Workgroup: Interdomain Routing Working Group
Internet-Draft: draft-wang-idr-cpr-03
Published: 20 October 2023
Intended Status: Informational
Expires: 22 April 2024
Authors: H. Wang
Huawei Technologies
J. Dong
Huawei Technologies
K. Talaulikar
Cisco Systems
T. Han
Huawei Technologies
R. Chen
ZTE Corporation
J. Xie
Huawei Technologies
X. Chen
Huawei Technologies

# BGP Colored Prefix Routing (CPR) for SRv6 based Services

## Abstract

This document describes a mechanism to advertise IPv6 prefixes in BGP which are associated with Color Extended Communities to establish end-to-end intent-aware paths for SRv6 services. Such IPv6 prefixes are called "Colored Prefixes", and this mechanism is called Colored Prefix Routing (CPR). In SRv6 networks, the Colored prefixes are the SRv6 locators associated with different intent. SRv6 services (e.g. SRv6 VPN services) with specific intent could be assigned with SRv6 SIDs under the corresponding SRv6 locators, which are advertised as Colored prefixes. This allows the SRv6 service traffic to be steered into end-to-end intent-aware paths simply based on the longest prefix matching of SRv6 Service SIDs to the Colored prefixes. In data plane, dedicated transport label or SID for the inter-domain path is not needed, thus the encapsulation efficiency could be optimized. The existing IPv6 Address Family could be used for the advertisement of IPv6 Colored prefixes, thus this mechanism is easy to interoperate and can be deployed incrementally in multi-domain networks.

## Status of This Memo

**Table of Contents**



## 1. Introduction

With the trend of using one common network to carry multiple types of services, each service type can have different requirements on the network. Such requirements are usually considered as the "intent" of the service or customer, and is represented as an abstract notion called "color".

In network scenarios where the services are delivered across multiple network domains, there is need to provide the services with different end-to-end paths to meet the intent. [I-D.hr-spring-intentaware-routing-using-color] describes the problem statements and requirements for inter-domain intent-aware routing.

The inter-domain path can be established using either MPLS or IP data plane. In MPLS based networks, the traditional inter-domain approach is to establish an end-to-end LSP based on the BGP-LU mechanisms as defined in [RFC8277]. Each domain or area border node needs to perform label swapping for the end-to-end BGP-LU LSP, and encapsulate the label stack which are used for the intra-domain LSP within the subsequent network domain or area.

While in IP based networks, the IP reachability information can be advertised to network nodes in different domains using BGP, so that all the domain or area border nodes have the routes to the IP prefixes of the destination node in other domains. With the introduction of SRv6 [RFC8402] [RFC8754] [RFC8986], BGP services are assigned with SRv6 Service SIDs [RFC9252], which are routable in the network according to its SRv6 locator prefix. Thus, the inter-domain path can be established simply based on the inter-domain prefix routes, and the BGP-LU inter-domain LSP mechanism is not necessary for IPv6 and SRv6 based networks.

This document describes a mechanism to advertise IPv6 prefixes which are associated with Color Extended Community to establish end-to-end intent-aware paths for SRv6 services. The color value in the Color Extended Community indicate the intent [RFC9256]. Such IPv6 prefixes are called "Colored Prefixes", and this mechanism is called Colored Prefix Routing (CPR). In SRv6 networks, the Colored prefixes are the SRv6 locators associated with different intent. BGP services over SRv6 (e.g. SRv6 VPN services) [RFC9252] with specific intent could be assigned with SRv6 SIDs under the corresponding SRv6 locators, which are advertised as Colored prefixes. This allows the SRv6 service traffic to be steered (as specified in [RFC9252]) into end-to-end intent-aware paths simply based on the longest prefix matching of SRv6 Service SIDs to the Colored prefixes. In data plane, the dedicated transport label or SID for the inter-domain path is not needed, thus the encapsulation efficiency could be optimized. The existing IPv6 Address Family could be used for the advertisement of IPv6 Colored prefixes, which makes this mechanism easy to interoperate and can be deployed incrementally in multi-domain networks.

## 2. BGP CPR

This section describes the BGP CPR mechanisms. More specifically, section 2.1 describes the allocation of the IPv6 Colored prefixes, section 2.2 describes the advertisement of Colored prefixes in BGP, section 2.3 describes the resolution of CPR routes to the intra-domain paths, and section 2.4 describes the steering of BGP SRv6 services to CPR routes.

### 2.1. Colored Prefix Allocation

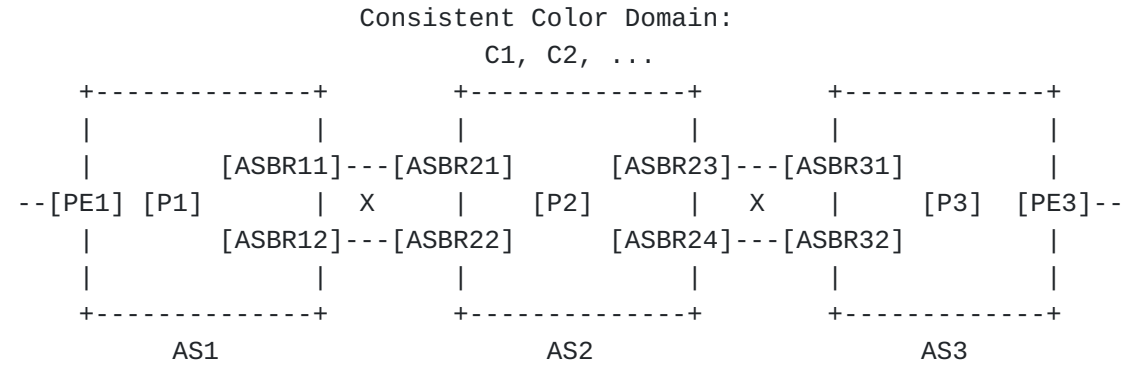
In SRv6 networks, an SRv6 locator needs to be allocated for each node. In order to distinguish N different intent, a PE node needs to be allocated with N SRv6 locators, each of which is associated a different intent that is identified by a color value. One way to achieve this is by splitting the base SRv6 locator of the node into N sub-locators, and these sub-locators are Colored prefixes associated with different intents.

For example, node PE2 is allocated with the base SRv6 Locator 2001:db8:aaaa:1::/64. In order to provide 16 different intents, this base SRv6 Locator is split into 16 sub-locators from 2001:db8:aaaa:1:0000::/68 to 2001:db8:aaaa:1:F000::/68, each of these sub-locators is associated with a different intent, such as low-delay, high-bandwidth, etc.

### 2.2. Colored Prefix Advertisement

After the allocation of Colored prefixes on a PE node, routes to these Colored prefixes need to be advertised both in the local domain and also to other domains using BGP, so that the BGP SRv6 services routes could be resolved using the corresponding CPR route.

In a multi-domain IPv6 network, the IPv6 unicast Address Family/ Subsequent Address Family (AFI/SAFI = 2/1) [RFC2545] is used for the advertisement of the Colored prefix routes. The color extended community [RFC9012] is carried with the Colored prefix route with the color value indicating the intent [RFC9256]. The procedure of Colored prefix advertisement is described using an example with the following topology:




Colored Prefixes of PE3:
Low delay: 2001:db8:aaaa:1:1000::/68
high bandwidth: 2001:db8:aaaa:1:2000::/68
...

Figure 1. Example Topology for CPR Route Illustration

Assume PE3 is provisioned with two different Colored prefixes CLP-1 and CLP-2 for two different intent such as "low-delay" and "high-bandwidth" respectively. In this example, It is assumed that the color representing a specific intent is consistent through all the domains.

- PE3 originates BGP IPv6 unicast (AFI/SAFI=2/1) route for the Colored prefixes PE3:CL1:: and PE3:CL2::. Each route should carry the corresponding color extended community C1 or C2. PE3 also advertises a route for the base SRv6 Locator prefix PE3:BL, and there is no color extended community carried with this route.

- ASBR31 and ASBR32 receive the CPR routes of PE3, and advertise the CPR routes further to ASBR23 and ASBR24 with next-hop set to itself.

- ASBR23 and ASBR24 receive the CPR routes of PE3. Since the color-to-intent mapping in AS2 is consistent with that in AS3, the Color Extended Community in the received CPR routes are kept unchanged. ASBR23 and ASBR 24 advertise the CPR routes further in AS2 with the next-hop set to itself.

- The behavior of ASBR21 and ASBR22 are similar to the behavior of ASBR31 and ASBR32.

- The behavior of ASBR11 and ASBR12 are similar to the behavior of ASBR31 and ASBR32.

In normal case, the color value in the color extended community associated with the CPR route is consistent through all the domains. While in some special cases, one intent may be represented as different color value in different domains, then the Color Extended Community in the CPR routes may be updated at the border nodes of the domains based on the color-mapping policy.

In network scenarios where some of the intermediate network domains are MPLS based, the CPR routes may still be advertised using the IPv6 unicast address family (AFI/SAFI=2/1) in the MPLS-based intermediate domains, and at the MPLS domain border nodes, some route resolution policy could be used to make the CPR routes resolved to intra-domain intent-aware MPLS LSPs. Another possible mechanism is to use the IPv6 LU address family (AFI/SAFI=2/4) to advertise the CPR routes in the MPLS domains, the detailed procedure is described in Section 7.1.2.1 of [I-D.agrawal-spring-srv6-mpls-interworking].

### 2.3. CPR to Intra-domain Path Resolution

A domain border node which receives a CPR route can resolve the CPR route to an intra-domain color-aware path based on the tuple (N, C), where N is the next-hop of the CPR route, and C is the color extended

community of the CPR route. The intra-domain color-aware path could be built with any of the following mechanisms:

*SRv6 or SR-MPLS Policy

*SRv6 or SR-MPLS Flex-Algo

*RSVP-TE

For example, PE1 receives a CPR route to PE3:CL1 with color C1 and next-hop ASBR11, it can resolve the CPR routes to an intra-domain SRv6 Policy based on the tuple (ASBR11, C1).

The intra-domain path resolution scheme could be based on any existing tunnel resolution policy, and new tunnel resolution mechanisms could be introduced if needed.

### 2.4. SRv6 Service Route Advertisement

For an SRv6 service which is associated with a specific intent, the SRv6 Service SID could be allocated under the corresponding Colored locator prefix. For example, on PE3 in the example topology, an SRv6 VPN service with the low delay intent can be allocated with the SRv6 End.DT4 SID 2001:db8:aaaa:1:1000::0100, where 2001:db8:aaaa:1:1000::/68 is the SRv6 Colored prefix for low delay service.

The SRv6 service routes are advertised using the mechanism defined in [RFC9252]. The inter-domain VPN Option C is used, which means the next-hop of the SRv6 service route is set to the originating PE and not changed. Since the intent of the service is embedded in the SRv6 service SID, the SRv6 service route does not need to carry the color extended community.

### 2.5. SRv6 Service Steering

With the CPR routing mechanism, the ingress PE node which receives the SRv6 service routes follows the behavior of SRv6 shortest path forwarding (refer to Section 5 and 6 of [RFC9252]). The SRv6 service SID carried in the service route is used as the destination address in the outer IPv6 header encapsulated to the service packet. If the corresponding CPR route has been received and installed, longest prefix matching of SRv6 service SIDs to the Colored prefixes is performed, then the intra-domain color-aware paths in each network domain which the CPR route is resolved to are used for forwarding the SRv6 service traffic.

## 3. Encapsulation and Forwarding Processes

This section describes the encapsulation and forwarding process of data packets which are matched with the corresponding CPR route.

### 3.1. CPR over SRv6 Intra-Domain Paths

Following is an illustration of the packet encapsulation and forwarding process of CPR over SRv6 Policy. The abstract representation of IPv6 and SRH in section 6 of [RFC8754] is used.

PE3 is provisioned with a Colored prefix PE3:C1 for "low-delay".

In AS1, the SRv6 Policy for (ASBR11, C1) is represented with SID list (P1, BR11).

In AS2, the SRv6 Policy for (ASBR23, C1) is represented with the SID list (P2, BR23).

In AS3, the SRv6 Policy for (PE3, C1) is represented with the SID list (P3, PE3).

For packets which belong to an SRv6 VPN service associated with the SRv6 Service SID PE3:CL1.DT, the packet encapsulation and forwarding process is shown as below:

```
PE1 ->P1  : (PE1, P1)(PE3:CL1.DT, BR11; SL=2)(C-pkt)
P1  ->BR11: (PE1, BR11)(PE3:CL1.DT, BR11; SL=1)(C-pkt)
BR11->BR21: (PE1, PE3:CL1.DT)(C-pkt)
BR21->P2  : (PE1, P2)(PE3:CL1.DT, BR23; SL=2)(C-pkt)
P2  ->BR23: (PE1, BR23)(PE3:CL1.DT, BR23; SL=1)(C-pkt)
BR23->BR31: (PE1, PE3:CL1.DT)(C-pkt)
BR31->P3  : (PE1, P3)(PE3:CL1.DT, PE3; SL=2)(C-pkt)
P3  ->PE3 : (PE1, PE3)(PE3:CL1.DT, PE3; SL=1)(C-pkt)
```

In some network domains, SRv6 Flex-Algo may be used to provide intent-aware intra-domain path. The encapsulation is similar to the case with SRv6 Policy.

### 3.2. CPR over MPLS Intra-Domain Paths

In network scenarios where some of the network domains use MPLS based data plane, the CPR route can be resolved over a color-aware intra-domain MPLS LSP. Such intra-domain MPLS LSP may be established using SR-MPLS Policy, SR-MPLS Flex-Algo or RSVP-TE.

The encapsulation and forwarding of SRv6 service packets (which are actually IPv6 packets) over an intra-domain MPLS LSP is based on the MPLS mechanisms as defined in [RFC3031] [RFC3032] and [RFC8660]. The behavior is similar to that of 6PE [RFC4798].

In AS1, the SR-MPLS Policy for (ASBR11, C1) is represented with Label-stack (P1, BR11).

In AS2, the SR-MPLS Flex-Algo for (ASBR23, C1) is represented with Label-stack (BR23).

In AS3, the SR-MPLS Policy for (PE3, C1) is represented with Label-stack (P3, PE3).

For packets which belong to an SRv6 VPN service associated with the SRv6 Service SID PE3:CL1.DT, the packet encapsulation and forwarding process is shown as below:

```
PE1 ->P1  : Label-stack (P1, BR11) (PE1, PE3:CL1.DT)(C-pkt)
P1  ->BR11:     Label-stack (BR11) (PE1, PE3:CL1.DT)(C-pkt)
BR11->BR21:                        (PE1, PE3:CL1.DT)(C-pkt)
BR21->P2  :     Label-stack (BR23) (PE1, PE3:CL1.DT)(C-pkt)
P2  ->BR23:     Label-stack (BR23) (PE1, PE3:CL1.DT)(C-pkt)
BR23->BR31:                        (PE1, PE3:CL1.DT)(C-pkt)
BR31->P3  :  Label-stack (P3, PE3) (PE1, PE3:CL1.DT)(C-pkt)
P3  ->PE3 :      Label-stack (PE3) (PE1, PE3:CL1.DT)(C-pkt)
```

## 4. Operational Considerations

When the Colored prefixes are assigned as the sub-locators of the node's base SRv6 locator, the IPv6 unicast route of the base locator prefix is the covering prefix of all the Colored locator prefixes. To make sure the Colored locator prefixes can be distributed to the ingress PE nodes along the border nodes, it is required that route aggregation be disabled for IPv6 unicast routes which carries the color extended community.

All the border nodes and the ingress PE nodes needs to install the Colored locator prefixes into the RIB and FIB. For transit domains which support the CPR mechanism, the border nodes can use the tuple (N, C) to resolve the CPR routes to intent-aware intra-domain paths. For transit domains which do not support the CPR mechanism, the border nodes would ignore the color extended community and resolve the CPR routes over a best effort intra-domain path to the next-hop N, while the CPR route will be advertised further to the downstream domains with only the next-hop changed to itself. This allows the CPR routes to be resolved to intent-aware intra-domain paths in any network domains which support the CPR mechanism, while can fall back to resolve over best-effort intra-domain path in the legacy network domains.

In this document, IPv6 Unicast address family is used for the advertisement of IPv6 Colored prefixes. The primary advantage of this approach is the improved interoperability with legacy networks that lack support for intent-aware paths, and the facilitation of incremental deployment of intent-aware routing mechanisms. One potential concern arises regarding the necessity of separating

Colored prefixes from public IPv6 unicast routes. While as the IP prefixes and SRv6 locators of network infrastructure are usually advertised as part of the IP unicast routes, and appropriate filters are configured at the boundaries of network administration, this is not considered to be a significant issue. The proposal in [I-D.ietf-idr-bgp-car] provides a complementary solution that is also based on the notion of color indicating the intent and where the SRv6 Locator prefix itself signifies the intent, the difference is that a separate SAFI is used.

[I-D.ietf-idr-bgp-ct] describes another mechanism for intent-aware routing, in which the SRv6 service SIDs are not directly associated with the intent, while additional SRv6 transport SIDs are required for steering traffic to the inter-domain intent-aware paths, and an SRv6 operation similar to MPLS label swapping is needed on the border nodes of network domains.

## 5. IANA Considerations

This document makes no request of IANA.

## 6. Security Considerations

The mechanism described in this document provide an approach for inter-domain intent-aware routing based on existing BGP protocol mechanisms. It does not introduce any additional security considerations than those described in [RFC4271] and [RFC4272].

## 7. Acknowledgements

The authors would like to thank Shunwan Zhuang, Zhibo Hu, Zhenbin Li and Dhananjaya Rao for the review and valuable discussion.

## 8. References

### 8.1. Normative References

**[RFC2545]** Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing", RFC 2545, DOI 10.17487/RFC2545, March 1999, <https://www.rfc-editor.org/info/rfc2545>.

**[RFC4271]** Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

**[RFC4272]** Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

**[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>.

**[RFC8754]** Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

**[RFC8986]** Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

**[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>.

**[RFC9252]** Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <https://www.rfc-editor.org/info/rfc9252>.

**[RFC9256]** Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

### 8.2. Informative References

**[I-D.agrawal-spring-srv6-mpls-interworking]** Agrawal, S., Ali, Z., Filsfils, C., Voyer, D., Dawra, G., Li, Z., Hegde, S., and S. R. Sangli, "SRv6 and MPLS interworking", Work in Progress, Internet-Draft, draft-agrawal-spring-srv6-mpls-interworking-12, 10 July 2023, <https://datatracker.ietf.org/doc/html/draft-agrawal-spring-srv6-mpls-interworking-12>.

**[I-D.hr-spring-intentaware-routing-using-color]** Hegde, S., Rao, D., Sangli, S. R., Agrawal, S., Filsfils, C., Talaulikar, K., Patel, K., Uttaro, J., Decraene, B., Bogdanov, A., Jalil, L., Xu, X., Gulko, A., Khaddam, M., Contreras, L. M., Steinberg, D., Guichard, J., Henderickx, W., and Co-authors, "Problem statement for Inter-domain Intent-aware Routing using Color", Work in Progress,

Internet-Draft, draft-hr-spring-intentaware-routing-using-color-02, 10 July 2023, <https://datatracker.ietf.org/doc/html/draft-hr-spring-intentaware-routing-using-color-02>.

**[I-D.ietf-idr-bgp-car]** Rao, D., Agrawal, S., and Co-authors, "BGP Color-Aware Routing (CAR)", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-car-02, 6 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-car-02>.

**[I-D.ietf-idr-bgp-ct]** Vairavakkalai, K. and N. Venkataraman, "BGP Classful Transport Planes", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-ct-16, 22 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ct-16>.

**[RFC3031]** Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

**[RFC3032]** Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

**[RFC4798]** De Clercq, J., Ooms, D., Prevost, S., and F. Le Faucheur, "Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)", RFC 4798, DOI 10.17487/RFC4798, February 2007, <https://www.rfc-editor.org/info/rfc4798>.

**[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>.

**[RFC8660]** Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

## Authors' Addresses

Haibo Wang
Huawei Technologies
China

Email: rainsword.wang@huawei.com

Jie Dong

Huawei Technologies
China

Email: jie.dong@huawei.com

Ketan Talaulikar
Cisco Systems
India

Email: ketant.ietf@gmail.com

Tao Han
Huawei Technologies
China

Email: hantao@huawei.com

Ran Chen
ZTE Corporation
China

Email: chen.ran@zte.com.cn

Jingrong Xie
Huawei Technologies
China

Email: xiejingrong@huawei.com

Xinjun Chen
Huawei Technologies
China

Email: ifocus.chen@huawei.com