Inter-Domain Routing   G. Dawra, Ed.
Internet-Draft   LinkedIn
Intended status: Standards Track   C. Filsfils
Expires: October 26, 2022   Cisco Systems
                                                    K. Talaulikar, Ed.
                                                           Arrcus Inc
                                                             F. Clad
                                                       Cisco Systems
                                                          D. Bernier
                                                          Bell Canada
                                                           J. Uttaro
                                                                AT&T
                                                         B. Decraene
                                                              Orange
                                                          H. Elmalky
                                                            Ericsson
                                                               X. Xu
                                                        Capitalonline
                                                         J. Guichard
                                                Futurewei Technologies
                                                               C. Li
                                                  Huawei Technologies
                                                      April 24, 2022

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

Abstract

   Service functions are deployed as, physical or virtualized elements
   along with network nodes or on servers in data centers. Segment
   Routing (SR) brings in the concept of segments which can be
   topological or service instructions. Service segments are SR
   segments that are associated with service functions. SR Policies are
   used for the setup of paths for steering of traffic through service
   functions using their service segments.

   BGP Link-State (BGP-LS) enables distribution of topology information
   from the network to a controller or an application in general so it
   can learn the network topology. This document specifies the
   extensions to BGP-LS for the advertisement of service functions along
   their associated service segments. The BGP-LS advertisement of
   service function information along with the network nodes that they
   are attached to, or associated with, enables controllers compute and
   setup service paths in the network.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   Segments are introduced in the SR architecture [RFC8402]. Segment
   Routing based Service chaining is well described in
   [I-D.ietf-spring-sr-service-programming] with an example of 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 [I-D.ietf-spring-segment-routing-policy] instantiation.

   Consider the network represented in Figure 1 below where:

   o A and B are two end hosts using IPv4.

   o S1 is an SR-aware firewall Service.

   o S2 is an SR-unaware DPI Service.

```
              SR-C                       --3--
                                        /     \
                                       /       \
                           A----1----2----4----5----6----B
                                 |         |
                                 |         |
                                S1        S2
```
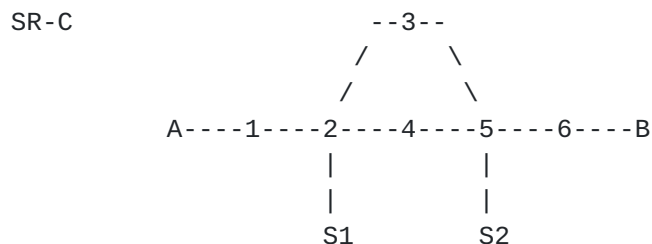
                   Figure 1: Network with Services

   SR Controller (SR-C) is connected to the network.

   SR-C can receive BGP-LS updates to discover topology, and calculate
   constrained paths between nodes 1 and 6.

   However, if SR-C is configured to compute 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 the scope of this document. There are no extensions
   required in SR-TE Policy SAFI.

1.1. 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.

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 be encoded as LOC:FUNCT:ARG as specified in [RFC8986].

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.ietf-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 introduces new TLVs for the SRv6 SID NLRI
[I-D.ietf-idr-bgpls-srv6-ext] and SR-MPLS SID/Label TLV [RFC9085] to
associate the Service SID value with Service-related Information
using Service Chaining(SC) Sub-TLV.

SRv6 SID Information TLV [I-D.ietf-idr-bgpls-srv6-ext] encodes
behavior along with associated SID Flags.

A Service Chaining (SC) 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)           |
+---------------------------------------+
```

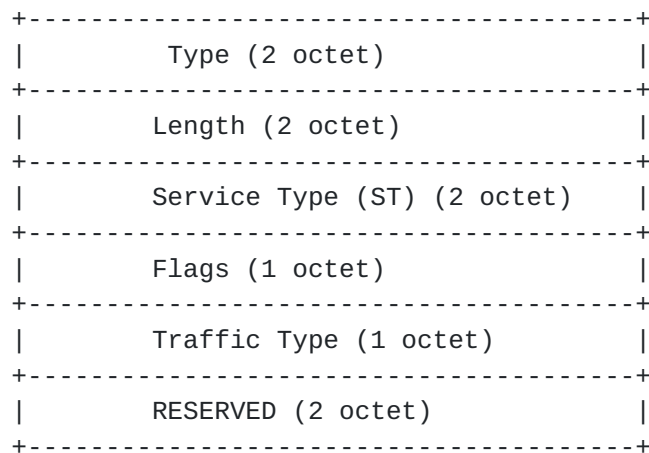   Figure 2: Service Chaining (SC) TLV

   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 TLV.

   There may be multiple instances of similar Services that need 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 the 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) TLV of SRv6 SID NLRI may encode
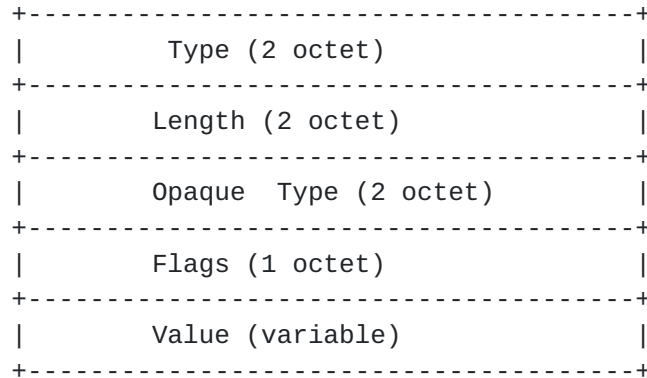vendor specific information. Multiple of OM TLVs may be encoded.

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

                Figure 3: Opaque Metadata(OM) 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) 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 include SRv6 SID NLRI with SRv6 SID
   Information TLV in the BGP-LS Attribute:

   o Service SID: C5::AD:F2 SID

   o Endpoint Behavior: END.AD

   The BGP-LS Attribute MUST contain a SC TLV with:

   o Service Type: Deep Packet Inspection(DPI)

   o Traffic Type: IPv4 Capable.

   The BGP-LS Attribute MAY contain a OM 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   |       Service        | Reference | Date       |
| Value(TBD) |                      |           |            |
+------------+----------------------+-----------+------------+
|  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.ietf-idr-bgpls-srv6-ext]. Details about the Service functions
   are defined in[I-D.ietf-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. Acknowledgements

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

9. References

9.1. Normative References

[I-D.ietf-idr-bgpls-srv6-ext]
          Dawra, G., Filsfils, C., Talaulikar, K., Chen, M.,
          Bernier, D., and B. Decraene, "BGP Link State Extensions
          for SRv6", draft-ietf-idr-bgpls-srv6-ext-09 (work in
          progress), November 2021.

[I-D.ietf-spring-sr-service-programming]
          Clad, F., Xu, X., Filsfils, C., Bernier, D., Li, C.,
          Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and
          S. Salsano, "Service Programming with Segment Routing",
          draft-ietf-spring-sr-service-programming-05 (work in
          progress), September 2021.

[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>.

[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>.

[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/info/rfc8174>.

   [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L.,
              Decraene, B., Litkowski, S., and R. Shakir, "Segment
              Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,
              July 2018, <https://www.rfc-editor.org/info/rfc8402>.

   [RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer,
              D., Matsushima, S., and Z. Li, "Segment Routing over IPv6
              (SRv6) Network Programming", RFC 8986,
              DOI 10.17487/RFC8986, February 2021,
              <https://www.rfc-editor.org/info/rfc8986>.

   [RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler,
              H., and M. Chen, "Border Gateway Protocol - Link State
              (BGP-LS) Extensions for Segment Routing", RFC 9085,
              DOI 10.17487/RFC9085, August 2021,
              <https://www.rfc-editor.org/info/rfc9085>.

9.2. Informative References

   [I-D.ietf-idr-segment-routing-te-policy]
              Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P.,
              Jain, D., and S. Lin, "Advertising Segment Routing
              Policies in BGP", draft-ietf-idr-segment-routing-te-
              policy-17 (work in progress), April 2022.

   [I-D.ietf-spring-segment-routing-policy]
              Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and
              P. Mattes, "Segment Routing Policy Architecture", draft-
              ietf-spring-segment-routing-policy-22 (work in progress),
              March 2022.

   [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>.

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

   [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>.

Authors' Addresses

   Gaurav Dawra (editor)
   LinkedIn
   USA

   Email: gdawra.ietf@gmail.com

   Clarence Filsfils
   Cisco Systems
   Belgium

   Email: cfilsfil@cisco.com

   Ketan Talaulikar (editor)
   Arrcus Inc
   India

   Email: ketant.ietf@gmail.com

   Francois Clad
   Cisco Systems
   France

   Email: fclad@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
Capitalonline

Email: xiaohu.xu@capitalonline.net

Jim Guichard
Futurewei Technologies
USA

Email: james.n.guichard@futurewei.com

Cheng Li
Huawei Technologies
China

Email: chengli13@huawei.com