

IDR Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: 7 September 2023

Z. Wu  
H. Wang  
L. Wang  
Z. Tan  
X. Ding  
Huawei Technologies  
6 March 2023

BGP Flowspec Redirect Load Balancing Group Community  
draft-wu-idr-flowspec-redirect-group-01

Abstract

This document defines an extension to "BGP Community Container Attribute" [draft-ietf-idr-wide-bgp-communities], which allows flowspec redirection to multiple paths. This extended community serves to redirect traffic to a load balancing group and supports both equal-cost multi-path (ECMP) and unequal-cost multi-path (UCMP) scenarios.

Requirements Language

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].

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 7 September 2023.

Copyright Notice

Copyright (c) 2023 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	3
1.1. Terminology . . . . .	3
2. Redirect Load Balancing Group Community . . . . .	3
2.1. Community Value . . . . .	4
2.2. Param TLV . . . . .	4
2.3. Sub-TLVs(Path-tlvs) . . . . .	5
2.3.1. Path-tlv Type 1: IPv4 Prefix Only . . . . .	6
2.3.2. Path-tlv Type 2: IPv4 Prefix with Weight . . . . .	6
2.3.3. Path-tlv Type 3: IPv4 Prefix with Color . . . . .	7
2.3.4. Path-tlv Type 4: IPv4 Prefix with Color and Weight . . . . .	8
2.3.5. Path-tlv Type 5: IPv6 Prefix Only . . . . .	8
2.3.6. Path-tlv Type 6: IPv6 Prefix with Weight . . . . .	9
2.3.7. Path-tlv Type 7: IPv6 Prefix with Color . . . . .	10
2.3.8. Path-tlv Type 8: IPv6 Prefix with Color and Weight . . . . .	10
3. Scenarios . . . . .	11
3.1. ECMP . . . . .	11
3.2. UCMP . . . . .	11
4. Validation Procedure . . . . .	12
5. Error Handling . . . . .	12
5.1. Redirect Group Wide Community Parameter TLV . . . . .	12
5.2. Redirect Group Wide Community Parameter Sub-TLVs . . . . .	12
6. Operational Considerations . . . . .	12
6.1. Configuration Control . . . . .	13
6.2. Parsing . . . . .	13
6.3. Formating . . . . .	13
7. IANA Considerations . . . . .	14
7.1. BGP Wide Communities Community Type : Redirect Group . . . . .	14
8. Security Considerations . . . . .	14
9. References . . . . .	14
9.1. Normative References . . . . .	14
9.2. References . . . . .	15
Authors' Addresses . . . . .	15

## 1. Introduction

"Redirect to IP Extended Community", defined in [I-D.ietf-idr-flowspec-redirect-ip], allows traffic to be redirected to a specific IPv4 or IPv6 address, and [I-D.ietf-idr-ts-flowspec-srv6-policy] defines the redirection action to a SRv6 tunnel by additionally carrying the "Color Extended Community" [RFC8955].

However, scenarios involving redirection load balancing are not described in both documents. Although in some implementations, Equal-cost multi-path (ECMP) of "Redirect to IP" action can be achieved by encoding multiple redirect Extended Communities, the current set of mechanisms can hardly support neither ECMP of SRv6 tunnels nor unequal-cost multi-path (UCMP) of either types.

This document defines an extension to "BGP Community Container Attribute" [I-D.ietf-idr-wide-bgp-communities], the "Redirect Load Balancing Group" community. It is a new type of wide community container attribute with encoding format of multiple redirection path TLVs. Each of these TLVs represents a different redirection action. It allows traffic redirection to a load balancing group and supports both ECMP and UCMP scenarios.

The "Redirect Load Balancing Group" community is intended to be used within flowspec-v1 scenarios, the compatibility and interactions with flowspec-v2 is outside the scope of this document.

### 1.1. Terminology

This document introduces the following terms:

ECMP: Equal-Cost Multi-Path

UCMP: Unequal-Cost Multi-Path

Redirect Group: Redirect Load Balancing Group Community, a new type of BGP Community Container Attribute defined by this document

Path-tlv: Sub-tlv of the BGP Wide Community Parameter TLV, each represents a redirection path

## 2. Redirect Load Balancing Group Community

This document defines a new type of "BGP Community Container Attribute", the "Redirect Load Balancing Group" community type. The format complies with "BGP Community Container Attribute" [I-D.ietf-idr-wide-bgp-communities] and is shown below:

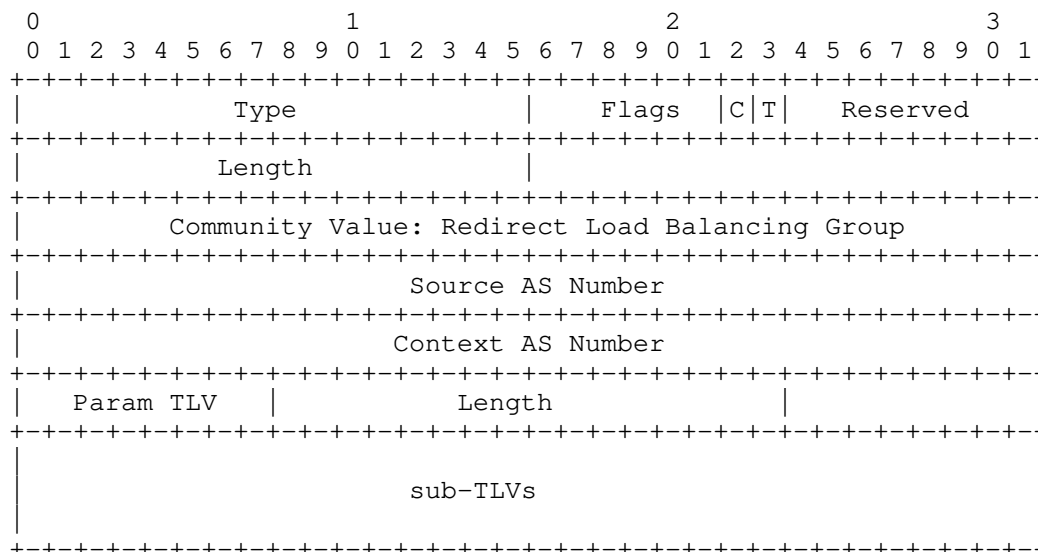


Figure 1: Redirect Load Balancing Group Community Format

The Type, Flags, Reserved and Length fields comply with the "BGP Community Container Attribute Common Header" definition.

The container type MUST be 1, which represents BGP Wide Community.

The Length field represents the total length of the container's contents in octets.

### 2.1. Community Value

The Community Value, Source AS Number and Context AS Number fields comply with the corresponding definition in "BGP Community Container Attribute".

Community Value: 4 octets value that represents the "Redirect Load Balancing Group" community type. The value is TBD and requires IANA registration; See Section 5.1.

### 2.2. Param TLV

The BGP Wide Community Parameter TLV (Sub-Type 3) contains a list of path-tlvs, comply with "BGP Wide Community Parameter(s) TLV" section of "BGP Community Container Attribute".

The Parameter TLV MUST present and SHOULD appear only once in a "Redirect Load Balancing Group" community container, no or multiple present SHOULD be considered malformed.

Sub-Type: Type 3 (BGP Wide Community Parameter TLV)

Length: Length of all the sub-TLVs in octets.

### 2.3. Sub-TLVs (Path-tlvs)

The list of path-tlvs that Param Tlv contains. Each path-tlv represents a different redirection path.

The general format of the sub-TLVs comply with path-tlvs' format defined in "BGP Community Container Attribute", as below:

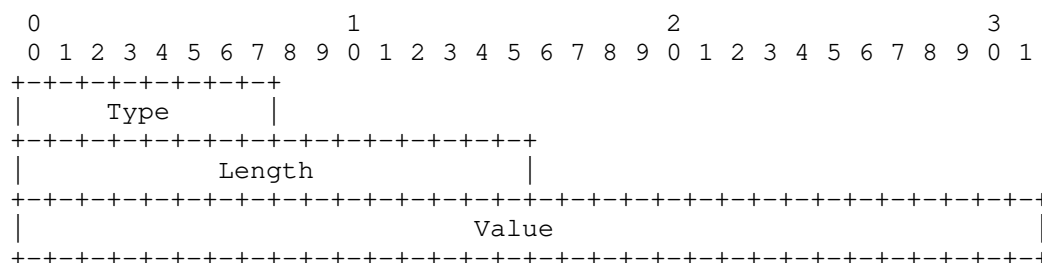


Figure 2: Param Sub-TLV Format

The Type field is an octet from 1~254 (0 and 255 are reserved). Supported type of the sub-TLVs includes:

Type 1: IPv4 Prefix Only

Type 2: IPv4 Prefix with Weight

Type 3: IPv4 Prefix with Color

Type 4: IPv4 Prefix with Color and Weight

Type 5: IPv6 Prefix Only

Type 6: IPv6 Prefix with Weight

Type 7: IPv6 Prefix with Color

Type 8: IPv6 Prefix with Color and Weight

These sub-TLV types SHOULD be used exclusively within "Redirect Load Balancing Group" community containers.

The Length represents the length of the "Value" field in octets, and it is fixed for each specific sub-TLV.

If the length and type of a sub-TLV do not match, the "Redirect Load Balancing Group" community container SHOULD be considered malformed.

If a sub-TLV is a total duplication of a previous one, the latter sub-TLV MUST be ignored.

In principle, sub-TLVs of different types may be combined in any mode. The supported combinations depend on the specific implementation.

### 2.3.1. Path-tlv Type 1: IPv4 Prefix Only

Indicating the redirection path is unweighted and to a IPv4 address. The format is shown below:

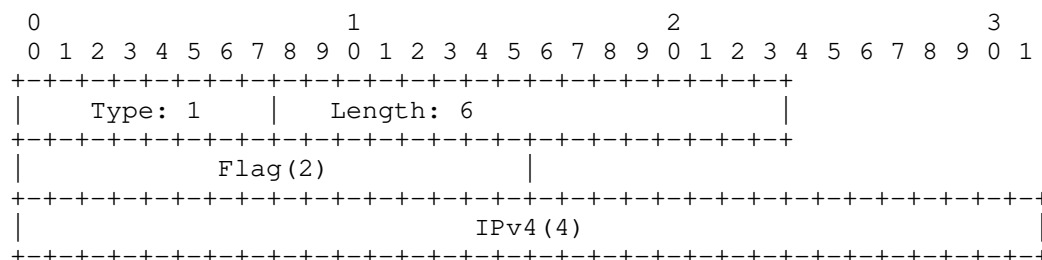


Figure 3: Path-tlv Type 1: IPv4 Prefix Only

Length: MUST be 6.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv4: 4-octet IPv4 address, redirection destination

### 2.3.2. Path-tlv Type 2: IPv4 Prefix with Weight

Indicating the redirection path is weighted and to a IPv4 address. The format is shown below:

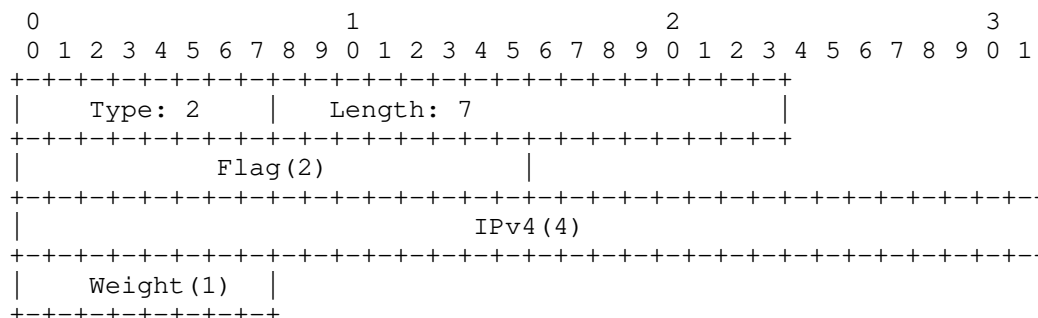


Figure 4: Path-tlv Type 2: IPv4 Prefix with Weight

Length: MUST be 7.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv4: 4-octet IPv4 address, redirection destination

Weight: 1 octet, values from 1~255, load balancing weight

### 2.3.3. Path-tlv Type 3: IPv4 Prefix with Color

Indicating the redirection path is unweighted and to a SR-TE tunnel.  
The format is shown below:

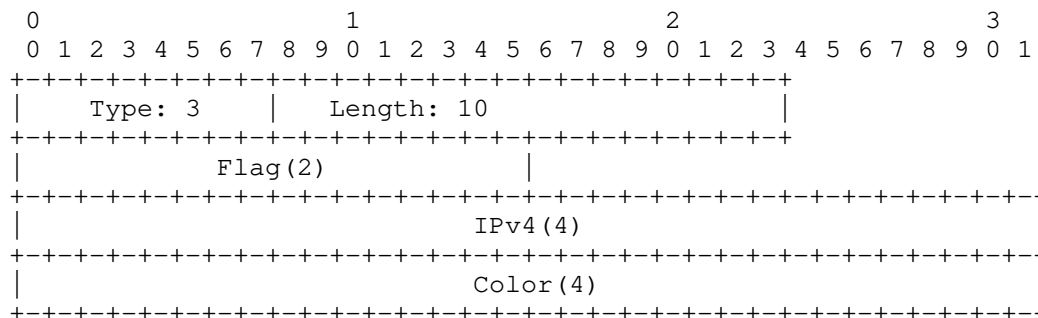


Figure 5: Path-tlv Type 3: IPv4 Prefix with Color

Length: MUST be 10.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv4: 4-octet IPv4 address, SR-TE tunnel Endpoint for redirection

Color: 4 octets, SR-TE tunnel Color for redirection

#### 2.3.4. Path-tlv Type 4: IPv4 Prefix with Color and Weight

Indicating the redirection path is weighted and to a SR-TE tunnel.  
The format is shown below:

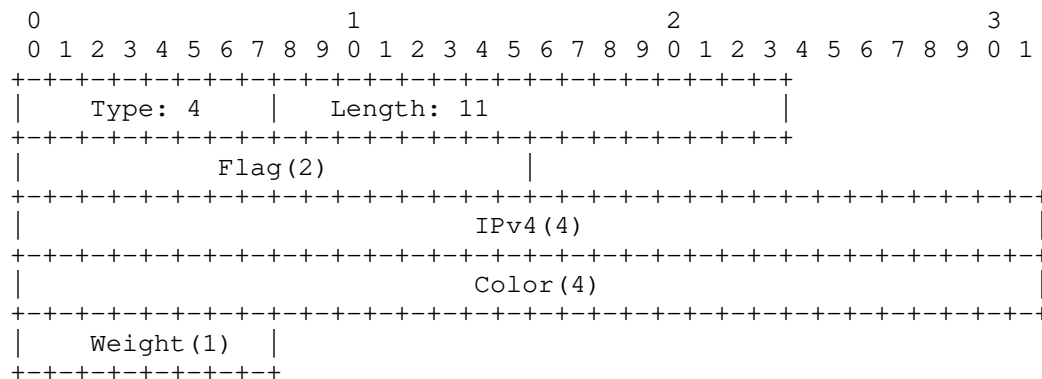


Figure 6: Path-tlv Type 4: IPv4 Prefix with Color and Weight

Length: MUST be 11.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv4: 4-octet IPv4 address, SR-TE tunnel Endpoint for redirection

Color: 4 octets, SR-TE tunnel Color for redirection

Weight: 1 octet, values from 1~255, load balancing weight

#### 2.3.5. Path-tlv Type 5: IPv6 Prefix Only

Indicating the redirection path is unweighted and to a IPv6 address.  
The format is shown below:



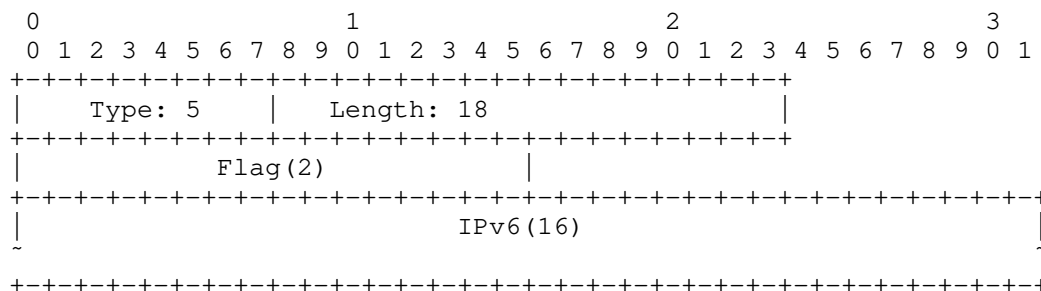


Figure 7: Path-tlv Type 5: IPv6 Prefix Only

Length: MUST be 18.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv6: 16-octet IPv6 address, redirection destination

#### 2.3.6. Path-tlv Type 6: IPv6 Prefix with Weight

Indicating the redirection path is weighted and to a IPv6 address.  
The format is shown below:

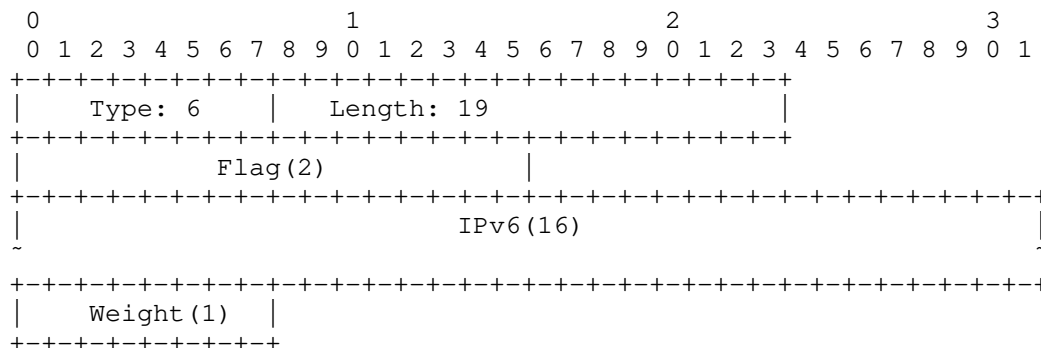


Figure 8: Path-tlv Type 6: IPv6 Prefix with Weight

Length: MUST be 19.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv6: 16-octet IPv6 address, redirection destination

Weight: 1 octet, values from 1~255, load balancing weight

### 2.3.7. Path-tlv Type 7: IPv6 Prefix with Color

Indicating the redirection path is unweighted and to a SRv6 tunnel.  
The format is shown below:

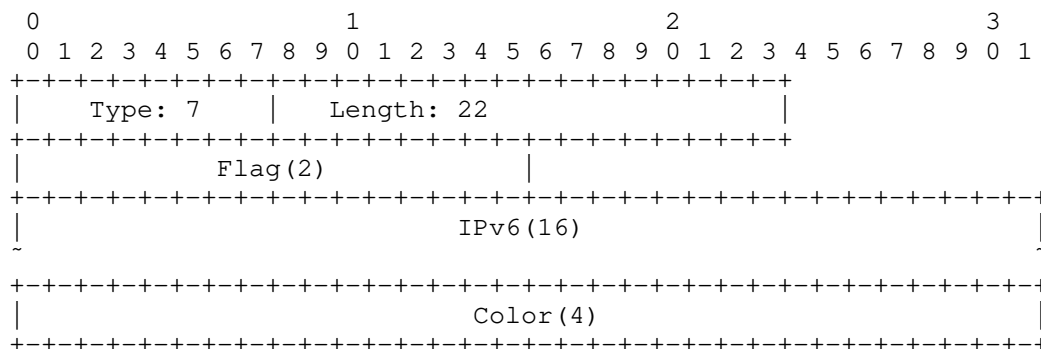


Figure 9: Path-tlv Type 7: IPv6 Prefix with Color

Length: MUST be 22.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv6: 16-octet IPv6 address, SRv6 tunnel Endpoint for redirection

Color: 4 octets, SRv6 tunnel Color for redirection

### 2.3.8. Path-tlv Type 8: IPv6 Prefix with Color and Weight

Indicating the redirection path is weighted and to a SRv6 tunnel.  
The format is shown below:

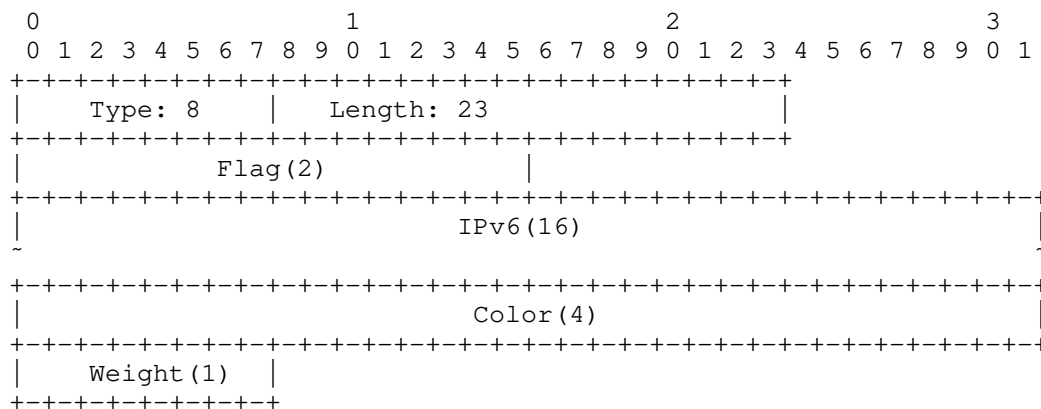


Figure 10: Path-tlv Type 8: IPv6 Prefix with Color and Weight

Length: MUST be 23.

Flags: 2 octets, reserved for future use, MUST be set to 0 upon the sender and MUST be ignored upon the receiver.

IPv6: 16-octet IPv6 address, SRv6 tunnel Endpoint for redirection

Color: 4 octets, SRv6 tunnel Color for redirection

Weight: 1 octet, values from 1~255, load balancing weight

### 3. Scenarios

This section describes a few use-case scenarios when deploying "Redirect Load Balancing Group" community type.

Weighted path-tlv types: Path-tlvs contain a Weight field, such as Type 2, 4, 6, 8

Unweighted path-tlv types: Path-tlvs do not contain a Weight field, such as Type 1, 3, 5, 7

#### 3.1. ECMP

A system that originates a flowspec route with a "Redirect Load Balancing Group" community, among which its parameter TLV contains more than 1 path-tlvs. If not all path-tlvs are of a weighted type, these path-tlvs will form a ECMP group.

Implementations MUST be prepared to accept a Parameter TLV with both weighted and unweighted path-tlvs. In this case, the Weight field of the weighted path-tlv SHOULD be ignored.

#### 3.2. UCMP

A system that originates a flowspec route with a "Redirect Load Balancing Group" community, among which its parameter TLV contains more than 1 path-tlvs. If all path-tlvs are of a weighted type, these path-tlvs will form a UCMP group.

In this case, the Weight field value of these path-tlvs SHOULD NOT be ignored, and the values are used as the ratio of the UCMP group.

#### 4. Validation Procedure

In the absence of explicit configuration, a Redirect Group attribute MUST be validated before it is used for redirection action or sent to a BGP peer.

The validation procedure for a Redirect Group attribute follows the following rules:

- \* Each Path-tlv of the Redirect Group attribute SHOULD be validated separately. The validation of each path follows the validation procedure of Redirect to IP Action [I-D.ietf-idr-flowspec-redirect-ip].
- \* A Redirect Group attribute SHOULD be considered verified, only after all path-tlvs in the Redirect Group attribute are verified.
- \* If any path-tlvs are invalid, these paths SHOULD NOT participate in load-balance calculation and used for redirection actions.
- \* If any path-tlvs are invalid, the Redirect Group attribute SHOULD NOT be sent to a BGP peer.

#### 5. Error Handling

Comply with Error Handling Procedure in "BGP Community Container Attribute" [I-D.ietf-idr-wide-bgp-communities].

In addition:

##### 5.1. Redirect Group Wide Community Parameter TLV

A "Redirect Load Balancing Group" community container with no or multiple parameter TLVs SHOULD be considered malformed, and a "treat as withdraw" behavior is expected.

##### 5.2. Redirect Group Wide Community Parameter Sub-TLVs

If the length and type of a sub-TLV do not match, the "Redirect Load Balancing Group" community container SHOULD be considered malformed, and a "treat as withdraw" behavior is expected.

#### 6. Operational Considerations

The Extended Community attributes for redirection mentioned in this section include:

- \* Redirect to IP Extended Community  
[I-D.ietf-idr-flowspec-redirect-ip]
- \* Redirect to IPv6 Extended Community  
[I-D.ietf-idr-flowspec-redirect-ip]
- \* Redirect to SRv6 Policy [I-D.ietf-idr-ts-flowspec-srv6-policy]

### 6.1. Configuration Control

There SHOULD be an explicit configuration to control whether the Redirect Group attribute is used for redirection actions. In the absence of the explicit configuration (by default), the Redirect Group attribute MAY NOT take precedence over Extended Community attribute. With the explicit configuration, the Redirect Group attribute MAY take precedence over Extended Community attribute for redirection.

For clarity, the first scenario, in which the Redirect Group attribute does not take precedence, is called configuration situation A. And the second scenario is called configuration situation B.

### 6.2. Parsing

While receiving a flowspec route with Redirect Group attribute from a BGP peer:

- \* In configuration situation A, the Redirect Group attribute SHOULD NOT be used for redirection actions. If the route carries Extended Community attributes for redirection, these attributes MAY be used to generate the redirection actions. The Redirect Group attribute SHOULD still be saved locally and advertised with the flowspec route to other appropriate peers.
- \* In configuration situation B, the Redirect Group attribute SHOULD take precedence over Extended Community attribute for redirection. If the route carries Extended Community attributes for redirection, these attributes SHOULD NOT be used to generate the redirection actions, but SHOULD still be saved locally and advertised with the flowspec route to other appropriate peers.

### 6.3. Formating

While encoding a local-generated flowspec route:

- \* In configuration situation A, a Redirect Group attribute SHOULD NOT be encoded. Appropriate Extended Community attributes MAY be used for specifying redirection actions.

- \* In configuration situation B, the Redirect Group attribute SHOULD be encoded for specifying redirection actions, despite of there is one or more paths. For the sake of compatibility, we MAY select the path with the lowest IP address from the paths of the Redirect Group attribute and encode it with appropriate Extended Community attributes. During this selection, an IPv4 address is preferred over an IPv6 address.

While encoding a flowspec route learned from other BGP peers:

- \* In configuration situation A, the Redirect Group attribute MUST be encoded without modification.
- \* In configuration situation B, the Redirect Group attribute MUST pass the validation procedure before it is encoded and sent to a BGP peer.

## 7. IANA Considerations

### 7.1. BGP Wide Communities Community Type : Redirect Group

This document requests a new community value under "Registered Type 1 BGP Wide Community Community Types" registry. This registry is defined and requested in "BGP Community Container Attribute" [I-D.ietf-idr-wide-bgp-communities].

Requested value:

Name	Type Value
-----	-----
Redirect Load Balancing Group	TBD

## 8. Security Considerations

A system that originates a flowspec route with a "Redirect Load Balancing Group" BGP wide community can cause many receivers of that route to redirect traffic to a single next-hop, overwhelming that next-hop and resulting in inadvertent or deliberate denial-of-service. This is also a concern about the "redirect to IP" extended community, therefore this document introduces no additional security considerations than those already covered in [RFC8955].

## 9. References

### 9.1. Normative References

- [I-D.ietf-idr-flowspec-redirect-ip]  
Uttaro, J., Haas, J., Texier, M., akarch@cisco.com, Ray, S., Simpson, A., and W. Henderickx, "BGP Flow-Spec Redirect to IP Action", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-redirect-ip-02, 5 February 2015, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-redirect-ip-02>>.
- [I-D.ietf-idr-wide-bgp-communities]  
Raszuk, R., Haas, J., Lange, A., Decraene, B., Amante, S., and P. Jakma, "BGP Community Container Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-wide-bgp-communities-10, 2 March 2023, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-wide-bgp-communities-10>>.
- [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>>.

## 9.2. References

- [I-D.ietf-idr-ts-flowspec-srv6-policy]  
Wenying, J., Liu, Y., Zhuang, S., Mishra, G. S., and S. Chen, "Traffic Steering using BGP FlowSpec with SRv6 Policy", Work in Progress, Internet-Draft, draft-ietf-idr-ts-flowspec-srv6-policy-01, 9 October 2022, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-ts-flowspec-srv6-policy-01>>.
- [RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <<https://www.rfc-editor.org/info/rfc8955>>.
- [RFC8956] Loibl, C., Ed., Raszuk, R., Ed., and S. Hares, Ed., "Dissemination of Flow Specification Rules for IPv6", RFC 8956, DOI 10.17487/RFC8956, December 2020, <<https://www.rfc-editor.org/info/rfc8956>>.

Authors' Addresses

Zhiwen Wu  
Huawei Technologies  
No. 156 Beiqing Road  
Beijing  
100095  
P.R. China  
Email: wuzhiwen1@huawei.com

Haibo Wang  
Huawei Technologies  
No. 156 Beiqing Road  
Beijing  
100095  
P.R. China  
Email: rainsword.wang@huawei.com

Lili Wang  
Huawei Technologies  
No. 156 Beiqing Road  
Beijing  
100095  
P.R. China  
Email: lily.wong@huawei.com

Zhen Tan  
Huawei Technologies  
No. 156 Beiqing Road  
Beijing  
100095  
P.R. China  
Email: tanzhen6@huawei.com

Xiangfeng Ding  
Huawei Technologies  
No. 156 Beiqing Road  
Beijing  
100095  
P.R. China  
Email: dingxiangfeng@huawei.com