Color Operation with BGP Label Unicast
draft-chan-idr-bgp-lu2-00.txt

Abstract

This document specifies how to carry colored path advertisement via an enhancement to the existing protocol BGP Label Unicast. It would allow backward compatibility with RFC8277.

The targeted solution is to use stack of labels advertised via BGP Label Unicast 2.0 for end to end traffic steering across multiple IGP domains. The operation is similar to Segment Routing.

This proposed protocol will convey the necessary reachability information to the ingress PE node to construct an end to end path.

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 http://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 May 4, 2020.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

The proposed protocol is aimed to solve interdomain traffic steering, with different transport services in mind. One application is low latency service across multiple IGP domains, which could scale up to 100k routers network.

BGP is a flexible protocol. With additional of color attribute to BGP Label Unicast, a path with specific color would be given a meaning in application - a low latency path, a fully protected path, or a path for diversity.

The stack of labels would mean an end to end path across domains through each ABR or ASBR. Each ABR or ASBR will take one label from the stack, and hence pick the forwarding path to next ABR, ASBR, or the final destination.

And the label in the stack may be derived from any of the below

- Prefix SID
- Binding SID for RSVP LSP
- Binding SID for SR-TE LSP
- Local assigned label

The enhancement to the original RFC8277 is to add color extended community, with multiple advertisement allowed. The result is similar to multi-topology BGP-LU with different colors.

A new [BGP-CAP] should be required to enable such slicing.

On the other hand, to enable the service prefixes to be mapped according, the L3VPN, L2VPN, EVPN and prefix with BGP signaling, the color extended community is also added there. In the PE node, the service prefixes with color will be matched to a transport tunnel with the same color.
The following is an example:

PE1----ABR1-----ABR2-----PE2

PE1 will send the following labels with a color 100 path:

[2001 13001 801 16], where:
- 2001 - SR label to reach ABR1
- 13001 - Binding-SID label to reach ABR2. Underlying tunnel type is RSVP-TE
- 801 - Binding-SID label to reach PE2. Underlying tunnel type is SR-TE
- 16 - a VPN label

If PE1 wants to reach PE2 with another colored path, say color 200, the label stack could be different.

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 RFC 2119 [RFC2119].

   In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying significance described in RFC 2119.

3. Carrying Label Mapping Information with Color and Label Stack

3.1. Color extended community for BGP Labeled Unicast

   The addition of Color Extended Community is an opaque extended community from RFC4360 and RFC5512. The draft allows multiple color values advertisement.

   ![Figure 1: Color value advertisement format](image)
Both in BGP update and MP_UNREACH_NLRI message, multiple color extended communities could be included. It means that multiple colors, indicating different kind of services, could share the same label stack.

If only one color extended community is specified, only prefix with that color value is updated or withdrawn.

If a MP_UNREACH_NLRI message without any color specified is received for a given prefix, that prefix with any color should not be affected.

If color extended community is not present in a BGP update message, it would be treated as normal BGP-LU without any color.

3 bits of XXX is reserved here for the draft.
Color value 0 is reserved for future interoperability purpose.

3.2. Color extended community for service prefixes

The same format of color extended community is advertised with service prefixes. The order of the color extended community could be interpreted as
- Order of primary and fallback colors
- Or, ECMP of equal split between color tunnels

The above would be interpreted by the receiving PE upon its local configuration.

4. Uniqueness of path entries

Use of color can be considered to slice into multiple BGP Label Unicast RIB. Therefore, it should be treated as unique entries for the <color, prefix>.

e.g. <color, prefix>, [labels]
<1, 10.1.1.1/32>, [100 200]
<2, 10.1.1.1/32>, [100 200]
<null, 10.1.1.1/32>, [100 200]

All these 3 NLRI are considered different but valid entries for different color instances.

5. AIGP consideration

AIGP (RFC7311) would be also used in here to embed certain metric across.

6. Explicit Withdraw of a <color, prefix>

According to RFC8277, MP_UNREACH_NLRI can be used to remove binding of a <color, prefix>.
Compatibility is set to 0xC00000 to specify the use of color. Multiple color extended communities could be applied here.

```
+--------+-----------------------------+
<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+--------+-----------------------------+
```

Figure 2: NLRI for Withdrawal

7. Error Handling Procedure

If BGP receiver could not handle the NLRI, it should silently discard with error logging.

8. Security Considerations

9. IANA Considerations

TBD. It will require a new BGP capability code to enable such color operation.

New SAFI might be required as well.

10. References

10.1. Normative References


10.2. Informative References


11. Acknowledgments

The following people have contributed to this document:

Jeff Haas, Juniper Networks

Shraddha Hedge, Juniper Networks

Santosh Kolenchery, Juniper Networks

Shihari Sangli, Juniper Networks

Krzysztof Szarkowicz, Juniper Networks

Yimin Shen, Juniper Networks
Authors Addresses

Louis Chan (editor)
Juniper Networks
2604, Cityplaza One, 1111 King’s Road
Taikoo Shing
Hong Kong

Phone: +8522587659
Email: louisc@juniper.net