

Network Working Group  
Internet Draft  
Intended status: Informational  
Expires: May 21, 2020  
Consulting

L. Dunbar  
Futurewei  
S. Hares  
Hickory Hill

November 21, 2019

SDWAN WAN Ports Property Advertisement in BGP UPDATE  
draft-dunbar-idr-sdwan-port-safi-06

Abstract

The document describes how the SDWAN SAFI, which is assigned by IANA in the First Come First Server range, is used for SDWAN edge nodes to propagate its WAN port properties to its controller.

In the context of this document, BGP Route Reflectors (RR) is the component of the SDWAN Controller that receives the BGP UPDATE from SDWAN edges and in turns propagate the information to a group of authorized SDWAN edges reachable via overlay networks.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on Dec 5, 2020.

## Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	<a href="#">Introduction.....</a>	<a href="#">3</a>
<a href="#">2.</a>	<a href="#">Conventions used in this document.....</a>	<a href="#">3</a>
<a href="#">2.1.</a>	<a href="#">Information to be propagated for SDWAN UPDATE.....</a>	<a href="#">4</a>
<a href="#">2.2.</a>	<a href="#">SAFI under the MP-NLRI.....</a>	<a href="#">6</a>
<a href="#">2.3.</a>	<a href="#">How about a new Path Attribute under BGP UPDATE?.....</a>	<a href="#">6</a>
<a href="#">3.</a>	<a href="#">SDWAN WAN Port Identifier encoding in the MP-NLRI Path Attribute6</a>	
<a href="#">4.</a>	<a href="#">WAN Port Properties encoding in the Tunnel Path Attribute.....</a>	<a href="#">8</a>
<a href="#">4.1.</a>	<a href="#">Port Ext SubTLV for NAT.....</a>	<a href="#">9</a>
<a href="#">4.2.</a>	<a href="#">IPsec Security Association Property.....</a>	<a href="#">10</a>
<a href="#">4.3.</a>	<a href="#">Remote Endpoint.....</a>	<a href="#">11</a>
<a href="#">5.</a>	<a href="#">Manageability Considerations.....</a>	<a href="#">12</a>
<a href="#">6.</a>	<a href="#">Security Considerations.....</a>	<a href="#">12</a>
<a href="#">7.</a>	<a href="#">IANA Considerations.....</a>	<a href="#">12</a>
<a href="#">8.</a>	<a href="#">References.....</a>	<a href="#">12</a>
<a href="#">8.1.</a>	<a href="#">Normative References.....</a>	<a href="#">12</a>
<a href="#">8.2.</a>	<a href="#">Informative References.....</a>	<a href="#">13</a>
<a href="#">9.</a>	<a href="#">Acknowledgments.....</a>	<a href="#">14</a>

## 1. Introduction

[Net2Cloud-Problem] introduces using SDWAN to reach dynamic workloads in multiple third-party data centers and aggregate multiple underlay paths, including public untrusted networks, provided by different service providers.

[SDWAN-BGP-USAGE] describes multiple SDWAN scenarios and illustrates how BGP is used as control plane for the SDWAN networks.

The document describes BGP UPDATE for SDWAN edge nodes to propagate its WAN port properties to RR.

## 2. Conventions used in this document

Cloud DC: Off-Premise Data Centers that usually host applications and workload owned by different organizations or tenants.

Controller: Used interchangeably with SDWAN controller to manage SDWAN overlay path creation/deletion and monitor the path conditions between sites.

CPE-Based VPN: Virtual Private Secure network formed among CPEs. This is to differentiate from most commonly used PE-based VPNs a la [RFC 4364](#).

MP-NLRI: The MP\_REACH\_NLRI Path Attribute defined in [RFC4760](#).

SDWAN End-point: An WAN port (logical or physical) of a SDWAN edge node. (If "endpoint" is used, it refers to a SDWAN End-point).

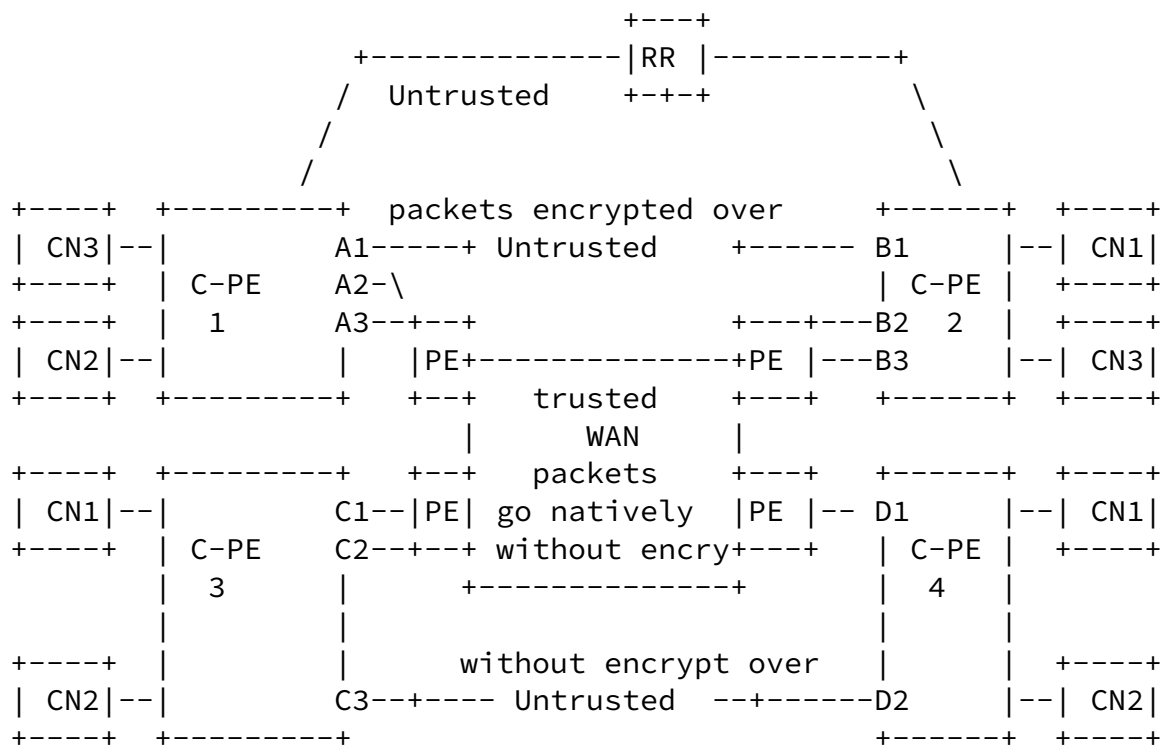
OnPrem: On Premises data centers and branch offices

SDWAN: Software Defined Wide Area Network. In this document, "SDWAN" refers to the solutions of pooling WAN bandwidth from multiple underlay networks to get better WAN bandwidth management, visibility & control. When the underlay networks are private networks, traffic can be

forwarded without additional encryption; when the underlay networks are public, such as Internet, some traffic needs to be encrypted when forwarding through those WAN ports (depending on user provided policies).

## 2.1. Information to be propagated for SDWAN UPDATE

Figure below shows the Hybrid SDWAN scenario:



CN: Client Network

Figure 1: Hybrid SDWAN

Using C-PE2 for illustration, C-PE2 needs to send out two separate BGP UPDATE messages.

BGP UPDATE #1 is to propagate C-PE2 attached routes, which are the regular VPN (L3VPN or EVPN) BGP Route UPDATE message,

MP-NLRI Path Attribute

Internet-Draft

SAFI for SDWAN WAN Properties

November 2019

Nexthop (C-PE2)

NLRI

10.1.x.x.

VLAN 15

12.1.1x

Tunnel-Encap Path Attribute

Details of any tunnels that applicable to the routes carried  
by the MP-NLRI Path Attribute

BGP UPDATE #2 is to propagate C-PE2's WAN port properties to RR,  
which should include:

- Identifier for the WAN Port
- The NAT property for the WAN Port
- The minimum IPsec information for establishing Port based IPsec.

Separating WAN port properties UPDATE from client routes UPDATE makes the implementation simpler, because the properties of a SDWAN node's WAN Port can change independent from the client routes attached to the C-PE2. WAN port properties change can be caused by many factors, such as ISP service agreement changes for the service connected to the WAN Port, the WAN port being disabled, or its IPsec property changes, etc. Since most SDWAN edges only have a small number of WAN ports, the disadvantage of multiple BGP UPDATE messages to advertise properties of those WAN ports is relatively small.

Following the same approach used by [idr-segment-routing-te-policy] where the SR Policy identifier is encoded in the MP-NLRI Path Attribute and the detailed SR Policies are encoded in the Tunnel Path attribute, the BGP UPDATE for SDWAN WAN port can have the WAN Port identifier encoded in the MP-NLRI Path Attribute and the associated WAN Port properties encoded in the Tunnel Path Attribute.

Receivers of the UPDATE can associate the SDWAN node identifier, site identifier with the node's WAN Port properties.

## 2.2. SAFI under the MP-NLRI

It is possible to continue using the same IP SAFI in the MP-NLRI [[RFC4760](#)] Path Attribute for advertising the SDWAN WAN port properties. If the same IP SAFI used, receiver needs extra logic to differentiate regular BGP MP-NLRI routes advertisement from the SDWAN WAN port properties advertisement and recognize the extra Site ID field added to the MP-NLRI. The benefit of using the same IP SAFI is that the UPDATE can traverse existing routers without being dropped. However, the SDWAN UPDATE is only between SDWAN edge and the RR, all the intermediate nodes treat the UPDATE message as regular IP data frame.

That is why it is simpler to follow the same approach used by [idr-segment-routing-te-policy] to have a unique SAFI (IANA assigned SDWAN SAFI = 74) mainly to differentiate the SDWAN UPDATE from regular route UPDATE.

This SDWAN SAFI is for a scenario where one SDWAN edge node has multiple WAN ports, some of which connected to private networks and others connected to public untrusted networks [Scenario #2 described in the [[SDWAN-BGP-USAGE](#)]]. The same routes attached to the SDWAN can be reached by the private networks without encryption (for better performance) or by the public networks with encryption.

## 2.3. How about a new Path Attribute under BGP UPDATE?

It is also possible to have a new Path Attribute, say SDWAN Path Attribute, combined with Tunnel Path Attribute to advertise SDWAN WAN Port properties. Besides having a different Path Attribute ID, everything else is same as using MP-NLRI & Tunnel Path Attributes.

## [3.](#) SDWAN WAN Port Identifier encoding in the MP-NLRI Path Attribute

SDWAN WAN Port Identifier needs the following attributes

- locally significant port number,
- the location of the SDWAN device, and

- the globally routable address for the WAN Port.

Here is the encoding for those attributes in the NLRI field within the MP\_REACH\_NLRI Path Attribute of [RFC4760](#), under a SDWAN SAFI (code = 74):

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

where:

- NLRI Length: 1 octet of length expressed in bits as defined in [\[RFC4760\]](#).
- SDWAN-Type: to define the encoding of the rest of the SDWAN NLRI. There could be different sub-TLVs for different SDWAN WAN ports and their associated policies.
- Port local ID: SDWAN edge node Port identifier, which can be locally significant. Each port can have unique properties. For example, some ports may get ISP or DHCP assigned IP addresses (IPv4 or IPv6), some may have private IP addresses that packets to/from those ports have to traverse NAT. The detailed properties about the port are further encoded in the subTLVs, e.g. Port-subTLV under the Tunnel Path Attribute.
- SDWAN-Site-ID: used to identify a common property shared by a set of SDWAN edge nodes, such as the property of a specific geographic location shared by a group of SDWAN edge nodes. The property is used to steer an overlay route to traverse specific





#### [4.1.](#) Port Ext SubTLV for NAT

NAT information is encoded is the Port Ext sub-TLV is for describing the NAT property if the port has private address and the network identifier to which the WAN port is connected, etc.

A SDWAN edge node can inquire STUN (Session Traversal of UDP Through Network Address Translation [RFC 3489](#)) Server to get the NAT property, the public IP address and the Public Port number to pass to peers.

```

 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Port Ext 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                            |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

- o Port Ext Type: indicate it is the Port Ext SubTLV.
- o PortExt subTLV Length: the length of the subTLV.
- o 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.

- 0 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.

- o 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).
- o 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)
- o Transport Network ID.Central Controller assign a global unique ID to each transport network.
- o RD ID.Routing Domain ID.Need to be global unique.
- o Local IP.The local (or private) IP address of the port.
- o Local Port.used by Remote SDWAN edge node for establishing IPsec to this specific port.
- o Public IP.The IP address after the NAT. If NAT is not used, this field is set to NULL.
- o Public Port.The Port after the NAT. If NAT is not used, this field is set to NULL.

#### [4.2.](#) IPsec Security Association Property

The IPsecSA sub-TLV is for the SDWAN edge node to establish IPsec security association with their peers via the port that face untrusted network. The minimum set of the IPsec information is from [\[CONTROLLER-IKE\]](#).

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		
IPsec-SA Type										IPsecSA Length										Flag													
Transform										Transport										AH						ESP							
Key Counter																																	

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| key1 length   |                               Public Key                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| key2 length   |                               Nonce                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| key3 length   |   key3 (for potential other keys   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               |   Duration   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

- o IPsec-SA SubTLV Type: to be assigned by IANA. The type value has to be between 128~255 because IPsec-SA subTLV needs 2 bytes for length to carry the needed information.
- o IPsec-SA subTLV Length (2 Byte): 25 (or more)
- o Flags: 1 octet of flags. None are defined at this stage. Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Transform (1 Byte): the value can be AH, ESP, or AH+ESP.
- o Transport (1 byte): the value can be Tunnel Mode or Transport mode
- o 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.
- o 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.
- o Rekey Counter: 4 bytes
- o Public Key: IPsec public key
- o Nonce: IPsec Nonce
- o Key3: other potential key
- o Duration: SA life span.

#### [4.3.](#) Remote Endpoint

The Remote Endpoint sub-TLV is not used for SDWAN NLRI because

- o The SDWAN Node ID and Site ID are already encoded in the SDWAN NLRI,
- o The network connected by the SDWAN WAN port might have identifier that is more than the AS number. SDWAN controller might use its own specific identifier for the network.

- o The Transport-Network-ID in the EncapExt sub-TLV represents the SDWAN unique network identifier.

If the Remote Endpoint Sub-TLV is present, it is ignored by other SDWAN edge nodes.

## [5. Manageability Considerations](#)

TBD - this needs to be filled out before publishing

## [6. Security Considerations](#)

The document describes the encoding for SDWAN edge nodes to advertise its SDWAN WAN ports properties to their peers via untrusted & unsecure networks.

The secure propagation is achieved by secure channels, such as TLS, SSL, or IPsec, between the SDWAN edge nodes and the local controller RR.

[More details need to be filled in here]

## [7. IANA Considerations](#)

This document requires the following IANA actions.

- o SDWAN Overlay SAFI = 74 assigned by IANA
- o SDWAN Route Type

## [8. References](#)

### 8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate

## 8.2. Informative References

- [RFC8192] S. Hares, et al, "Interface to Network Security Functions (I2NSF) Problem Statement and Use Cases", July 2017
- [RFC5521] P. Mohapatra, E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", April 2009.
- [CONTROLLER-IKE] D. Carrel, et al, "IPsec Key Exchange using a Controller", [draft-carrel-ipsecme-controller-ike-01](#), work-in-progress.
- [Tunnel-Encap] E. Rosen, et al, "The BGP Tunnel Encapsulation Attribute", [draft-ietf-idr-tunnel-encaps-09](#), Feb 2018.
- [VPN-over-Internet] E. Rosen, "Provide Secure Layer L3VPNs over Public Infrastructure", [draft-rosen-bess-secure-l3vpn-00](#), work-in-progress, July 2018
- [DMVPN] Dynamic Multi-point VPN:  
<https://www.cisco.com/c/en/us/products/security/dynamic-multipoint-vpn-dmvpn/index.html>
- [DSVPN] Dynamic Smart VPN:  
<http://forum.huawei.com/enterprise/en/thread-390771-1-1.html>
- [ITU-T-X1036] ITU-T Recommendation X.1036, "Framework for creation, storage, distribution and enforcement of policies for network security", Nov 2007.
- [Net2Cloud-Problem] L. Dunbar and A. Malis, "Seamless Interconnect Underlay to Cloud Overlay Problem Statement", [draft-dm-net2cloud-problem-statement-02](#), June 2018

[Net2Cloud-gap] L. Dunbar, A. Malis, and C. Jacquenet, "Gap Analysis of Interconnecting Underlay with Cloud Overlay", [draft-dm-net2cloud-gap-analysis-02](#), work-in-progress, Aug 2018.

Dunbar, et al.

Expires Dec 21, 2020

[Page 13]

---

Internet-Draft

SAFI for SDWAN WAN Properties

November 2019

[Tunnel-Encap] E. Rosen, et al "The BGP Tunnel Encapsulation Attribute", [draft-ietf-idr-tunnel-encaps-10](#), Aug 2018.

## [9](#). Acknowledgments

Acknowledgements to Wang Haibo, Hao Weiguo, and ShengCheng for implementation contribution; Many thanks to Jim Guichard, John Scudder, Darren Dukes, Andy Malis, Rachel Huang and Donald Eastlake for their review and contributions.

This document was prepared using 2-Word-v2.0.template.dot.

Internet-Draft

SAFI for SDWAN WAN Properties

November 2019

#### Authors' Addresses

Linda Dunbar  
Futurewei  
Email: [ldunbar@futurewei.com](mailto:ldunbar@futurewei.com)

Sue Hares  
Hickory Hill Consulting  
Email: [shares@ndzh.com](mailto:shares@ndzh.com)

Dunbar, et al.

Expires Dec 21, 2020

[Page 15]