Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: July 20, 2019

G. Dawra, Ed.
LinkedIn
C. Filsfils
Cisco Systems
D. Bernier
Bell Canada
J. Uttaro
AT&T
B. Decraene
Orange
H. Elmalky
Ericsson
X. Xu
Alibaba
F. Clad
K. Talaulikar
Cisco Systems
January 16, 2019

BGP-LS Advertisement of Segment Routing Service Segments
draft-dawra-idr-bgp-ls-sr-service-segments-01

Abstract

BGP Link-State (BGP-LS) enables distribution of topology information from the network to a Path Computation Engine (PCE) or any controller/application in general so it can learn the network topology. Service functions are deployed as virtualized elements along with network elements or on servers in data centers. The advertisement of such attached service capabilities along with the network nodes that they are attached to or associated with enable their discovery and for programming of service paths that use these service functions. Segment Routing (SR) bring in the concept of segments which can be topological or service instructions. SR Policies enable setup of paths which are a mix of topological and service segments.

This document specifies the extensions to BGP-LS for discovery and advertisement of service segments so as to enable setup of service programming paths using Segment Routing.

Status of This Memo

Table of Contents

1. Introduction

Segments are introduced in the SR architecture [I-D.ietf-spring-segment-routing]. Segment Routing based Service chaining is well described in Section 6 of [I-D.xuclad-spring-sr-service-programming] document with an example network and services.

This document extend the example to add a Segment Routing Controller (SR-C) to the network, for the purpose of service discovery and SR policy instantiation.

Consider the network represented in Figure 1 below where:

- A and B are two end hosts using IPv4.

- S1 is an SR-aware firewall Service.

- S2 is an SR-unaware DPI Service.

```
      SR-C      --3--
        |      /     \
        |     /       \
  A----1----2----4----5----6----B
            |         |
            |         |
            S1        S2
```

Figure 1: Network with Services

SR Controller (SR-C) is connected to Node 1, but may be attached to any node 1-6 in the network.

SR-C is capable of receiving BGP-LS updates to discover topology, and calculating constrained paths between 1 and 6.

However, if SR-C is configured to computation a constrained path from 1 and 6, including a DPI service (i.e., S2) it is not yet possible due to the lack of service distribution. SR-C does not know where a DPI Service is nor the SID for it. It does not know that S2 is a service it needs.

This document proposes an extension to BGP-LS for Service Chaining to distribute the service information to SR-C. There may be other alternate mechanisms to distribute service information to SR-C and are outside of scope of this document. There are no extensions required in SR-TE Policy SAFI.

2. BGP-LS Extensions for Service Chaining

For an attached service, following data needs to be shared with SR-C:

- o Service SID value (e.g. MPLS label or IPv6 address). Service SID MAY only be encoded as LOC:FUNCT, where LOC is the L most significant bits and FUNCT is the 128-L least significant bits[I-D.filsfils-spring-srv6-network-programming]. ARGs bits, if any, MAY be set to 0 in the advertised service SID.

- o Function Identifier (Static Proxy, Dynamic Proxy, Shared Memory Proxy, Masquerading Proxy, SR Aware Service etc).

- o Service Type (DPI, Firewall, Classifier, LB etc).

- o Traffic Type (IPv4 OR IPv6 OR Ethernet)

- o Opaque Data (Such as brand and version, other extra information)

[I-D.xuclad-spring-sr-service-programming]defines SR-aware and SR-unaware services. This document will reuse these definitions. Per [RFC7752] Node Attributes are ONLY associated with the Node NLRI. All non-VPN information SHALL be encoded using AFI 16388 / SAFI 71. VPN information SHALL be encoded using AFI 16388 / SAFI 72 with associated RTs.

This document extends SRv6 Node SID TLV [I-D.dawra-idr-bgpls-srv6-ext] and SR-MPLS SID/Label TLV [I-D.ietf-idr-bgp-ls-segment-routing-ext] to associate the Service SID Value with Service-related Information using Service Chaining(SC) Sub-TLV.

Function Sub-TLV [I-D.dawra-idr-bgpls-srv6-ext] of Node SID TLV encodes Identifier(Function ID) along with associated Function Flags.

A Service Chaining (SC) Sub-TLV in Figure 2 is defined as:

```
+--------------------------------------+
|         Type (2 octet)               |
+--------------------------------------+
|        Length (2 octet)              |
+--------------------------------------+
|        Service Type(ST) (2 octet     |
+--------------------------------------+
|        Flags (1 octet)               |
+--------------------------------------+
|        Traffic Type(1 octet)         |
+--------------------------------------+
|        RESERVED (2 octet)            |
+--------------------------------------+
```

Figure 2: Service Chaining(SC) Sub-TLV

Where:

Type: 16 bit field. TBD

Length: 16 bit field. The total length of the value portion of the TLV.

Service Type(ST): 16bit field. Service Type: categorizes the Service: (such as "Firewall", "Classifier" etc).

Flags: 8 bit field. Bits SHOULD be 0 on transmission and MUST be ignored on reception.

Traffic Type: 8 Bit field. A bit to identify if Service is IPv4 OR IPv6 OR L2 Ethernet Capable. Where:

Bit 0(LSB): Set to 1 if Service is IPv4 Capable

Bit 1: Set to 1 if Service is IPv6 Capable

Bit 2: Set to 1 if Service is Ethernet Capable

RESERVED: 16bit field. SHOULD be 0 on transmission and MUST be ignored on reception.

Service Type(ST) MUST be encoded as part of SC Sub-TLV.

There may be multiple instances of similar Services that needs to be distinguished. For example, firewalls made by different vendors A and B may need to be identified differently because, while they have similar functionality, their behavior is not identical.

In order for SDN Controller to identify the categories of Services and their associated SIDs, this section defines the BGP-LS extensions required to encode these characteristics and other relevant information about these Services.

Another Optional Opaque Metadata(OM) Sub-TLV of Node SID TLV may encode vendor specific information. Multiple of OM Sub-TLVs may be encoded.

```
+--------------------------------------+
|         Type (2 octet)               |
+--------------------------------------+
|        Length (2 octet)              |
+--------------------------------------+
|        Opaque  Type (2 octet)        |
+--------------------------------------+
|        Flags (1 octet)               |
+--------------------------------------+
|        Value (variable)              |
+--------------------------------------+
```

Figure 3: Opaque Metadata(OM) Sub-TLV

- o Type: 16 bit field. TBD.
- o Length: 16 bit field. The total length of the value portion of the TLV.
- o Opaque Type: 8-bit field. Only publishers and consumers of the opaque data are supposed to understand the data.
- o Flags: 8 bit field. Bits SHOULD be 0 on transmission and MUST be ignored on reception.
- o Value: Variable Length. Based on the data being encoded and length is recorded in length field.

Opaque Metadata(OM) Sub-TLV defined in Figure 3 may encode propriety or Service Opaque information such as:

- o Vendor specific Service Information.
- o Traffic Limiting Information to particular Service Type.
- o Opaque Information unique to the Service
- o Propriety Enterprise Service specific Information.

3. Illustration

In our SRv6 example above Figure 1 , Node 5 is configured with an SRv6 dynamic proxy segments (End.AD) C5::AD:F2 for S2.

The BGP-LS advertisement MUST contain and Node SID TLV:

o Service SID: C5::AD:F2 SID

o Function ID: END.AD

The BGP-LS advertisement MUST contain a SC Sub-TLV with:

o Service Type: Deep Packet Inspection(DPI)

o Traffic Type: IPv4 Capable.

The BGP-LS advertisement MAY contain a OM Sub-TLV with:

o Opaque Type: Cisco DPI Version

o Value: 3.5

In our example in Figure 1, using BGP SR-TE SAFI Update [I-D.ietf-idr-segment-routing-te-policy], SR Controller computes the candidate path and pushes the Policy.

SRv6 encapsulation policy < CF1::, C3::, C5::AD:F2, C6::D4:B > is signaled to Node 1 which has mix of service and topological segments.

4. IANA Considerations

This document requests assigning code-points from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs".

4.1. Service Type Table

IANA is request to create a new top-level registry called "Service Type Table (STT)". Valid values are in the range 0 to 65535. Values 0 and 65535 are to be marked "Reserved, not to be allocated".

Service Value(TBD)	Service	Reference	Date
32	Classifier	ref-to-set	date-to-set
33	Firewall	ref-to-set	date-to-set
34	Load Balancer	ref-to-set	date-to-set
35	DPI	ref-to-set	date-to-set

Figure 4

4.2. Segment routing function Identifier(SFI)

IANA is request to extend a top-level registry called "Segment Routing Function Identifier(SFI)" with new code points. This document extends the SFI values defined in [I-D.dawra-idr-bgpls-srv6-ext]. Details about the Service functions are defined in[I-D.xuclad-spring-sr-service-programming].

Function	Function Identifier
Static Proxy	8
Dynamic Proxy	9
Shared Memory Proxy	10
Masquerading Proxy	11
SRv6 Aware Service	12

5. Manageability Considerations

This section is structured as recommended in[RFC5706]

6. Operational Considerations

6.1. Operations

Existing BGP and BGP-LS operational procedures apply. No additional operation procedures are defined in this document.

7. Security Considerations

Procedures and protocol extensions defined in this document do not affect the BGP security model. See the 'Security Considerations' section of [RFC4271]for a discussion of BGP security. Also refer to[RFC4272]and[RFC6952]for analysis of security issues for BGP.

8. Conclusions

This document proposes extensions to the BGP-LS to allow discovery of Services using Segment Routing.

9. Acknowledgements

The authors would like to thank Krishnaswamy Ananthamurthy for his review of this document.

10. References

10.1. Normative References

[I-D.dawra-idr-bgpls-srv6-ext]
Dawra, G., Filsfils, C., Talaulikar, K., Chen, M., daniel.bernier@bell.ca, d., Uttaro, J., Decraene, B., and H. Elmalky, "BGP Link State extensions for IPv6 Segment Routing(SRv6)", draft-dawra-idr-bgpls-srv6-ext-04 (work in progress), September 2018.

[I-D.xuclad-spring-sr-service-programming]
Clad, F., Xu, X., Filsfils, C., daniel.bernier@bell.ca, d., Li, C., Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and S. Salsano, "Service Programming with Segment Routing", draft-xuclad-spring-sr-service-programming-01 (work in progress), October 2018.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC5706] Harrington, D., "Guidelines for Considering Operations and Management of New Protocols and Protocol Extensions", RFC 5706, DOI 10.17487/RFC5706, November 2009, <https://www.rfc-editor.org/info/rfc5706>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <https://www.rfc-editor.org/info/rfc6952>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

10.2. Informative References

[I-D.dawra-bgp-srv6-vpn]
(Unknown), (., Dawra, G., Filsfils, C., Dukes, D., Brissette, P., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Decraene, B., and S. Matsushima, "BGP Signaling of IPv6-Segment-Routing-based VPN Networks", draft-dawra-bgp-srv6-vpn-00 (work in progress), March 2017.

[I-D.filsfils-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-06 (work in progress), October 2018.

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Previdi, S., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-15 (work in progress), October 2018.

[I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in EVPN", draft-ietf-bess-evpn-prefix-advertisement-11 (work in progress), May 2018.

[I-D.ietf-idr-bgp-ls-segment-routing-ext]
Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H., and M. Chen, "BGP Link-State extensions for Segment Routing", draft-ietf-idr-bgp-ls-segment-routing-ext-11 (work in progress), October 2018.

[I-D.ietf-idr-bgp-prefix-sid]
Previdi, S., Filsfils, C., Lindem, A., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix SID extensions for BGP", draft-ietf-idr-bgp-prefix-sid-27 (work in progress), June 2018.

[I-D.ietf-idr-segment-routing-te-policy]
Previdi, S., Filsfils, C., Jain, D., Mattes, P., Rosen, E., and S. Lin, "Advertising Segment Routing Policies in BGP", draft-ietf-idr-segment-routing-te-policy-05 (work in progress), November 2018.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-22 (work in progress), December 2018.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-02 (work in progress), October 2018.

[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/info/rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, DOI 10.17487/RFC4659, September 2006, <https://www.rfc-editor.org/info/rfc4659>.

[RFC5549] Le Faucheur, F. and E. Rosen, "Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop", RFC 5549, DOI 10.17487/RFC5549, May 2009, <https://www.rfc-editor.org/info/rfc5549>.

Authors' Addresses

Gaurav Dawra (editor)
LinkedIn
USA

Email: gdawra.ietf@gmail.com

Clarence Filsfils
Cisco Systems
Belgium

Email: cfilsfil@cisco.com

Daniel Bernier
Bell Canada
Canada

Email: daniel.bernier@bell.ca

Jim Uttaro
AT&T
USA

Email: ju1738@att.com

Bruno Decraene
Orange
France

Email: bruno.decraene@orange.com

Hani Elmalky
Ericsson
USA

Email: hani.elmalky@gmail.com

Xiaohu Xu
Alibaba

Email: xiaohu.xxh@alibaba-inc.com

Francois Clad
Cisco Systems
France

Email: fclad@cisco.com

Ketan Talaulikar
Cisco Systems
India

Email: ketant@cisco.com