Authors: Z. Wu   H. Wang
         Huawei Technologies   Huawei Technologies
         L. Wang   Z. Tan
         Huawei Technologies   Huawei Technologies

BGP Flowspec Redirect Load Balancing Group Community

Abstract

This document defines an extension to "BGP Community Container
Attribute" [I-D.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

Copyright Notice

Table of Contents

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.jiang-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.

1.1. Terminology

This document introduces the following terms:

ECMP: Equal-Cost Multi-Path

UCMP: Unequal-Cost Multi-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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type            |   Flags   |C|T|    Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Community Value: Redirect Load Balancing Group       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Source AS Number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Context AS Number                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Param TLV  |              Length          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                             |
|                        sub-TLVs                             |
|                                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
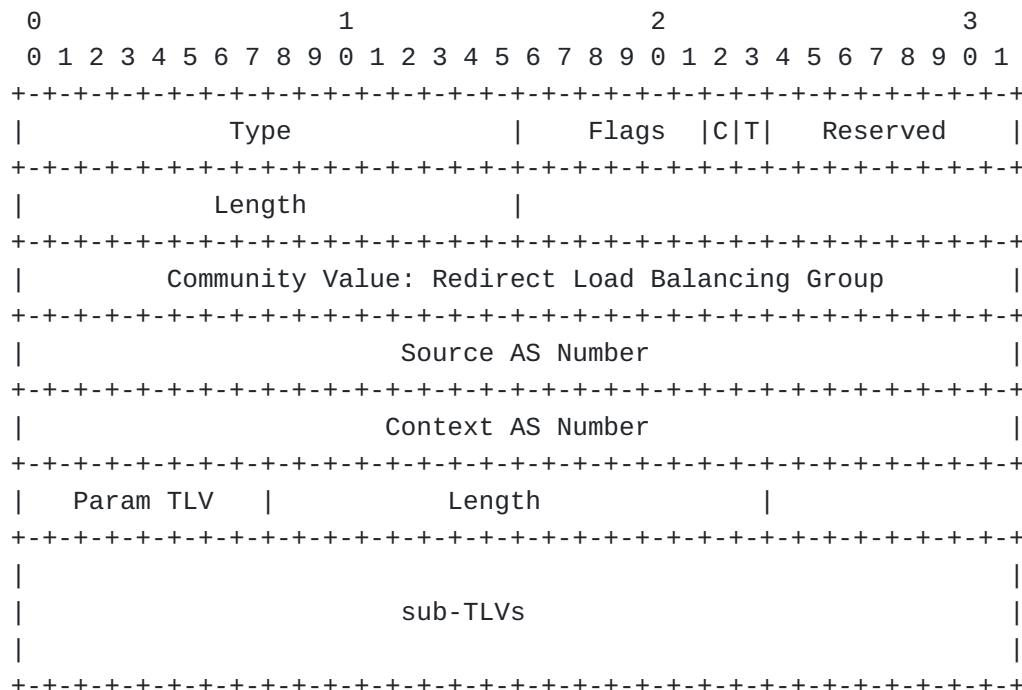
Figure 1: Redirect Load Balancing Group Community Format

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 atoms, 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(Atoms)

The list of atoms that Param Tlv contains. Each atom represents a different redirection path.

The general format of the sub-TLVs comply with atoms' format defined in "BGP Community Container Attribute", as below:
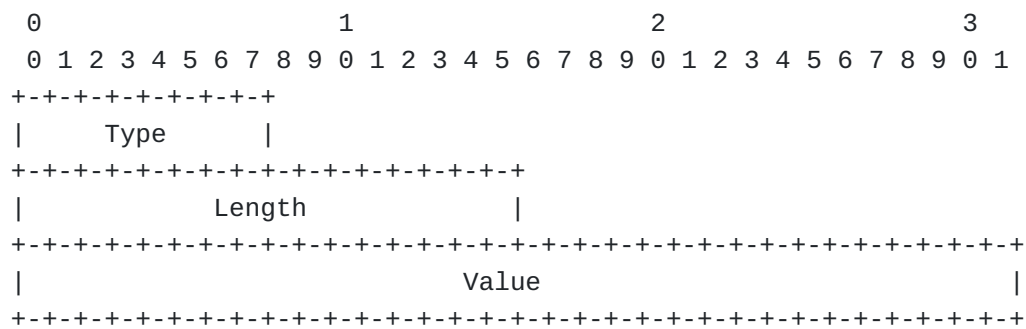
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|     Type      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Value                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                   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 dupilication 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. Atom Type 1: IPv4 Prefix Only

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 1    |   Length: 6                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Flag(2)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4(4)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
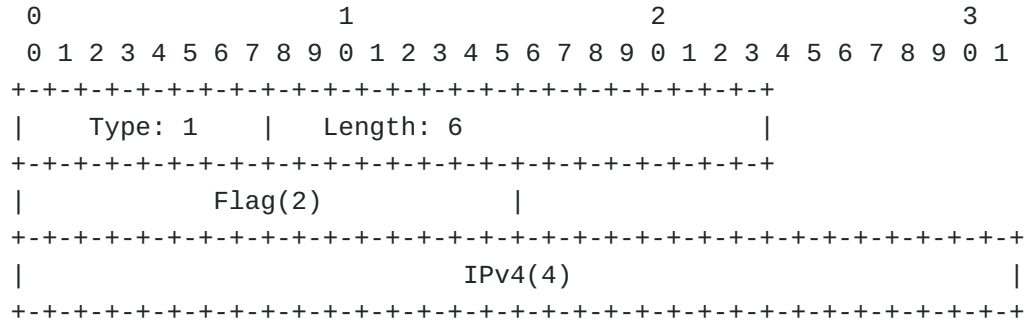
                 Figure 3: Atom Type 1: IPv4 Prefix Only

   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 reciever.

   IPv4: 4-octet IPv4 address, redirection destination

2.3.2. Atom Type 2: IPv4 Prefix with Weight

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 2    |   Length: 7                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Flag(2)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4(4)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Weight(1)   |
+-+-+-+-+-+-+-+-+-+
```
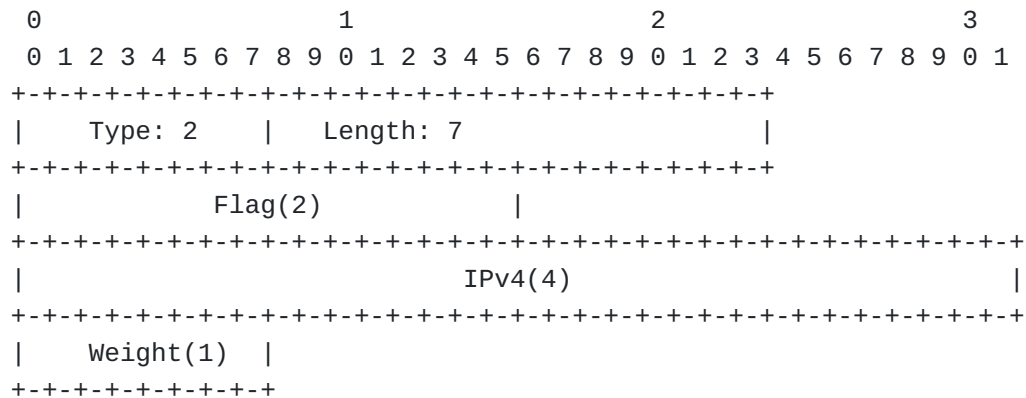
                 Figure 4: Atom Type 2: IPv4 Prefix with Weight

   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 reciever.

IPv4: 4-octet IPv4 address, redirection destination

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

2.3.3. Atom Type 3: IPv4 Prefix with Color

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 3   |   Length: 10              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Flag(2)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4(4)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Color(4)                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
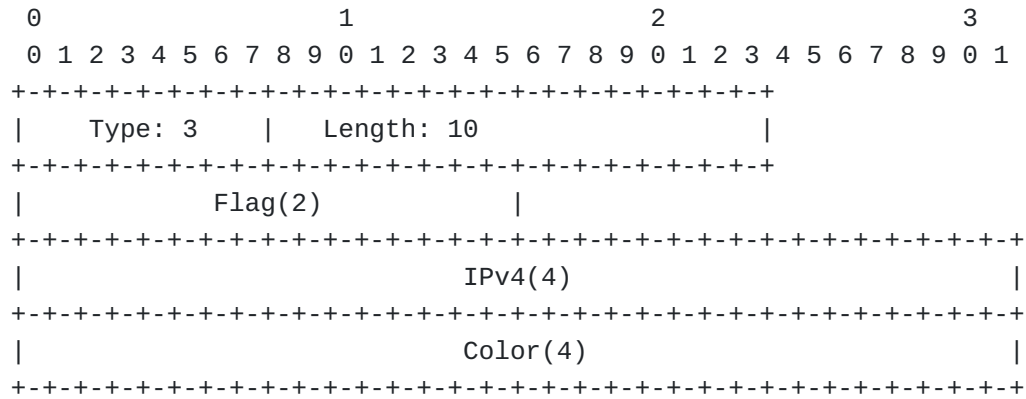
Figure 5: Atom Type 3: IPv4 Prefix with Color

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 reciever.

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

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

2.3.4. Atom 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 4     |   Length: 11               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Flag(2)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv4(4)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Color(4)                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Weight(1)  |
+-+-+-+-+-+-+-+-+
```
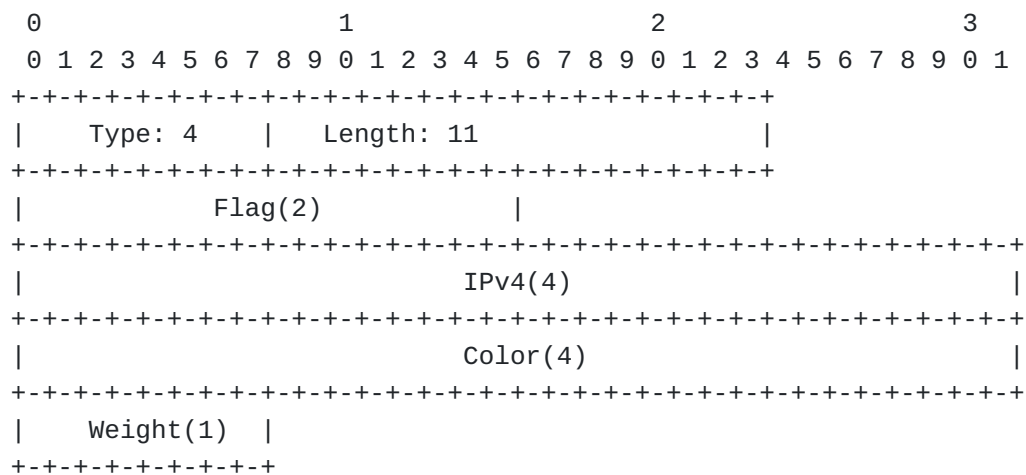
Figure 6: Atom Type 4: IPv4 Prefix with Color and Weight

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 reciever.

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. Atom Type 5: IPv6 Prefix Only

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 5     |   Length: 18               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Flag(2)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv6(16)                           |
~                                                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
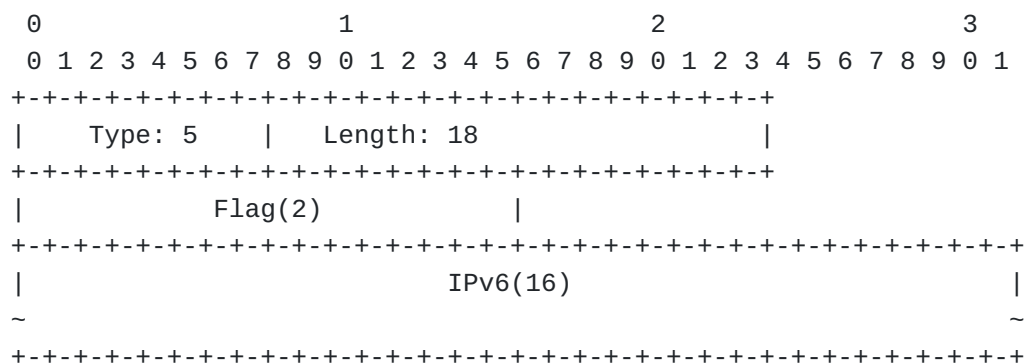
Figure 7: Atom Type 5: IPv6 Prefix Only

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 reciever.

```
    IPv6: 16-octet IPv6 address, redirection destination
```

2.3.6. Atom Type 6: IPv6 Prefix with Weight

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type: 6    |    Length: 19                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Flag(2)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv6(16)                            |
~                                                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Weight(1)  |
+-+-+-+-+-+-+-+-+-+
```
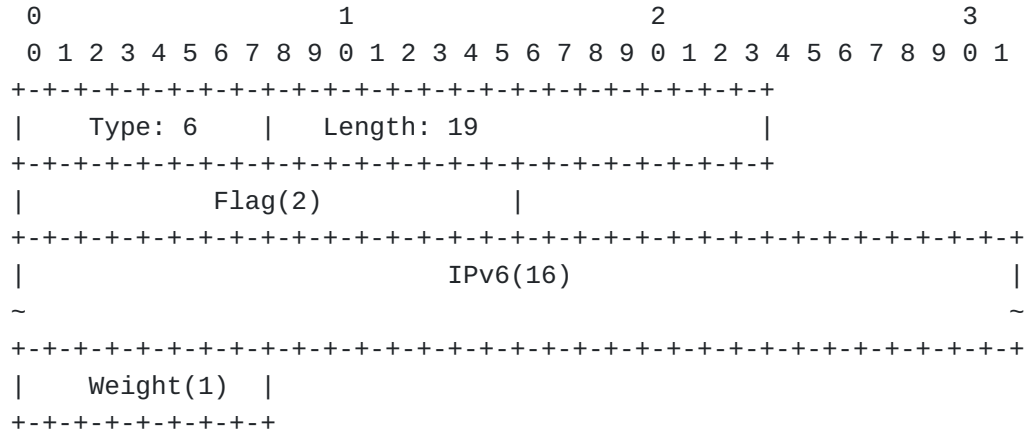
                Figure 8: Atom Type 6: IPv6 Prefix with Weight

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 reciever.

IPv6: 16-octet IPv6 address, redirection destination

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

2.3.7. Atom Type 7: IPv6 Prefix with Color

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type: 7    |    Length: 22                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Flag(2)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          IPv6(16)                            |
~                                                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Color(4)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
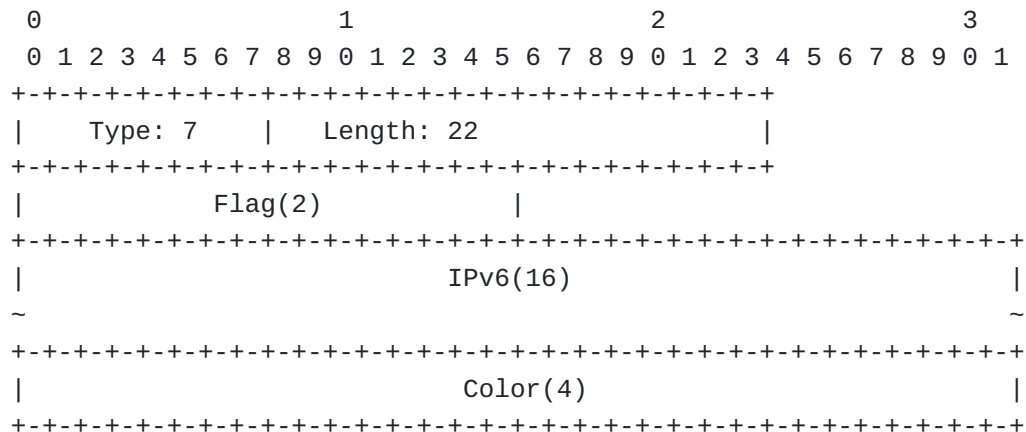
Figure 9: Atom Type 7: IPv6 Prefix with Color

   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 reciever.

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

   Color: 4 octets, SRv6 tunnel Color for redirection

2.3.8. Atom Type 8: IPv6 Prefix with Color and Weight

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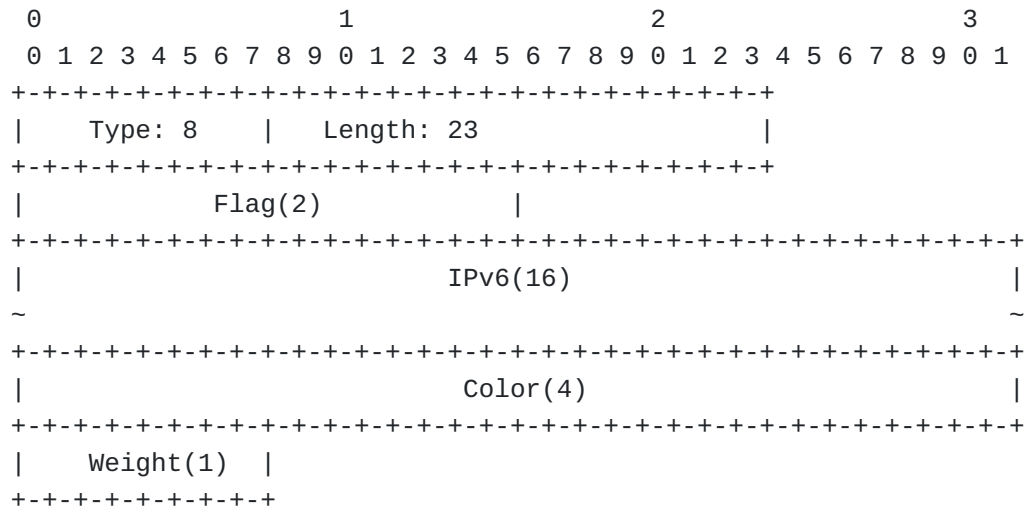   The format is shown below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type: 8   |   Length: 23                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Flag(2)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         IPv6(16)                             |
~                                                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Color(4)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Weight(1)  |
+-+-+-+-+-+-+-+-+
```

        Figure 10: Atom 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 reciever.

   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 atom types:
Atoms contain a Weight field, such as Type 2,
4, 6, 8

Unweighted atom types: Atoms 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 atoms. If not all atoms are of a weighted type, these
atoms will form a ECMP group.

Implementations MUST be prepared to accept a Parameter TLV with both
weighted and unweighted atoms. In this case, the Weight field of the
weighted atom 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 atoms. If all atoms are of a weighted type, these atoms
will form a UCMP group.

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

4. Error Handling

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

In addition:

4.1. 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.

4.2. 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.

5. IANA Considerations

5.1. BGP Wide Communities Community Type : Redirect Group

This document requests a new community value under "Registered Type 1 BGP Wide Community Community Types" registery. This registery 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
```

6. 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].

7. References

7.1. Normative References

[I-D.ietf-idr-flowspec-redirect-ip]
          Uttaro, J., Haas, J., Texier, M., Karch, A., 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://www.ietf.org/archive/id/draft-ietf-idr-flowspec-redirect-ip-02.txt>.

[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-06, 10 January 2022, <https://www.ietf.org/archive/id/draft-ietf-idr-wide-bgp-communities-06.txt>.

[I-D.jiang-idr-ts-flowspec-srv6-policy] Jiang, W., Liu, Y., and S. Chen, "Traffic Steering using BGP Flowspec with SRv6 Policy", Work in Progress, Internet-Draft, draft-jiang-idr-ts-flowspec-srv6-policy-04, 11 July 2021, <https://www.ietf.org/archive/id/draft-jiang-idr-ts-flowspec-srv6-policy-04.txt>.

[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>.

7.2. References

[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