Workgroup: LSR Working Group
Internet-Draft:
draft-wang-lsr-stub-link-attributes-05
Published: 21 October 2022
Intended Status: Standards Track
Expires: 24 April 2023
Authors:
A. Wang, China Telecom
Z. Hu, Huawei Technologies
A. Lindem, Cisco Systems
G. Mishra, Verizon Inc.
J. Sun, ZTE Corporation

Advertisement of Stub Link Attributes

Abstract

This document describes the mechanism that can be used to advertise the stub link attributes within the IS-IS or OSPF domain.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Stub links are used commonly within enterprise or service provider networks. One common use case is the inter-AS routing scenario where there are no IGP adjacencies between the adjacent BGP domains, another use case is at the network boundary that the interfaces are used to connect to the application servers.

For operators that have multiple ASes interconnecting with each other via the stub links, there is a requirement to obtain the inter-AS topology information as described in [I-D.ietf-idr-bgpls-inter-as-topology-ext]. To achieve such goal, it is required that the BGP-LS to be enabled on every router that has the stub links, which is challenging for the network operation. It is desirable to advertise the stub link info into the IGP to ease the deployment of BGP-LS on any router in the IGP domain.

For stub links that are used to connect the servers, knowing the status of these stub links can facilitate the routers within the IGP to accomplish TE tasks in some scenarios.

But OSPF and IS-IS have no capability to identify such stub links and their associated attributes now.

This document defines the protocol extension for OSPFv2/v3 and IS-IS to indicate the stub links and their associated attributes.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] .

3. Consideration for Identifying Stub Link

OSPF[RFC5392] defines the Inter-AS-TE-v2 LSA and Inter-AS-TE-v3 LSA to carry the TE information about inter-AS links. IS-IS[RFC5316] defines the Inter-AS Reachability TLV to carry the TE information about inter-AS links. But they are normally being used under RSVP-TE, especially inter-domain RSVP-TE scenarios. As illustrated in the potential scenarios that described in Appendix A, there is still the need for a generic solution which also covers non inter-AS stub links.

Then, to solve the problems that described in the applied scenarios, this document defines the Stub-Link TLV to identify the stub link and transmit the associated attributes for OSPF and IS-IS respectively.

4. Protocol Extension for Stub Link Attributes

The following sections define the protocol extension to indicate the stub link and its associated attributes in OSPFv2/v3 and IS-IS.

4.1. OSPF Stub-Link TLV

This document defines the Stub-Link TLV to describe stub link of a single router.

For OSPFv2, the newly defined Stub-Link TLV is named as OSPFv2 Extended Stub-Link TLV, which is included in the OSPFv2 Extended Link Opaque LSA [RFC7684]

For OSPFv3, the newly defined Stub-Link TLV is named as Router-Stub-Link TLV, which is included in the OSPFv3 Router-LSA [RFC8362]

OSPFv2 Extended Stub-Link TLV and OSPFv3 Router-Stub-Link TLV has the following same format:

```
 0                   1                   2                   3
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type(Stub-Link)            |       Length                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|U|        Flags                |     Reserved                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Link Prefix Sub-TLVs                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Existing Sub-TLVs (variable)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: OSPF Stub-Link TLV

Type: The TLV type. The value is 2(TBD) for OSPFv2 Extended Link Stub-Link TLV under the IANA codepoint "OSPFv2 Extended Link Opaque LSA TLVs", and is 10(TBD) for OSPFv3 Router-Stub-Link TLV under the IANA codepoint "OSPFv3 Extended-LSA TLVs"

Length: Variable, dependent on sub-TLVs

Flags: Define the type of the stub-link:

* U bit(bit 0): Identify the unnumbered stub link if this bit is set.

* bit 1-bit 15: Reserved

Link Prefix Sub-TLV: The prefix of the stub-link. It's format is defined in Section 4.3 and Section 4.4.

Existing Sub-TLVs: Sub-TLV that defined within "OSPFv2 Extended Link TLV Sub-TLVs" and "OSPFv3 Extended-LSA Sub-TLVs" can be included if necessary.

If the stub-link is identified as unnumbered stub link (U bit is set), then the "Remote IPv4 Address sub-TLV" or "Remote Interface IPv6 Address sub-TLV", which should be set to the identifier value of remote router, SHOULD be included to facilitate the pairing of inter-AS link.

This document creates a registry for Stub-Link attributes in Section 6.

4.2. IS-IS Stub-link TLV

This document defines the IS-IS Stub-Link TLV to describes stub link of a single router.

The IS-IS Stub-Link TLV has the following format:

```
 0                   1                   2                   3
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type(Stub-Link)            |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|U|       Flags                 |          Reserved             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Link Prefix Sub-TLV                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Existing Sub-TLVs(Variable)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: IS-IS Stub-Link TLV

Type: IS-IS TLV codepoint. Value is 151 (TBD) for stub-link TLV.

Length: Variable, dependent on sub-TLVs

Flags: Define the type of the stub-link:

* 0: U bit(bit 0): Identify the unnumbered stub link if this bit is set.

* bit 1-bit 15: Reserved

Link Prefix Sub-TLV: The prefix of the stub-link. It's format is defined in Section 4.3 and Section 4.4.

Existing Sub-TLVs: Sub-TLVs that defined within "IS-IS Sub-TLVs for TLVs Advertising Neighbor Information " can be included if necessary.

If the stub-link is identified as unnumbered stub link type (U bit is set), then the "IPv4 Remote ASBR ID" or "IPv6 Remote ASBR ID" sub-TLV SHOULD be included to facilitate the pairing of inter-AS link.

4.3. IPv4 Prefix Sub-TLV

The IPv4 Prefix Sub-TLV has the following format:

```
 0                   1                   2                   3
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type                       |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv4 Prefix                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: IPv4 Prefix Sub-TLV

Type: IPv4 Prefix Sub-TLV codepoint. Value is 25(TBD) for OSPFv2 (under "OSPFv2 Extended Link Sub-TLVs")

30(TBD) for OSPFv3(under OSPFv3 Extended-LSA Sub-TLVs)

45(TBD) for IS-IS(under "IS-IS Sub-TLVs for TLVs Advertising Neighbor Information")

Length: Netmask length value of the IPv4 Prefix. Value should be in 2-32.

IPv4 Prefix: The value of 4-octet IPv4 Prefix address, the host part should be zero.

4.4. IPv6 Prefix Sub-TLV

The IPv6 Prefix Sub-TLV has the following format:

```
 0                   1                   2                   3
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type                       |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv6 Prefix                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv6 Prefix                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv6 Prefix                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv6 Prefix                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: IPv6 Prefix Sub-TLV

Type: IPv6 Prefix Sub-TLV codepoint. Value is 31(TBD) for OSPFv3. (under OSPFv3 Extended-LSA Sub-TLVs)

46(TBD) for IS-IS(under "IS-IS Sub-TLVs for TLVs Advertising Neighbor Information")

Length: Netmask length value of the IPv6 Prefix. Value should be in 2-128.

IPv6 Prefix: The value of 16-octet IPv6 Prefix address, the host part should be zero.

5. Security Considerations

Security concerns for IS-IS are addressed in [RFC5304] and[RFC5310]

Security concern for OSPFv3 is addressed in [RFC4552]

Advertisement of the additional information defined in this document introduces no new security concerns.

6. IANA Considerations

IANA is requested to the allocation in following registries:

Registry	Type	Meaning
OSPFv2 Extended Link Opaque LSA TLVs	2	OSPFv2 Extended Stub-Link T
OSPFv3 Extended-LSA TLVs	10	Router-Stub-Link TLV
IS-IS Top-Level TLV	151	IS-IS Stub-Link TLV
OSPFv2 Extended Link TLV	25	IPv4 Prefix Sub-TLV
OSPFv3 Extended-LSA Sub-TLVs	30	IPv4 Prefix Sub-TLV
OSPFv3 Extended-LSA Sub-TLVs	31	IPv6 Prefix Sub-TLV
IS-IS Sub-TLVs for TLVs Advertising Neighbor Information	45	IPv4 Prefix Sub-TLV
IS-IS Sub-TLVs for TLVs Advertising Neighbor Information	46	IPv6 Prefix Sub-TLV

Figure 5: IANA Allocation for newly defined TLVs and Sub-TLVs

7. Acknowledgement

Thanks Ketan Talaulikar, Shunwan Zhang, Peter Psenak, Tony Li, Les Ginsberg, Dhruv Dhody, Jeff Tantsura and Robert Raszuk for their suggestions and comments on this idea.

8. References

8.1. Normative References

[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>.

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, DOI 10.17487/RFC4552, June 2006, <https://www.rfc-editor.org/info/rfc4552>.

[RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <https://www.rfc-editor.org/info/rfc5304>.

[RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <https://www.rfc-editor.org/info/rfc5310>.

[RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <https://www.rfc-editor.org/info/rfc5316>.

[RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <https://www.rfc-editor.org/info/rfc5392>.

[RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <https://www.rfc-editor.org/info/rfc7684>.

[RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", RFC 8362, DOI 10.17487/RFC8362, April 2018, <https://www.rfc-editor.org/info/rfc8362>.

8.2. Informative References

[I-D.ietf-idr-bgpls-inter-as-topology-ext] Wang, A., Chen, H., Talaulikar, K., and S. Zhuang, "BGP-LS Extension for Inter-AS Topology Retrieval", Work in Progress, Internet-Draft, draft-ietf-idr-bgpls-inter-as-topology-ext-11, 12 May 2022, <https://www.ietf.org/archive/id/draft-ietf-idr-bgpls-inter-as-topology-ext-11.txt>.

Appendix A. Applied Scenarios

The following sections describe the scenarios that knowing the stub link related attributes information can help solve the corresponding necessity in questions.

A.1. Inter-AS topology recovery

Figure 1 describes the scenario that the necessity of inter-AS topology recovery for Native IP point-to-point stub link scenario.

R10, R11 and R12 are located in AS1. R20, R21,R22 are located in AS2. The controller runs BGP-LS with R10 in AS1 and R20 in AS2 respectively.

There is one BGP session among the border router R11 and R21, which are connected by several stub links(passive interfaces) between them. The situation within the R21 and R22 are the same.

Since the links between the border routers are passive, there will be no IGP neighbors between them. The BGP-LS information carried in each AS will not report these stub links,and the controller can't recovery the inter-AS topology automatically.

```
                           +----------+
            +--------------+Controller+---------------+
            ^              +----------+               ^
            |                                         |
      BGP-LS|                                         |BGP-LS
            |            +---+BGP+----+               |
       +-----------------+            +--------------------+
       |    |            |            |               |    |
       |    |        +---+-+        +-+---+           |    |
       |    |        |     +--------+     |           |    |
       |    +--------+R11  |--------|  R21+-----------+    |
       |    |        |     +--------+     |           |    |
       |    |        +---+-+        +-+---+           |    |
       | +--+--+         |            |           +--+--+  |
       | |     |         |            |           |     |  |
       | |R10  |         |            |           |R20  |  |
       | |     |         |            |           |     |  |
       | +--+--+     +---+-+        +-+---+       +--+--+  |
       |    |        |     +--------+     |           |    |
       |    +--------+R12  |--------|  R22+-----------+    |
       |             |     +--------+     |                |
       |             +---+-+        +-+---+                |
       |                 |            |                    |
       +--------AS1------+            +----------AS2-------+
                         +---+BGP+----+
```

Figure 1: Inter-AS Topology Recovery(P2P Scenario)

Figure 2 describes the similar situation but in LAN environment. The border routers of AS1, AS2 and AS3 are connected via one LAN interfaces(that is to say, the corresponding interfaces on R1, R2 and R3 are on the same subnet). There are three different BGP sessions from the loopback address of the border routers among them respectively. It is necessary to recovery the underlying inter-AS topology automatically.

```
          +---------+                      +---------+
         |         |                      |         |
         |      +--+--+               +--+--+       |
         |      |     |               |     |       |
         |      |R1   +-------+ +------+R2   |       |
         |      |     |       | |      |     |       |
         |      +--+--+       | |      +--+--+       |
         |         |          | |         |          |
         +---AS1---+          | |         +----AS2--+
                            +-+-+-+
                            |     |
                            | SW  |
                            |     |
                            +--+--+
                               |
                            +--+--+
                         +--+     +--+
                         |  |  R3 |  |
                         |  +-----+  |
                         |           |
                         |           |
                         +-----AS3---+
```

Figure 2: Inter-AS Topology Recovery(LAN Scenario)

A.2. Egress Engineering for Anycast Servers

Figure 3 describes the scenario that the stub link information can be used for egress engineering for Anycast servers that connected to the network. In the example, the R1, R2 and R3 are border routers which are connected directly the server S1, S2 and S3 that have the same IP address IPa. The characteristics of the stub links that connected to these Anycast servers are different. It will be help for the router R0, to know the attributes of the stub links and select the optimal Anycast server to serve the customer's application.

```
        +----------------+
        |                |
        |            +---+-+      +-----+
        |            |     |      | S1  |
        |    +-------+R1   +-------+     |
        |    |       |     |      |(IPa)|
        |    |       +---+-+      +-----+
        |    |           |
        | +--+--+    +---+-+      +-----+
        | |     |    |     |      | S2  |
        | |R0   +----+R2   +-------+     |
        | |     |    |     |      |(IPa)|
        | +--+--+    +---+-+      +-----+
        |    |           |
        |    |       +---+-+      +-----+
        |    |       |     |      | S3  |
        |    +-------+R3   +-------+     |
        |            |     |      |(IPa)|
        |            +---+-+      +-----+
        |                |
        +----------------+
```

Figure 3: Egress Engineering for Anycast Server

A.3. Optimized BGP Next-hop Selection

Figure 4 describes the scenario that the stub link information can facilitate the optimized BGP next hop selection. The router R10 and R20 which are located in different AS establish the BGP session directly, with the explicit route set on each other which point to the egress stub interface between the border routers. The attributes of the stub links among the border routers are vary. It is certainly will be helpful for the router R10 and R20 to select the optimized BGP next hop, that is via the stub links among them, to reach each other.

```
               |<---------------+BGP+------------------>|
               |                                        |
               |                                        |
          +-----------------+            +-------------------+
          |    |            |            |             |     |
          |    |        +---+-+        +-+---+         |     |
          |    |        |     +--------+     |         |     |
          |    +--------+R11  |--------|  R21+----------+     |
          |    |        |     +--------+     |         |     |
          |    |        +---+-+        +-+---+         |     |
          | +--+--+         |            |           +--+--+ |
          | |     |         |            |           |     | |
          | |R10  |         |            |           |R20  | |
          | |     |         |            |           |     | |
          | +--+--+     +---+-+        +-+---+       +--+--+ |
          |    |        |     +--------+     |          |    |
          |    +--------+R12  |--------|  R22+----------+    |
          |             |     +--------+     |               |
          |             +---+-+        +-+---+               |
          |                 |            |                   |
          +-------+AS1+-----+            +---------+AS2+-----+
```

Figure 4: Optimized BGP next hop selection

Authors' Addresses

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing
102209
China

Email: wangaj3@chinatelecom.cn

Zhibo Hu
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing
100095
China

Email: huzhibo@huawei.com

Acee Lindem
Cisco Systems
No. 301 Midenhall Way
Cary, NC 27513
United States of America

Email: acee@cisco.com

Gyan S. Mishra
Verizon Inc.
13101 Columbia Pike

Silver Spring, MD 20904
United States of America

Email: gyan.s.mishra@verizon.com

Jinsong Sun
ZTE Corporation
No. 68, Ziijnhua Road
Nan Jing
210012
China

Email: sun.jinsong@zte.com.cn