# draft-dunbar-idr-sdwan-port-safi-03

Linda Dunbar

Sue Hares

IETF 105 July 2019

Key Characteristics of Hybrid SDWAN: Multi-players

IP or Ethernet

Direct Connect

Cloud DC

MPLS

PE1

PE 2

GW 1

P1

C-PE1

vR

P2

Internet

GW2

Via untrusted public network

RR

IPsec SA configurations and Key refreshment
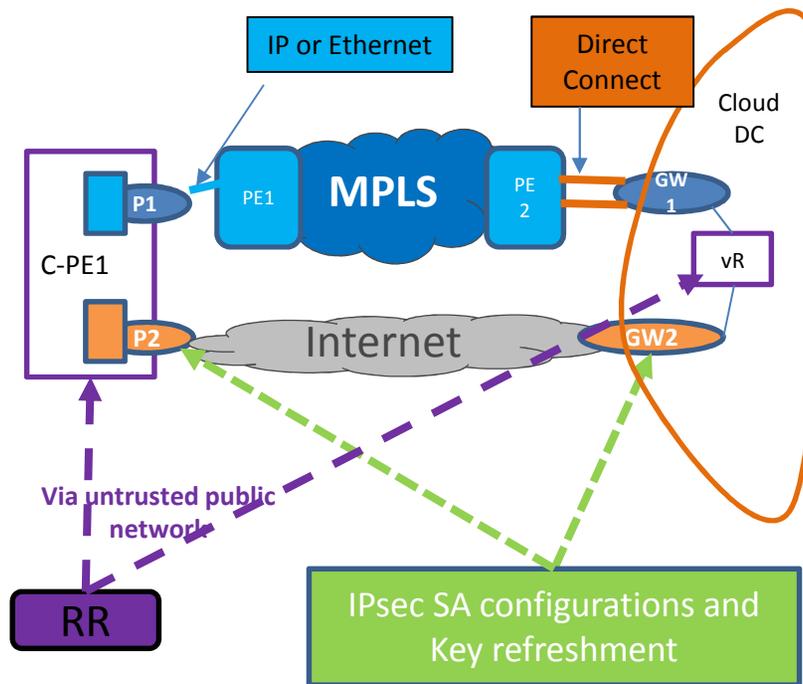
- Multiple parties: Enterprises, VPN providers and Cloud Providers
- Multiple paths with different performance & authentication parameters.
- the connections between BGP RR & CPEs are over public network, therefore requires TLS, DTLS, or IPsec.

Application components can be available from multiple Cloud DCs.
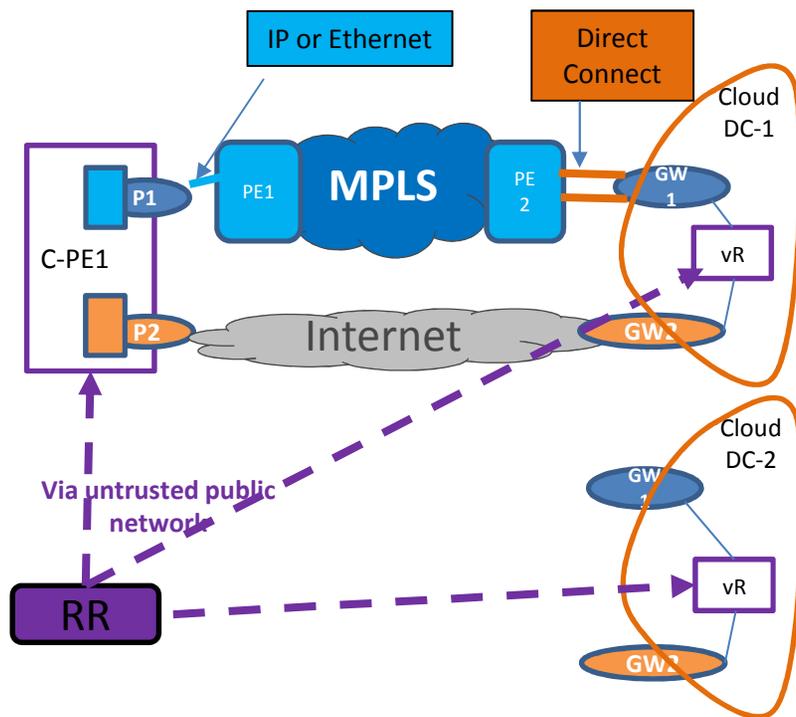
SD-WAN Scenarios

- **Homogeneous SD-WAN:**
  - edge encrypting all sensitive traffic to other edge nodes, regardless if the underlay is private or public.

- **SD-WAN over Hybrid Networks**
  - Traffic over Private VPN networks (e.g. MPLS) can go natively without encryption to achieve better performance, and
  - traffic over internet are carried by IPsec tunnels.
  - User specified policy governs a flow
- Already in practice in Networks

SDWAN Ctrl

MPLS Private Network

C-PE A Underlay PE-B PE-Z

PE-C

Cloud DC

IPsec tunnels at different granularities: Per PE, Per Tenant, Per subnet, etc.

Non sensitive data that don't need encryption

SDWAN Ctrl

ISP-1 Untrusted Internet

C-PE 1 A1 B1 C-PE 2

Cloud DC

Per WAN port IPsec Tunnel

A2 MPLS: trusted network, no encryption needed B2

CPE 3

4
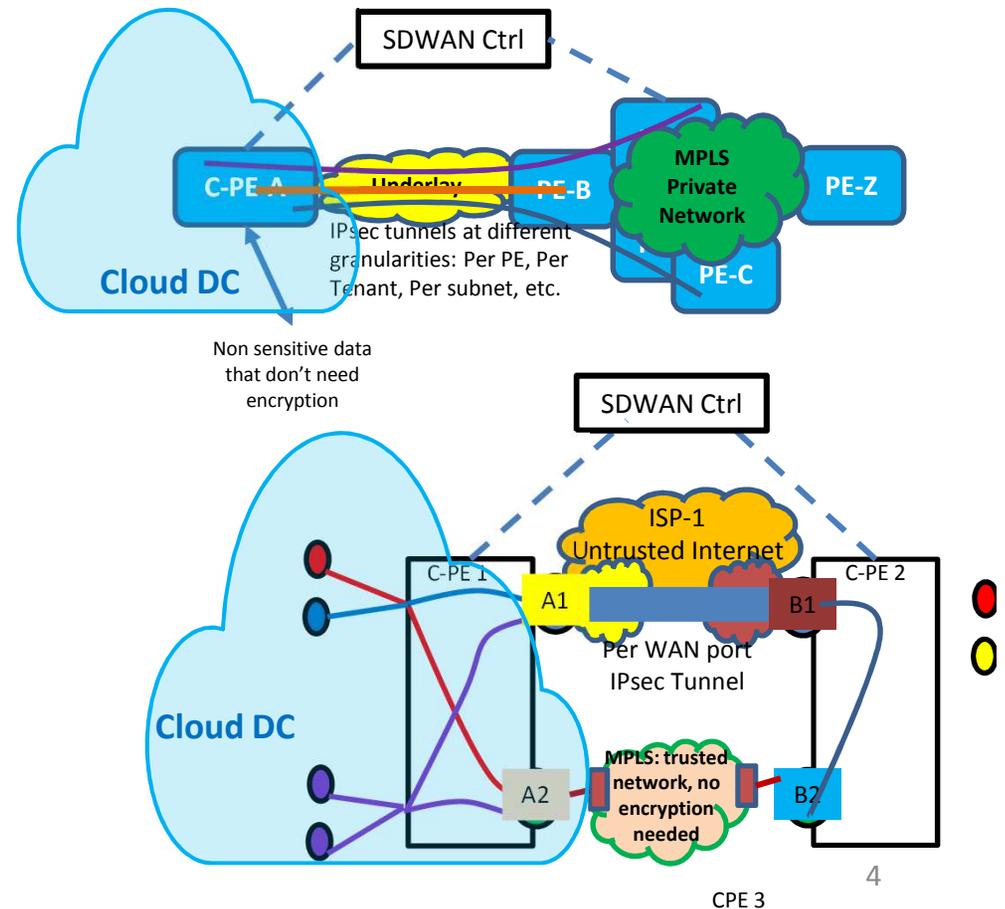
Security Sequence SDWAN Control Plane

Controller- Device Exchange

- Device (Router R1) sends routes to Controller (BGP RR) [Tunnel-Encap]

- Controller (RR) advertise tunnels for clients routes via Secure link (R1)

- The WAN ports properties are registered to Controller, and controller propagates.

- SDWAN edges pairwise secure channel requires Controller managed re-key schedule and distribution

- IPsec SA management independent from the attached client routes. i.e. IPsec parameters negotiation, public key exchange, and re-key schemes detached from client routes.
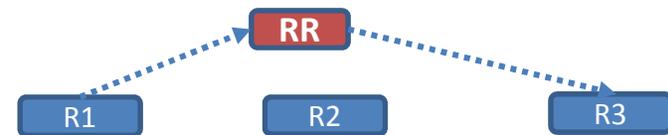
Controller manages the authorized recipients
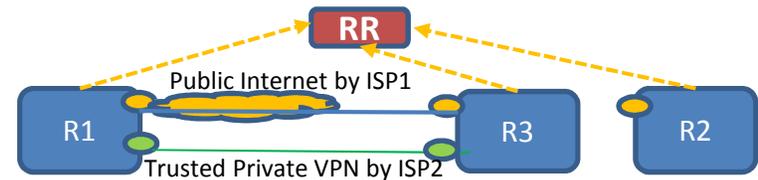
All SDWAN nodes advertise their WAN port properties to Controller

Sequence similar to I2NSF Controller

5

SDWAN Ports scoped Advertisement

- Tenant Separation Method :
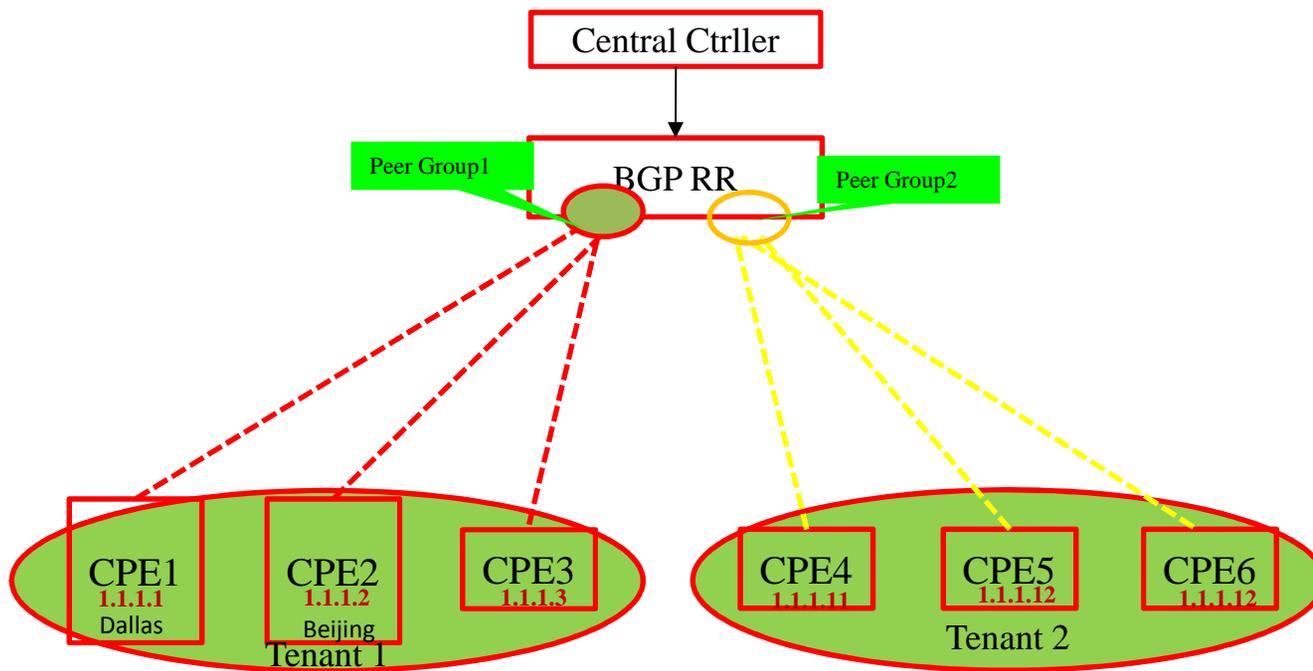
BACKUP SLIDES

WAN Ports Registration

SD-WAN node's private address and WAN Ports/Addresses registration.

SD-WAN Ctrl (BGP RR)

registration

10.1.1.1 via A1
10.1.1.1 via A2

10.1.2.1 via B1
10.1.2.1 via B2

ISP1   ISP3   A1   B1
ISP4   B2

C-PE 1
10.1.1.1

C-PE 2
10.1.2.1

Cloud DC

MPLS: trusted network, no encryption needed

A3   B3

10.1.2.1 via C1
10.1.2.1 via C2

C1   C2

C-PE 3
10.4.2.1
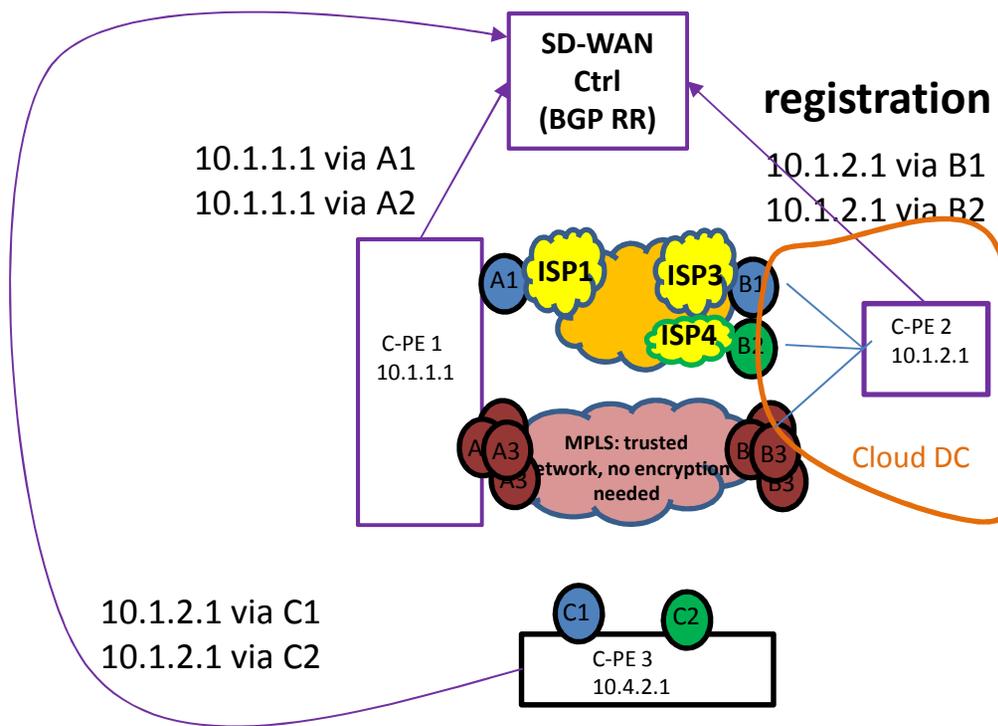
- WAN ports can be from different network providers (A1/A2/A3/B1/B2/B3)
- Each PE advertise its WAN ports to Controller, which then propagate the advertisements to authorized peers.
- **PEs Loopback addresses & routes attached are not visible to some ISPs**

A1/A2/A3/B1/B2/B3 are logical address that can be applied to a set of ports

Why BGP as Control Plane for SDWAN WAN Ports Registration

- Compelling reasons of using BGP:
  - BGP already widely deployed as sole protocol (see RFC 7938)
  - Robust and simple implementation
 - Wide acceptance – minimal learning
  - Reliable transport
  - Guaranteed in-order delivery
  - Incremental updates
  - No flooding and selective filtering
  - RR already has the capability to apply policies to communications among peers.

- Alternative: NHRP, DSVPN/DMVPN
  - In addition to more proposal changes needed, it doesn't scale well
- Prior extension of BGP for non client routes reachability:
  - Flowspec, BGP LS, Segment routing policies, etc

Multiple ways to skin the cat…

Extending Tunnel-Encap with SAFI as IP to achieve WAN port registration?

RR

Tunnel Update 1
- Tunnel-Type Public; IPsec
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: A1;
  2) SubTLV: Ipsec properties

Tunnel Update 2
- Tunnel-Type: Private network (no encap)
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: A2;
  2) subTLV: properties associated with the private network;

Tunnel Update 3
- Tunnel-Type: Private Addr (no encap)
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: A3;
  2) subTLV: properties associated with NAT;

C-PE 1
10.1.1.1

A1
A2
C3

B1
B2
B3

C-PE 2
10.1.1.2

Tunnel Update 1
- Tunnel-Type Public; IPsec
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: B1;
  2) SubTLV: Ipsec properties

Tunnel Update 2
- Tunnel-Type: Private network (no encap)
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: B2;
  2) subTLV: properties associated with the private network;

Tunnel Update 3
- Tunnel-Type: Private Addr (no encap)
- NLRI: 10.1.1.2
  1) subTLV: Remote endpoint: B3;
  2) subTLV: properties associated with NAT;

Site ID

Tunnel-Encap with SDWAN NLRI for SDWAN WAN Ports Prosperities & Policies

```
+------------------+
|   NLRI Length    |  1 octet
+------------------+
|   SDWAN-Type     |  2 Octets
+------------------+
|Port-Distinguisher|  4 octets
+------------------+
|   SDWAN-Site-ID  |  4 octets
+------------------+
|   SDWAN-Node-ID  |  4 or 16 octets
+------------------+
```

SDWAN in cloud can have different TYPE

Can be identifiers to Port # or Cloud GW

The new SAFI=74 has been assigned by IANA for advertising properties of WAN ports that face untrusted networks

Similar to SR Policy NLRI/SAFI to specify specific polices

```
+------------------+
|   NLRI Length    |
+------------------+
|   Distinguisher  |
+------------------+
|   Policy Color   |
+------------------+
|   Endpoint       |
+------------------+
```

- SDWAN NLRI Length: expressed in bits as defined in [RFC4760].

- SDWAN-Type: to define the encoding of the rest of the SDWAN NLRI.

- Port Distinguisher: SDWAN node Port identifier.

- SDWAN-Site-ID: used to identify a common property shared by a set of SDWAN nodes.

- SDWAN Node ID: the SDWAN node identifier, (e.g. system ID or the loopback address (IPv4 or IPv6)).

Advantage of new NLRI: to represent different address space: SDWAN WAN port, which can be port number & not IP address
Disadvantage of new NLRI: intermediate Routers can drop the UPDATE due to not recognizing the new NLRI.
> Not applicable to SDWAN overlay, as the UPDATE to RR is simple IP forwarding, not terminated by any routers/switches in between

11

Extended Port Property

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|EncapExt Type  |    EncapExt subTLV Length      |I|O|R|R|R|R|R|R|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| NAT Type      |   Encap-Type   |Trans networkID|    RD ID     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local  IP Address                         |
            32-bits for IPv4, 128-bits for Ipv6
                      ~~~~~~~~~~~~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local  Port                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Public IP                                 |
            32-bits for IPv4, 128-bits for Ipv6
                      ~~~~~~~~~~~~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Public Port                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
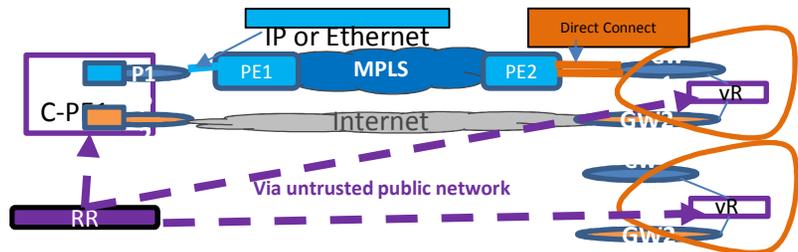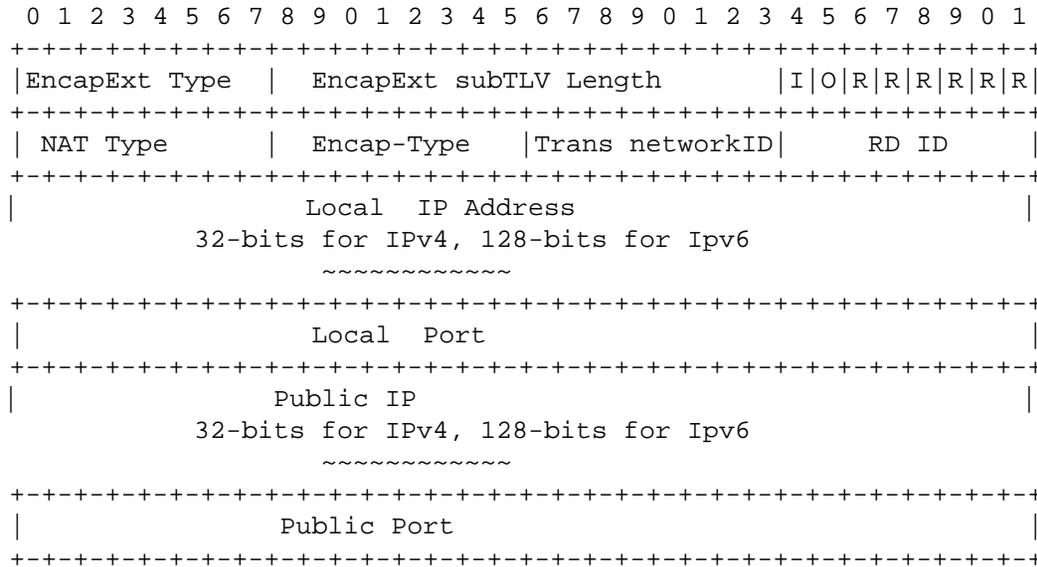
Flags:
-I bit (CPE port address or Inner address scheme)
  If =0 → inner addr is IPv4.
  If =1 → inner address is IPv6.
-O bit (Outer address scheme):
  If =0 → the public (outer) address is IPv4.
  If =1 → the public (outer) address is IPv6.
-R bits: reserved for future use. Must be set to 0 now.
NAT Type: without NAT; 1:1 static NAT; Full Cone; Restricted Cone; Port Restricted Cone; Symmetric; or Unknown (i.e. no response from the STUN server).

Encap Type：the supported encap types for the port facing public network, such as IPsec+GRE, IPsec+VxLAN, IPsec without GRE, GRE (when packets don't need encryption)
Transport Network ID: Central Controller assign a global unique ID to each transport network；
RD ID: Routing Domain ID，Need to be global unique.
Local IP: The local (or private) IP address of the port；
Local Port: used by Remote SDWAN node for establishing IPsec to this specific port.
Public IP: The IP address after the NAT. If NAT is not used, this field is set to NULL.
Public Port: The Port after the NAT. If NAT is not used, this field is set to NULL.