IDR Working Group
Internet-Draft
Intended status: Informational
Expires: July 22, 2022

R. Chen
Ch. Dai
Sh. Peng
ZTE Corporation

January 18, 2022

Inter-domain Network Slicing via BGP-LU
draft-zhou-idr-inter-domain-lcu-03

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/multiple prefix to stitch multiple network segments.

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 July 22, 2022.

Table of Contents

1. Introduction

In the traditional end to end inter-domain network slicing, BGP-LU is used to build inter-domain MPLS LSP, and overlay service will be directly over BGP-LU LSP. For an E2E BGP-LU LSP, if overlay service has TE requirements that defined by a color, the BGP-LU LSP need also have a sense of color, i.e., BGP-LU label could be allocated per color.

[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, which contains two options:

- One is to define the multiple paths for the same destination prefix and advertise in BGP UPDATE message, and each UPDATE message can contain the color Extended Community [RFC9012] with different color value, which helps to select the underlying resources.This mode require additional path function defined in [RFC7911].

- The other is that multiple prefixs and multiple colors are configured on PE. One prefixs corresponds to one color. This mode does not require to additional path function.

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].

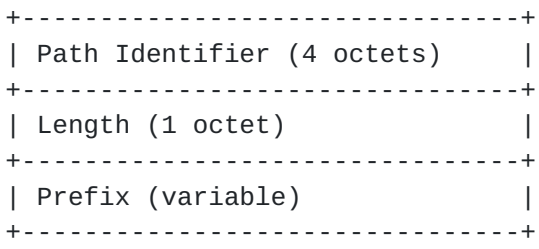
3. Color

[RFC9012] 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.

TO help routing decisions , each UPDATE may 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].



The Path Identifier only identifies a path, not carrying any particular semantics.In this document, it can be generated by the <Prefix,Color> tuple. The assignment to the Path Identifier for a path by a BGP speaker is purely a local matter.

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)

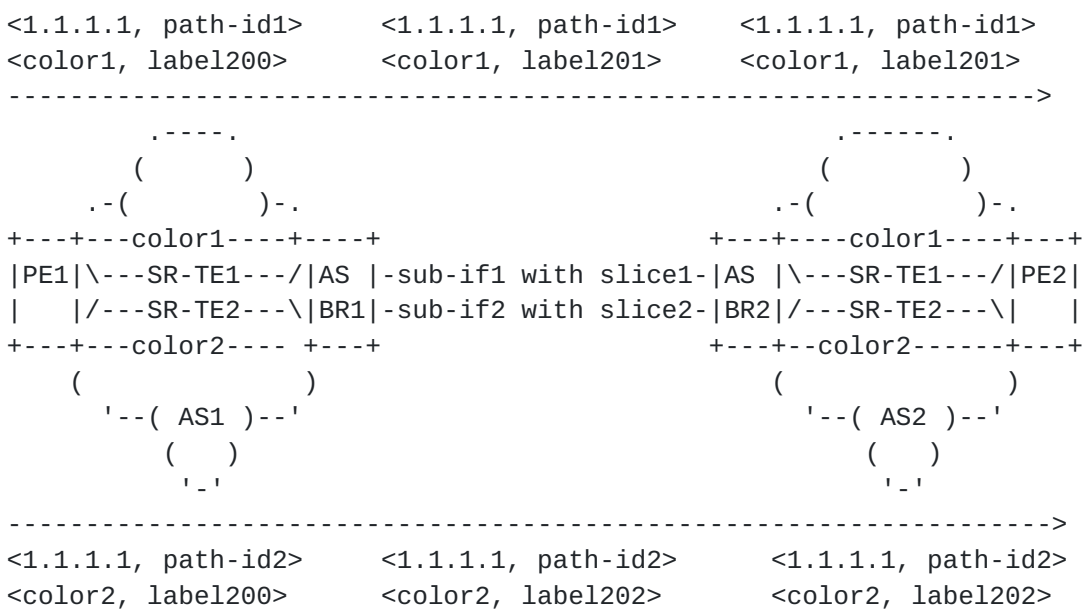
[RFC8277] describes how to use BGP to bind MPLS label(s) to the address prefixes. The specific format of the UPDATE message is detailed in Section 2 of [RFC8277].

[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. Inter-domain Network Slicing via BGP-LU

[RFC7911] defined that multiple paths for a particular address prefix by a Path identifier can be advertised in BGP UPDATE message, and each UPDATE message can contain the Color Extended Community [RFC9012] with different color value. That is a simple existing way to realize BGP-LU color function, and no protocol extension required.

Consider the following example of establishing multiple BGP-LU LSPs per different colors in a cross-domain scenario.



Label Exchange Tables:

ASBR1:

inLabel	outLabel	nextHop
201	200	SR-TE1
202	200	SR-TE2

ASBR2:

inLabel	outLabel	nextHop
201	201	sub-if1
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: Example of Inter-domain Network Slicing via BGP-LU

In figure 1, PE1 advertises two paths: <1.1.1.1, path-id1> carries the color1 attribute and <1.1.1.1, path-id2> carries the color2 attribute 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.

ASBR1 receives these two paths from PE1, and when sending to ASBR2, it modifies the next hop to itself. And allocate two new labels based on <prefix, path-id, color>. As shown in Figure 1, ASBR1 sends two paths to ASBR2, <1.1.1.1, path-id1> carries color1+label201, and <1.1.1.1, path-id2> carries color2+label202.

Similarly, ASBR2 also generates two different labels based on the <prefix, path-id, color>. As shown in Figure 1, multiple end to end BGP-LU LSPs are established.Different BGP-LU LSPs select the underlay SR-BE/TE tunnels according to their colors.

7. SRv6 support

Colored BGP-LU can be also used to setup end-to-end color-aware connectivity using Segment Routing over IPv6 (SRv6) [RFC8402].

As defined in [I-D.ietf-bess-srv6-services],to provide SRv6 service with underlay SRv6 policy connectivity, the egress PE signals the BGP overlay service route with SRv6 Service SID and color extended community . The ingress PE encapsulates the payload in an outer IPv6 header which contains the underlay SRv6 policy segment list and the overlay Service SID.

In addition, another solution is to provide SRv6 servcie with underlay SRv6 best-effort connectivity that is created by global IPv6 (AFI/SAFI 2/1) with color extended community. The underlay SRv6 SID is allocated based on <global IPv6, path-id, color>. The ingress PE encapsulates the payload in an outer IPv6 header which contains the underlay SRv6 SID and the Service SID.

8. 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.

9. Security Considerations

TBD

10. IANA Considerations

TBD.

11. Normative References

[I-D.ietf-bess-srv6-services]
Dawra, G., Filsfils, C., Talaulikar, K., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "SRv6 BGP based Overlay Services", draft-ietf-bess-srv6-services-08 (work in progress), November 2021.

[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-14 (work in progress), October 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>.

[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>.

[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>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

Authors' Addresses

Ran Chen
ZTE Corporation

Email: chen.ran@zte.com.cn

Chunning Dai
ZTE Corporation

Email: dai.chunning@zte.com.cn

Shaofu Peng
ZTE Corporation

Email: peng.shaofu@zte.com.cn