

BESS Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 24, 2015

H. Jeng
AT&T
L. Jalil
Verizon
R. Bonica
Juniper Networks
K. Patel
Cisco Systems
L. Yong
Huawei Technologies
February 20, 2015

Covering Prefixes Outbound Route Filter for BGP-4
draft-ietf-bess-orf-covering-prefixes-04

Abstract

This document defines a new Outbound Route Filter (ORF) type, called the "Covering Prefixes ORF (CP-ORF)". CP-ORF is applicable in Virtual Hub-and-Spoke VPNs. It also is applicable in BGP/MPLS Ethernet VPN (EVPN) networks.

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 <http://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 24, 2015.

Copyright Notice

Copyright (c) 2015 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 (<http://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
1.1.	Terminology	3
2.	CP-ORF Encoding	3
3.	Processing Rules	6
4.	Applicability In Virtual Hub-and-Spoke VPNs	9
4.1.	Multicast Considerations	12
5.	Applicability In BGP/MPLS Ethernet VPN (EVPN)	12
6.	Clean-up	16
7.	IANA Considerations	16
8.	Security Considerations	17
9.	Contributors	17
10.	Acknowledgements	17
11.	References	17
11.1.	Normative References	17
11.2.	Informative References	18
	Authors' Addresses	19

[1.](#) Introduction

A BGP [[RFC4271](#)] speaker can send Outbound Route Filters (ORF) [[RFC5291](#)] to a peer. The peer uses ORFs to filter routing updates that it sends to the BGP speaker. Using ORF, a BGP speaker can realize a "route pull" paradigm, in which the BGP speaker, on demand, pulls certain routes from the peer.

This document defines a new ORF-type, called the "Covering Prefixes ORF (CP-ORF)". CP-ORF is applicable in Virtual Hub-and-Spoke VPNs [[RFC7024](#)] [[RFC4364](#)]. It also is applicable BGP/MPLS Ethernet VPN (EVPN) [[I-D.ietf-l2vpn-evpn](#)] networks.

1.1. Terminology

This document uses the following terms:

- o Address Family Identifier (AFI) - defined in [[RFC4760](#)]
- o Subsequent Address Family Identifier (SAFI) - defined in [[RFC4760](#)]
- o VPN IP Default Route - defined in [[RFC7024](#)].
- o V-Hub - defined in [[RFC7024](#)].
- o V-Spoke - defined in [[RFC7024](#)].
- o BGP/MPLS Ethernet VPN (EVPN) - defined in [[I-D.ietf-l2vpn-evpn](#)]
- o EVPN Instance (EVI) - defined in [[I-D.ietf-l2vpn-evpn](#)]
- o Unknown MAC Route (UMR) - A regular EVPN MAC/IP Advertisement route where the MAC Address Length is set to 48 and the MAC address to 00:00:00:00:00:00
- o Default MAC Gateway (DMG) - An EVPN PE that advertises a UMR

2. CP-ORF Encoding

[RFC 5291](#) augments the BGP ROUTE-REFRESH message so that it can carry ORF entries. When the ROUTE-REFRESH message carries ORF entries, it includes the following fields:

- o AFI [[IANA.AFI](#)]
- o SAFI [[IANA.SAFI](#)]
- o When-to-refresh (IMMEDIATE or DEFERRED)
- o ORF Type
- o Length (of ORF entries)

The ROUTE-REFRESH message also contains a list of ORF entries. Each ORF entry contains the following fields:

- o Action (ADD, REMOVE, or REMOVE-ALL)
- o Match (PERMIT or DENY)

The ORF entry may also contain Type-specific information. Type-specific information is present only when the Action is equal to ADD or REMOVE. It is not present when the Action is equal to REMOVE-ALL.

When the BGP ROUTE-REFRESH message carries CP-ORF entries, the following conditions MUST be true:

- o ORF Type MUST be equal to CP-ORF (65).
- o The AFI MUST be equal to IPv4, IPv6 or L2VPN
- o If the AFI is equal to IPv4 or IPv6, SAFI MUST be equal to MPLS-labeled VPN address
- o If the AFI is equal to L2VPN, the SAFI MUST be equal to BGP EVPN
- o Match field MUST be equal to PERMIT

Figure 1 depicts the encoding of the CP-ORF type-specific information.

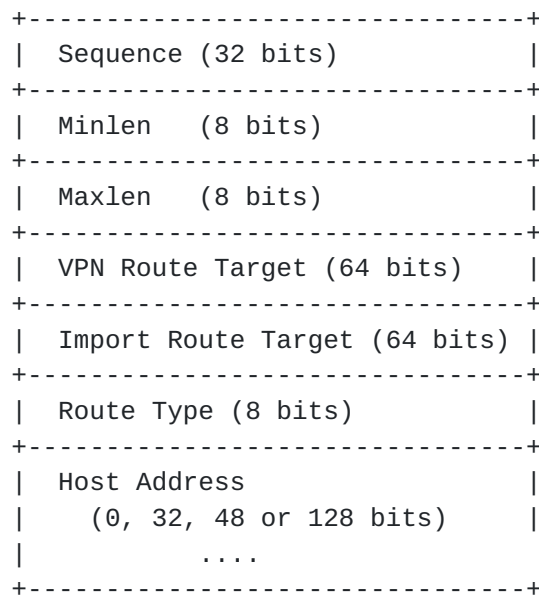


Figure 1: CP-ORF Type-specific Encoding

The Sequence field specifies the relative ordering of the entry among all CP-ORF entries.

The CP-ORF recipient uses the following fields to identify routes that match the CP-ORF:

- o Minlen (measured in bits)

- o Maxlen (measured in bits)
- o VPN Route Target
- o Route Type
- o Host Address

See [Section 3](#) for details.

The CP-ORF recipient marks routes that match CP-ORF with the Import Route Target before advertising those routes to the CP-ORF originator. See [Section 3](#) for details.

If the ROUTE-REFRESH AFI is equal to IPv4:

- o The value of Minlen MUST be less than or equal to 32
- o The value of Maxlen MUST be less than or equal to 32
- o The value of Minlen MUST be less than or equal to the value of Maxlen
- o The value of Route Type MUST be 0 (i.e., RESERVED)
- o The Host Address MUST contain exactly 32 bits

If the ROUTE-REFRESH AFI is equal to IPv6:

- o The value of Minlen MUST be less than or equal to 128
- o The value of Maxlen MUST be less than or equal to 128
- o The value of Minlen MUST be less than or equal to the value of Maxlen
- o The value of Route Type MUST be 0 (i.e., RESERVED)
- o The Host Address MUST contain exactly 128 bits

If the ROUTE-REFRESH AFI is equal to L2VPN, the value of Route Type MUST be one of the following values, taken from IANA EVPN Registry [[IANA.EVPN](#)]:

- o 1 - Ethernet Autodiscovery Route
- o 2 - MAC/IP Advertisement Route

- o 3 - Inclusive Multicast Route
- o 4 - Ethernet Segment Route

If the ROUTE-REFRESH AFI is equal to L2VPN and the value of Route Type is equal to Ethernet Autodiscovery Route, Inclusive Multicast Route, or Ethernet Segment Route:

- o The value of Minlen MUST be equal to 0
- o The value of Maxlen MUST be equal to 0
- o The Host Address MUST be absent (i.e., contain 0 bits)

If the ROUTE-REFRESH AFI is equal to L2VPN and the value of Route Type is equal to MAC/IP Advertisement Route:

- o The value of Minlen MUST be less than or equal to 48
- o The value of Maxlen MUST be less than or equal to 48
- o The value of Minlen MUST be less than or equal to the value of Maxlen
- o The Host Address MUST contain exactly 48 bits.

3. Processing Rules

According to [[RFC4271](#)], every BGP speaker maintains a single Loc-RIB. For each of its peers, the BGP speaker also maintains an Outbound Filter and an Adj-RIB-Out. The Outbound Filter defines policy that determines which Loc-RIB entries are processed into the corresponding Adj-RIB-Out. Mechanisms such as RT-Constraint [[RFC4684](#)] and ORF [[RFC5291](#)] enable a router's peer to influence the Outbound Filter. Therefore, the Outbound Filter for a given peer is constructed using a combination of the locally configured policy and the information received via RT-Constraint and ORF from the peer.

Using this model we can describe the operations of CP-ORF as follows:

When a BGP speaker receives a ROUTE-REFRESH message that contains a CP-ORF, and that ROUTE-REFRESH message violates any of the encoding rules specified in [Section 2](#), the BGP speaker MUST ignore the entire ROUTE-REFRESH message. It SHOULD also log the event. However, an implementation MAY apply logging thresholds to avoid excessive messaging or log file overflow.

Otherwise, the BGP speaker processes each CP-ORF entry as indicated by the Action field. If the Action is equal to ADD, the BGP speaker adds the CP-ORF entry to the Outbound Filter associated with the peer in the position specified by the Sequence field. If the Action is equal to REMOVE, the BGP speaker removes the CP-ORF entry from the Outbound Filter. If the Action is equal to REMOVE-ALL, the BGP speaker removes all CP-ORF entries from the Outbound Filter.

Whenever the BGP speaker applies an Outbound Filter to a route contained in its Loc-RIB, it evaluates the route in terms of the CP-ORF entries first. It then evaluates the route in terms of the remaining, non-CP-ORF entries. The rules for the former are described below. The rules for the latter are outside the scope of this document.

The following route types can match a CP-ORF:

- o IPv4-VPN
- o IPv6-VPN
- o L2VPN

In order for an IPv4-VPN route or IPv6-VPN route to match a CP-ORF, all of the following conditions MUST be true:

- o the route carries an RT whose value is the same as the CP-ORF VPN Route Target
- o the route prefix length is greater than or equal to the CP-ORF Minlen plus 64 (i.e., the length of a VPN Route Distinguisher)
- o the route prefix length is less than or equal to the CP-ORF Maxlen plus 64 (i.e., the length of a VPN Route Distinguisher)
- o ignoring the Route Distinguisher, the leading bits of the route prefix are identical to the leading bits of the CP-ORF Host Address. CP-ORF Minlen defines the number of bits that must be identical.
- o Loc-RIB does not contain a more specific route that also satisfies all of the above listed conditions.

The BGP speaker ignores Route Distinguishers when determining whether a prefix matches a host address. For example, assume that a CP-ORF carries the following information:

- o Minlen equal to 1

- o Maxlen equal to 32
- o Host Address equal to 192.0.2.1

Assume also that Loc-RIB contains routes for the following IPv4-VPN prefixes, and that all of these routes carry an RT whose value is the same as the CP-ORF VPN Route Target:

- o 1:0.0.0.0/64.
- o 2:192.0.2.0/88
- o 3:192.0.2.0/89

Only the prefix 3:192.0.2.0/89 matches the CP-ORF. The prefix 1:0.0.0.0/64 does not match, because its length (64) is less than the CP-ORF Minlen (1) plus the length of an L3VPN Route Distinguisher (64). If Loc-RIB did not contain the prefix 3:192.0.2.0/89, 2:192.0.2.0/88 would match the CP-ORF. However, because Loc-RIB also contains a more specific covering route (3:192.0.2.0/89), 2:192.0.2.0/88 does not match. Only 3:192.0.2.0/89 satisfies all of the above listed match criteria. Note that the matching algorithm ignored Route Distinguishers.

In order for an EVPN route to match a CP-ORF, all of the following conditions MUST be true:

- o the EVPN route type is equal to the CP-ORF Route Type
- o the route carries an RT whose value is equal to the CP-ORF VPN Route Target

In addition, if the CP-ORF Route Type is equal to MAC/IP Advertisement Route, the following conditions also MUST be true:

- o the EVPN Route MAC Address Length is greater than or equal to the CP-ORF Minlen plus 64 (i.e., the length of a VPN Route Distinguisher)
- o the EVPN Route MAC Address Length is less than or equal to the CP-ORF Maxlen plus 64 (i.e., the length of a VPN Route Distinguisher)
- o ignoring the Route Distinguisher, the leading bits of the EVPN Route MAC Address are identical to the leading bits of the CP-ORF Host Address. CP-ORF Minlen defines the number of bits that must be identical.

If a route matches the selection criteria of a CP-ORF entry, and it does not violate any subsequent rule specified by the Outbound Filter (e.g., rules that reflect local policy, or rules that are due to RT-Constrains), the BGP speaker places the route into the Adj-RIB-Out. In Adj-RIB-Out, the BGP speaker adds the CP-ORF Import Route Target to the list of Route Targets that the route already carries. The BGP speaker also adds a Transitive Opaque Extended Community [[RFC4360](#)] with subtype equal to CP-ORF (0x03). As a result of being placed in Adj-RIB-Out, the route is advertised to the peer associated with the Adj-RIB-Out.

Receiving CP-ORF entries with REMOVE or REMOVE-ALL Actions may cause a route that has previously been installed in a particular Adj-RIB-Out be excluded from that Adj-RIB-Out. In this case, as specified in [[RFC4271](#)], "the previously advertised route in that Adj-RIB-Out MUST be withdrawn from service by means of an UPDATE message".

[RFC 5291](#) states that a BGP speaker should respond to a ROUTE REFRESH message as follows:

"If the When-to-refresh indicates IMMEDIATE, then after processing all the ORF entries carried in the message the speaker re-advertises to the peer routes from the Adj-RIB-Out associated with the peer that have the same AFI/SAFI as what is carried in the message, and taking into account all the ORF entries for that AFI/SAFI received from the peer. The speaker MUST re-advertise all the routes that have been affected by the ORF entries carried in the message, but MAY also re-advertise the routes that have not been affected by the ORF entries carried in the message."

When the ROUTE-REFRESH message includes one or more CP-ORF entries, the BGP speaker MUST re-advertise routes that have been affected by ORF entries carried by the message. While the speaker MAY also re-advertise the routes that have not been affected by the ORF entries carried in the message, this memo RECOMMENDS not to re-advertise the routes that have not been affected.

4. Applicability In Virtual Hub-and-Spoke VPNs

In a Virtual Hub-and-Spoke environment, VPN sites are attached to Provider Edge (PE) routers. For a given VPN, a PE router acts in exactly one of the following roles:

- o As neither a V-hub nor a V-Spoke
- o As a V-hub
- o As a V-spoke

To illustrate CP-ORF operation in conjunction with Virtual Hub-and-Spoke assume the following:

- o One of the sites in a particular VPN, RED-VPN, is connected to a PE that acts as neither a V-hub nor a V-Spoke for RED-VPN. We refer to this PE as PE1.
- o Another site in RED-VPN is connected to another PE, and that PE acts as a V-hub for RED-VPN. We refer to this PE as V-hub1.
- o Yet another site in RED-VPN is connected to another PE, and that PE acts as a V-spoke for RED-VPN. We refer to this PE as V-spoke1.

All of these PEs advertise RED-VPN routes to a route reflector (RR). They mark these routes with a route target, which we will call RT-RED. In particular, PE1 advertises a RED-VPN route to a prefix that we will call P. P covers a host address, that we will call H.

For the purpose of illustration also assume that the PEs and the RRs use Route Target Constraint [[RFC4684](#)].

V-hub1 serves the RED-VPN. Therefore, V-hub1 advertises a VPN IP default route for the RED-VPN to the RR, carrying the route target RT-RED-FROM-HUB1.

V-spoke1 establishes a BGP session with the RR, negotiating the CP-ORF capability, as well as the Multiprotocol Extensions Capability [[RFC4760](#)]. Upon establishment of the BGP session, the RR does not advertise any routes to V-spoke1. The RR will not advertise any routes until it receives either a ROUTE-REFRESH message or a BGP UPDATE message containing a Route Target Membership NLRI [[RFC4684](#)].

Immediately after the BGP session is established, V-spoke1 sends the RR a BGP UPDATE message containing a Route Target Membership NLRI. The Route Target Membership NLRI specifies RT-RED-FROM-HUB1 as its route target. In response to the BGP-UPDATE message, the RR advertises the VPN IP default route for the RED-VPN to V-spoke1. This route carries the route target RT-RED-FROM-HUB1. V-spoke1 subjects this route to its import policy and accepts it because it carries the route target RT-RED-FROM-HUB1.

Now, V-spoke1 begins normal operation, sending all of its RED-VPN traffic through V-hub1. At some point, V-spoke1 determines that it might benefit from a more direct route to H. (Criteria by which V-spoke1 determines that it needs a more direct route to H are beyond the scope of this document.)

In order to discover a more direct route, V-spoke1 assigns a unique numeric identifier to H. V-spoke1 then sends a ROUTE-REFRESH message to the RR, containing the following information:

- o AFI is equal to IPv4 or IPv6, as appropriate
- o SAFI is equal to "MPLS-labeled VPN address"
- o When-to-refresh is equal IMMEDIATE
- o Action is equal to ADD
- o Match is equal to PERMIT
- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal to the identifier associated with H
- o CP-ORF Minlen is equal to 1
- o CP-ORF Maxlen is equal to 32 or 128, as appropriate
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED-FROM-HUB1
- o CP-ORF Route Type is equal to 0 (i.e., undefined)
- o CP-ORF Host Address is equal H

Upon receipt of the ROUTE-REFRESH message, the RR MUST ensure that it carries all routes belonging to the RED-VPN. In at least one special case, where all of the RR clients are V-spokes and none of the RR clients are V-hubs, the RR will lack some or all of the required RED-VPN routes. So, the RR sends a BGP UPDATE message containing a Route Target Membership NLRI for VPN-RED to all of its peers. This causes the peers to advertise VPN-RED routes to the RR, if they have not done so already.

Next, the RR adds the received CP-ORF to the Outbound Filter associated with V-spoke1. Using the procedures in [Section 3](#), the RR determines whether any of the routes in its Loc-RIB satisfy the selection criteria of the newly updated Outbound Filter. If any routes satisfy the match criteria, they are added to the Adj-RIB-Out associated with V-spoke1. In Adj-RIB-Out, the RR adds RT-RED-FROM-HUB1 to the list of Route Targets that the route already carries. The RR also adds a Transitive Opaque Extended Community [[RFC4360](#)] with subtype equal to CP-ORF. Finally, RR advertises the newly added

routes to V-spoke1. In this example, the RR advertises P to V-Spoke1 with a next-hop of PE1.

V-spoke1 subjects the advertised routes to its import policy and accepts them because they carry the route target RT-RED-FROM-HUB1.

V-spoke1 may repeat this process whenever it discovers another flow that might benefit from a more direct route to its destination.

4.1. Multicast Considerations

When applying Multicast VPN [[RFC6513](#)][RFC6514] procedures, routes bearing a Transitive Opaque Extended Community [[RFC4360](#)] with subtype equal to CP-ORF MUST NOT be used to determine Eligible Upstream Multicast Hops (UMH).

5. Applicability In BGP/MPLS Ethernet VPN (EVPN)

In a EVPN environment, CE devices are attached to Provider Edge (PE) routers. A CE can be a host, a router or a switch. For a given EVPN Instance (EVI), a PE router acts in exactly one of the following roles:

- o As neither a Default MAC Gateway (DMG) nor a Spoke
- o As a DMG
- o As a Spoke

To illustrate CP-ORF operation in the EVPN environment assume the following:

- o A CE device in a particular EVI, RED-EVI, is connected to a PE that acts as neither a DMG nor a Spoke for RED-EVI. We refer to this PE as PE1.
- o Another CE device in RED-EVI is connected to another PE, and that PE acts as a DMG for RED-EVI. We refer to this PE as DMG1.
- o Yet another CE device in RED-EVI is connected to another PE, and that PE acts as a Spoke for RED-EVI. We refer to this PE as Spoke1.

All of these PEs advertise RED-EVI routes to a RR. They mark these routes with a route target, which we will call RT-RED. In particular, PE1 advertises a RED-EVI route to a MAC Address that we will call M.

The RED-EVI VRFs on all of these PEs are provisioned to import EVPN routes that carry RT-RED.

Since DMG1 acts as a DMG for RED-EVI, DMG1 advertises a Unknown MAC Route (UMR) for the RED-EVI to the RR, carrying the route target RT-RED. The UMR is characterized as follows:

- o EVPN Route Type is equal to MAC/IP Advertisement Route
- o MAC address length is equal to 0
- o IP address length is equal to 0

Spoke1 establishes a BGP session with the RR, negotiating the CP-ORF capability, as well as the Multiprotocol Extensions Capability [[RFC4760](#)]. Upon establishment of the BGP session, the RR does not advertise any routes to Spoke1. The RR will not advertise any routes until it receives a ROUTE-REFRESH message.

Immediately after the BGP session is established, Spoke1 sends the RR a ROUTE REFRESH message containing the following information:

- o AFI is equal to L2VPN
- o SAFI is equal to BGP EVPN
- o When-to-refresh is equal IMMEDIATE
- o Action is equal to ADD
- o Match is equal to PERMIT

The ROUTE REFRESH message also contains four ORF entries. The first ORF entry contains the following information:

- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal 1
- o CP-ORF Minlen is equal to 0
- o CP-ORF Maxlen is equal to 0
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED
- o CP-ORF Route Type is equal to 1 (Ethernet Autodiscovery Route)

The second ORF entry contains the following information:

- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal 2
- o CP-ORF Minlen is equal to 0
- o CP-ORF Maxlen is equal to 0
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED
- o CP-ORF Route Type is equal to 2 (MAC/IP Advertisement Route)

The third ORF entry contains the following information:

- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal 3
- o CP-ORF Minlen is equal to 0
- o CP-ORF Maxlen is equal to 0
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED
- o CP-ORF Route Type is equal to 3 (Inclusive Multicast Route)

The fourth ORF entry contains the following information:

- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal 4
- o CP-ORF Minlen is equal to 0
- o CP-ORF Maxlen is equal to 0
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED
- o CP-ORF Route Type is equal to 4 (Ethernet Segment Route)

In response to the ROUTE REFRESH message, the RR advertises the following to V-spoke1:

- o All Ethernet Autodiscovery Routes belonging to RED-EVI
- o A UMR advertised by DMG1 and belonging to RED-EVI
- o All Inclusive Multicast Routes belonging to RED-EVI
- o All Ethernet Segment Routes belonging to RED-EVI

All of these routes carries the route target RT-RED. Spoke1 subjects these routes to its import policy and accepts them because they carry the route target RT-RED.

Now, Spoke1 begins normal operation, sending all of its RED-VPN traffic through DMG1. At some point, Spoke1 determines that it might benefit from a more direct route to M. (Criteria by which Spoke1 determines that it needs a more direct route to M are beyond the scope of this document.)

In order to discover a more direct route, Spoke1 assigns a unique numeric identifier to M. V-spoke1 then sends a ROUTE-REFRESH message to the RR, containing the following information:

- o AFI is equal to L2VPN
- o SAFI is equal to BGP EVPN
- o When-to-refresh is equal IMMEDIATE
- o Action is equal to ADD
- o Match is equal to PERMIT
- o ORF Type is equal to CP-ORF
- o CP-ORF Sequence is equal to the identifier associated with M
- o CP-ORF Minlen is equal to 1
- o CP-ORF Maxlen is equal to 48
- o CP-ORF VPN Route Target is equal to RT-RED
- o CP-ORF Import Route Target is equal to RT-RED
- o CP-ORF Route Type is equal to 2 (i.e., MAC/IP Advertisement Route)

- o CP-ORF Host Address is equal M

Next, the RR adds the received CP-ORF to the Outbound Filter associated with Spoke1. Using the procedures in [Section 3](#), the RR determines whether any of the routes in its Loc-RIB satisfy the selection criteria of the newly updated Outbound Filter. If any routes satisfy the match criteria, they are added to the Adj-RIB-Out associated with Spoke1. The RR adds a Transitive Opaque Extended Community [[RFC4360](#)] with subtype equal to CP-ORF. Note that as these routes are added to the Adj-RIB-Out, the RR does not change the list of Route Targets that the route already carries. Finally, RR advertises the newly added routes to V-spoke1. In this example, the RR advertises M to V-Spoke1 with a next-hop of PE1.

Spoke1 subjects the advertised routes to its import policy and accepts them because they carry the route target RT-RED.

Spoke1 may repeat this process whenever it discovers another flow that might benefit from a more direct route to its destination.

Note that in general an EVI may have more than one DMG, in which case each spoke would receive a UMR from each of them. The spoke should follow its local route selection procedures to select one of them as the "best", and use the selected one.

6. Clean-up

Each CP-ORF consumes memory and compute resources on the device that supports it. Therefore, in order to obtain optimal performance, BGP speakers periodically evaluate all CP-ORFs that they have originated and remove unneeded CP-ORFs. The criteria by which a BGP speaker identifies unneeded CP-ORF entries is a matter of local policy, and is beyond the scope of this document.

7. IANA Considerations

This memo uses code points from the first-come-first-served range of the following registries:

+-----+-----+	
Registry	Code Point
+-----+-----+	
BGP Outbound Route Filtering (ORF) Types	CP-ORF (65)
Transitive Opaque Extended Community Sub-Type	CP-ORF (0x03)
+-----+-----+	

IANA is requested to update the above mentioned registry entries so that they include a stable reference to this memo.

8. Security Considerations

Each CP-ORF consumes memory and compute resources on the device that supports it. Therefore, a device supporting CP-ORF takes the following steps to protect itself from oversubscription:

- o When negotiating the ORF capability, advertise willingness to receive the CP-ORF only to known, trusted iBGP peers
- o Enforce a per-peer limit on the number of CP-ORFs that can be installed at any given time. Ignore all requests to add CP-ORFs beyond that limit

[Section 5 of RFC 5291](#) describes how BGP speakers use the Outbound Route Filtering Capability to advertise their intention to send CP-ORFs and their willingness to accept CP-ORFs.

Security considerations for BGP are presented in [RFC4271](#) while further security analysis of BGP is found in [[RFC6952](#)].

9. Contributors

The following individuals contributed to the development of this document:

- o Yakov Rekhter
- o Xiaohu Xu

10. Acknowledgements

The authors wish to acknowledge Han Nguyen and James Uttaro for their comments and contributions.

11. References

11.1. Normative References

- [I-D.ietf-l2vpn-evpn]
Sajassi, A., Aggarwal, R., Bitar, N., Isaac, A., and J. Uttaro, "BGP MPLS Based Ethernet VPN", [draft-ietf-l2vpn-evpn-11](#) (work in progress), October 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), January 2006.

- [RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", [RFC 4360](#), February 2006.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), February 2006.
- [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](#), November 2006.
- [RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", [RFC 4760](#), January 2007.
- [RFC5291] Chen, E. and Y. Rekhter, "Outbound Route Filtering Capability for BGP-4", [RFC 5291](#), August 2008.
- [RFC6513] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", [RFC 6513](#), February 2012.
- [RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", [RFC 6514](#), February 2012.
- [RFC7024] Jeng, H., Uttaro, J., Jalil, L., Decraene, B., Rekhter, Y., and R. Aggarwal, "Virtual Hub-and-Spoke in BGP/MPLS VPNs", [RFC 7024](#), October 2013.

[11.2.](#) Informative References

- [IANA.AFI] IANA, "Address Family Numbers", <<http://www.iana.org/assignments/address-family-numbers/address-family-numbers.xhtml>>.
- [IANA.EVPN] IANA, "Ethernet VPN (EVPN)", <<http://www.iana.org/assignments/evpn/evpn.xhtml>>.
- [IANA.SAFI] IANA, "Subsequent Address Family Identifiers (SAFI) Parameters", <[http://www.iana.org/assignments/safi-namespace/safi-namespace-2](http://www.iana.org/assignments/safi-namespace/safi-namespace.xhtml#safi-namespace-2)>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", [RFC 6952](#), May 2013.

Authors' Addresses

Huajin Jeng
AT&T

Email: hj2387@att.com

Luay Jalil
Verizon

Email: luay.jalil@verizon.com

Ron Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, Virginia 20170
USA

Email: rbonica@juniper.net

Keyur Patel
Cisco Systems
170 W. Tasman Drive
San Jose, California 95134
USA

Email: keyupate@cisco.com

Lucy Yong
Huawei Technologies
Austin, Texas
USA

Email: lucy.yong@huawei.com

