IDR Working Group Loui
INTERNET-DRAFT
Intended status: Experimental Juniper Ne
Expires: Oct 15, 2020 Apr 16

Color Operation with BGP Label Unicast
draft-chan-idr-bgp-lu2-01.txt

Abstract

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

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

This proposed protocol will convey the necessary reachability information
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
and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Fo
(IETF). Note that other groups may also distribute working documents as
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
updated, replaced, or obsoleted by other documents at any time. It is
inappropriate to use Internet-Drafts as reference material or to cite the
than as "work in progress."

This Internet-Draft will expire on Oct 15, 2020.

Table of Contents

1. Introduction

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

BGP is a flexible protocol. With additional of color attribute to BGP Lab
Unicast, a path with specific color would be given a meaning in applicati
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
or ASBR. Each ABR or ASBR will take one label from the stack, and hence p
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 communit
multiple advertisement allowed. The result is similar to multi-topology B
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 accordingl
L3VPN, L2VPN, EVPN and prefix with BGP signaling, the color extended comm

---

also added there. In the PE node, the service prefixes with color will be
to a transport tunnel with the same color.

The following is an example. Between PE1 and PE2, there is a VPN service
with label 16, which is associated with color 100.

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

PE1 will send the following labels with a color 100 path plus VPN label

[2001 13001 801 16], where

2001 - SR label to reach ABR1

13001 - a Binding-SID label for ABR1-ABR2 tunnel. Underlying tunnel type

801 - a Binding-SID label for ABR2-PE2 tunnel. Underlying tunnel type is

16 - a VPN label, which is signaled via other means

[2001 13001 801] denotes the label stack for this color 100 path to reac

The document here is going to describe how PE1 gains enough information t
this label stack across routing domains.

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

At the same time, this architecture is also controller friendly, since al
notation is Segment Routing compatible, like use of Binding-SID.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOU
"SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are t
interpreted as described in RFC 2119 [RFC2119].

In this document, these words will appear with that interpretation only
ALL CAPS. Lower case uses of these words are not to be interpreted as
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
RFC4360 and RFC5512. The draft allows multiple color values advertisement

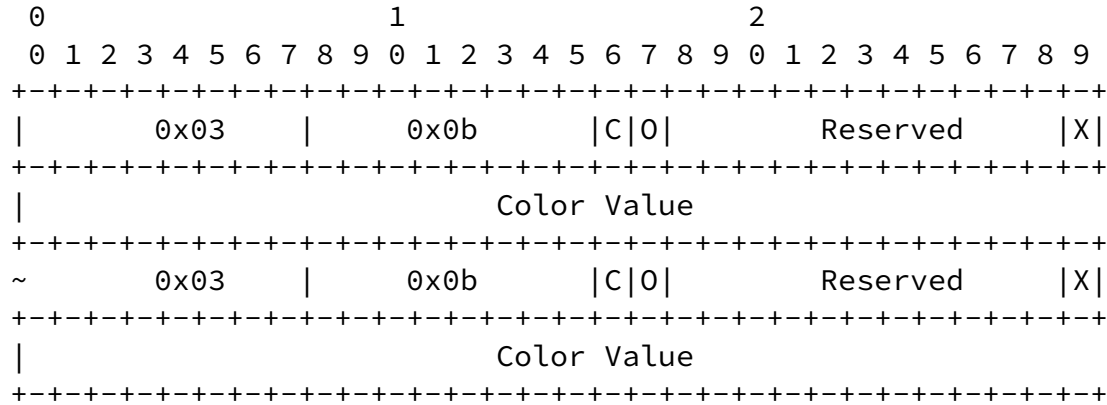

Figure 1: Color value advertisement format

Both in BGP update and MP_UNREACH_NLRI message, multiple color extended c
could be included. It means that multiple colors, indicating different ki
services, could share the same label stack.

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

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

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

3 bits of XXX is reserved here for the draft.

The meaning for XXX is interpreted as sub-slice of color, with 0 to 7 in or 000b and 111b in binary. These sub-slice could be used in either of th following case.

a) Primary path and fallback paths in order of preference
  0 primary path
  1 first and most preferred backup path

  7 least preferred backup path

b) ECMP paths up to 8, since all paths should be active in forwarding pla

Color value 0 is reserved for future interoperability purpose.

Color value 1 - 31 are not recommended to use, and this range is reserved future use.

3.2. Color extended community for service prefixes

The same format of color extended community is advertised with service pr
The order of the color extended community could be interpreted as

- Order of primary and fallback colors
- Or, ECMP of equal split between color paths

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

It is optional to enable sub-slice notation.

But if sub-slice bits are used, it will be used to map directly to each o slice path. If sub-slice path is not available for mapping, it should jus to resolving by color.

4. Uniqueness of path entries

a) Use of color can be considered to slice into multiple BGP Label Unicas
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
instances.

b) With sub-slice notation
<color-sub, prefix>, [labels]

<1-0, 10.1.1.1/32>, [100 200]

<1-1, 10.1.1.1/32>, [101 300]

<1-7, 10.1.1.1/32>, [102 400]

These 3 NLRI are distinct, and the second and third NLRI could be used
backup or ECMP purpose.

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
prefix>.

Compatibility is set to 0xC00000 to specify the use of color. Multiple co
extended communities could be applied here.

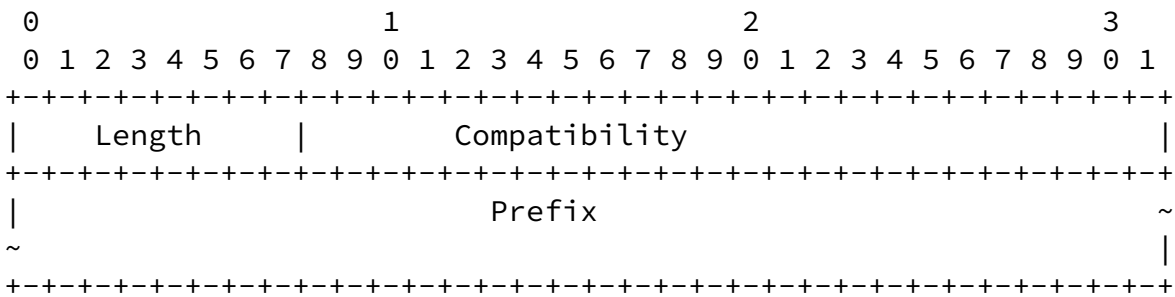

Figure 2: NLRI for Withdrawal

7. Error Handling Procedure

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

8. Controller Compatibility

The proposed architecture is compatible with controller for end to end
provisioning. Persistent label, like Binding-SIS is recommended to be use
controller could learn these labels from the network, and program specifi
end path.

Controller could also be deployed based on domain by domain perspective.
Optimizing latency of a RSVP LSP, or maintain the bandwidth and loading b
TE LSPs.

9. Security Considerations

10. IANA Considerations

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

New SAFI might be required as well.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement
BCP 14, RFC 2119, March 1997.

11.2. Informative References

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4"
3107, DOI 10.17487/RFC3107, May 2001,

<https://www.rfc-editor.org/info/rfc3107>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended
Communities Attribute", RFC 4360, February 2006

<https://www.rfc-editor.org/info/rfc4360>.

[RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent A
Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribu
5512, April 2009.

<https://www.rfc-editor.org/info/rfc5512>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and
McPherson, "Dissemination of Flow Specification Rules", RFC 5575,
10.17487/RFC5575, August 2009,

<http://www.rfc-editor.org/info/rfc5575>.

[RFC7311] Mohapatra, P., Fernando, R., Rosen, E., and J. Uttaro,
"The Accumulated IGP Metric Attribute for BGP", RFC 7311,
DOI 10.17487/RFC7311, August 2014,

<https://www.rfc-editor.org/info/rfc7311>.

[RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisemen
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",
DOI 10.17487/RFC8277, October 2017,

<https://www.rfc-editor.org/info/rfc8277>.

[BGP-CAP] Chandra, R. and J. Scudder, "Capabilities Advertisement

with BGP-4", RFC 2842, May 2000.

12. 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: +85225876659
Email: louisc@juniper.net