Workgroup: idr
Internet-Draft:
draft-zzhang-idr-bgp-rt-constrains-
extension-03
Updates: 4684 (if approved)
Published: 26 June 2023
Intended Status: Standards Track
Expires: 28 December 2023
Authors: Z. Zhang
Juniper Networks
J. Haas
Juniper Networks

Generic Route Constraint Distribution Mechanism for BGP

Abstract

This document defines a mechanism based upon Constrained Route Distribution for BGP (RFC 4684) that works with various types of BGP Community-like Path Attributes. Similar to RFC 4684, this mechanism can be used to build a route distribution graph to limit the propagation of BGP Routes. Unlike RFC 4684, this mechanism is not restricted to BGP Extended Communities (RFC 4360).

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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 28 December 2023.

Copyright Notice

Table of Contents

1. Introduction

1.1. Constrained Route Distribution

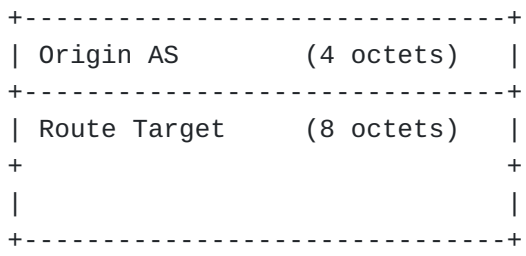
In BGP/MPLS Layer 3 VPNs [RFC4364], Route Target Extended Communities [RFC4360] are used to control VPN membership. Networks providing VPN services may be large. In such networks, VPN routes for a given VPN may be only needed at a small subset of Provider Edge (PE) routers.

The Constrained Route Distribution feature [RFC4684] assists in scaling such large VPN networks by building a distribution graph of VPN routes through the BGP routing infrastructure. Much of the benefit of this feature comes from BGP routers, such as Route Reflectors [RFC4456], avoiding the work of sending all VPN routes to a PE that may simply discard unneeded routes. Instead, the PE may receive only the VPN routes for VPNs located on that PE.

1.2. Brief Summary of Constrained Route Distribution Procedure

BGP Speakers implementing [RFC4684] advertise their interest in receiving VPN routes that contain specific Route Target Extended Communities by advertising Route Target membership NLRI.

The format of the Route Target membership NLRI in [RFC4684] follows. It may be of length from 0 to 96 bits.



The Origin AS contains the Autonomous System number of the originator of this NLRI.

The Route Target contains a BGP Route Target Extended Community, or a prefix of a BGP Route Target Extended Community.

Route Target membership NLRI act as a filter mechanism on VPN routes. The BGP Speaker receiving these Route Target membership NLRI from another BGP Speaker will propagate VPN routes that match these membership NLRI. VPN routes that do not match these membership NLRI will not be propagated.

The propagation of Route Target membership NLRI from an originating PE router to other interested BGP Speakers builds a distribution graph for VPN routes matching the desired Route Targets.

1.3. Need for a Generic Route Constraint Distribution Mechanism

Since BGP/MPLS Layer 3 VPNs were introduced, many new BGP VPN features have been created that leverage the original concepts in [RFC4364]. While many of these new features similarly use Route Target Extended Communities for VPN membership, some use other Extended Communities. That is, they utilize a different Type/Sub-Type code than those defined in [RFC4360].

While [RFC4684] is explicit about being utilized for Route Targets, the definition of a Route Target has become more fluid as VPN features have been introduced; for example, ES-Import from [RFC7432]. It could be observed that that [RFC4684] is capable of being used on any type of [RFC4360] BGP Extended Community, for any VPN route type. However, other attributes are coming to be used for idenitifying VPN routes and a procedure that is only applicable to Extended Communities cannot be used.

[RFC5701] introduced the IPv6 Address Specific BGP Extended Community Attribute. This type of BGP Community permits the encoding of an IPv6 address as the Global Administrator of a route. Similar to the [RFC4360] Extended Communities, the IPv6 Address Specific type carries a Type and Sub-Type field. One of the Type/Sub-Type allocations is for an IPv6 address specific Route Target. This permits operators to leverage IPv6 addressing when building their VPNs.

IPv6 Extensions for Route Target Distribution [I-D.ietf-idr-bgp-ipv6-rt-constrain] proposes to permit matching for IPv6 address specific Extended Communities using [RFC4684] by overloading the NLRI length for Route Target membership NLRI for NLRI longer than 96 bits. (See [RFC4684], Section 4.) However, this doesn't account for Route Target membership NLRI length shorter than 96 bits. These shorter prefixes permit matching of many more specific Route Targets from a less specific Route Target membership BGP Route. Therefore, a different mechanism is needed for safely matching IPv6 address specific Route Targets.

The simplest change would be to utilize a new AFI/SAFI for IPv6 Route Target Distribution that only matches IPv6 address specific Route Targets. It can be further observed that various forms of BGP "Community" types continue to evolve to suit a variety of BGP route filtering needs, including those not intended for VPN services. Examples of these include BGP Large Communities [RFC8092], BGP Wide Communities [I-D.ietf-idr-wide-bgp-communities], and Bitmask Route Targets [I-D.zzhang-idr-bitmask-route-target].

This document proposes a mechanism to match arbitrary BGP Community-like attributes, including those with Route Target-like semantics,

for building Constrained Route Distribution graphs for BGP routes containing those attributes.

2. Community-like Attributes

2.1. Definition of Community-like Attributes

BGP Communities were originally introduced in [RFC1997]. That RFC contains the definition, "A community is a group of destinations which share some common property." Recall that in BGP-4 [RFC4271], a BGP Route is defined as a pairing of destinations (NLRI) with Path Attributes.

In practice, a Community is implemented as an element of a BGP Path Attribute that is used to mark a prefix in a way that protocol and BGP policy mechanisms may be used to interact with that BGP Route.

Since [RFC1997], this idea of marking BGP Routes has been extended to other mechanisms such as BGP Extended Communities [RFC4360], and BGP Large Communities [RFC8092]. Other similar mechanisms are regularly considered for standardization.

For purposes of this document, a Community-like Attribute (CLA) has the semantics of being an attribute of a BGP Path Attribute that is intended to interact with protocol mechanisms and may enable policy mechanisms to interact with that BGP Route. Thus, classic [RFC1997] BGP Communities, BGP Extended Communities, and Large BGP Communities are all CLAs.

2.2. Prefix Structure of BGP Community-like Attributes

[RFC4684] provides for matching less-specific BGP Extended Communities by utilizing a shorter NLRI length for the Route Target membership NLRI. To highlight situations where such summarization is useful, consider the various forms of Route Target extended community from [RFC4360]. In each of those types, the Sub-Type field is 0x02, with the Type selecting the format:

* 0x00 - 2-octet Global Administrator field, 4-octet Local Administrator field.

* 0x01 - 4-octet Global Administrator field, 2-octet Local Administrator field.

* 0x02 - 4-octet IPv4 Address Global Administrator field, 2-octet Local Administrator field.

The Global Administrator field for Route Targets is typically an Autonomous System number.

Summarization offers several useful options where the Sub-Type of the Route Target Extended Community is 0x02. Examples include:

- Type = 0x00 and NLRI length = 48: Match all 2-octet Global Administrator fields of a given value; for example Origin AS 64511:Route Target 64496:*.

- Type = 0x01 and NLRI length = 64: Match all 4-octet Global Administrator fields of a given value; for example Origin AS 64511:Route Target 65551:*.

- Type = 0x03 and NLRI length = 88: Match all IPv4 Address Global Administrator fields of a given value; for example Origin AS 64511:Route Target 192.0.2.*:*.

Similarly, for inter-domain purposes, matching all Route Target Membership NLRI for a given Origin AS may be useful:

- NLRI length = 32; for example Origin AS 64511:*. This matches all classes of Extended Community originated from AS 64511.

- NLRI length = 44; for example Origin AS 64511:0x0002:*. This matches all Extended Communities originated from AS 64511 that have the first two octets as 0x0002, which includes the class of Extended Communities that are 2-octet Global Administrator Route Target types.

It's even possible to utilize a Prefix Length that splits a well defined field. When the structure of that field is understood, clever operators may be able to generate summaries. It should be noted that understanding the intent of such summarization may be difficult to discern from the NLRI in question. Some examples:

- NLRI length = 31; for example Origin AS 6451[01]:*. This matches all classes of Extended Community originated from Origin ASes 64510 and 64511.

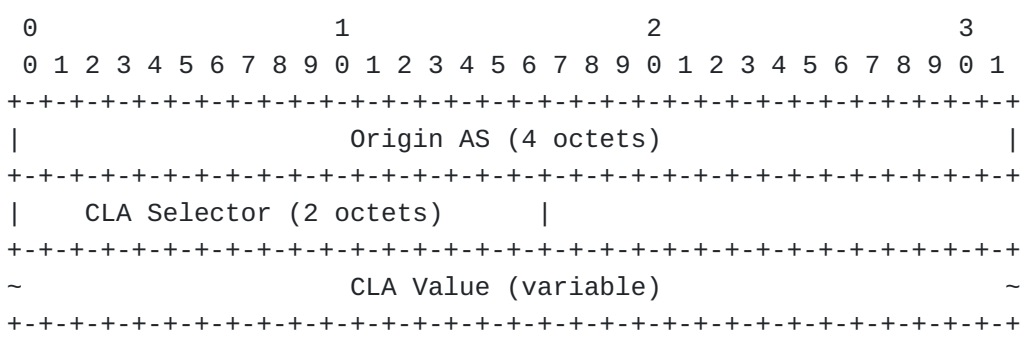
- NLRI length = 47; for example Origin AS 64511:0x0002:*. This matches all two-octet AS-Specific Extended Communities originated from AS 64511 that include Route Targets (0x02) and Route Origins (0x03).

The purpose of highlighting that a variable NLRI length can be applied in these ways is to demonstrate the flexibility of summarization. This is most true when the structure of that attribute is arranged most general to most specific; that is, Global to Local Admin as we have in Extended Communities.

3. Specification

3.1. NLRI Definition

To support applying Constrained Route Distribution procedures to BGP Community-like attributes, the following NLRI is defined. The "Generic Route Constraint Distribution Mechanism" NLRI uses a new SAFI (TBD) with the following format:



It can be observed that the format of this NLRI emulates the format of the Route Target membership NLRI from [RFC4684], with the addition of the CLA selector to permit the recipient to correctly interpret the CLA value.

3.2. NLRI Length Encoding

To support potentially large Community-like Values, the NLRI length field is encoded using 1 or 2 octets using the same mechanism as [RFC5575], Section 4. The text from that RFC is copied here:

> If the NLRI length value is smaller than 240 (0xf0 hex), the length field can be encoded as a single octet. Otherwise, it is encoded as an extended-length 2-octet value in which the most significant nibble of the first byte is all ones.
>
> In the figure above, values less-than 240 are encoded using two hex digits (0xnn). Values above 240 are encoded using 3 hex digits (0xfnnn). The highest value that can be represented with this encoding is 4095. The value 241 is encoded as 0xf0f1.

3.3. Operation

The two-octet CLA Selector identifies the type of Community-like attribute in a BGP route to apply the Constrained Route Distribution procedures to. The value of this field, registered with IANA, may identify Community-like attributes that exist in a given BGP Path

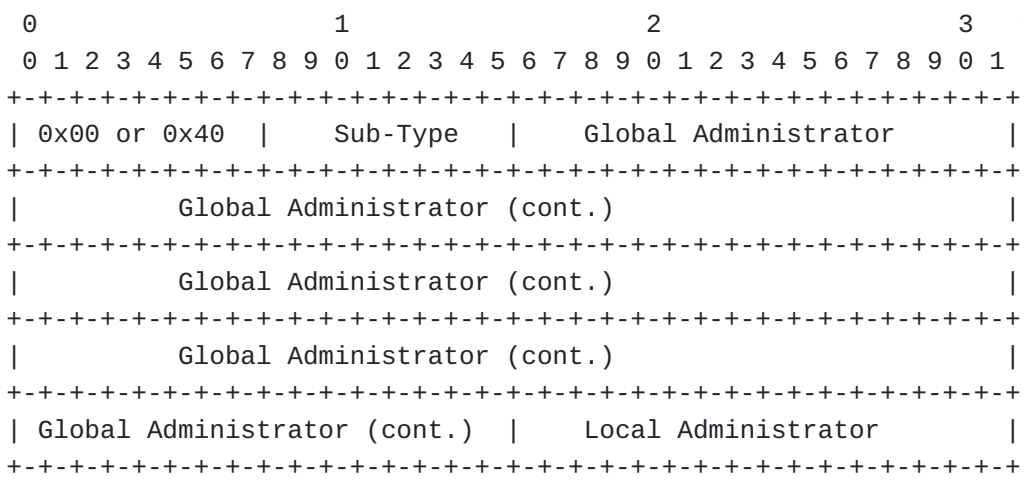
Attribute, or internal fields of structured BGP Path Attributes. Examples of a stand-alone BGP Path Attribute may be [RFC1997] classic BGP Communities or [RFC8092] Large BGP Communities. Examples of internal community values may be Bitmask Route Targets [I-D.zzhang-idr-bitmask-route-target] defined inside a BGP Wide Community Container, or newly defined sub-TLVs in a BGP Tunnel Encapsulation Attribute [I-D.ietf-idr-tunnel-encaps].

The Community-like Attribute is encoded in the CLA Value field. Sufficient octets are encoded for the Prefix Length of this NLRI.

4. Examples

4.1. IPv6 Specific Extended Communities

[RFC5701] defines IPv6 Specific Extended Communities. Its structure, from the RFC is:



Where Global Administrator is 16 octets in length, and Local Administrator is 2 octets in length. The community is a fixed length of 20 octets.

The Community Selector for Large BGP Communities is assigned 1, per this document.

The encoding for a Generic Route Constraint Distribution Mechanism NLRI for an IPv6 Specific Extended Community for an Origin AS of 64511, for the IPv6 Specific Extended Community [2001:DB8::2]:100 would be:

```
NLRI length          = 0xd0 (208)
Origin AS            = 0x0000fbff (64511)
Community Selector   = 0x0001 (2)         # IPv6 Specific
                                          # Extended Community
Community-like Value = 0x0001000f (65551) # Global Administrator
                       0x2001 0DB8 0000 0000 0000 0000 0000 0000
                       0x0000 0000 0000 0000 0000 0000 0000 0002
                                          # Global Administrator
                       0x00000064 (100)   # Local Administrator
```

4.2. Large BGP Communities

[RFC8092] defines Large BGP Communities. Its structure, from the RFC is:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Global Administrator                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Local Data Part 1                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Local Data Part 2                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where each of the fields Global Administrator, Local Data Part 1, and Local Data Part 2 are 4 octets in length. The community is a fixed length of 12 octets.

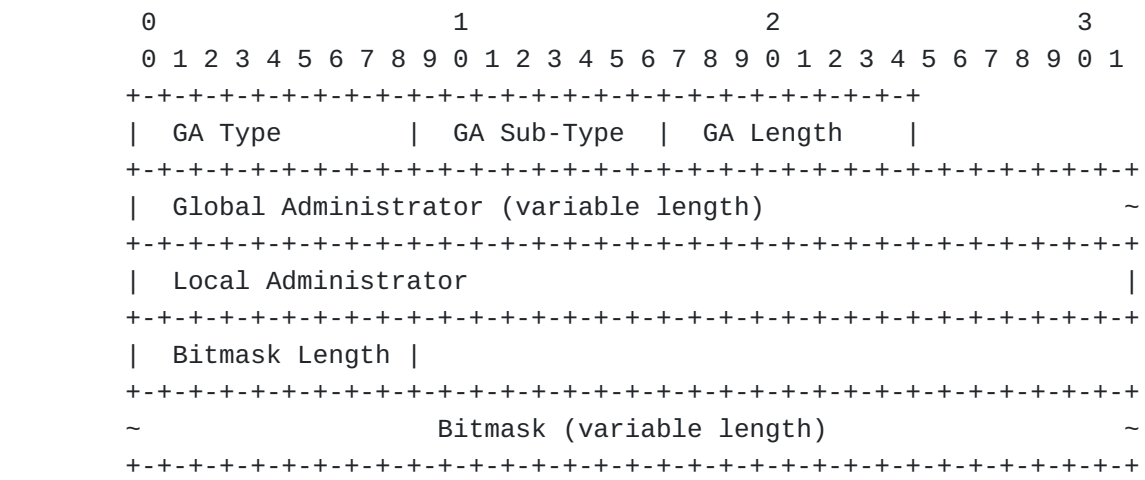
The Community Selector for Large BGP Communities is assigned 2, per this document.

The encoding for a Generic Route Cosntraint Mechanism NLRI for Large BGP Communities for an Origin AS of 64511, for Large BGP Community 65551:100:16777215 would be:

```
NLRI length          = 0x90 (144)
Origin AS            = 0x0000fbff (64511)
Community Selector   = 0x0001 (2)           # Large BGP Community
Community-like Value = 0x0001000f (65551)   # Global Administrator
                       0x00000064 (100)     # Local Data Part 1
                       0x00ffffff (16777215) # Local Data Part 2
```

4.3. Bitmask Route Target

[I-D.zzhang-idr-bitmask-route-target] defines Bitmask Route Targets. Bitmask Route Targets are encoded within the BGP Community Container Path Attribute, which is defined in [I-D.ietf-idr-wide-bgp-communities]. The structure of the Bitmask Route Target, from the Internet-Draft, is:



GA Type, GA Sub-Type, and GA Length are 1 octet in length.

Local Administrator is 4 octets in length.

The Bitmask is a number of octets that will fit the Bitmask Length.

The following GA Types and corresponding lengths are defined:

- 1: AS Number, 4 octets

- 2: IPv4 Address, 4 octets

- 3: IPv6 Address, 16 octets

The Community Selector for Bitmask Route Targets is assigned 3, per this document.

The Bitmask Route Target, a Community-like attribute, is carried as the payload (that is, the value portion) of another Path Attribute. The Generic Route Constraint Distribution Mechanism NLRI is not constructed to match any of the outer portions of the Community Container; rather it matches only the payload, that is, the Bitmask Route Target itself.

4.3.1. AS Number Bitmask Route Target

The encoding for a Generic Route Constraint Distribution Mechanism NLRI for Origin AS 64511 for an AS-Number based Bitmask Route Target for AS 65551 with Local Administrator value 100 and a bitmask of 0xc0ffee (3 octets) would be:

```
NLRI length          = 0xa8 (168)
Origin AS            = 0x0000fbff (64511)
Community Selector   = 0x0002 (3)         # Bitmask Route Target
Community-Like Value = 0x01 (1)           # GA Type AS Number
                       0x02 (2)           # GA Sub-Type (Route Target)
                       0x04 (4)           # GA Length
                       0x0001000f (65551) # Global Administrator
                       0x00000064 (100)   # Local Administrator
                       0x03 (3)           # Bitmask Length
                       0xc0ffee           # Bitmask
```

4.3.2. IPv6 Address Bitmask Route Target

The encoding for a Generic Route Constraint Distribution Mechanism NLRI for Origin AS 64511 for an AS-Number based Bitmask Route Target for 2001:DB8::2 with Local Administrator value 100 and a bitmask of 0xc0ffee (3 octets) would be:

```
NLRI length          = 0xf108 (264)
Origin AS            = 0x0000fbff (64511)
Community Selector   = 0x0002 (2)         # Bitmask Route Target
Community-Like Value = 0x01 (1)           # GA Type IPv6 Address
                       0x02 (2)           # GA Sub-Type (Route Target)
                       0x10 (16)          # GA Length
                       0x2001 0DB8 0000 0000 0000 0000 0000 0000
                       0x0000 0000 0000 0000 0000 0000 0000 0002
                                          # Global Administrator
                       0x00000064 (100)   # Local Administrator
                       0x03 (3)           # Bitmask Length
                       0xc0ffee           # Bitmask
```

5. Security Considerations

This document does not change security aspects discussed in [RFC4684].

6. IANA Considerations

This document requests IANA to assign a new SAFI, the "Generic Route Constraint Distribution Mechanism" from the First Come First Served "Subsequent Address Family Identifiers (SAFI) Parameters" registry.

This documument requests IANA to create a new registry, the Generic Route Constraint CLA Selector Registry. It should have the following initial values and registration policies assigned:

Value	Description	Defining Specification for Community-like attribute (CLA)	Reference for this Value
0	RESERVED	-	This document
1	IPv6 Address Specific BGP Extended Communities	RFC 5701	This document
2	Large BGP Communities	RFC 8092	This document
3	Bitmask Route Targets	draft-zzhang-idr-bitmask-route-target	This document
4..64511	Available for first come, first served allocation.		
255	RESERVED	-	This document

Table 1

7. Acknowledgements

The authors would like to thank John Scudder for his comments and suggestions.

8. References

8.1. Normative 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>.

[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>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, DOI 10.17487/RFC4684, November 2006, <https://www.rfc-editor.org/info/rfc4684>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <https://www.rfc-editor.org/info/rfc5575>.

[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>.

8.2. Informative References

[I-D.ietf-idr-bgp-ipv6-rt-constrain] Patel, K., Raszuk, R., Djernaes, M., Dong, J., and M. Chen, "IPv6 Extensions for Route Target Distribution", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-ipv6-rt-constrain-12, 26 April 2018, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ipv6-rt-constrain-12>.

[I-D.ietf-idr-rtc-no-rt] Rosen, E. C., Patel, K., Haas, J., and R. Raszuk, "Route Target Constrained Distribution of Routes with no Route Targets", Work in Progress, Internet-Draft, draft-ietf-idr-rtc-no-rt-12, 1 October 2019, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-rtc-no-rt-12>.

[I-D.ietf-idr-tunnel-encaps] Patel, K., Van de Velde, G., Sangli, S. R., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-tunnel-encaps-22, 7 January 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-tunnel-encaps-22>.

[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-11, 9 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-wide-bgp-communities-11>.

[I-D.zzhang-idr-bitmask-route-target] Zhang, Z. J., Sangli, S. R., and J. Haas, "Bitmask Route Target", Work in Progress, Internet-Draft, draft-zzhang-idr-bitmask-route-target-01, 21 April 2021, <https://datatracker.ietf.org/doc/html/draft-zzhang-idr-bitmask-route-target-01>.

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <https://www.rfc-editor.org/info/rfc1997>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC5398] Huston, G., "Autonomous System (AS) Number Reservation for Documentation Use", RFC 5398, DOI 10.17487/RFC5398, December 2008, <https://www.rfc-editor.org/info/rfc5398>.

[RFC5701] Rekhter, Y., "IPv6 Address Specific BGP Extended Community Attribute", RFC 5701, DOI 10.17487/RFC5701, November 2009, <https://www.rfc-editor.org/info/rfc5701>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC8092] Heitz, J., Ed., Snijders, J., Ed., Patel, K., Bagdonas, I., and N. Hilliard, "BGP Large Communities Attribute", RFC 8092, DOI 10.17487/RFC8092, February 2017, <https://www.rfc-editor.org/info/rfc8092>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

Appendix A. Open Issues

* How should BGP Routes with no communities of a given type be handled? The scenario covered in [I-D.ietf-idr-rtc-no-rt] becomes potentially far more common.

Appendix B. Change Log

* 01 -> 02: Update text for bitmask route target definition. Add open issues.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

Jeffrey Haas
Juniper Networks

Email: jhaas@juniper.net