SDWAN Edge Discovery


Linda Dunbar
Sue Hares
Robert Raszuk
Kausik Majumdar
Nov 2020
Intention of the draft

• Purpose:
  – Tag the client routes with Extended Communities for tunnel-encaps + Color
  – Using BGP UPDATE to advertise properties of hybrid underlays of SDWAN
    • Properties: Security Associations, port in SDWAN

• Based on:
  – draft-ietf-idr-tunnel-encaps-20
Basic Schemes

• **UPDATE U1: Client routes UPDATE:**
  – NLRI: Prefix: 10.1.1.1
  – Next-hop set to the address of the Edge node, e.g. C-CPE2 (for recursive lookup)
  – Encapsulation Extended Community
    (Tunnel-Type=SDWAN-Hybrid or other existing tunnel types);
  – Color Extended Community;

• **UPDATE U2: Detailed tunnel attributes of the underlay networks:**
  – Tunnel-egress-endpoint Sub-TLV
  – Extended Port Sub-TLV (Includes the ISP property Sub-TLV (Optional))
  – Geo Location Sub-TLV
  – IPSec Sub-TLVs or
  – IPsec SA IDs (for preconfigured IPSec SSs)
DETAILS – AT HIGH LEVEL
UPDATE #1: Client Routes IP SAFI =1 NLRI

**BGP Route Controller - RR**

**MP-NLRI:** (AFI/SAFI = 1/1)
- Prefix: 10.1.1.x; 20.1.1.x
- Nexthop 2.2.2.2 /* C-PE-2*/

**Encapsulation Extended Community:** TunnelType = SDWAN-Hybrid
**Color Extended Community:** Color = RED

**NLRI:** (AFI/SAFI = 1/1)
- Prefix: 30.1.1.x
- Nexthop 2.2.2.2 /* C-PE-2*/

**Encap Extended Community:** TunnelType = SDWAN-Hybrid
**Color Extended Community:** Color = Purple

30.1.1.x can only go to C-PE3

To indicate multiple types of underlay networks: MPLS VPN, IPsec, etc. Maybe “Hybrid” underlay is a better name?

**L3VPN path**

**C-PE1**
- lo: 1.1.1.1
- ISP: 192.10.0.10

**C-PE2**
- lo: 2.2.2.2
- ISP: 192.0.0.1

**C-PE3**
- lo: 3.3.3.3

Client Route: 11.1.1.x
Client Route: 21.1.1.x

Client Route: 10.1.1.x
Client Route: 20.1.1.x
Client Route: 30.1.1.x

Port 170.0.0.1 has multiple IPsec SAs
- IPsec SA (RED) -> C-PE1
- IPsec SA (Purple) -> C-PE3
BGP UPDATE #2: Detailed Tunnel attributes: “Color” encoded in SDWAN NLRI

BGP Route Controller - RR

ISP1: 192.0.0.1
ISP2: 160.0.0.1
ISP3: 160.0.0.1 (L3VPN)

Port 170.0.0.1 has multiple IPsec SAs
IPsec SA (RED) -> C-PE1
IPsec SA (Purple) -> C-PE3

Client Route: 11.1.1.1
Client Route: 21.1.1.1
ISP: 192.10.0.10
ISP: 160.0.0.1
ISP: 170.0.0.1

L3VPN path

Client Route: 20.1.1.1
Client Route: 30.1.1.1
ISP: 192.0.0.1
ISP: 160.0.0.1
ISP: 170.0.0.1

Port address (170.0.0.1) is encoded in the Port Sub-TLV
IPsec SAs for 192.0.0.1
IPsec SA (RED) for 170.0.0.1
IPsec SA (Purple) for 170.0.0.1
Color-SubTLV = Purple

Nexthop: x.x.x.x
NLRI: SAFI = SDWAN
Tunnel Encap Attr (SDWAN-Hybrid)
Tunnel-egress-endpoint SubTLV
Port Sub-TLV/ ISP-sub-TLV
IPSec SA-IDs

Color RED encoded in the SDWAN NLRI Header

ISP3: 192.10.0.10
ISP1: 192.0.0.1
ISP2: 170.0.0.1
ISP3: 160.0.0.1 (L3VPN)

Color SubTLV = Purple

IPsec SAs for 172.0.0.1
IPsec SA (RED) for 170.0.0.1
IPsec SA (Purple) for 170.0.0.1

C-PE1
C-PE2
C-PE3

Client Route: 10.1.1.1
Client Route: 20.1.1.1
Client Route: 30.1.1.1
ISP: 192.0.0.1
ISP: 160.0.0.1
ISP: 170.0.0.1

Port 170.0.0.1 has multiple IPsec SAs
IPsec SA (RED) -> C-PE1
IPsec SA (Purple) -> C-PE3
When IPsec SAs are pre-configured, IPsec SA-IDs can be included in the SDWAN-Hybrid Tunnel Path Attribute.

- **Site-Type= 1 (No GeoLoc SubTLV);**
- **Port-Local-ID =* (apply to all ports. Each IPSec SA has it is unique End-Point-TLV);**
- **SDWAN-Site-ID = Color-extended Community (carried by the client routes),**
- **SDWAN-Node-ID**

**Tunnel Type = SDWAN-Hybrid**

- **Tunnel Egress Endpoint Sub-TLV**
- **Extended Port SubTLV (with option to include ISP Sub-TLV)**
  - IPsec SA #1 Identifier
  - IPsec SA #2 Identifier
- **Color Sub-TLV: Color Extended Community**

*Multiple underlays tunnels*
Tunnel Type ( =SDWAN-Hybrid)

Tunnel-Egress-Endpoint-SubTLV

- Extended Port Sub-TLV (Optional) or use Tunnel end point sub-TLV to describe the WAN Port Address

- IPsec SA Nonce Sub-TLV
- IPsec SA Public Key Sub-TLV
- IPsec SA Proposal Sub-TLV with Num Transforms
  - {Transforms Substructure Sub-sub TLV}

Site-Type=1 (No GeoLoc SubTLV);
Port-Local-ID =* (apply to all ports. Each IPsec SA has its own unique End-Point-TLV);
SDWAN-Site-ID = Color-extended Community (carried by the client routes),
SDWAN-Node-ID

NLRI SAFI = SDWAN: Tunnel Encap Attribute: with option to include IPsec SA sub-TLVs
SDWAN SAFI (=74) NLRI Encoding Format for Underlay Network Properties Advertisement

MP-REACH-NLRI Path Attribute
  AFI = 1; SAFI = SDWAN
  \langle Site-Type, Port-Local-ID, SDWAN-Color, SDWAN-Node-ID\rangle
  Geo-location Sub-TLV (Optional)
  Tunnel Encaps Attribute (23)
    Tunnel Type: SDWAN-Hybrid (to indicate hybrid underlay tunnels)
    Tunnel-egress-endpoint Sub-TLV
    Extended Port Sub-TLV (with or without NAT, Optional, including the ISP-subTLVs)
    VxLAN, GRE, etc. for Secure VPN underlays

IPsec SA Nonce Sub-TLV
IPsec SA Public Key Sub-TLV
IPsec SA Transform Sub-TLV
  \{Transforms Substructure Sub-sub TLV\}

List of IPsec-SA Attributes.
Or
IPSec-SA IDs
NLRI: SDWAN SAFI = 74

<table>
<thead>
<tr>
<th>NLRI Length</th>
<th>Site-Type</th>
<th>Port-Local-ID</th>
<th>SDWAN-Color = RED</th>
<th>Node-ID = 2.2.2.2</th>
</tr>
</thead>
</table>

- **Site Type**: 1 octet value. The SDWAN Site Type defines the different types of Site IDs to be used in the deployment. The draft defines the following types:
  
  Site-Type = 1: For simple deployment, such as all edge nodes under one SDWAN management system, a simple identifier is enough for the SDWAN management to map the site to its precise geolocation.
  
  Site-Type = 2: to indicate that the value in the site-ID is locally significant, therefore, need a Geo-Loc Sub-TLV to fully describe the accurate location of the node. This is for large SDWAN heterogeneous deployment where Site IDs has to be described by proper Geo-location of the Edge Nodes [LISP-GEOLoc].

- **Port local ID**: SDWAN edge node Port identifier, which can be locally significant. The detailed properties about the network connected to the port are further encoded in the Tunnel Path Attribute. If the SDWAN NLRI applies to multiple ports, this field is NULL.

- **SDWAN-Color**: used to identify a common property shared by a set of SDWAN edge nodes or a group of WAN ports, such as the Color encoded in the Client Routes’ “Color Extended Community” or property of a specific geographic location shared by a group of prefixes.
We want your feedback!
DETAILS – AT FIELDS IN SUBTLV
Tunnel Path Attributes and Sub-TLVs inside the SDWAN NLRI
## Extended Port (NAT) Sub-TLV

| Port Ext Type | EncapExt subTLV Length | I | O | R | R | R | R | R | R | R | R |
|---------------|------------------------|---|---|---|---|---|---|---|---|---|---|---|
| NAT Type      | Encap-Type             | Trans networkID | RD ID |
| Local IP Address | Local Port | Public IP | Public Port | ISP-Sub-TLV |
| 32-bits for IPv4, 128-bits for IPv6 | 32-bits for IPv4, 128-bits for IPv6 | 32-bits for IPv4, 128-bits for IPv6 |

### 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; If NAT is not used, this field is set to NULL.

### Local Port:
- used by Remote SDWAN edge node for establishing IPsec to this specific port. If NAT is not used, this field is set to NULL.

### Public IP:
- The IP address after the NAT.

### Public Port:
- The Port after the NAT.

Edge node get NAT properties via STUN requests/responds. Peers may not be able to access the STUN server.
ISP of the Underlay Network Sub-TLV

- **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
  - LTE – 3
  - 5G – 4
- **Port Type**: There are different types of ports. They are listed below as:
  - Ethernet – 1
  - Fiber Cable – 2
  - Coax Cable – 3
  - Cellular – 4
- **Port Speed**: The port seed is defined as 2 octet value. The values are defined as Gigabit speed.

```
+-----------------+-----------------+-----------------+-----------------+
|       Type      |      Length     |      Flag      |    Reserved     |
+-----------------+-----------------+-----------------+-----------------+
|Connection Type  |   Port Type     |        Port Speed |
+-----------------+-----------------+-----------------+
```
Encoding for IPsec Properties
Two Types of IPsec SA attributes (only use one)
Sub-Sub-TLV

• Full Set: with multiple Sub-TLVs 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

• Simple Set: Simple Deployment with limited number of parameters
  – One Sub-TLV to represent Public Key, Nonce, ReKey, Transform
## Nonce Sub-TLV, Public Key Sub-TLV

### Nonce Sub-TLV:

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | Type=Nonce | length |
| +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| |       ID Length |       Nonce Length | | Flags |
| +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| |                             Rekey                             |
| +---------------------------------------------------------------+
| |                            Counter                            |
| +---------------------------------------------------------------+
| |           IPsec SA ID         |           Reserved            |
| +---------------------------------------------------------------+
| |                                                               |
| ~                          Nonce Data                           ~
| |                                                               |
| +---------------------------------------------------------------+

**IPsec SA ID** - The 2 bytes IPSec SA ID could 0 or non-zero values. It is cross referenced by client route's IPSec Tunnel Encap IPSec-SA-ID or IPSec-SA-Group Sub-TLV in Section 5 of the Draft. When there are multiple IPsec SAs terminated at one address, such as WAN port address or the node address, they are differentiated by the different IPsec SA IDs.

### Public Sub-TLV:

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

+---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | Diffie-Hellman Group Num | | Reserved |
| +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | ~                       Key Exchange Data                       |
| +---------------------------------------------------------------+
| |                             Duration                            |
| +---------------------------------------------------------------+
### Simplified IPsec SA attributes advertisement

<table>
<thead>
<tr>
<th>IPsec-simType</th>
<th>IPsecSA Length</th>
<th>Flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transform</td>
<td>Mode</td>
<td>AH</td>
</tr>
<tr>
<td>ReKey Counter</td>
<td>Public Key</td>
<td>Nonce</td>
</tr>
<tr>
<td>Duration</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **IPsec-simType**: to be assigned by IANA.
- **Flags**: for future usage.
- **Transform (1 Byte)**: the value can be AH, ESP, or AH+ESP.
- **Mode (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.
- **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 (Security Parameter Index)**
- **Public Key**: IPsec public key
- **Nonce**: IPsec Nonce
- **Duration**: SA life span.