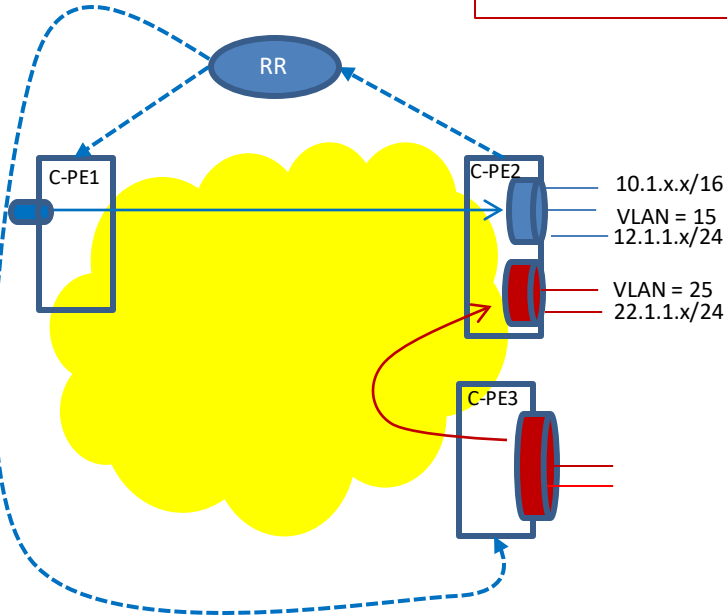
One BGP UPDATE Message to C-PE1 to indicate the BLUE tunnel:
- multiple routes encoded in the MP-NLRI Path Attribute
    - 10.1.x.x/16
    - VLAN #15
    - 12.1.1.x/24
- IPsec attributes are encoded in the Tunnel-Encap Path Attribute
    - IPsec attributes for C-PE1 to C-PE2

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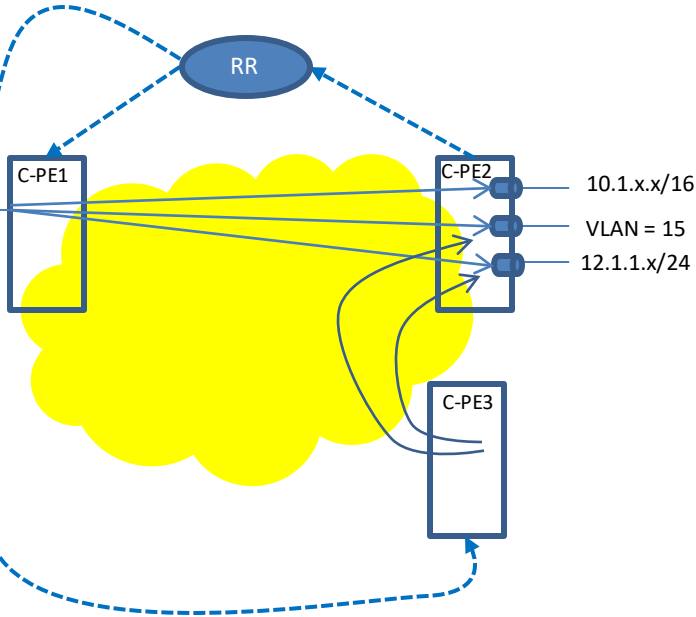
Three Separate BGP UPDATE messages from C-PE2 to RR:

UPDATE 1:
- MP-NLRI Path Attribute
    - 10.1.x.x/16 encoded
- 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

UPDATE 2:
- MP-NLRI Path Attribute
    - VLAN #15
- 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

UPDATE 3:
- MP-NLRI Path Attribute
    - 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

BGP Walk Through for Scenario #2: C-PE with multiple Underlays

Prefix 11.1.1.x
NH: C-PE 1
Tunnel Encap Attr (Type=IP)
MPLS-in-GRE

Client route 11.1.1.1 can only use MPLS-in-GRE path

Extended community: RT for SDWAN Instance 1, 2
Prefix 21.1.1.x
NH: C-PE 1
Tunnel Encap Attr (Type=IP:AFI/SAFI = 1/1)
MPLS-in-GRE
IPSec SA for 192.10.0.10
IPSec SA for 160.0.0.1

Extended community: RT for SDWAN Instance 1,
Prefix: 10.1.x.x/16; 12.1.1
NH: C-PE 2
Tunnel Encap Attr (AFI/SAFI = 1/1)
MPLS-in-GRE
IPSec SA for 192.0.0.1 (Need a new tunnel type)
IPSec SA for 170.0.0.1
IPsec SA for 2.2.2.2

BGP Route Controller - RR

OF: Permit for SDWAN Instance 1, 2
Deny All

IP SAFI BGP Session

PE

MPLS VPN

PE

ISP3: 192.10.0.10

ISP1: 192.0.0.1

C-PE 1

C-PE 2

Client Route: 11.1.1.x

Client Route: 21.1.1.x

ISP2: 160.0.0.1

ISP2: 170.0.0.1

10.1.x.x/16

VLAN = 15

12.1.1.x/24

BGP AS 1
lo: 1.1.1.1

SDWAN Instance 1
Import RT 1:1
Export RT 1:1
SDWAN Instance 2
Import RT 1:2
Export RT 1:2

BGP AS 1
lo: 2.2.2.2

SDWAN Instance 1
Import RT 1:1
Export RT 1:1
SDWAN Instance 2
Import RT 1:2
Export RT 1:2

BGP Walk Through for Applications Based Segmentation in SDWAN
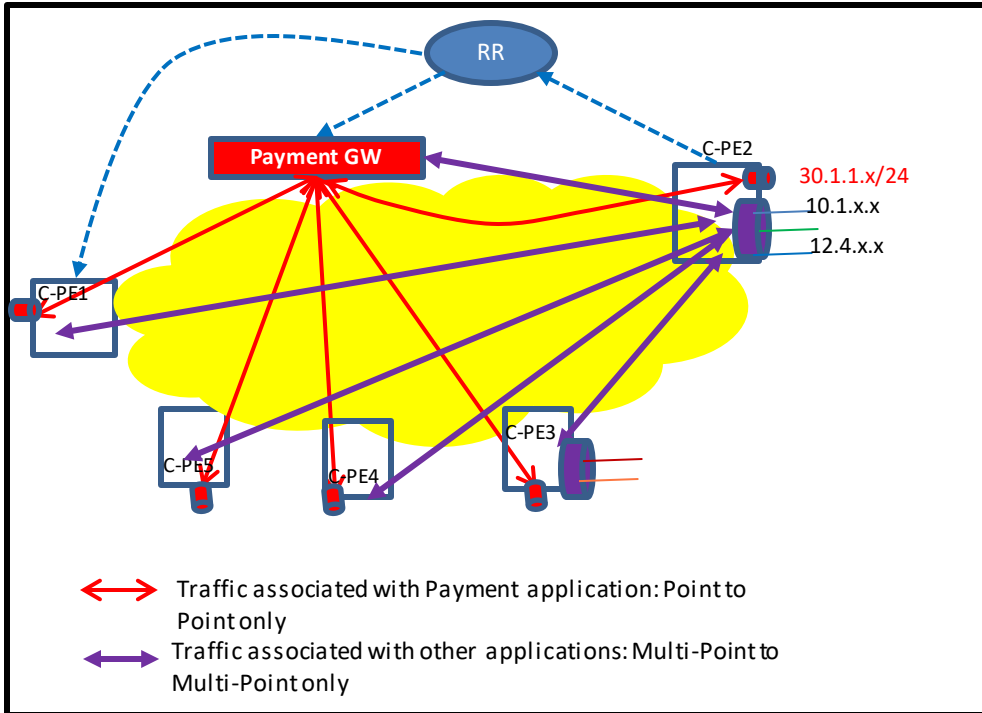
Assume Payment Application has different IP address than other segments, e.g. 30.1.1.x/24 for Payment application

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
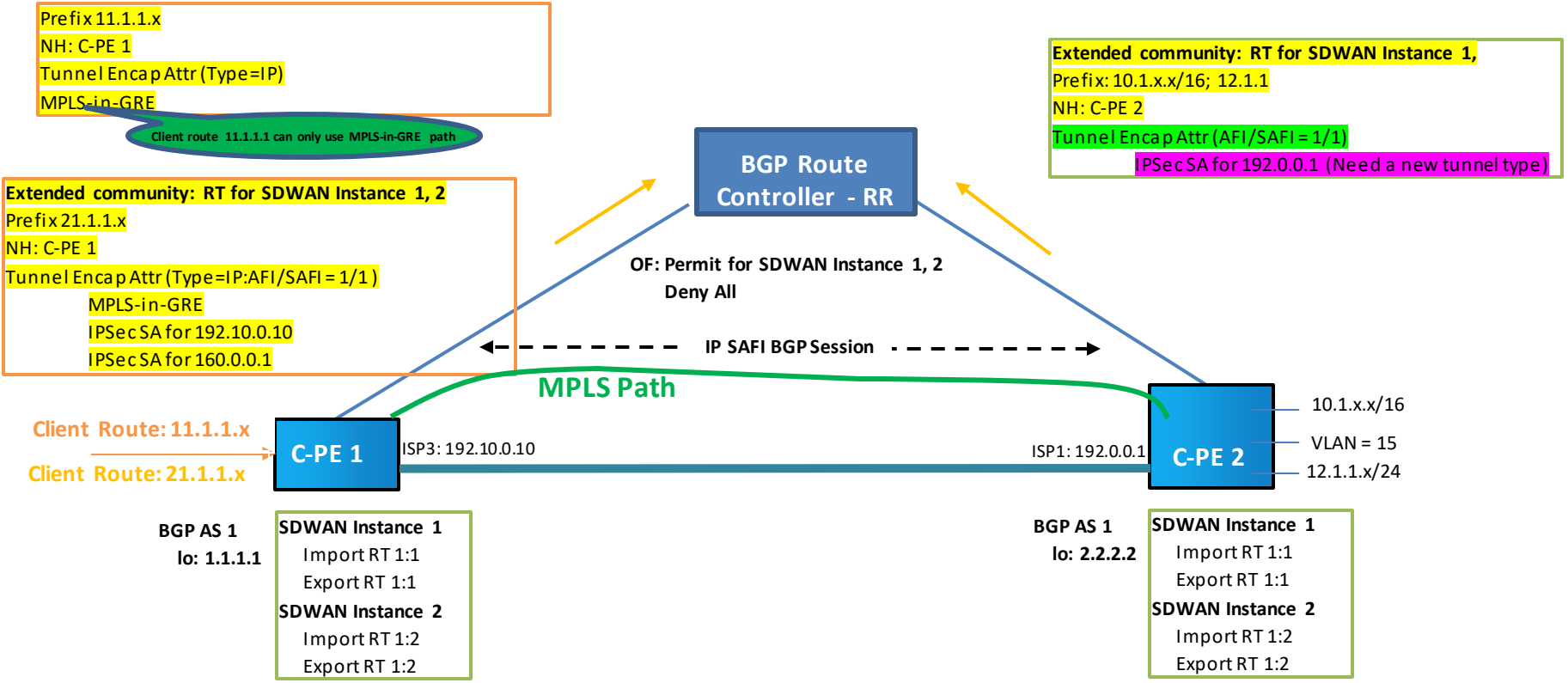   - 12.4.x.x

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

BGP Walk Through for Scenario #3: PE based SDWAN

Prefix 11.1.1.x
NH: C-PE 1
Tunnel Encap Attr (Type=IP)
MPLS-in-GRE

Client route 11.1.1.1 can only use MPLS-in-GRE path

Extended community: RT for SDWAN Instance 1,
Prefix: 10.1.x.x/16; 12.1.1
NH: C-PE 2
Tunnel Encap Attr (AFI/SAFI = 1/1)
          IPSec SA for 192.0.0.1 (Need a new tunnel type)

Extended community: RT for SDWAN Instance 1, 2
Prefix 21.1.1.x
NH: C-PE 1
Tunnel Encap Attr (Type=IP:AFI/SAFI = 1/1)
     MPLS-in-GRE
     IPSec SA for 192.10.0.10
     IPSec SA for 160.0.0.1

**BGP Route
Controller - RR**

OF: Permit for SDWAN Instance 1, 2
Deny All

← - - - - - - - - - - - IP SAFI BGP Session - - - - - - - - - - - →

MPLS Path

Client Route: 11.1.1.x

Client Route: 21.1.1.x

C-PE 1

ISP3: 192.10.0.10

ISP1: 192.0.0.1

C-PE 2

10.1.x.x/16

VLAN = 15

12.1.1.x/24

BGP AS 1

lo: 1.1.1.1

SDWAN Instance 1
  Import RT 1:1
  Export RT 1:1
SDWAN Instance 2
  Import RT 1:2
  Export RT 1:2

BGP AS 1

lo: 2.2.2.2

SDWAN Instance 1
  Import RT 1:1
  Export RT 1:1
SDWAN Instance 2
  Import RT 1:2
  Export RT 1:2

Next Step

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