LSVR
Internet-Draft
Intended status: Standards Track
Expires: 18 November 2025

C. Sheng
H. Shi
J. Dong, Ed.
Huawei

17 May 2025

Usage of BGP-LS-SPF in Multi-segment SD-WAN
draft-sheng-lsvr-bgp-spf-for-sdwan-03

Abstract

This document introduces the usage of BGP-LS-SPF protocol in multi-segment SD-WAN scenarios. It allows SD-WAN tunnels to be published as logical links, which can cross the internet, MPLS networks, and various operator network. The BGP-LS-SPF protocol can construct an overlay network topology for logical links and physical links across these heterogeneous networks, and calculate the reachability routes of overlay network nodes based on this topology.

Discussion Venues

This note is to be removed before publishing as an RFC.

Discussion of this document takes place on the Link State Vector Routing Working Group mailing list (lsvr@ietf.org), which is archived at https://mailarchive.ietf.org/arch/browse/lsvr/.

Source for this draft and an issue tracker can be found at https://github.com/VMatrix1900/draft-lsvr-bgp-spf-for-sdwan.

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). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on 18 November 2025.

Table of Contents

1. Introduction

As pointed out in [I-D.draft-ietf-rtgwg-net2cloud-problem-statement], enterprises are migrating their workloads to cloud service. The enterprise branch interconnection and enterprise site to cloud DC connection may cross heterogeneous network such as operator networks, enterprise-owned backbone networks or direct connection lines.

For large enterprises to access the cloud service and interconnect their branches, a PoP GWs network can be built to provide multi-cloud, multi-tenant, and multi-branch interconnection. Depending on the geographical distribution of the enterprise branches, the PoP GWs network may be a cross-regional or even a global network. The PoP GW

can be connected to the operator network or the enterprise-owned backbone network. The PoP GWs devices can also be directly connected through dedicated lines.

According to [I-D.draft-ietf-bess-bgp-sdwan-usage], SD-WAN tunnels can be established between two GWs devices connected to the operator network, MPLS VPN network, or internet network through the WAN ports of the two PoP GWs devices. All GWs are under the control of one BGP instance. [I-D.draft-ietf-idr-sdwan-edge-discovery] defines the mechanism for SD-WAN edges to discover each other’s properties via BGP update through RR. This allows the interconnection between enterprise branches and multi-cloud to pass through multiple SD-WAN tunnels or direct connection lines, as shown in Figure 1.

This draft provides a way to use the BGP-LS-SPF protocol to collect the identification of PoP GW device node and the topology of SD-WAN tunnel and direct connection lines. In this way, each PoP GW device can learn the PoP GWs network topology, and calculate the route to any other PoP GW.

2. Terminology

This specification reuses terms defined in Section 5.2 of [I-D.draft-ietf-lsvr-bgp-spf] including BGP-LS-SPF Node NLRI, BGP-LS-SPF Link NLRI, Dijkstra Algorithm.

* PoP GW: Point of Presence Gateway

* SD-WAN: Software Defined Wide Area Network. In this document, "SD-WAN" refers to policy-driven transporting IP packets over multiple different underlay networks to get better WAN bandwidth management, visibility and control.

* RR: Route Reflector

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

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Usage of BGP-LS-SPF in Multi-segment SD-WAN

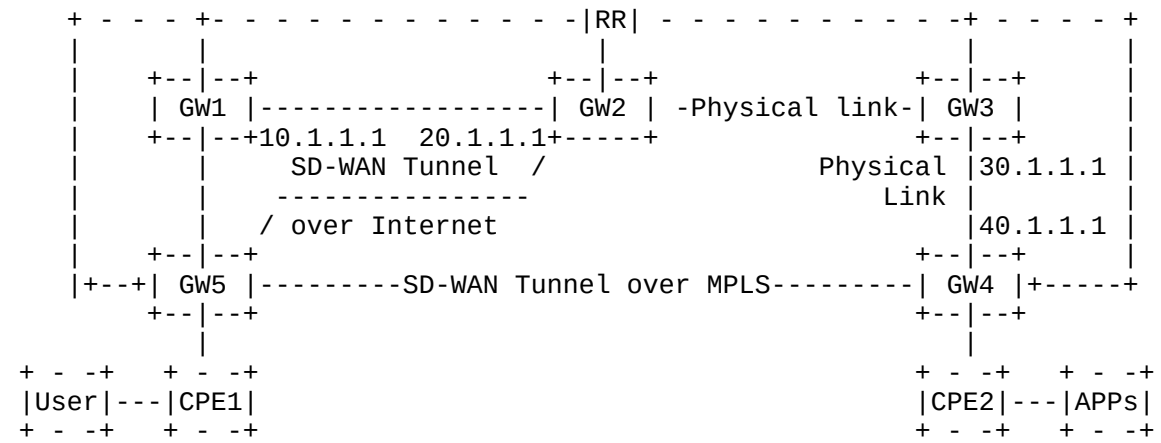

Figure 1: PoP GWs network

As shown in Figure 1, GW1, GW2, GW5 are connected to the same internet/ISP network. The GW2 and GW3 are connected through direct dedicated links. GW5 and GW4 are connected by MPLS VPN. BGP-SD-WAN neighbors are established between GWs through RR. BGP-LS-SPF neighbors are established between each GW and RR. SD-WAN tunnel links are established between GWs through BGP-SD-WAN neighbors reflecting SD-WAN routes(see [I-D.draft-ietf-idr-sdwan-edge-discovery]), as shown in the SD-WAN Tunnel between GW1 and GW2 with WAN port IP addresses of 10.1.1.1 and 20.1.1.1, respectively. GW nodes reflect the SD-WAN tunnel topology information to all GWs, including dedicated line-connected GWs, through BGP-LS-SPF neighbors with RR.

GW2-GW3-GW4 are connected through dedicated lines. BGP-LS-SPF neighbors are established between GWs through dedicated lines, and also between GWs and RR. The BGP-LS-SPF neighbors between dedicated lines are used to discover the topology information of the dedicated lines, such as the direct link with port IP addresses of 30.1.1.1 and 40.1.1.1 between GW3 and GW4 shown in the figure. The dedicated line topology information is reflected to all GWs, including SD-WAN tunnel-connected GWs, through BGP-LS-SPF neighbors with RR.

BGP-LS-SPF can be used in two scenarios in Multi-segment SD-WAN: 1. TE. When TE is used, SLA of all SD-WAN tunnels will be collected to calculate shortest path. The protocol ID of BGP-LS is BGP. The BGP-LS-SPF LINK NLRI is used to carry the two endpoint IP address of the SD-WAN tunnel or dedicated lines. The BGP-LS-SPF NODE NLRI is used to carry PoP GW device node identification. 2. BE. When BE is used, only reachability of a SD-WAN site is collected. An SD-WAN site may contains multiple GWs. There is no need to collect the SLA of every

SD-WAN tunnels between two sites. In this case, a new BGP-LS Protocol-ID is used and new Node Descriptor sub-tlv is defined to carry the site ID.

In both scenarios, BGP-LS-SPF LINK NLRI and NODE NLRI are advertised to other GWs through the RR. In this way, all GW learns the topology of whole PoP GWs network and can calculate the next hop to any other GW using Dijkstra Algorithm.

4. Extensions to BGP-LS

4.1. SDWAN Protocol ID

This document specifies the advertisement of SDWAN topology information via BGP-LS-SPF Link NLRI type and Node NLRI type, which requires use of a new BGP-LS Protocol-ID (value 10). The use of a new Protocol-ID allows separation and differentiation between the BGP-LS NLRIs carrying SDWAN topology information from the BGP-LS NLRIs carrying other link-state information defined in [RFC9552].

4.2. Node Descriptor Sub-tlv

This document introduces a new Node Descriptor Sub-TLV to carry the SDWAN Site ID to identify an SDWAN site. A site may contains multiple GWs. This field has the same meaning of SD-WAN-Color in Section 6.1 of [I-D.draft-ietf-idr-sdwan-edge-discovery], representing a group of tunnels terminated at SD-WAN GWs co-located at the site.

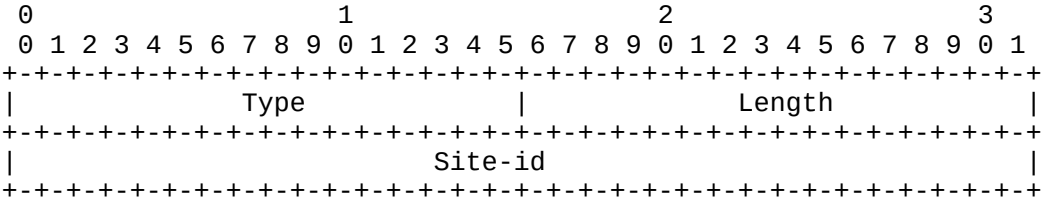

Figure 2: Node Descriptor Sub-TLV Format

4.3. Link-Type TLV

The link could be Overlay link (Such as Internet, MPLS, LTE etc.,) and Underlay/Physical link (Such as Dedicated line, Direct link etc.,). Different customer may require different types of link. For example, FinTech customer has very high security requirement and would like to exclude Internet and LTE, only use MPLS or Dedicated line; some customer only wants to use the Dedicated line/Direct link to get the highest quality path; some customer prefers to use LTE

only as backup link to save the cost. The calculation of these customized SD-WAN path needs to include or exclude one or more specific link types, therefore, when SD-WAN link information is advertised through BGP-LS-SPF Link NLRI, the SD-WAN link type needs to be explicitly indicated.

In this document, a new BGP-LS-SPF Attribute TLV of the BGP-LS-SPF Link NLRI is added to identify a SD-WAN link type, called Link-Type TLV. The format of the Link-Type TLV is defined as follows:

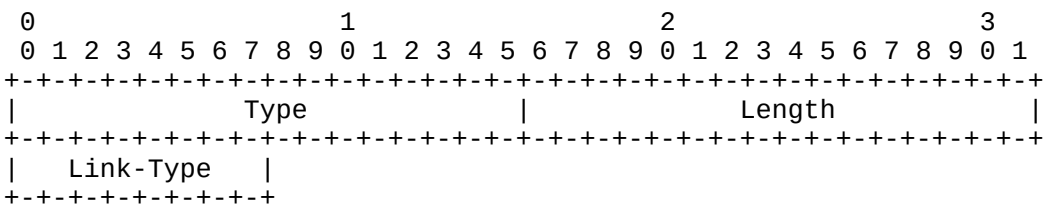

Figure 3: Link-Type TLV Format

where: Type: TBA

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 1.

Link-Type:

* 0: Reserved

* 1: Physical/Dedicated Line/Direct link

* 2: Internet

* 3: MPLS

* 4: LTE

This BGP-LS-SPF Attribute TLV of the BGP-LS-SPF Link NLRI is defined to indicate the Link-Type of the SD-WAN link.

5. Security Considerations

This document does not introduce any new security considerations.

6. IANA Considerations

6.1. BGP-LS Protocol-IDs

IANA maintains a registry called "BGP-LS Protocol-IDs" in the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group.

This document requests IANA to allocate the following Protocol-ID codepoint:

Protocol ID	NLRI information source protocol	Reference
10	SDWAN	this document

Table 1

6.2. BGP-LS TLVs

IANA maintains a registry called "BGP-LS NLRI and Attribute TLVs" in the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry group.

This document requests IANA to allocate the following TLV codepoint:

TLV Code Point	Description	Reference
TBD	SDWAN Node Descriptors	this document
TBD	Link-Type	this document

Table 2

7. References

7.1. Normative References

[I-D.draft-ietf-bess-bgp-sdwan-usage]
Dunbar, L., Sajassi, A., Drake, J., Najem, B., and S. Hares, "BGP Usage for SD-WAN Overlay Networks", Work in Progress, Internet-Draft, draft-ietf-bess-bgp-sdwan-usage-25, 20 December 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-bgp-sdwan-usage-25>.

[I-D.draft-ietf-idr-sdwan-edge-discovery]
Dunbar, L., Hares, S., Majumdar, K., Raszuk, R., and V. Kasiviswanathan, "BGP UPDATE for SD-WAN Edge Discovery", Work in Progress, Internet-Draft, draft-ietf-idr-sdwan-edge-discovery-22, 17 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-sdwan-edge-discovery-22>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

[RFC9552] Talaulikar, K., Ed., "Distribution of Link-State and Traffic Engineering Information Using BGP", RFC 9552, DOI 10.17487/RFC9552, December 2023, <https://www.rfc-editor.org/rfc/rfc9552>.

7.2. Informative References

[I-D.draft-ietf-rtgwg-net2cloud-problem-statement]
Dunbar, L., Malis, A. G., Jacquenet, C., Toy, M., and K. Majumdar, "Dynamic Networks to Hybrid Cloud DCs: Problems and Mitigation Practices", Work in Progress, Internet-Draft, draft-ietf-rtgwg-net2cloud-problem-statement-42, 17 January 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-net2cloud-problem-statement-42>.

[I-D.draft-ietf-lsvr-bgp-spf]
Patel, K., Lindem, A., Zandi, S., and W. Henderickx, "BGP Link-State Shortest Path First (SPF) Routing", Work in Progress, Internet-Draft, draft-ietf-lsvr-bgp-spf-51, 23 January 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-lsvr-bgp-spf-51>.

Appendix A. Acknowledgements

The authors would like to thank Donglei Pang for his contribution to the document.

Appendix B. Contributors

Shunwan Zhuang Huawei Email: zhuangshunwan@huawei.com

Authors' Addresses

Cheng Sheng
Huawei
Beiqing Road
Beijing
Email: shengcheng@huawei.com

Hang Shi
Huawei
Beiqing Road
Beijing
China
Email: shihang9@huawei.com

Jie Dong (editor)
Huawei
China
Email: jie.dong@huawei.com