SDWAN Edge Discovery

draft to IDR

Linda Dunbar
Sue Hares
Robert Raszuk
Kausik Majumdar
June 2020

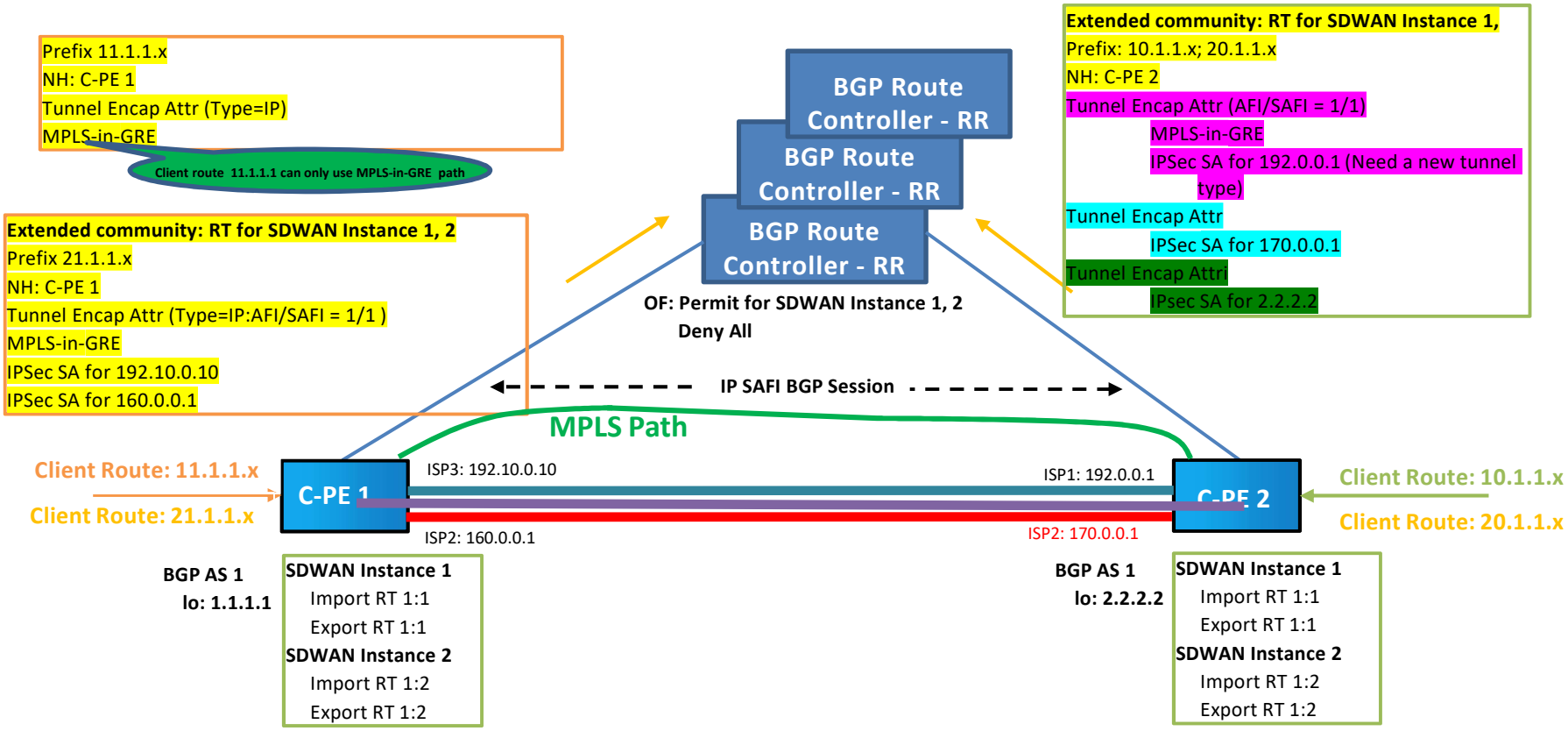
Intention of the draft

- Purpose:
  - **OVERLAY SIGNALLING ONLY**
  - Using BGP UPDATE to advertise properties of SDWAN Edge
- We would like to hear your feedback.

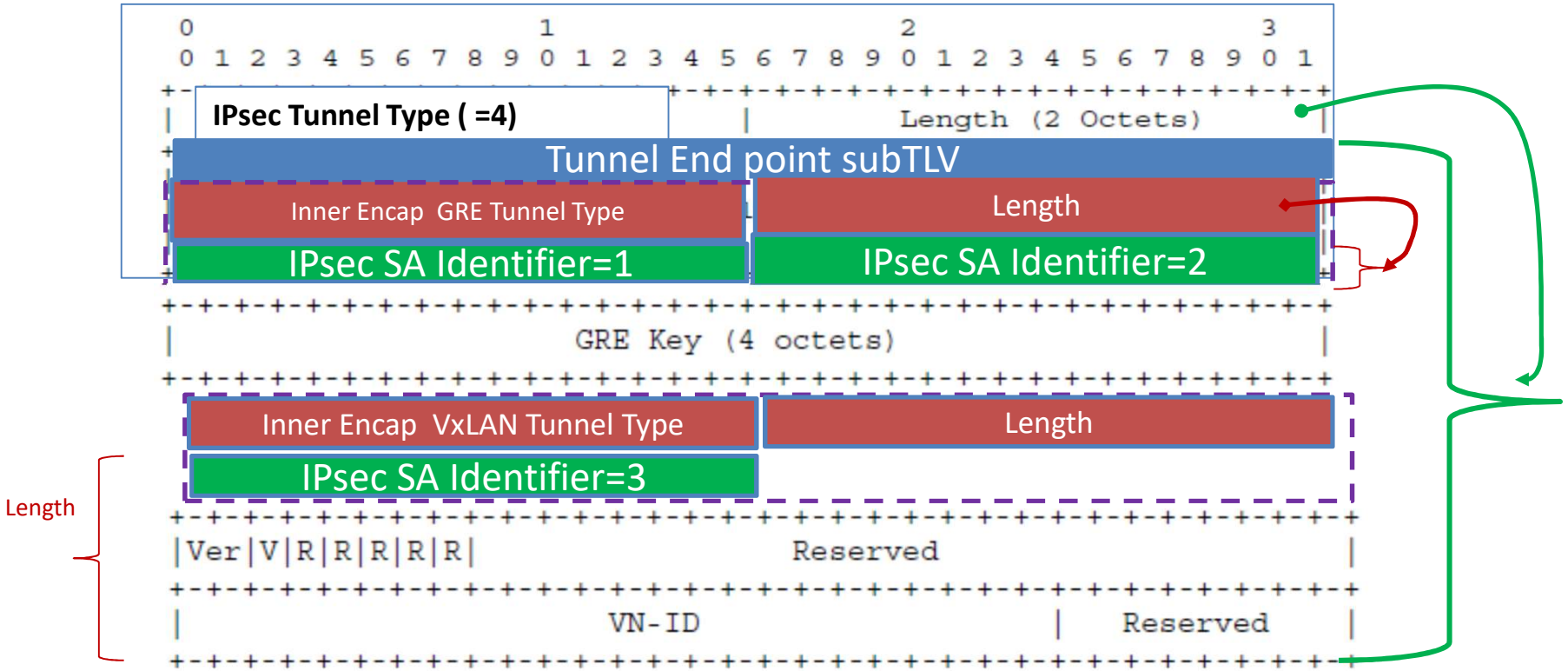
Basic Schemes

- **Two types of BGP Update: advertised at different pace**
  - SAFI = 1 regular IP client routes properties propagation
  - SAFI = 74 Underlay network properties propagation
- **Client routes UPDATE: same as before, e.g.**
  - NLRI:
    - Prefix: 10.1.1.1
  - Tunnel-encap path attributes to describe all the tunnels that the prefix can be carried, e.g.
    - MPLS, GRE, VxLAN
    - IPsec tunnel Identifiers if there is an IPsec SA especially established for this prefix
    - Or identifier of the WAN port based IPsec tunnel
- **Underlay network properties UPDATE:**
  - WAN port address
  - Underlay ISP property TLV
  - NAT TLV (Optional)
  - IPSec TLV

Client Routes IP SAFI =1 NLRI Update



Encoding Example for a client route that can be carried by two IPsec SA with GRE inner encapsulation and another IPsec SA with VxLAN inner encapsulation Using IPsec-SA-Group



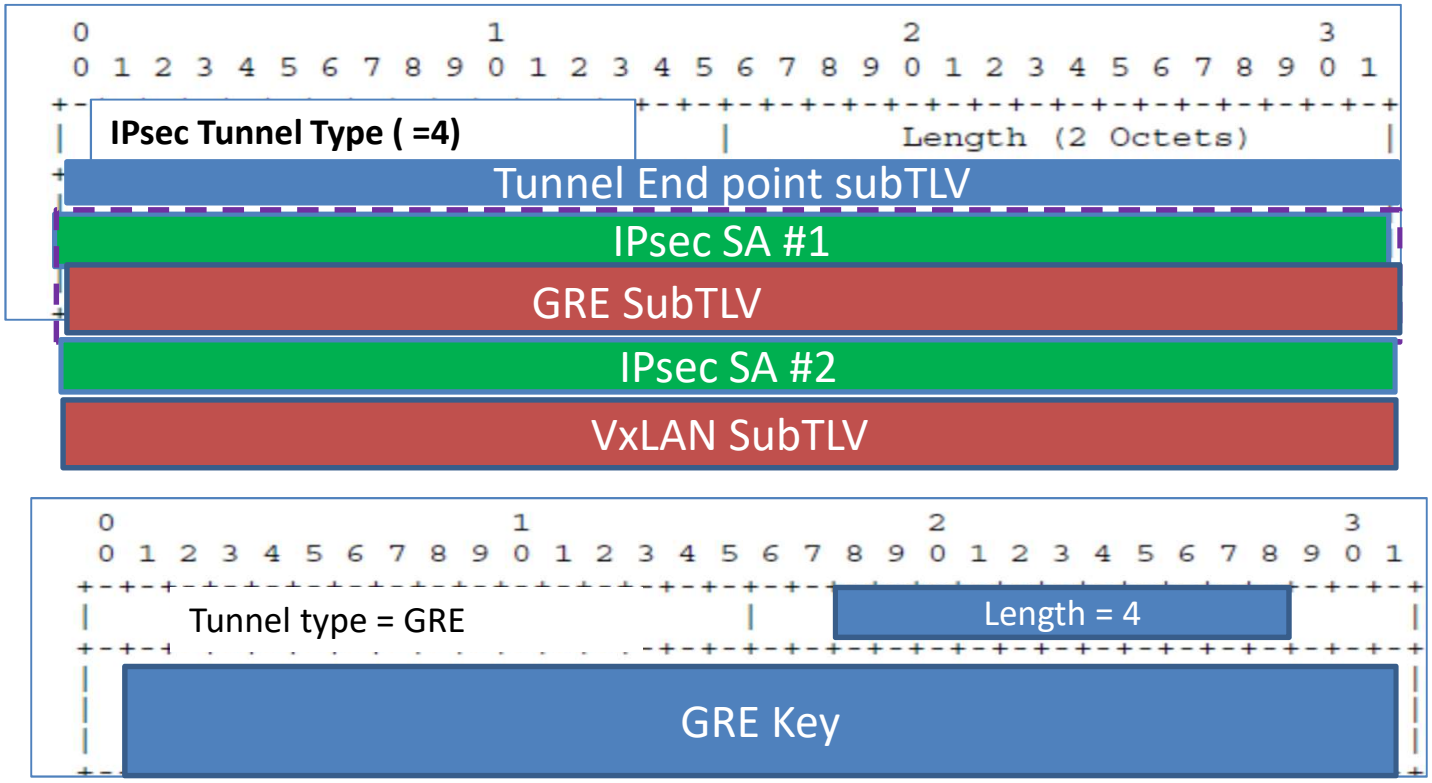
The Inner Encap Type value will take the value specified by Tunnel Encap Section 12.5:

types 8 (VXLAN), 9 (NVGRE), 11 (MPLS-in-GRE), and 12 (VXLAN-GPE) in the "BGP Tunnel Encapsulation Tunnel Types" registry.

types 1 (L2TPv3), 2 (GRE), and 7 (IP in IP) in the "BGP Tunnel Encapsulation Tunnel Types" registry.

Encoding Example for a client route that can be carried by two IPsec SA with GRE inner encapsulation and another IPsec SA with VxLAN inner encapsulation

Using IPsec-SA-ID SubTLV



SDWAN SAFI (=74) NLRI Encoding Format
Only for Underlay Network Properties Advertisement

NLRI

192.0.0.1

SDWAN SAFI NLRI: <Site-Type, IPSec-SA-Type, Port-Local-ID, SDWAN-Site-ID, SDWAN-Node-ID>

Attributes:

Geo-location Sub-TLV

Tunnel Encaps Attribute (23)

Tunnel Type: SDWAN-Underlay (to be assigned by IANA)

NAT Sub-TLV (Optional)

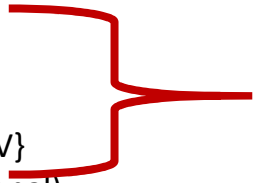
IPsec SA Nonce Sub-TLV (Mandatory)

IPsec SA Public Key Sub-TLV

IPsec SA Transform Sub-TLV

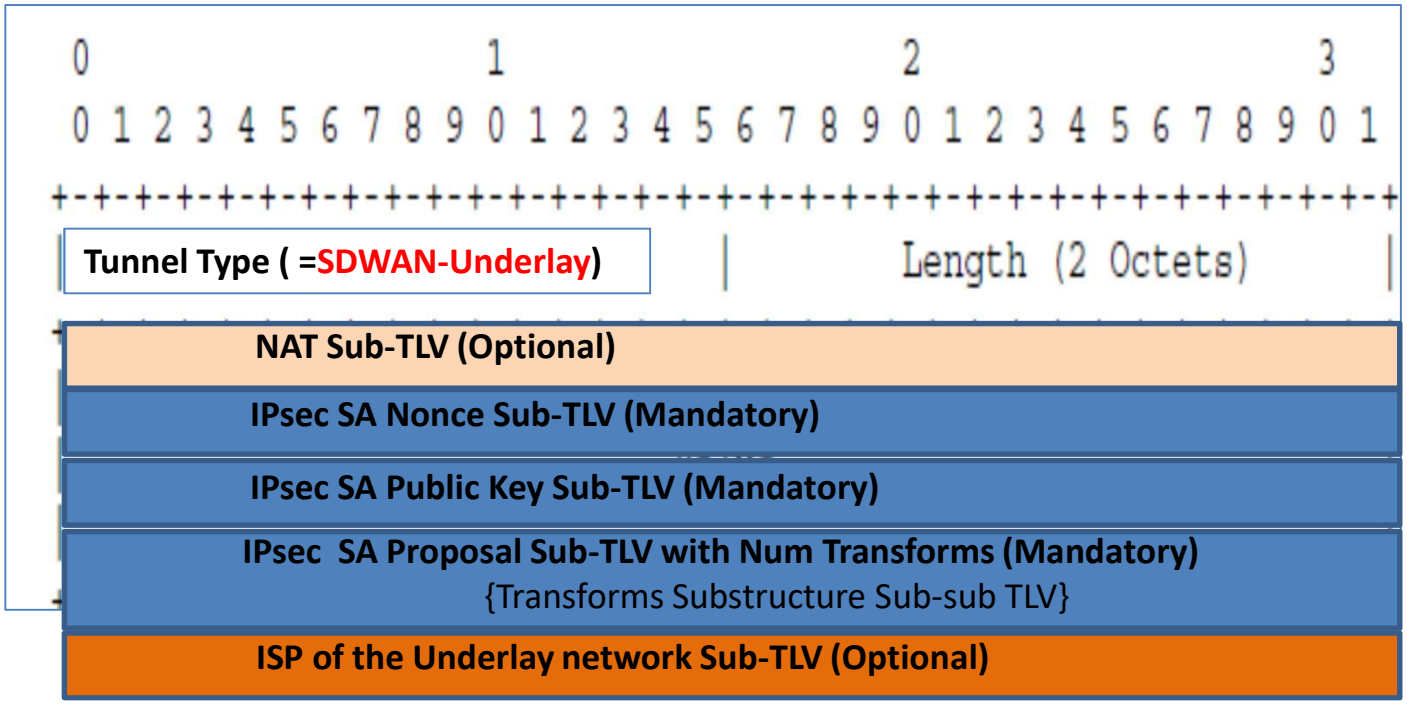
{Transforms Substructure Sub-sub TLV}

ISP of the Underlay network Sub-TLV (Optional)

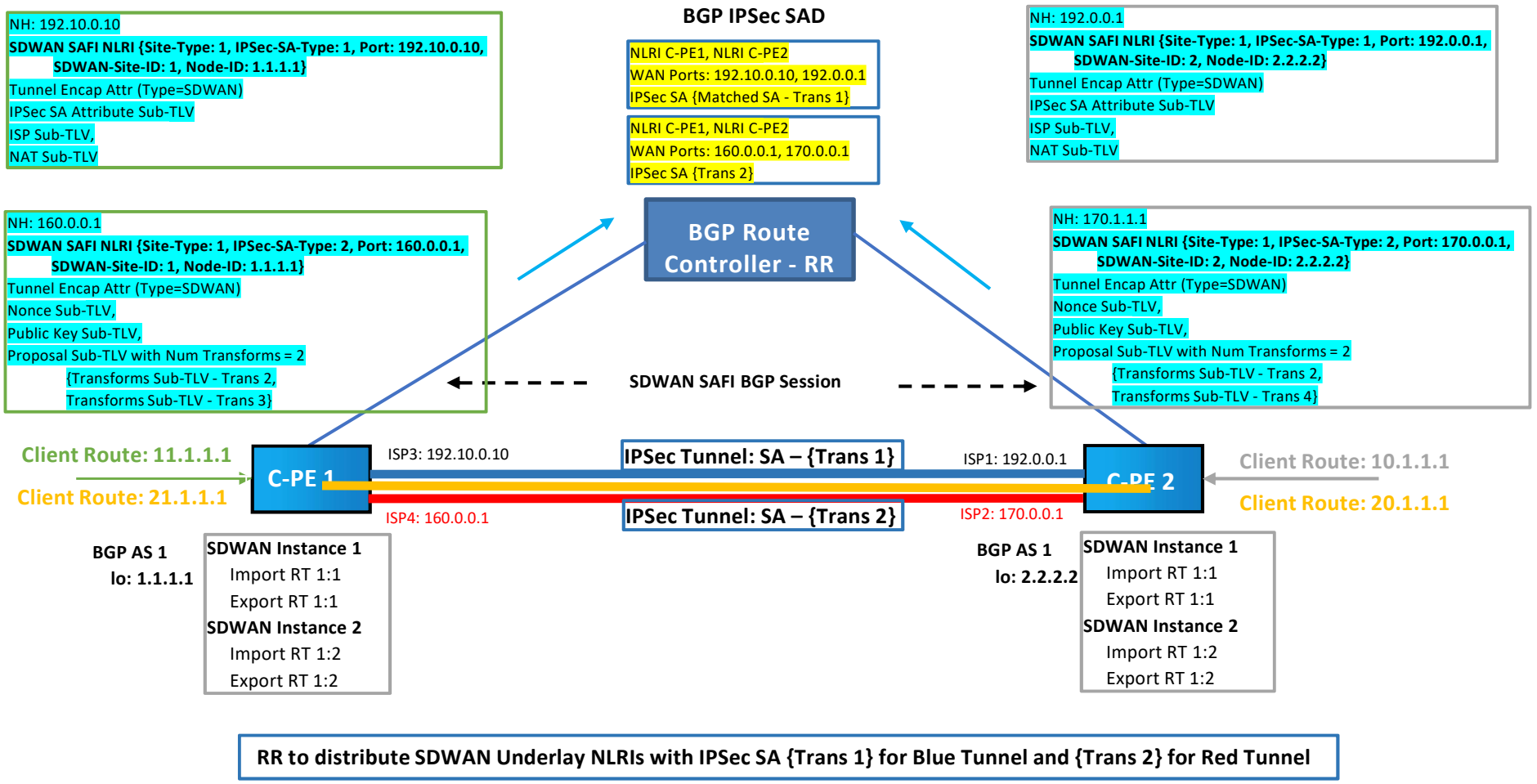


IPsec-SA Attributes

NLRI SAFI = SDWAN: Tunnel Encap Attribute to carry IPsec SA Property

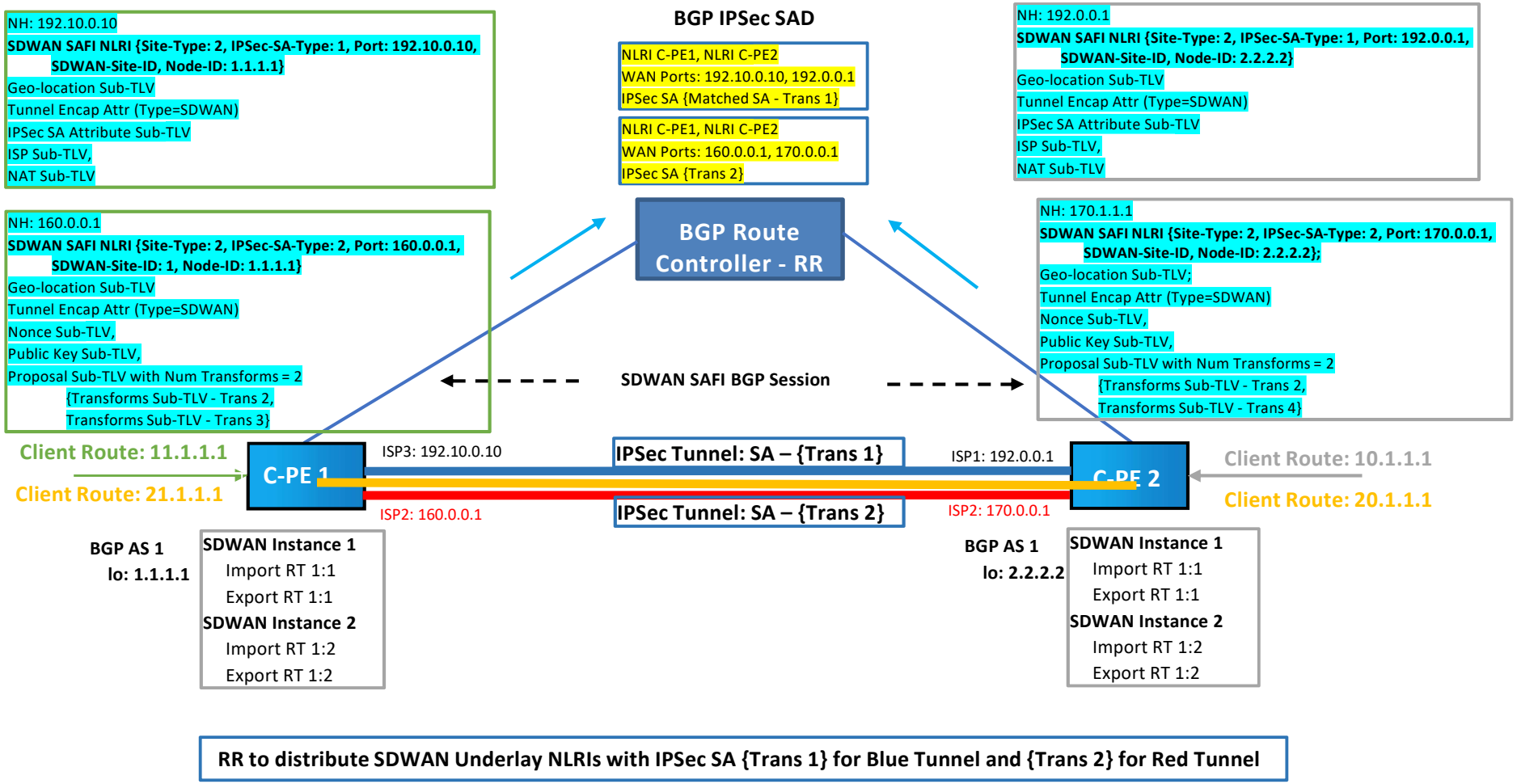


SDWAN SAFI NLRI Exchange – Establish Port based IPSec SA



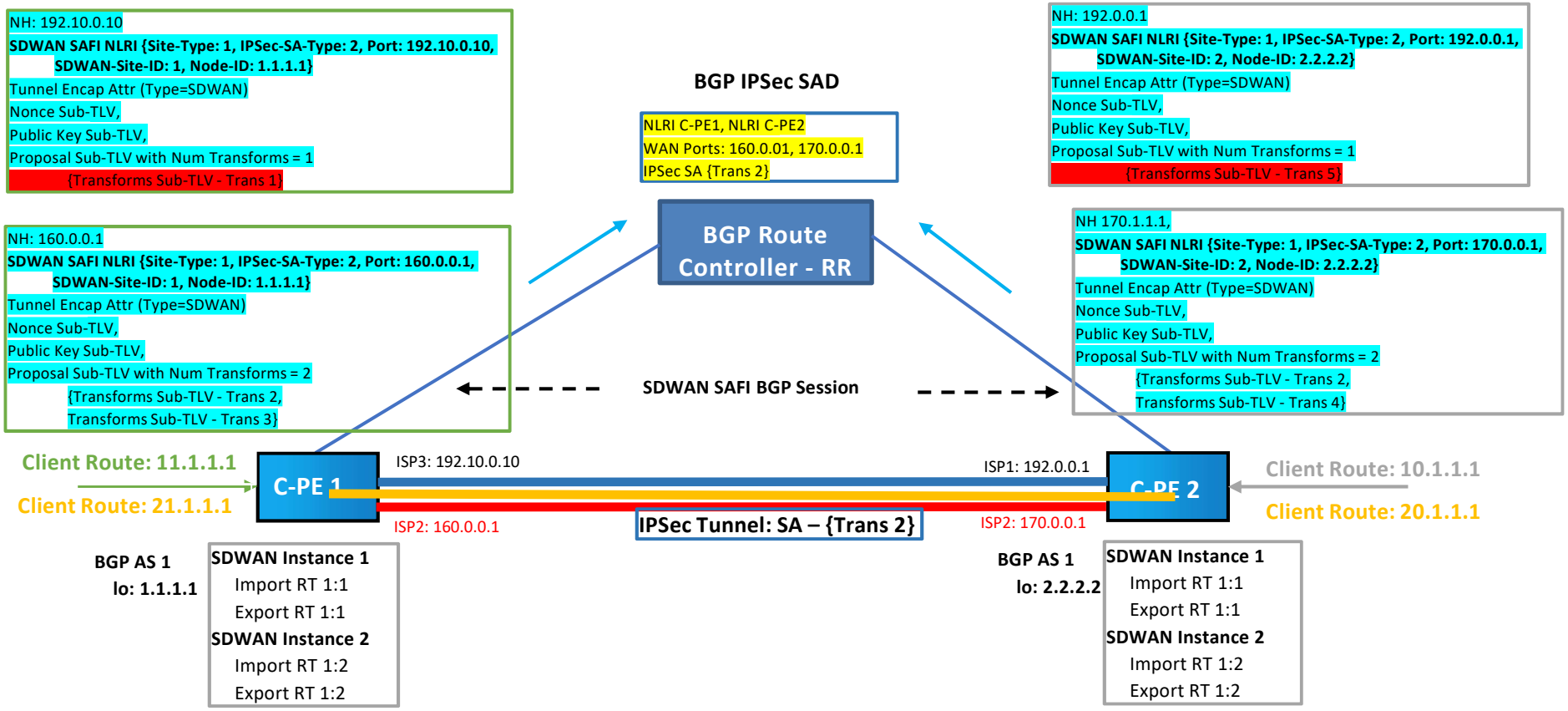
RR to distribute SDWAN Underlay NLRIs with IPSec SA {Trans 1} for Blue Tunnel and {Trans 2} for Red Tunnel

SDWAN SAFI NLRI Exchange – Includes Geo-Location Sub-TLV



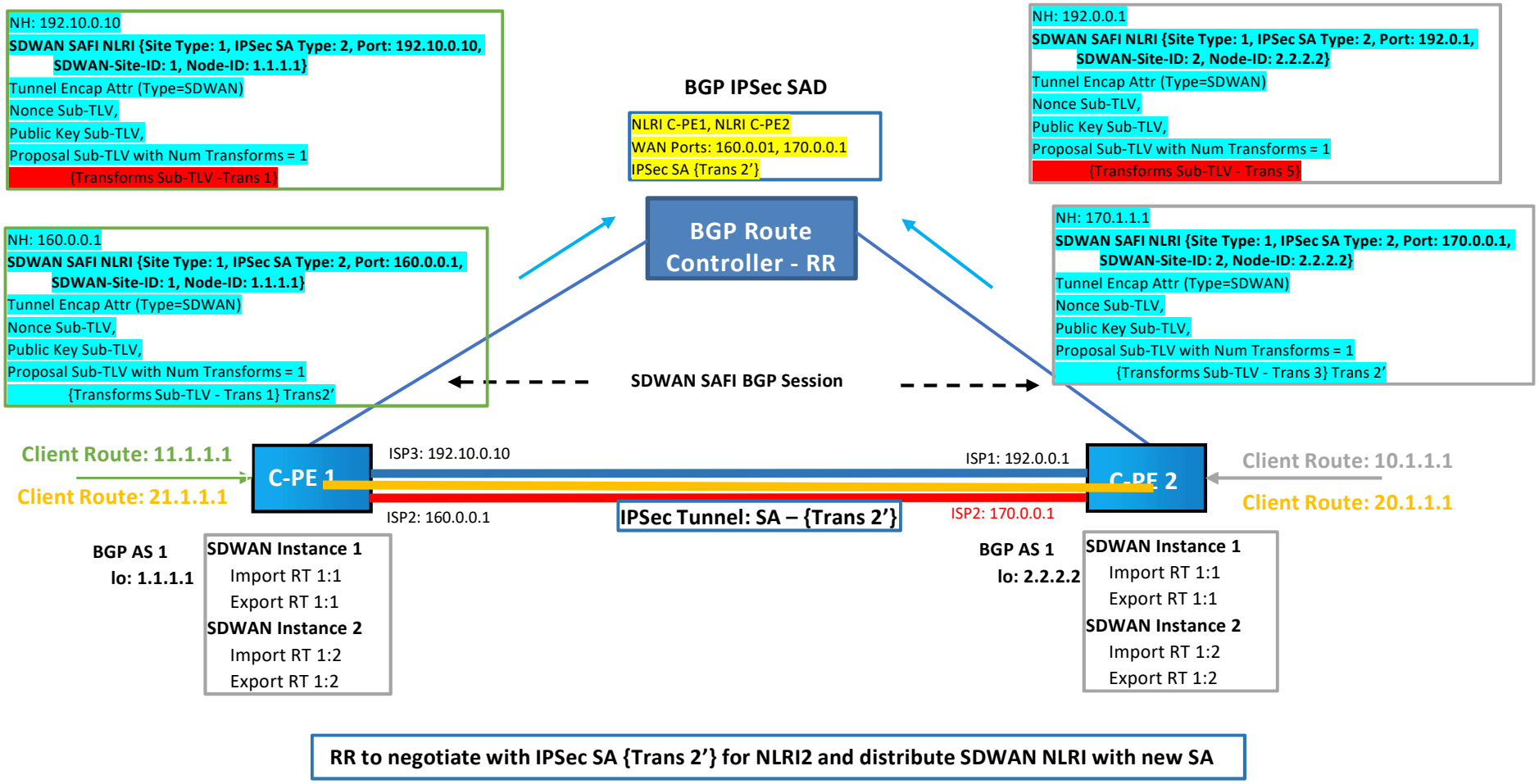
RR to distribute SDWAN Underlay NLRIs with IPSec SA {Trans 1} for Blue Tunnel and {Trans 2} for Red Tunnel

SDWAN SAFI NLRI Exchange – Mismatch in IPSec SA



RR to distribute SDWAN Underlay NLRIs with IPSec SA {Trans 2} for Red Tunnel. It won't distribute IPSec SA for the Blue Tunnel as there is a mismatch.

SDWANSAFI NLRI Exchange – Negotiates New IPSec SA (Future Scope)



Propagation Method

- Background:
  - SDWAN peers can be far apart and can be connected by untrusted networks.
  - A SDWAN edge might not know their authorized communication peers.
- A SDWAN Edge node “A” will only advertise its own properties to its designated Controller (RR) via secure tunnel (TLS, IPsec, or others).
- The Controller (RR) relay the received information further if needed
  - E.g. propagates the received UPDATE messages from one edge (say A) to the authorized peers.
  - using RFC 4684 to constrain Route distributions to authorized peers, or
  - By static configuration of the Peer Policies on RR (for small number of SDWAN edge nodes)
- The authorized peers can exchange IKEv2, establish the secure data channels (IPsec), and exchange more information among each other.

Comparison with SECURE-EVPN

- The SDWAN use cases are defined in the draft-dunbar-bess-bgp-sdwan-usage-06.
- The SECURE-EVPN draft tries to address the solution for the use cases defined in the above draft.
- It defines the solution using two new Tunnel Types **ESP-Transport** and **ESP-in-UDP-Transport** defined in Tunnel Encapsulation Attribute draft along with IPSec SA Sub-TLVs to get negotiated through Route Reflector.
- In this approach, it exchanges the client routes for different BGP AFI/SAFI along with the new Tunnel Types and the associated IPSec SA Sub-TLVs.
- Although this approach works, but may not be scalable solution for large enterprises as more client route updates need to be exchanged along with ESP-Transport or ESP-in-UDP-Transport Tunnel Type through RR. It increases the BGP client route update message size in the wire as it requires to carry several IPSec SA Sub-TLVs along with new Tunnel Type.
- Besides the incremented size for each client routes, there are some other additional functionality is missing in the SECURE-EVPN draft.
- The BGP based SDWAN Edge Discovery draft addresses the use cases defined in draft-dunbar-bess-bgp-sdwan-usage-06 in a different way than SECURE-EVPN approach. This approach is scalable in the network. It also tries to address some of the missing functionalities that is not covered under SECURE-EVPN solution.

Comparison with SECURE-EVPN

- The BGP based SDWAN Edge Discovery draft takes the approach of separating out of the client routes carrying the SDWAN underlay properties.
- The new SDWAN SAFI (=74) has been defined by IANA and that is used here to update the underlay properties using SDWAN SAFI.
- In this approach, the client routes need to carry new Tunnel Type called **SDWAN-Underlay** along with 16 bytes IPSec Sub-TLV to cross-connect the client routes with the SDWAN underlay routes.
- There is no additional changes in the client routes besides carrying the new IPSec Sub-TLV. Any type of client routes carrying the new IPSec SA Sub-TLV can associate with the SDWAN SAFI underlay route.
- The SDWAN SAFI underlay route UPDATE is for an edge node to advertise the properties of the directly attached underlay networks, including the underlay network ISP information, NAT information, the supported IPSec SA properties, the Geo-location of the edge node etc.
- This UPDATE is for peers to discover remote node's properties to establish IPsec tunnels and to traverse NAT. Once established IPsec tunnel between the two peers, it should have a unique tunnel identifier or WAN port for the Tunnel-Encap Path Attribute to indicate the tunnel can carry the attached client routes in the NLRI. By using the reference, the client routes update can be advertised independently from the underlay network properties, IPsec SA key attributes, and IPsec SA rekey exchange process.

Comparison with SECURE-EVPN

- The BGP based SDWAN Edge Discovery draft also defines two different mechanism to exchange IPSec SA properties between the SDWAN SAFI peers:
  - One approach is to use the Simple IPSec Security Association properties to negotiate through IPSec SA Sub-TLV.
  - The other approach is to use the full set of IPSec Sub-TLVs including Nonce, Public Key, Proposal and Transform Sub-TLVs. This draft doesn't try to re-define these sub-TLVs, rather they are referenced through SECURE-EVPN draft.
- As a result, it brings additional flexibility for the edge node to use IPSec SA properties to negotiate through simple IPSec SA Sub-TLV and overall less overhead for the application to use the full set of IPSec SA Sub-TLVs unless that is really required.
- As the client routes update are separated out from the underlay properties using SDWAN SAFI NLRI, this method would scale better in the network. The SDWAN SFAI underlay properties to be negotiated between the peers through RR only once. After that, only when the duration of the key expires the re-negotiation of the IPSec SA attributes happens. Hence, it is expected only minimal BGP SDWAN SAFI updates to happen in the network through RR.
- On the other hand, as the SDWAN SAFI underlay properties are loosely coupled with the client routes. The advertisement of the client routes are completely independent of the SDWAN SAFI routes update. This method doesn't impose any additional overhead other than carrying **SDWAN-Underlay** Tunnel Type along with 16 bytes IPSec Sub-TLV. Hence, client routes scale is not impacted by the underlay behavior.

Comparison with SECURE-EVPN

- Besides separating out the overlay and underlay route properties, in this approach SDWAN SAFI NLRI can carry optional sub-TLVs like underlay network ISP information, NAT information, the Geo-location of the edge node etc. That provides additional underlay network properties and geo-location to the IPSec tunnel end-points. As a result, the BGP SDWAN SAFI peers can take better decision on what policies to apply for the associated client routes as it traverses through the IPSec tunnel. These functionalities are not addressed in the SECURE-EVPN draft.
- The BGP based SDWAN Edge Discovery draft also brings the fine-grain network segmentation using the concept called SDWAN Instance Identifier. The SDWAN network can be segmented to multiple instances. Each SDWAN edge node may need to support multiple SDWAN instances. It allows client route to map to a certain SDWAN segmentation based on the client's policy. This method is similar to IP VPN VRF or EVPN EVI concept.
- As a result, when SDWAN edge node is acting as CE device the SDWAN Instance brings the L3 VRF concept in the CE device to provide the fine-grain segmentation for the client routes. It brings the constrained propagation of the client routes using SDWAN Instance route targets.

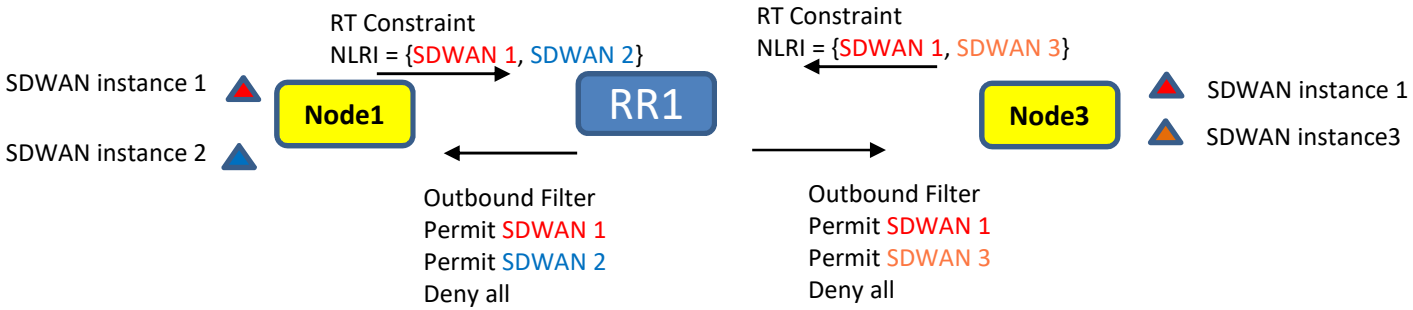
Comparison with SECURE-EVPN

- This method also brings additional flexibility for the client routes that are associated with different SDWAN Instances to have different IPSec SA properties aka IPSec Tunnels. Even though different client routes that are part of different SDWAN Instances tied with same underlay WAN port but they can have their separate IPSec Tunnels to solve different use cases as defined in draft-dunbar-bess-bgp-sdwan-usage-06. Again, this functionality is missing in the SECURE-EVPN draft.
- In summary, BGP based SDWAN Edge Discovery draft not only scale better in the network through separating out the client routes from the SDWAN SAFI underlay routes, but it also brings many more flexibility in the network through the complete solution for the SDWAN use cases. This solution also takes the reference of pre-defined IPSec SA Sub-TLVs from the SECURE-EVPN draft, thus make it a part of the overall complete solution .

SDWAN Properties
Constrained Distribution

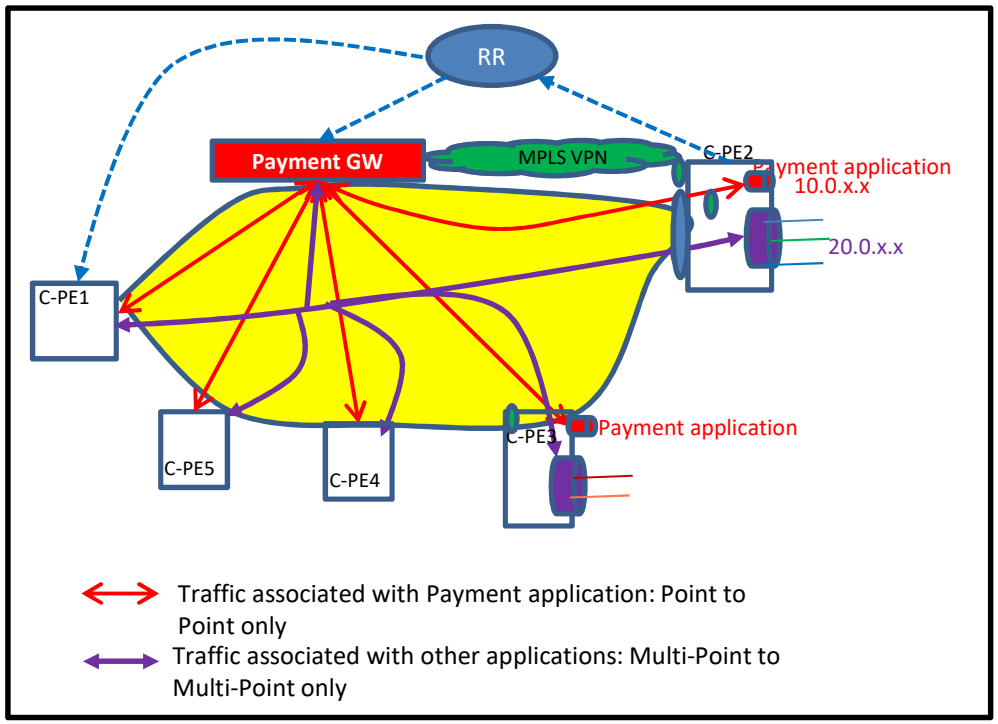
To prevents the propagation of SDWAN Update to edges that are not interested or not authorized

- **Static (**For small number of SDWAN edge nodes)
  - Polices can be configured on RR for RR to build outbound route filter, making RR as Route Controller (RC)
- **Dynamic:** using BGP RT constrained distribution (RFC4684)
  - A SDWAN edge sends RT Constraint (RTC) NLRI to the RR.
  - The RR installs an outbound route filter



Encoding
To differentiate SDWAN instances

Applications Based Segmentation in SDWAN



Characteristics:

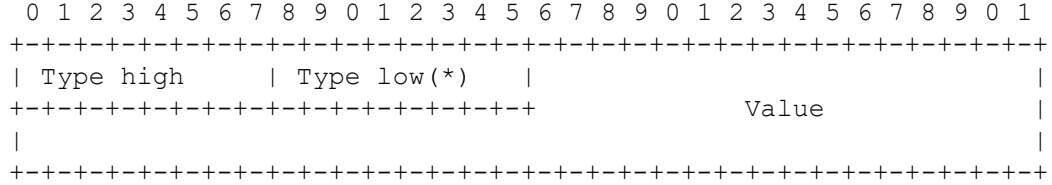
- RED route (payment applications) can only be propagated to "Payment GW"
- Purple routes needs to be propagated to all other nodes

Very similar to VPNs.
But need to differentiate from traditional MPLS VPNs

Encoding for SDWAN Segmentation for Control Plane

- Here is an approach to differentiate BGP Update by different SDWAN Instances:
- 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.
  - Question:
    - Use RFC7153, especially for 4 octet AS, use the Low 6 bit value or the Low Type

RFC4360: Extended Community for SDWAN Route Target



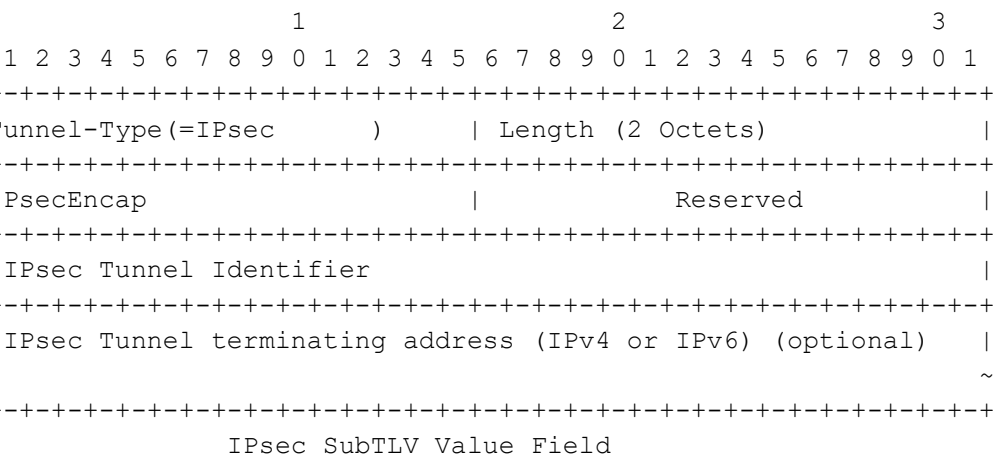
question:
- assign a new 6-bit VALUE for “SDWAN Instance Target” after the I & T bits?
- Or Request a new value of the low- order octet of the Type field for this community (different from the VPN Route Target 0x02)?

SDWAN Instance Identifier in Data Plane

- Packets from different SDWAN network instances (or segmentations) need to have their corresponding SDWAN instance identifier encoded in the header.
- For a SDWAN edge node which can be reached by both MPLS and IPsec path, the client packets reached by MPLS network will be encoded with the MPLS Labels based on the scheme specified by RFC8277.
- For GRE Encapsulation within IPsec tunnel, the GRE key field can be used to carry the SDWAN Instance ID. For NVO (VxLAN, GENEVE, etc.) encapsulation within the IPsec tunnel, Virtual Network Identifier (VNI) field is used to carry the SDWAN Instance ID.
- [Note: the SDWAN Instance ID is same as EVI in EVPN, or VNI if VxLAN is used].

Client routes advertisement

- One new IPsec Tunnel Type for binding client's routes with a prior established IPsec Tunnel needs to be added to The Tunnel Encapsulation Attribute [TUNNEL-ENCAP].
- The Client routes can be IP AFI/SAFI prefix, L2VPN-EVPN AFI/SAFI.



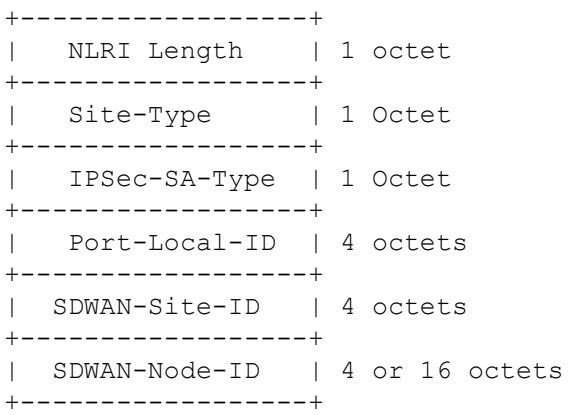
- IPsecEncap (2 octet) indicates the encapsulation type for payload:
  - Ox0000: no extra encapsulation with the IPsec ESP tunnel;
  - 0x0001: GRE encapsulation within the IPsec ESP tunnel;
  - 0x0002: VxLAN encapsulation within the IPsec ESP Tunnel;
  - Ox0003: GENEVE encapsulation within the IPsec ESP Tunnel;
  - More values will be added later.

- The IPsec Tunnel identifier (4 Octet) is for cross reference the IPsec tunnel attributes being advertised by the Underlay Network Properties Advertisement UPDATE. Since IPsec tunnel has a lot of attributes, the IPsec Tunnel Identifier is used instead of listing all those attributes in the client routes update.

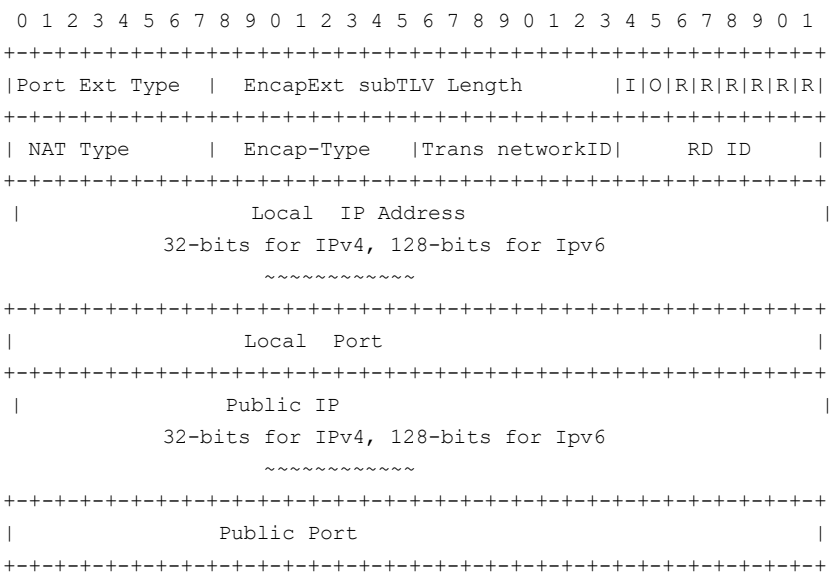
Encoding
for
advertising attached Underlay network properties

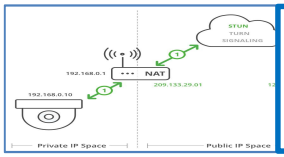
Encoding for Underlay Properties in MP-NLRI

- Site Type:
  - 1: Represents a type of deployment that a numeric number as Site-ID is enough for the Management system to map to the accurate location. E.g. the all edge nodes are managed by a single Management system.
  - 2: Represents large SDWAN heterogeneous deployment where Site IDs has to be represented by proper Geo-location of the Edge Nodes [LISP-GEOLoc].
- IPSec SA Type:
  - 1 – Simple IPSec Security Association properties defined in IPSec SA Sub-TLV.
  - 2 – The full set of IPSec Sub-TLVs including Nonce, Public Key, Proposal and Transform Sub-TLVs.
- Port local ID: can be locally significant. The detailed properties about the network connected to the port are further encoded in the Tunnel Path Attribute.
- SDWAN-Site-ID: used to identify a common property shared by a set of SDWAN edge nodes
- SDWAN Edge Node ID: (IPv4 or IPv6)



Ext SubTLV for NAT





Edge node get NAT properties via STUN requests/responds.
Peers may not be able to access the STUN server.

- Flags:

-I bit (CPE port address or Inner address scheme)

If set to 0, indicate the inner (private) address is IPv4.
If set to 1, it indicates the inner address is IPv6.

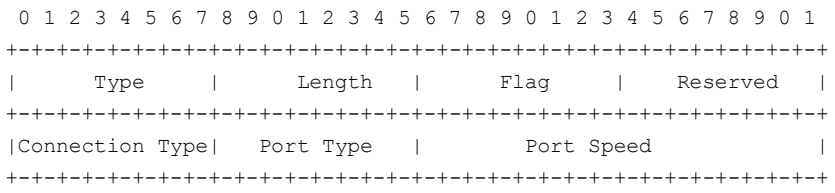
- O bit (Outer address scheme):

If set to 0, indicate the public (outer) address is IPv4.
If set to 1, it indicates 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 encapsulation 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 edge 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.

ISP of the Underlay Network SubTLV



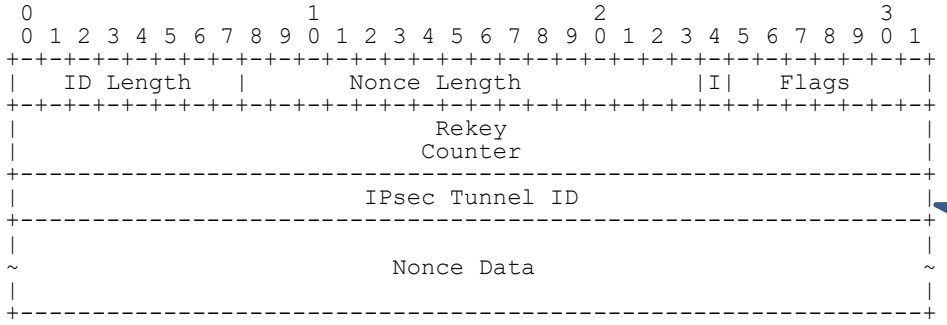
- Type: To be assigned by IANA
- Length: 6 bytes.
- Flag: a 1 octet value.
- Reserved: 1 octet of reserved bits. It SHOULD be set to zero on transmission and MUST be ignored on receipt.
- Connection Type: There are two different types of WAN Connectivity. They are listed below as:
  - Wired – 1
  - Wireless – 2
- Port Type: There are different types of ports. They are listed below as:
  - Fiber Cable – 1
  - Coax Cable – 2
- Port Speed: The port seed is defined as 2 octet value. The values are defined as Gigabit speed.

Two Types of IPsec SA attributes (only use one) Sub-Sub-TLV

- Full set: with multiple subTLVs for full suite of IPsec SA attributes
  - Nonce Sub-TLV
  - Public Key Sub-TLV
  - Proposal Sub-TLV: to indicate the number of Transform subTLVs to be included
    - Transforms Substructure Sub-TLV
- Or, Simple Deployment with limited number of parameters
  - One SubTLV to represent Public Key, Nonce, ReKey, Transform

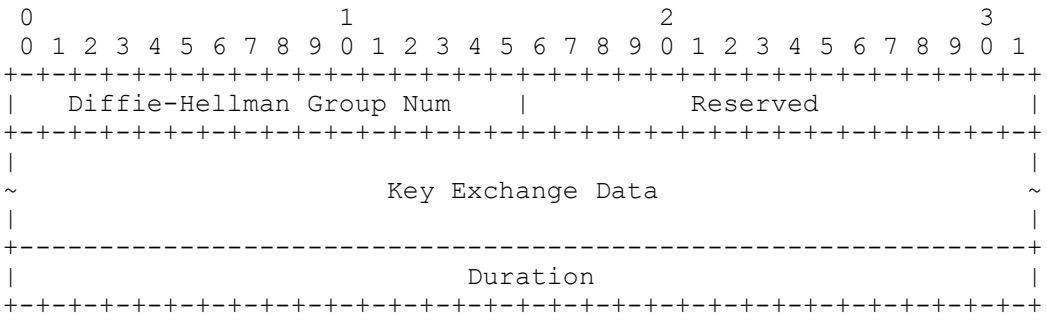
Nonce Sub-TLV, Public Key Sub-TLV

- Nonce Sub-TLV:



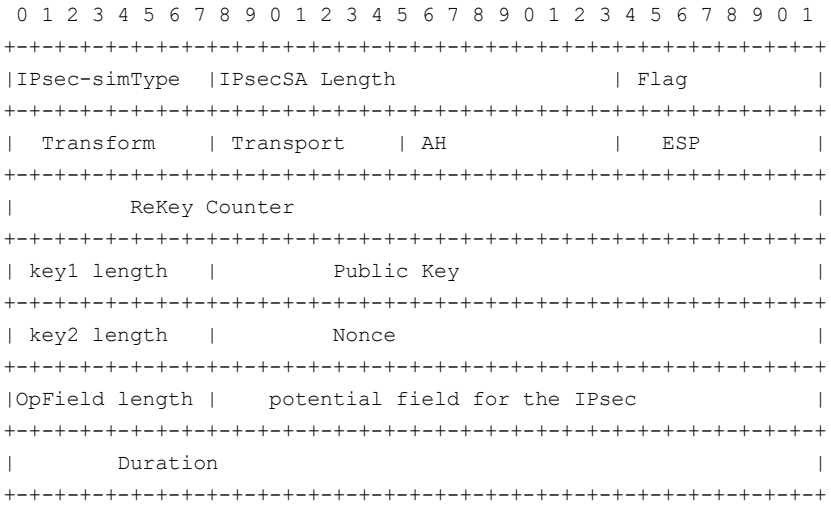
IPsec Tunnel ID is for cross reference by the client route NLRI Tunnel Encap Path Attribute for IPsec tunnel. For simple case of deployment, there is no need to differentiate different IPsec tunnels for different client routes, the IPsec Tunnel ID is 0.

- Public Sub-TLV:



Simplified IPsec SA attributes advertisement

Multiple client flows can be carried by the Port Based Secure Tunnel



- IPsec-simType: to be assigned by IANA.
- Flags: for future usage.
- Transform (1 Byte): the value can be AH, ESP, or AH+ESP.
- Transport (1 byte): Indicate Tunnel Mode or Transport mode
- AH (1 byte): AH authentication algorithms supported, which can be md5 | sha1 | sha2-256 | sha2-384 | sha2-512 | sm3. Each SDWAN edge node can have multiple authentication algorithms; send to its peers to negotiate the strongest one.
- ESP (1 byte): ESP authentication algorithms supported, which can be md5 | sha1 | sha2-256 | sha2-384 | sha2-512 | sm3. Each SDWAN edge node can have multiple authentication algorithms; send to its peers to negotiate the strongest one. Default algorithm is AES-256.
- When node supports multiple authentication algorithms, the "Transform Sub-TLV" described by [SECURE-EVPN] can be used to describe the additional algorithms supported by the node.
- Rekey Counter: 4 bytes
- Public Key: IPsec public key
- Nonce: IPsec Nonce
- Optional Field: other potential information associated with IPsec
- Duration: SA life span.