IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 21, 2020

J. Zhou
Chunning. Dai
Shaofu. Peng
ZTE Corp.

Feb 18, 2020

Inter-domain Network Slicing via BGP-LU
draft-zhou-idr-inter-domain-lcu-01

Abstract

This document aims to solve inter-domain network slicing problems using existing technologies. It attempts to establish multiple BGP-LU LSPs of different colors for a prefix to stitch multiple network segments.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction 2
2. Requirements notation 2
3. Color 2
4. Advertising multiple paths 3
5. Assigning Label(s) 4
6. Establishing BGP-LU LSPs 4
7. Deploy Considerations 6
8. Security Considerations 6
9. References 6
9.1. Normative References 6
9.2. Informative References 7
Authors' Addresses 7

1. Introduction

As described in [I-D.peng-teas-network-slicing], in the traditional end to end inter-domain network slicing, BGP-LU is used as a distributed slicing scheme. [RFC8277] specifies a set of procedures for using BGP to advertise that a specified router has bound a specified MPLS label to a specified address prefix. It's an effective way for inter-domain labels, but it does not have the ability to select the underlying network resources.

This document describes the colored BGP-LU LSP, in which the routing prefix not only carries a label(or a sequence of labels), but also carries an unique color attribute which helps to select the underlying logic slice.

[RFC7911] defines a BGP extension that allows the advertisement of multiple paths of the same address prefix. It can help with optimal routing and in a network by providing potential alternate or backup paths. In this document, it is not used for the link backup. It is only used to advertise multiple paths, whether intra-domain or inter-domain.

2. Requirements notation

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. Color

[I-D.ietf-idr-tunnel-encaps] introduces the concept of color, which is used as one of the KEY of SR policy [I-D.ietf-spring-segment-routing-policy]. The color of SR policy

defines a TE purpose, which includes a set of constraints such as bandwidth, delay, TE metric, etc.

Color is used a policy keyword to help routing decisions and can be carried through Community attribute [I-D.ietf-idr-tunnel-encaps] in BGP. Each UPDATE SHOULD contain a Color Extended Community with a specific color value, the Color Sub-TLV is only an opaque extended community.

4. Advertising multiple paths

A BGP speaker can advertise multiple paths for a particular address prefix by a Path identifier in the Extended NLRI Encoding as defined in [RFC7911].

```
+--------------------------------+
| Path Identifier (4 octets)     |
+--------------------------------+
| Length (1 octet)               |
+--------------------------------+
| Prefix (variable)              |
+--------------------------------+
```

The Path Identifier only identifies a path, not carrying any particular semantics. It MAY be generated by the <Prefix,Color> tuple. It MAY be generated from the information containing Prefix and Color. It MIGHT be generated by the prefix without color. The assignment to the Path Identifier for a path by a BGP speaker is purely a local matter. However, The Path Identifier MUST be assigned in a way that the BGP speaker is able to use the (Prefix, Path Identifier) to uniquely identify a path.

Therefore, if a BGP speaker has two colors for the prefix P, which correspond to two different paths, it may advertise two UPDATE NLRIs, <prefix, pathid1> with color1 extended community and <prefix, pathid2> with color2 extended community. Pathid1 and pathid2 in two

UPDATE NLRIs MUST be different.

Note that in this document, BGP speakers acting as border routers that interact with external neighbors need to support advertising multiple paths corresponding to the same prefix. Although multiple paths have differnet path ids, they have the same next hop. As for the procedures of mutual backup paths with the same prefix and the differnet next hops, refer to [RFC7911].

5. Assigning Label(s)

[RFC8277] describes how to use BGP to bind MPLS label(s) to address prefixes. The specific format of the UPDATE message is detailed in Section 2 of [RFC8277].

[RFC8277] Section 3.2 details the process of modifying the Label field during propagation. When propagating a SAFI-4 or SAFI-128 route, if the Network Address of Next Hop field has never changed, the label field must remain unchanged. Otherwise, if the Network Address of Next Hop field is changed, the label field(s) of the propagating route must contain the label(s) that is (are) bound to the prefix at the new next hop. What the label changes to depends on the local policy. However, LSPs with different color paths need to have different label(s).

6. Establishing BGP-LU LSPs

Consider the following LSR topology in Figure 1: PE1---ASBR1---ASBR2 ---PE2. Packet transfers from PE2 to PE1. PE1, as one of the tail routers, has a loopback IP 1.1.1.1. The overlay service of PE1 to PE2 for prefix 1.1.1.1 has two colors, color1 and color2.

```
   <1.1.1.1, path-id1>     <1.1.1.1, path-id1>    <1.1.1.1, path-id1>
   <color1, label200>      <color1, label201>     <color1, label201>
   ------------------------------------------------------------------>
            .----.                                      .------.
           (      )                                    (        )
        .-(        )-.                              .-(          )-.
   +---+---color1----+----+                     +---+----color1----+---+
   |PE1|\---SR-TE1---/|AS |-sub-if1 with slice1-|AS |\---SR-TE1---/|PE2|
   |   |/---SR-TE2---\|BR1|-sub-if2 with slice2-|BR2|/---SR-TE2---\|   |
   +---+---color2---- +---+                     +---+--color2------+---+
       (              )                              (             )
         '--( AS1 )--'                                '--( AS2 )--'
             (   )                                        (   )
              '-'                                          '-'
   ------------------------------------------------------------------->
   <1.1.1.1, path-id2>     <1.1.1.1, path-id2>      <1.1.1.1, path-id2>
   <color2, label200>      <color2, label202>       <color2, label202>

     Label Exchange Tables:
     ASBR1:                          ASBR2:
     inLabel outLabel nextHop        inLabel outLabel nextHop
     201      200     SR-TE1         201      201     sub-if1
```

```
202       200     SR-TE2         202       202     sub-if2

PE2:
prefix    color  outLabel   nextHop
1.1.1.1    1       201       SR-TE1
1.1.1.1    2       202       SR-TE2
```

Figure 1: Details of Network Slicing Example

As we can see, PE1 advertises two paths <1.1.1.1, path-id1> and <1.1.1.1, path-id2> to ASBR1. PE1 advertises the binding between the prefix 1.1.1.1 and label 200. Because of the end node, both paths have the same label value 200.

When ASBR1 receives these two paths from PE1, ASBR1 automatically performs the next hop address to self, and allocate two new labels based on these two paths. As shown in Figure 1, ASBR1 generates a label 201 by replacing the label 200 with <1.1.1.1, color1>, and generates a label 202 by replacing the label 200 with <1.1.1.1, color2>. Both of them are the local label exchange entries.

Similarly, ASBR2 also generates two different labels based on the two paths. As shown in Figure 1, multiple end to end BGP-LU LSPs are established.

So, at the entry node of per domain, BGP-LU LSP generated for specific color is over intra-domain SR-TE or SR Best-effort path generated for the color again. At exit node of per domain, BGP-LU LSP generated for specific color selects inter-domain forwarding resource per color.

7. Deploy Considerations

All BGP routers (PE1--ASBR1, ASBR1---ASBR2, ASBR2---PE2) SHOULD be BGP-LU neighbors in advance. There may be multiple border routers to ensure multipath backup. All routers require the ADD-PATH Capability which is described in chapter 4 of [RFC7911].

This document not only supports interprovider VPNs while the customer sites belong to different ASs, but also supports the Carrier-of-

Carriers VPNs while the customer site belong to the same AS. Multiple operators are involved, so AS border routers may involve color mapping, color namespaces, or color service chains. These services can be delivered by the controller configurations or the local configurations.

8. Security Considerations

TBD

9. References

9.1. Normative References

[I-D.ietf-idr-tunnel-encaps]
Patel, K., Velde, G., and S. Ramachandra, "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-15 (work in progress), December 2019.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-06 (work in progress), December 2019.

[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>.

[RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", RFC 7911, DOI 10.17487/RFC7911, July 2016, <https://www.rfc-editor.org/info/rfc7911>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

9.2. Informative References

[I-D.peng-teas-network-slicing]
Peng, S., Chen, R., Mirsky, G., and F. Qin, "Packet Network Slicing using Segment Routing", draft-peng-teas-network-slicing-02 (work in progress), December 2019.

Authors' Addresses

Jin Zhou
ZTE Corp.

Email: zhou.jin6@zte.com.cn

Chunning Dai
ZTE Corp.

Email: dai.chunning@zte.com.cn

Shaofu Peng
ZTE Corp.

Email: peng.shaofu@zte.com.cn