

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: August 6, 2020

Yimin Shen
Juniper Networks
Ravi Singh
Individual Contributor
Yuji Kamite
NTT Communications
February 3, 2020

BGP Flexible Color-Based Tunnel Selection
draft-shen-idr-flexible-color-tunnel-selection-01

Abstract

This document discusses color-based tunnel selection for BGP payload prefixes. It defines a set of extended mapping modes, and describes how to use them to construct tunnel selection schemes for flexible tunnel selection. Tunnel selection schemes can be implemented as policies on routers performing tunnel selection, or signaled by next hop routers or a central controller by using the BGP extensions specified in this document.

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 August 6, 2020.

Copyright Notice

Copyright (c) 2020 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 (<https://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.

Table of Contents

| | | |
|------------------------|--|--------------------|
| 1. | Introduction | 2 |
| 2. | Specification of Requirements | 4 |
| 3. | Extended Mapping Modes | 4 |
| 4. | Tunnel Selection Scheme and Operation | 6 |
| 5. | Provisioning of Tunnel Selection Schemes | 7 |
| 6. | BGP Tunnel Encapsulation Attribute Extensions | 8 |
| 6.1. | Wildcard Tunnel Type | 8 |
| 6.2. | Color Tunnel Selection Scheme Sub-TLV | 8 |
| 6.2.1. | Extended Mapping Mode Sub-sub-TLV | 9 |
| 6.2.2. | Encoding | 10 |
| 6.2.3. | Decoding | 10 |
| 6.3. | Association between Color Tunnel Selection Scheme Sub-TLV and Tunnel Type | 11 |
| 7. | Relationship with Color-Only Bits of Color Extended Community | 11 |
| 8. | IANA Considerations | 11 |
| 9. | Security Considerations | 12 |
| 10. | Acknowledgements | 12 |
| 11. | References | 12 |
| 11.1. | Normative References | 12 |
| 11.2. | Informative References | 13 |
| | Authors' Addresses | 13 |

[1.](#) Introduction

In an overlay network using BGP for payload prefix distribution, transporting the packets of a payload prefix from a BGP router to the next BGP router relies on the selection of a transport tunnel. This selection may be based on various attributes of the prefix, such as BGP next hop, color, and information in the Tunnel Encapsulation Attribute [[BGP-TUNNEL-ENCAP](#)], etc.

In one tunnel selection model, color is used as a primary criterion along with BGP next hop or tunnel endpoint address (contained in a Tunnel Endpoint sub-TLV of the Tunnel Encapsulation Attribute). This model is referred to as "color-based" tunnel selection in this document. The model is gaining many use cases today due to its general applicability. In particular, color as a generic notion may be used to represent a broad range of network attributes, such as virtual topology, network slice, path computation algorithm, TE constraint, administrative profile, etc. For some of the attributes,

there may not be a convenient mechanism to associate them with payload prefixes or tunnels, or to distribute them in a network, especially across domains or to a central controller. When these attributes need to be considered in tunnel selection, mapping them to colors and performing color-based tunnel selection will provide a generic solution.

The procedures in this document is relevant to color-based tunnel selection. In general, payload prefixes may be associated with colors through configuration or a Color Extended Community [[RFC5512](#)]. Transport tunnels may also be associated with colors through configuration (e.g. RSVP and LDP tunnels), a Color Extended Community (e.g. BGP LU), or a color embedded in BGP NLRI (e.g. BGP SR-TE policy [[BGP-SR-POLICY](#)]), etc. These payload prefixes and tunnels are called "colored payload prefixes" and "colored tunnels", respectively.

Normally, a payload prefix of color X is intended to be mapped to a tunnel of the same color X. This is considered as the default mapping mode of color-based tunnel selection. In some cases, when a tunnel of color X cannot be found or established, or when a previously mapped tunnel of color X fails, a network operator may want to map the payload prefix by attempting other modes, e.g. selecting a tunnel of another color Y, a tunnel without a color, a tunnel of color X but with an IPv4-mapped IPv6 endpoint address, and so on. Note that the colors may represent network slices, virtual topologies, path computation algorithms, etc. Hence, these modes will provide the flexibility and enable the operator to take the full transport capability of the network. In this document, these modes are called "extended mapping modes", and the procedure of automatically executing them in a user-defined order is called "fallback".

This document defines a set of extended mapping modes to complement the default mapping mode. It introduces the notion of "tunnel selection scheme". A tunnel selection scheme is an ordered list of extended mapping modes to be automatically executed in tunnel selection. When a tunnel is not selected by using the first mode in the list, fallback is performed by attempting the second mode, the third mode, and so on, until a tunnel is selected or the list is exhausted.

Color-based tunnel selection for uncolored payload prefixes is also considered in this document as a special case. By using a tunnel selection scheme, a colored or uncolored tunnel may be selected for an uncolored payload prefix in a flexible manner.

2. Specification of Requirements

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](#)] and [[RFC8174](#)].

3. Extended Mapping Modes

This document defines a set of extended mapping modes for flexible color-based tunnel selection. Each mode specifies how a payload prefix's endpoint IP address (derived from BGP next hop or a Tunnel Endpoint sub-TLV in the Tunnel Encapsulation Attribute [[BGP-TUNNEL-ENCAP](#)]) and color are used to select a tunnel. The document assumes that each payload prefix SHOULD have a single color for tunnel selection purpose or no color, and each tunnel SHOULD have a single color or no color.

In the definitions of the extended mapping modes below, N represents a payload prefix's endpoint IP address, and C represents its color. An uncolored payload prefix does not have a color. An extended mapping mode may have multiple steps for sub-level fallback within it. The steps are executed sequentially. The mode is completed as soon as a tunnel is successfully selected in one of the steps, and the rest steps are skipped.

(1) IP-color, optionally with a fallback color list of {C1, ..., Cn}

- If the payload prefix has a color C, select a tunnel with endpoint address N and color C.
- Select a tunnel with endpoint address N and color C1.
- ...
- Select a tunnel with endpoint address N and color Cn.

(2) Color-only, optionally with a fallback color list of {C1, ..., Cn}

- If the payload prefix has a color C, select a tunnel with color C, regardless of the tunnel's endpoint address.
- Select a tunnel with color C1, regardless of tunnel's endpoint address.
- ...

- Select a tunnel with color Cn, regardless of tunnel's endpoint address.

(3) IP-any-color

- Select a tunnel with endpoint address N and any color.

(4) IP-only

- Select a tunnel with endpoint address N and without a color.

(5) Converted-IPv6

This mode is applicable when N is an IPv4 address. Assume N' is the IPv6 address mapped from N.

- Select a tunnel with endpoint address N' and without a color.

(6) Converted-IPv6-color, optionally a fallback color list of {C1, ..., Cn}

This mode is applicable when N is an IPv4 address. Assume N' is the IPv6 address mapped from N.

- If the payload prefix has a color C, select a tunnel with endpoint address N' and color C.
- Select a tunnel with endpoint address N' and color C1.
- ...
- Select a tunnel with endpoint address N' and color Cn.

(7) Converted-IPv6-any-color

This mode is applicable when N is an IPv4 address. Assume N' is the IPv6 address mapped from N.

- Select a tunnel with endpoint address N' and any color.

(8) Color-profile

- If the payload prefix has a color C, use C as key to look up a profile to construct tunnel selection criteria and select a tunnel.

As shown above, the IP-color, Color-only, and Converted-IPv6-color modes may have a fallback color list for sub-level fallback across the colors.

This list is not exhaustive. More modes MAY be defined in the future.

4. Tunnel Selection Scheme and Operation

A tunnel selection scheme is defined as an ordered list of extended mapping modes ([Section 3](#)) to be automatically executed in tunnel selection. The first mode in the list is called a "primary" mode, and all the subsequent modes are called "fallback" modes. A scheme MUST have a primary mode, but MAY or MAY not have any fallback mode.

When a scheme is executed for a payload prefix, the modes in the list are executed sequentially, and within each mode, the steps of sub-level fallback are executed sequentially. When a tunnel is selected in a particular step in a particular mode, the scheme is completed, and all subsequent steps of the mode and all the subsequent modes in the list are skipped. If no tunnel is selected when the list is exhausted, the payload prefix will remain in unresolved state for transport.

In the case where a previously selected tunnel becomes inoperative, the scheme SHOULD be run to reselect a tunnel. In the case where a tunnel was previously selected and later another tunnel of higher preference (in the tunnel selection scheme or in a fallback color list) becomes available, the new tunnel MAY be selected to replace the current tunnel. This procedure is called a reversion. It may be performed manually by a network operator, or triggered automatically by the situation.

The following are some examples of tunnel selection schemes.

Example 1:

A payload prefix has a tunnel endpoint IPv4 address 203.0.113.1 and a color RED. It is associated with the following tunnel selection scheme:

- (1) IP-color
- (2) Converted-IPv6-color
- (3) IP-only

The intended tunnel selection procedure is:

(1) Find a tunnel with endpoint IPv4 address 203.0.113.1 and color RED.

(2) If the above is unsuccessful, convert the IPv4 address to an IPv6 address 2002:cb00:7101::/64. Find a tunnel with endpoint IPv6 address 2002:cb00:7101::/64 and color RED.

(3) If the above is unsuccessful, find a tunnel with endpoint IPv4 address 203.0.113.1 and without a color.

Example 2:

A prefix has a tunnel endpoint IPv4 address 203.0.113.1 and a color RED. It is associated with the following tunnel selection scheme:

(1) IP-color, with a fallback color list = {BLUE, GREEN}

(2) Converted-IPv6-color, with a fallback color list = {WHITE}

(3) IP-only

The intended tunnel selection procedure is:

(1) Find a tunnel with endpoint IPv4 address 203.0.113.1 and color RED. If it is unsuccessful, find a tunnel with endpoint IPv4 address 203.0.113.1 and color BLUE. If it is unsuccessful, find a tunnel with endpoint IPv4 address 203.0.113.1 and color GREEN.

(2) If the above is unsuccessful, convert the IPv4 address to an IPv6 address 2002:cb00:7101::/64. Find a tunnel with endpoint IPv6 address 2002:cb00:7101::/64 and color RED. If it is unsuccessful, find a tunnel with endpoint IPv6 address 2002:cb00:7101::/64 and color WHITE.

(3) If the above is unsuccessful, find a tunnel with endpoint IPv4 address 203.0.113.1 and without a color.

5. Provisioning of Tunnel Selection Schemes

A tunnel selection scheme MAY be provisioned for a payload prefix on a router which performs tunnel selection. In this case, the scheme may be implemented as a policy on the router. The configuration of such policy varies by vendors, and hence is out of the scope of this document.

A tunnel selection scheme MAY also be provisioned on a router or a central controller which originates the UPDATE message of a payload prefix, and then distributed to a router(s) which will perform tunnel

selection. To facilitate this, the document introduces a new "Color Tunnel Selection Scheme" sub-TLV ([Section 6](#)) to the Tunnel Encapsulation Attribute to carry the information. As color-based tunnel selection is typically across all tunnel types, the document also introduces a new "Wildcard" tunnel type to the Tunnel Encapsulation Attribute. When the tunnel selection scheme contained in a Color Tunnel Selection Scheme sub-TLV is applicable to all tunnel types, the top-level Tunnel Encapsulation Attribute TLV SHOULD set tunnel type to Wildcard.

In the case where a payload prefix has one scheme configured as a policy on a router, and another scheme received in a Color Tunnel Selection Scheme sub-TLV, the router SHOULD treat the policy in preference to the received information.

If a payload prefix does not have a tunnel selection scheme, the default mapping mode applicable to colored or non-colored payload prefixes SHOULD be used accordingly.

6. BGP Tunnel Encapsulation Attribute Extensions

This section specifies a new "Wildcard" tunnel type and a new Color Tunnel Selection Scheme sub-TLV for the Tunnel Encapsulation Attribute.

[6.1.](#) Wildcard Tunnel Type

The Wildcard tunnel type has the semantics of "any tunnel types". It allows a Tunnel Encapsulation Attribute TLV to carry information which is regardless of tunnel type or applicable to all tunnel types. In this document, it is used when a Tunnel Encapsulation Attribute TLV contains a Color Tunnel Selection Scheme sub-TLV which is applicable to all tunnel types.

[6.2.](#) Color Tunnel Selection Scheme Sub-TLV

The Color Tunnel Selection Scheme sub-TLV is used to carry the information of a tunnel selection scheme. The sub-TLV is contained in a Tunnel Encapsulation Attribute TLV. If the Tunnel Encapsulation Attribute TLV's tunnel type is Wildcard, the tunnel selection scheme is regardless of tunnel type. If the Tunnel Encapsulation Attribute TLV's tunnel type is a specific type, the tunnel selection scheme is applicable to tunnels of that type. In any case, a Tunnel Encapsulation Attribute TLV MUST not contain more than one Color Tunnel Selection Scheme sub-TLV.

The sub-TLV's Type is TBD (to be allocated by IANA). The sub-TLV's Length is the number of the octets of the sub-TLV's Value field. The

sub-TLV's Value field is composed of one or multiple Extended Mapping Mode sub-sub-TLVs.

6.2.1. Extended Mapping Mode Sub-sub-TLV

An Extended Mapping Mode sub-sub-TLV contains the information of an extended mapping mode. Its encoding is shown in Figure 1.

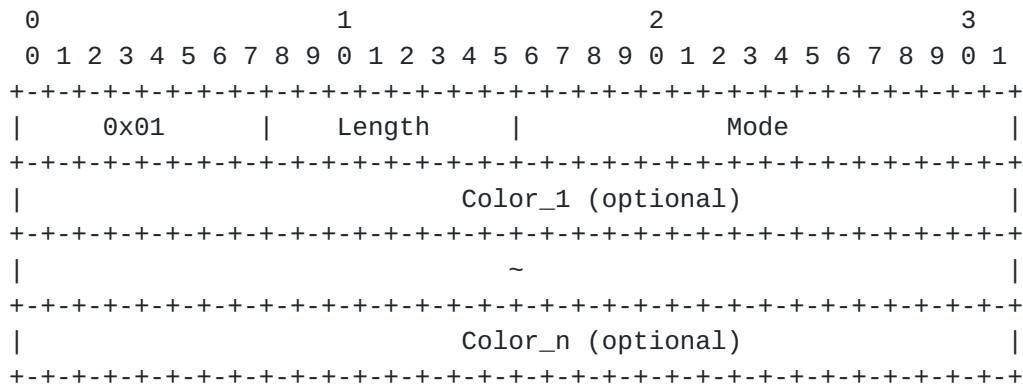


Figure 1

The Extended Mapping Mode sub-sub-TLV's Type is 0x01.

The Extended Mapping Mode sub-sub-TLV's Length is the total number of octets of the sub-sub-TLV's Value field.

The Extended Mapping Mode sub-sub-TLV's Value field contains a 2-octet extended mapping mode defined as below, and an optional fallback color list.

- 1 - IP-color
- 2 - Color-only
- 3 - IP-any-color
- 4 - IP-only
- 5 - Converted-IPv6
- 6 - Converted-IPv6-color
- 7 - Converted-IPv6-any-color
- 8 - Color-profile

The IP-color, Color-only and Converted-IPv6-color modes MAY have an optional fallback color list. The list contains one or multiple 4-octet color values, i.e. Color_1, ..., Color_n, in the order from the highest preference to the lowest preference.

6.2.2. Encoding

Given a tunnel selection scheme, a Color Tunnel Selection Scheme sub-TLV is constructed in the following manner:

- o First, an Extended Mapping Mode sub-sub-TLV containing the primary mode is added. If this mode is IP-Color, Color-Only, or Converted-IPv6-Color, and if cross-color fallback is applicable to this mode, a fallback color list is added to the sub-sub-TLV.
- o If there is one or multiple desired fallback modes, an Extended Mapping Mode sub-sub-TLV containing the first fallback mode is added. If this mode is IP-Color, Color-Only, or Converted-IPv6-Color, and if cross-color fallback is applicable to this mode, a fallback color list is added to the sub-sub-TLV.
- o This process continues, until an Extended Mapping Mode sub-sub-TLV containing the last fallback mode is added. If this mode is IP-Color, Color-Only, or Converted-IPv6-Color, and if cross-color fallback is applicable to this mode, a fallback color list is added to the sub-sub-TLV.

6.2.3. Decoding

When decoding a Color Tunnel Selection Scheme sub-TLV, a receiving router MUST interpret the preference of the contained Extended Mapping Mode sub-sub-TLVs as the order in which they are encoded. If an Extended Mapping Mode sub-sub-TLV contains a mode which is not IP-Color, Color-Only, or Converted-IPv6-Color but has a fallback color list, the entire Color Tunnel Selection Scheme sub-TLV SHOULD be considered as malformed and ignored.

A receiving router MUST consider a payload prefix as having a modified tunnel selection scheme in any of the following situations, and perform tunnel selection accordingly:

- o The payload prefix did not have a valid Color Tunnel Selection Scheme sub-TLV in the previous UPDATE message, and it has one in the latest UPDATE message. Tunnel selection MUST be performed based on the latest tunnel selection scheme.
- o The payload prefix had a valid Color Tunnel Selection Scheme sub-TLV in the previous UPDATE message, but it does not have one in

the latest UPDATE message. Tunnel selection MUST revert to the default color or non-color mapping mode.

- o The payload prefix had a valid Color Tunnel Selection Scheme sub-TLV in the previous UPDATE message, and it has one with different content in the latest UPDATE message. Tunnel selection MUST be performed based on the latest tunnel selection scheme.

6.3. Association between Color Tunnel Selection Scheme Sub-TLV and Tunnel Type

A Color Tunnel Selection Scheme sub-TLV MAY be contained in a Tunnel Encapsulation Attribute TLV of Wildcard tunnel type, indicating that the scheme SHOULD be performed regardless of tunnel type. The sub-TLV MAY also be contained in a Tunnel Encapsulation Attribute TLV of a specific tunnel type, indicating that the scheme SHOULD consider only the tunnels of that type. In the case where a Tunnel Encapsulation Attribute contains a TLV of Wildcard tunnel type and another TLV of a specific tunnel type, and both TLVs contain a Color Tunnel Selection Scheme sub-TLV, tunnel selection for that specific tunnel type SHOULD be based on the corresponding Color Tunnel Selection Scheme sub-TLV.

7. Relationship with Color-Only Bits of Color Extended Community

[RFC8402] and [BGP-SR-POLICY] define two "Color-Only" bits (i.e. CO bits) in the BGP Color Extended Community for color-based tunnel selection in the context of segment routing. Each of the four combinations of the CO bits corresponds to a predefined fallback scheme.

This document complements these documents by supporting more generic and flexible fallback schemes which are user definable. In fact, the predefined fallback schemes of the CO bits can be fully supported by using the Color Tunnel Selection Scheme sub-TLV. When a router advertises an UPDATE message, it SHOULD NOT use a Color Extended Community with CO bits and a Color Tunnel Selection Scheme sub-TLV at the same time, in order to avoid collision between them. If a router receives an UPDATE message containing both objects, it SHOULD give preference to CO bits, and ignore the other. If the tunnel selection scheme is implemented as a policy on the receiving router, the router SHOULD give the preference to the policy.

8. IANA Considerations

This document requires the IANA to allocate a value for the Wildcard tunnel type as a new BGP Tunnel Encapsulation Attribute Type, and a type value for the new Color Tunnel Selection Scheme sub-TLV.

9. Security Considerations

This document introduces procedures for color-based tunnel selection to use tunnel selection schemes. The procedures can cause traffic to be diverted from default path(s). This requires measures to be taken at a number of levels to avoid undesirable transport behaviors and security risks. First, network coloring (i.e. color assignment to network resources, attributes, payload prefixes, tunnels, etc.) MUST be carefully planned and validated at a global level to avoid errors and collisions. Second, tunnel selection schemes MUST be legitimate and always select valid tunnels leading to desired endpoints. For schemes implemented as policies, this SHOULD be ensured by policy configuration. For schemes distributed via Color Tunnel Selection Scheme sub-TLV of BGP Tunnel Encapsulation Attribute, receivers SHOULD use BGP security procedures to validate each originator's identity and detect unauthorized content modification during distribution.

10. Acknowledgements

This document leverages work done by Junan Chen. Thanks to Jeff Hass and Srihari Sangli for their kind reviews and comments which helped to improve the clarity of this document.

11. References

11.1. Normative References

- [RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", [RFC 5512](https://www.rfc-editor.org/info/rfc5512), DOI 10.17487/RFC5512, April 2009, <<https://www.rfc-editor.org/info/rfc5512>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [RFC 8402](https://www.rfc-editor.org/info/rfc8402), DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [BGP-SR-POLICY] Previdi, S., Filsfils, C., Mattes, P., Rosen, E., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", [draft-previdi-idr-segment-routing-te-policy](#) (work in progress), 2019.

[BGP-TUNNEL-ENCAP]

Patel, K., Velde, G., and S. Sangli, "The BGP Tunnel Encapsulation Attribute", [draft-vandeveld-idr-remote-next-hop](#) (work in progress), 2019.

11.2. Informative 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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Authors' Addresses

Yimin Shen
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Phone: +1 9785890722
Email: yshen@juniper.net

Ravi Singh
Individual Contributor

Email: ravi.singh.ietf@gmail.com

Yuji Kamite
NTT Communications

Email: y.kamite@ntt.com

