RTG Working Group
Internet Draft
Intended status: Standard
Expires: October 18, 2021

L. Dunbar
Futurewei
Mehmet Toy
Verizon

May 18, 2021

SRv6 across SDWAN paths
draft-dunbar-sr-sdwan-over-hybrid-networks-07

Abstract

This document describes the mechanism of steering packets across SDWAN segments based on the metadata carried by the SRv6 packets.

Some of the SDWAN segments are untrusted networks, and some are private networks. The goal is to achieve the optimal E2E quality.

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/1id-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 April 23, 2021.

Table of Contents

1. Introduction

SRv6 has many advantages. This document defines using the metadata encoded in the SRH for SRv6 packets to cross an SDWAN network.

SDWAN is about pooling WAN bandwidth from multiple service providers to get better WAN bandwidth management, visibility & control. There could be multiple underlay paths between a pair of edge nodes,

potentially managed by different service providers, such as MPLS paths and paths over the public internet.

This document describes the SRv6 SRH metadata encoding for the SRv6 packets to cross SDWAN for the scenarios described by the [BGP-SDWAN-Usage]:

1) Homogeneous WAN, with edge nodes encrypting all traffic over the WAN to other edge nodes, regardless of whether the underlay is private or public.

2) Hybrid WAN Underlay, in which traffic over IP VPN is forwarded natively without IPsec protection and carried by IPsec tunnels when forwarded over the public Internet.

3) Private VPN PE-based SDWAN, which is about existing VPN (e.g., EVPN or IPVPN) being expanded by the additional ports facing the untrusted Internet for PEs to offload low-priority traffic when the VPN paths are congested.

2. Conventions used in this document

BSID - Binding SID

DC - Data Center

DN - Data Network (5G)

SD-WAN - Software-Defined Wide Area Network

SID - Segment Identifier

SR - Segment Routing

3. Use Cases

3.1. SDWAN as Last Mile for Accessing Cloud Services

Digital Transformation is propelling more and more enterprises to utilize the rich Cloud services, such as virtual machines, remote databases, analytic tools, machine learning APIs, etc. Cloud services enable enterprises to run their workloads/Apps at locations geographically close to their end-users and provide advanced analytic tools and APIs for the applications and data hosted in the Clouds.

The wide availability at any location, which is one of the advantages of Cloud Services, can impose challenges to connect enterprises' on-premises applications with their Cloud services securely. The SRv6 domain that interconnect the enterprises' locations may not reach the Cloud DCs where the Cloud services are hosted. SDWAN is positioned as a flexible choice as the last mile to bridge the enterprise's SR domain to its desired Cloud services.

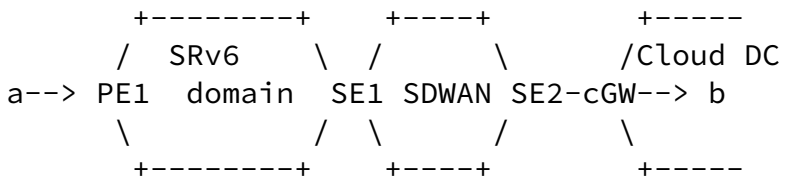

Figure 1: SDWAN as last mile

3.2. SRv6 Domain separated by SDWAN

SRv6 deployment is incremental. Some services do need to cross segments that do not support SRv6, as shown in the Figure below.

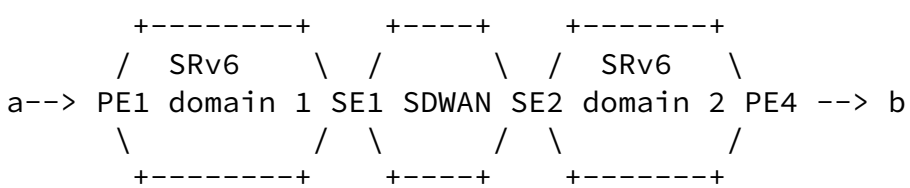

Figure 2: SDWAN connecting two SRv6 domains

4. SDWAN Path Programming in SRv6

An SDWAN path between two edge nodes can be an IPsec tunnel, an MPLS path, an IPv4, or an IPv6 path over a private IP network. For an

SRv6 packet to cross an SDWAN domain, the edge node, such as SE1 & SE2 in Figure 1 & Figure 2, can make the local decision in choosing an SDWAN path between the two edge nodes. Alternatively, the controller can instruct the SR domain head node, like the PE1 in Figure 2, to encode the metadata in the SRH that can indicate the SDWAN path for the SDWAN ingress node SE1.

4.2. End.SDWANv4

End.SDWANv4 is an End function for the receiving node to locally select an SDWAN path destined towards the IPv4 destination address encoded in the SRH.

The SDWAN tunnel information are encoded in another 128-bit value following the SID or SRH TLVs.

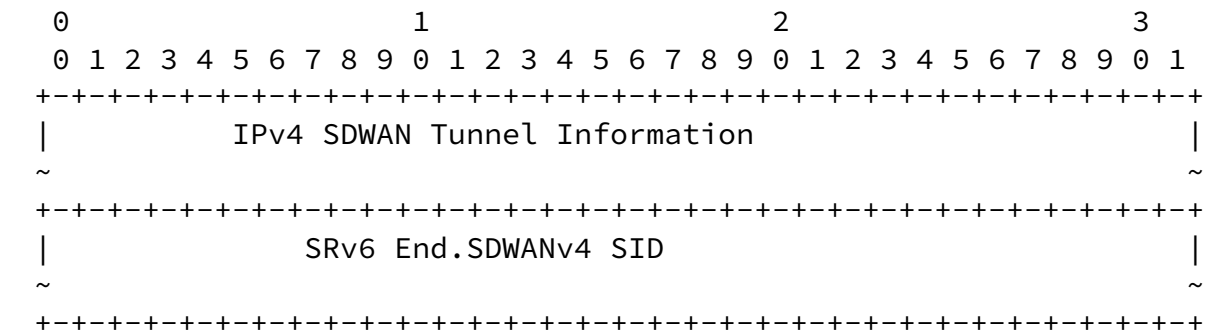

Figure 3. IPv4 SDWAN Sub-Path Encoding in SRH

The SDWAN Tunnel Information encoding follows the format from [SDWAN-Edge-Discovery]:

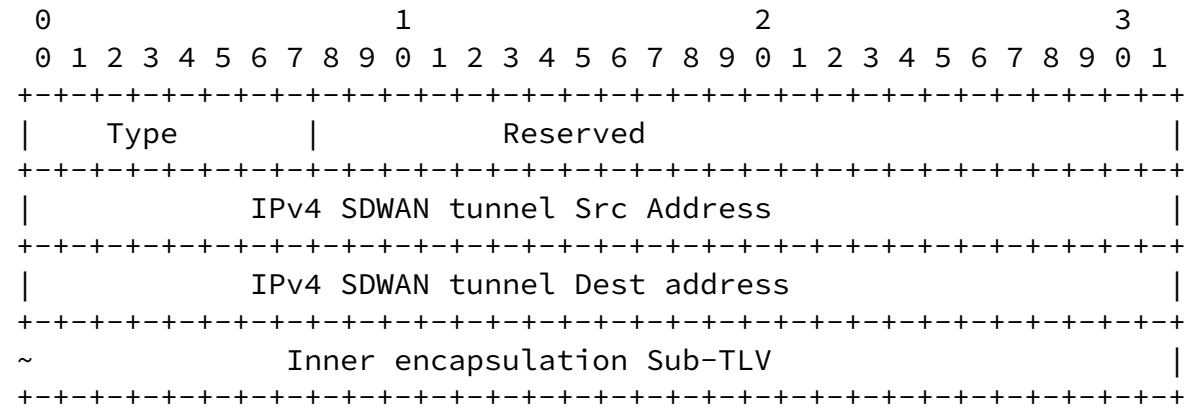

Figure 4. SDWAN Tunnel Encoding

Type = SDWAN-Hybrid: for the end node to locally select an SDWAN path with inner encapsulation type to carry the packet.

The inner encapsulation Sub-TLV can be GRE Sub-TLV or VxLAN Sub-TLV as specified in the [Tunnel-Encap].

For SRH to indicate exact SDWAN MPLS path to forward the packet, the SRH encoding should follow the encoding described in the [SRv6-traverse-MPLS].

This document's Section 4.2 and 4.3 describes the encoding for SR head node to indicate the IPsec IPv4 or IPv6 tunnels in SRH, respectively.

4.3. End.IPsecV4

End.IPsecV4 is an End function with IPv4 IPsec tunnel instantiation, i.e., instructing the receiving node to encapsulate the packet with an IPsec tunnel and forward to the IPv4 destination. The IPsec tunnel information can be encoded following the SID or SRH TLVs.

An End.IPsecV4 SID MUST be encoded preceding the IPsec tunnel information encapsulation.

The SRv6 path of crossing IPv4 IPsec tunnel is called IPv4 IPsec sub-path. The IPsec tunnel attributes are encoded by an END.IPsecV4 SID and the following IPv4 IPsec tunnel information encapsulation as shown in the following figure.

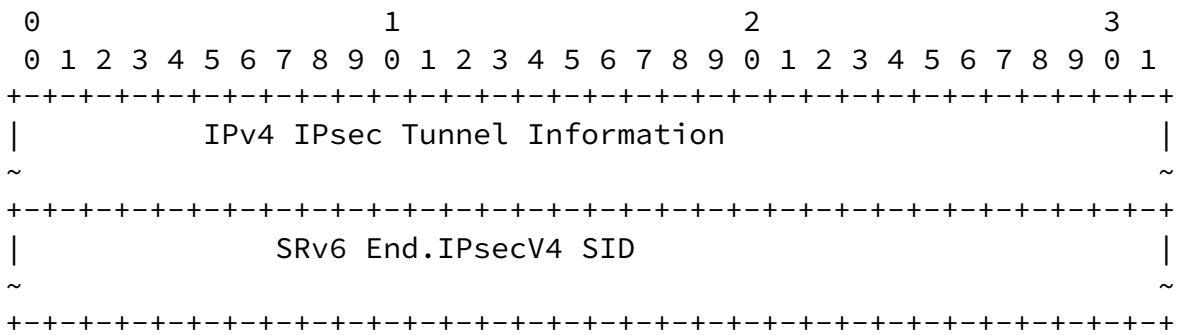

Figure 5. IPv4 IPsec Sub-Path Encoding in SRH

The IPv4 IPsec tunnel should be ESP Tunnel mode. For ESP Tunnel mode, there can be additional inner encapsulation. SDWAN edge nodes can also encapsulate the ESP IPsec packet inside UDP for NAT traversal and better ECMP [RFC3948].

Here is the IPsec tunnel information encoding:

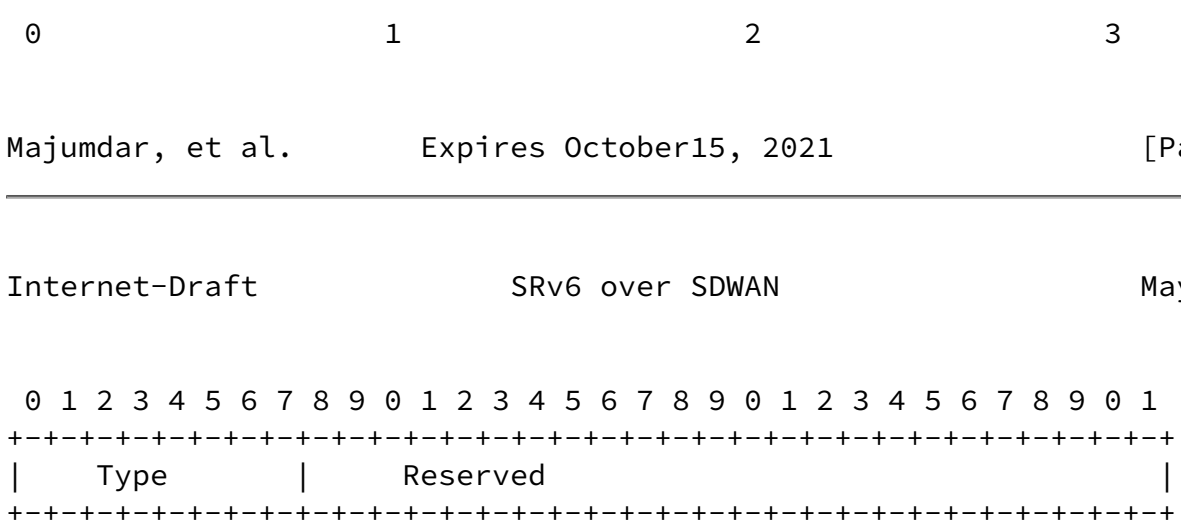

Figure 6. IPv4 IPsec Tunnel Encoding

Type = IPv4 IPsec

The IPsec SA sub-TLV, specified by [SDWAN-Edge-Discovery], lists the identifiers of the pre-established IPsec tunnels between the SDWAN Src Address and the Dest Address. One or multiple identifiers are listed in the IPsec-SA-ID Sub-TLV for the IPsec tunnels between the Source and the Destination addresses.

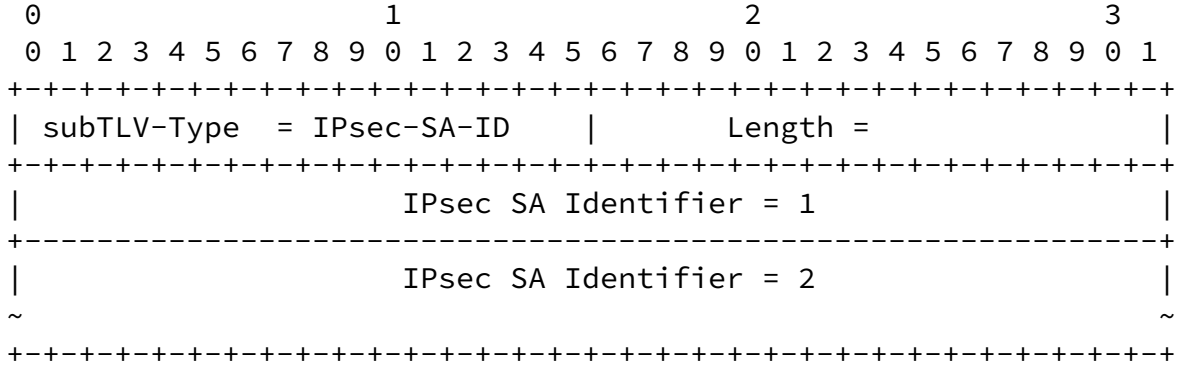

Figure 7. IPsec SA Sub-TLV

The inner encapsulation Sub-TLV can be GRE Sub-TLV or VxLAN Sub-TLV as specified in the [Tunnel-Encap].

When node N receives a packet whose IPv4 DA is S and S is a local End.IPsecV4 SID, the line S15 - S16 from the End processing [RFC8986] is replaced by the following:

S15. Encapsulates the SRv6 packet with a new IPsec tunnel encapsulation bound to the End.IPsecV4 SID S.

S16. Submit the IPsec encapsulated packet to the egress IPv4 FIB lookup for transmission to the IPsec end point IPv4 destination.

S17. }

4.4. End.IPsecV6

End.IPsecV6 is an End function with IPv6 IPsec tunnel instantiation, i.e., instructing the receiving node to encapsulate the packet with IPsec tunnel and forwarded to the IPv6 destination.

End.IPsecV6 behavior is very much like End.IPsecV4 except the destination and source address of the IPsec tunnel are IPv6 addresses.

When node N receives a packet whose IPv6 DA is S and S is a local End.IPsecV6 SID, the line S15 - S16 from the End processing [RFC8986] is replaced by the following:

S15. Encapsulates the SRv6 packet with a new IPsec tunnel encapsulation bound to the End.IPsecV6 SID S.

S16. Submit the IPsec encapsulated packet to the egress IPv6 FIB lookup for transmission to the IPsec end point IPv6 destination.

S17. }

The SRv6 path of crossing IPv6 IPsec tunnel is called IPv6 IPsec sub-path. The IPsec tunnel attributes are encoded by an END.IPsecV6 SID and the following IPv6 IPsec tunnel information encapsulation as shown in the following figure.

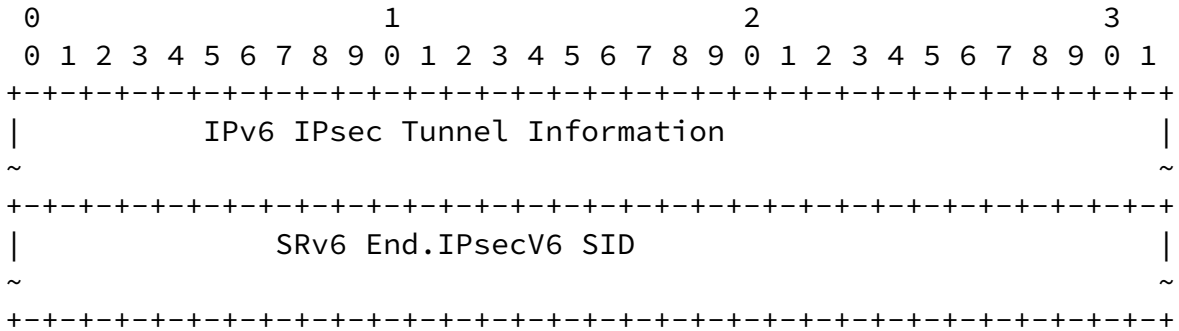

Figure 8. IPv6 IPsec Sub-Path Encoding in SRH

The IPv6 IPsec tunnel can be ESP Transport mode or Tunnel mode. For ESP Tunnel mode, there can be additional inner encapsulation. SDWAN edge nodes can also encapsulate the ESP IPsec packet inside UDP for NAT traversal and better ECMP [RFC3948].

Here is the IPv6 IPsec tunnel information encoding:

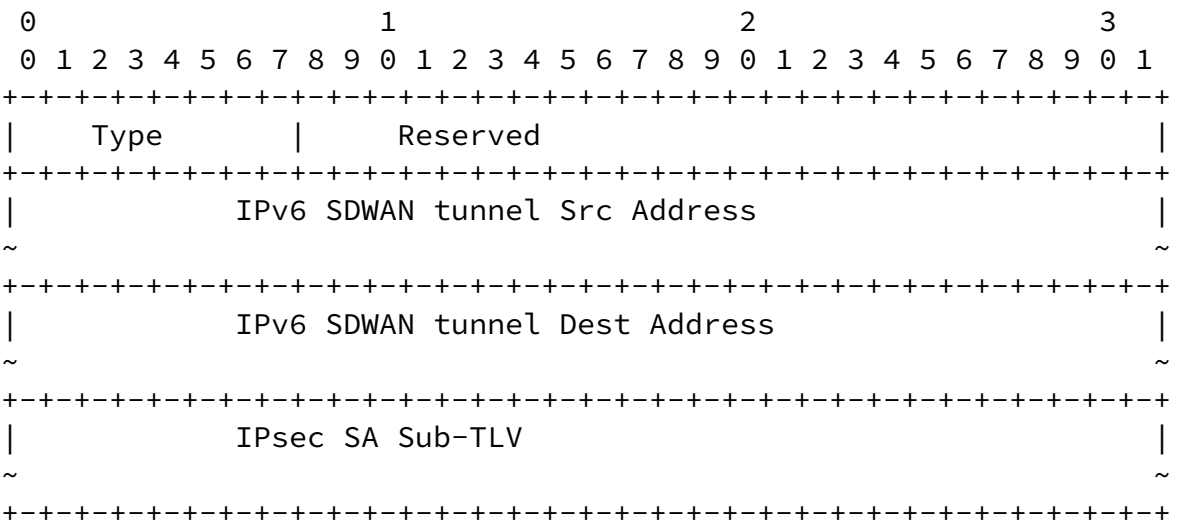

Figure 9. IPv6 IPsec Tunnel Encoding

Type = IPv6 IPsec

5. Packets from SDWAN to SRv6 Domain

For the SDWAN as Last Mile use case illustrated in Figure 1, packets from "b" -> "a" traverse from SDWAN domain to SRv6 domain. A Binding SID needs to be inserted by the SDWAN Edge node SE2 so that the SR domain ingress can replace the Binding SID with a list of SIDs across the SRv6 domain.

For an SDWAN path over an IPsec Tunnel, the Binding SID is encoded in the GRE key field for the GRE inner encapsulation or encoded in the VNID field for the VxLAN inner encapsulation.

For an SDWAN path over an MPLS underlay, the last MPLS label is used as the Binding SID for the SRv6 edge node to convert to a list of SRv6 SIDs across the SRv6 domain.

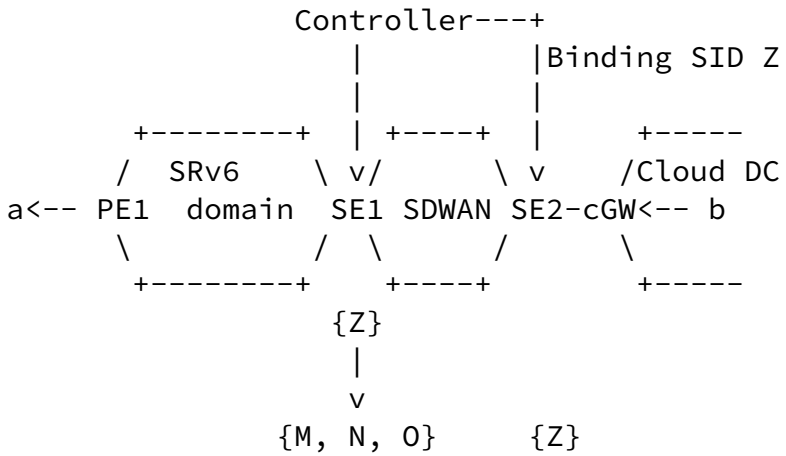

Figure 10: Binding SID inserted by SDWAN Edge

For the SRv6 domain separated by SDWAN use case illustrated in Figure 2, the End.SDWANv4/v6 or End.IPsecv4/v6 SID should not be the last SID in the SRH. After the SDWAN egress node decapsulates the SDWAN header (IPsec header or MPLS header), the remaining SIDs in the packet's SRH can forward the packet across the remaining SRv6 domain.

6. Illustration

To Be Added

7. IANA Considerations

TBD.

8. Security Considerations

Allowing traffic from untrusted network brings the following security risks:

1) Potential DDoS attack to the PEs with ports facing the untrusted network. I.e. the PE resource being attacked by unwanted traffic.
2) Potential risk of provider VPN network bandwidth being stolen by the entities who spoofed the addresses of SDWAN end nodes.

To mitigate security risk of 1) above, it is necessary for ports facing internet to enable Anti-DDoS feature to prevent major DDoS attack to those PEs.
To mitigate the security risk of 2) above, RFC7510 defines the use of DTLS to authenticate and encrypt the RFC7510 encapsulation.

9. Contributors

10. References

10.1. Normative References

[RFC2890] G. Dommety "Key and Sequence Number Extensions to GRE". Sep. 2000.

10.2. Informative References

[ITU-T-X1036] ITU-T Recommendation X.1036, "Framework for creation, storage, distribution and enforcement of policies for network security", Nov 2007.

[RFC6071] S. Frankel and S. Krishnan, "IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap", Feb 2011.

[RFC4364] E. Rosen and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", Feb 2006

[RFC4664] L. Andersson and E. Rosen, "Framework for Layer 2 Virtual Private Networks (L2VPNs)", Sept 2006.

[SR-SDWAN] D. Dukes, et al, "SR for SDWAN: VPN with Underlay SLA", draft-dukes-sr-for-sdwan-00, in progress, Oct 2017

[SRv6-SRH] S. Previdi, et al, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-13, in progress, April 2018.

[MPLS-SR] A. Bashandy, et al, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-13, in progress, April 2018.

[RFC7510] X. Xu, et al, "Encapsulating MPLS in UDP", April 2015.

[RFC8086] L. Yong, et al, "GRE-in-UDP Encapsulation", March 2017.

[BGP-SDWAN-Usage] L. Dunbar, et al, "BGP Usage for SDWAN Overlay Networks, draft-dunbar-bess-bgp-sdwan-usage-01, in progress, July 2019.

[SDWAN-Net2Cloud] L. Dunbar, et al, "Dynamic Networks to Hybrid Cloud DCs Problem Statement", draft-ietf-rtgwg-net2cloud-problem-statement-04, in progress, July 2019.

[MEF-Cloud] "Cloud Services Architecture Technical Specification",

Work in progress, April 2018

[SDWAN-BGP-USAGE] L. Dunber, et al, "BGP Usage for SDWAN Overlay Networks", draft-dunbar-bess-bgp-sdwan-usage-08, January 2021

[BGP-IPSEC-Discover] L. Dunber, et al, "BGP UPDATE for SDWAN Edge Discovery", draft-dunbar-idr-sdwan-edge-discovery-00, January 2021

[Tunnel-Encap] E. Rosen, et al "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-19, March 2021.

11. Acknowledgments

TBD.

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

Authors' Addresses

Linda Dunbar
Futurewei
2330 Central Expressway
Santa Clara, CA 95050
Email: linda.dunbar@futurewei.com

Mehmet Toy
Verizon

One Verizon Way
Basking Ridge, NJ 07920
Email: mehmet.toy@verizon.com