

Cooperative Route Filtering Capability for BGP-4

[draft-ietf-idr-route-filter-09.txt](#)

1. Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#) except that the right to produce derivative works is not granted.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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.''

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

2. Specification of Requirements

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 [RFC2119](#) [[RFC2119](#)].

[3.](#) Abstract

This document defines a BGP-based mechanism that allows a BGP speaker to send to its BGP peer a set of route filters that the peer would use to constrain/filter its outbound routing updates to the speaker.

[4.](#) Introduction

Currently it is not uncommon for a BGP speaker to receive, and then filter out some unwanted routes from its peers based on its local routing policy. Since the generation and transmission of routing updates by the sender, as well as the processing of routing updates by the receiver consume resources, it may be beneficial if the generation of such unwanted routing updates can be avoided in the first place.

This document defines a BGP-based mechanism that allows a BGP speaker to send to its BGP peer a set of Outbound Route Filters (ORFs). The peer would then apply these filters, in addition to its locally configured outbound filters (if any), to constrain/filter its outbound routing updates to the speaker.

[5.](#) Outbound Route Filter (ORF)

Conceptually an ORF entry is a tuple of the form <AFI/SAFI, ORF-Type, Action, Match, ORF-value>; an ORF consists of one or more ORF entries that have a common AFI/SAFI and ORF-Type. An ORF is identified by <AFI/SAFI, ORF-Type>.

The "AFI/SAFI" component provides a coarse granularity control by limiting the ORF to only the routes whose NLRI matches the "AFI/SAFI" component of the ORF.

The "ORF-Type" component determines the content of the ORF-value.

The "Action" component controls handling of the ORF Request by the remote peer. Action can be one of ADD, REMOVE, REMOVE-ALL. ADD adds an ORF entry to the ORF on the remote peer; REMOVE deletes a previously installed ORF entry on the remote peer; REMOVE-ALL deletes the previously installed entries in the specified ORF on the remote peer.

The "Match" component is used if support matching granularity on a per ORF entry basis is needed, in which case the "Match" component can be one of PERMIT or DENY. The semantics of PERMIT is to ask the peer to pass updates for the set of routes that match the ORF entry.

The semantics of DENY is to ask the peer not to pass updates for the set of routes that match the ORF entry.

[5.1.](#) Communities ORF-Type

The Community ORF-Type allows to express ORFs in terms of BGP Communities [[BGP-COMMUNITIES](#)]. That is, the Communities ORF-Type provides Communities-based route filtering.

Conceptually the ORF-value of the Communities ORF-Type consists of a single Community.

The sender SHOULD set the value of the Match field to PERMIT; the receiver SHOULD ignore the value of the Match field.

The remote peer should consider only those routes whose Communities attribute has at least one Community in common with the Communities list specified in the ORF.

[5.2.](#) Extended Communities ORF-Type

The Extended Community ORF-Type allows to express ORFs in terms of BGP Extended Communities [[BGP-EXT-COMMUNITIES](#)]. That is, the Extended Communities ORF-Type provides Extended Communities-based route filtering.

Conceptually the ORF-value of the Extended Communities ORF-Type consists of a single Extended Community.

The sender SHOULD set the value of the Match field to PERMIT; the receiver SHOULD ignore the value of the Match field.

The remote peer should consider only those routes whose Extended Communities attribute has at least one Extended Community in common

with the Extended Communities list specified in the ORF.

6. Carrying ORF entries in BGP

ORF entries are carried in the BGP ROUTE-REFRESH message [[BGP-RR](#)].

A BGP speaker can distinguish an incoming ROUTE-REFRESH message that carries one or more ORF entries from an incoming plain ROUTE-REFRESH message by using the Message Length field in the BGP message header.

A single ROUTE-REFRESH message could carry multiple ORF entries, as long as all these entries share the same AFI/SAFI.

From the encoding point of view each ORF entry consists of a common part and type-specific part.

The common part consists of <AFI/SAFI, ORF-Type, Action, Match>, and is encoded as follows:

The AFI/SAFI component of an ORF entry is encoded in the AFI/SAFI field of the ROUTE-REFRESH message.

Following the AFI/SAFI component is the one-octet When-to-refresh field. The value of this field can be one of IMMEDIATE (0x01) or DEFER (0x02). The semantics of IMMEDIATE and DEFER are discussed in the "Operation" section of this document.

Following the When-to-refresh field is a collection of one or more ORFs, grouped by ORF-Type.

The ORF-Type component is encoded as a one-octet field.

The Length of ORFs component is a two-octets field that contains the length (in octets) of the ORF entries that follows.

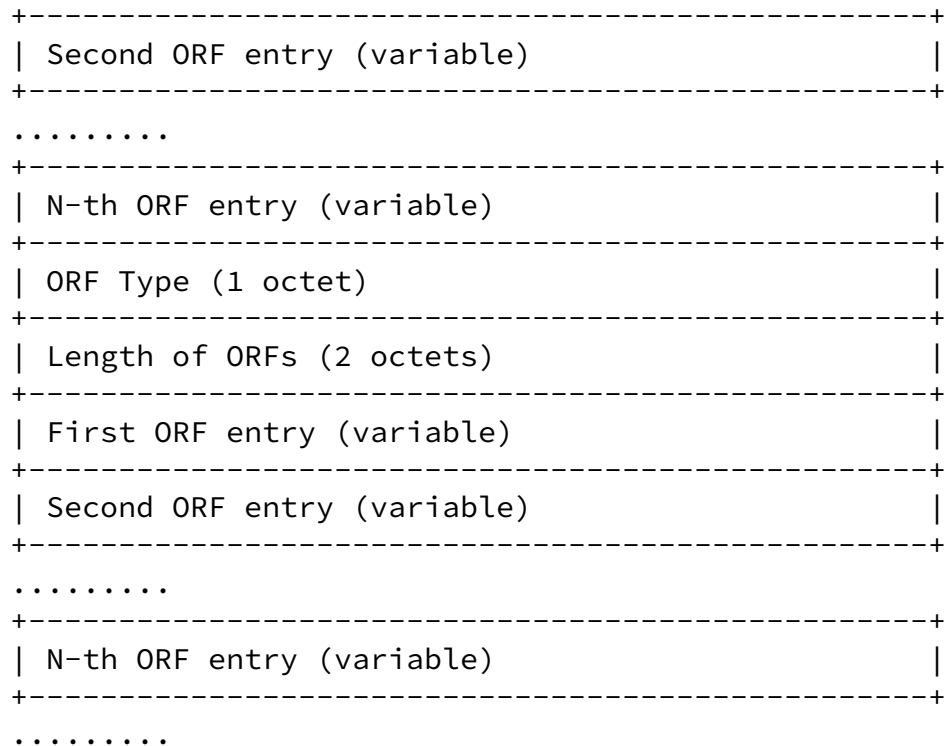
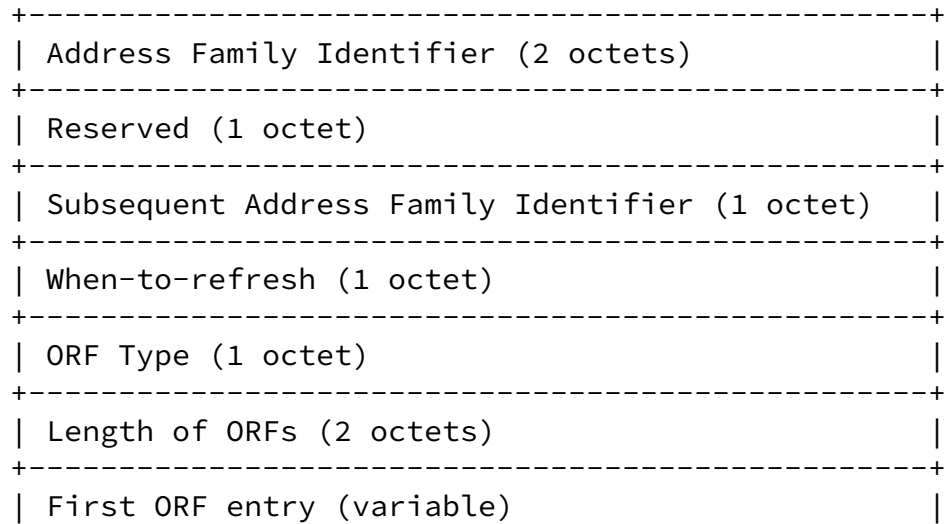


Fig 1. Carrying ORF entries in the ROUTE-REFRESH message

The rest of the components in the common part are encoded in first octet of each ORF-entry as follows (from the most significant to the least significant bit):

Action is a two-bit field. The value of this field is 0 for ADD, 1 for REMOVE, and 2 for REMOVE-ALL.

Match is a one-bit field. The value of this field is 0 for PERMIT and 1 for DENY. This field is significant only when the value of the Action field is either ADD or REMOVE.

Reserved is a 5-bit field. It is set to 0 on transmit and ignored on receive.

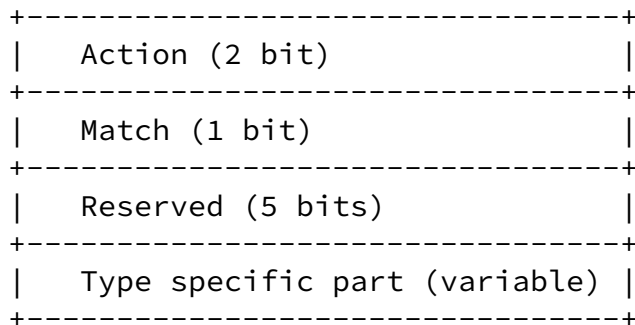


Fig 2. ORF entry encoding

When the Action component of an ORF entry specifies REMOVE-ALL, the entry consists of only the common part.

6.1. Type specific encoding (Communities ORF-Type)

The value of the ORF-Type for the Communities ORF-Type is 2.

The type-specific part of Communities ORF-Type consists of single Community encoded as a four-octets field.

6.2. Type specific encoding (Extended Communities ORF-Type)

The value of the ORF-Type for the Extended Communities ORF-Type is 3.

The type-specific part of Extended Communities ORF-Type consists of a single Extended Community encoded as an eight-octets field.

7. Cooperative Route Filtering Capability

A BGP speaker that is willing to receive ORF entries from its peer, or a BGP speaker that would like to send ORF entries to its peer advertises this to the peer by using the Cooperative Route Filtering Capability, as described below.

The Cooperative Route Filtering Capability is a new BGP capability [[BGP-CAP](#)] defined as follows:

Capability code: 3

Capability length: variable

Capability value: one or more of the following entries:

```
+-----+
| Address Family Identifier (2 octets) |
+-----+
| Reserved (1 octet) |
+-----+
| Subsequent Address Family Identifier (1 octet) |
+-----+
| Number of ORFs (1 octet) |
```

```
+-----+
| ORF Type (1 octet) |
+-----+
| Send/Receive (1 octet) |
+-----+
| ... |
+-----+
| ORF Type (1 octet) |
```

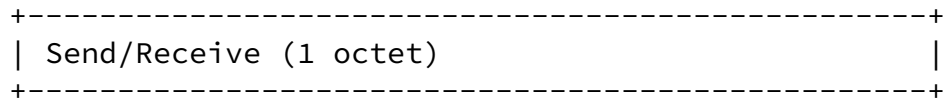


Fig 4. Capability encoding

The use and meaning of these fields are as follows:

Address Family Identifier (AFI):

This field carries the identity of the Network Layer protocol associated with the Network Address that follows. Presently defined values for this field are specified in [RFC1700](#) (see the Address Family Numbers section).

Subsequent Address Family Identifier (SAFI):

This field provides additional information about the type of the Network Layer Reachability Information carried in the attribute.

Number of ORF Types:

This field contains the number of Filter Types to be listed in the following fields.

ORF Type:

This field contains the value of an ORF Type.

Send/Receive:

This field indicates whether the sender is (a) willing to receive ORF entries from its peer (value 1), (b) would like to send ORF entries to its peer (value 2), or (c) both (value 3) for the ORF Type that follows.

A BGP speaker that is willing to receive ORF entries from its peer, or would like to send ORF entries to its peer SHOULD advertise the Cooperative Route Filtering Capability to the peer using BGP Capabilities advertisement [[BGP-CAP](#)].

A BGP speaker that implements the Cooperative Route Filtering Capability must support BGP ROUTE-REFRESH message, as defined in [[BGP-RR](#)]. A BGP speaker that advertises the Cooperative Route Filtering Capability to a peer using BGP Capabilities advertisement [[BGP-CAP](#)] doesn't have to advertise the BGP Route Refresh capability to that peer.

Consider a BGP speaker that advertises the Cooperative Route Filtering Capability indicating its willingness to receive a particular set of <AFI, SAFI, ORF-Type> from its peer, and that receives the Cooperative Route Filtering Capability indicating the desire of the peer to send a particular set <AFI, SAFI, ORF-Type> to the speaker. If for a given <AFI, SAFI> the intersection between these two sets are not-empty, the speaker SHOULD NOT advertise to the peer any routes with that <AFI, SAFI> prior to receiving from the peer any ROUTE-REFRESH message carrying that <AFI, SAFI>, where the message could be either without any ORF entries, or with one or more ORF entry and When-to-refresh field set to IMMEDIATE. If, on the other hand, for a given <AFI, SAFI> the intersection between these two sets is empty, the speaker SHOULD follow normal BGP procedures.

A BGP speaker may send a ROUTE-REFRESH message with one or more ORF entries to its peer only if the peer advertises to the speaker the Cooperative Route Filtering Capability indicating its willingness to receive ORF entries from the speaker, and the speaker advertises to the peer the Cooperative Route Filtering Capability indicating its desire to send ORF entries to the peer. The message may contain only ORF entries of <AFI, SAFI, ORF-type> that the peer is willing to receive, as advertised to the speaker in the Cooperative Route Filtering Capability.

When a BGP speaker receives a ROUTE-REFRESH message with one or more ORF entries from its peer, then the speaker performs the following actions. If the <AFI, SAFI, ORF-type> carried by the message doesn't match <AFI, SAFI, ORF-type> that the speaker is willing to receive from the peer (as advertised to the peer in the Cooperative Route Filtering Capability), the specified ORF is ignored. Otherwise, the speaker modifies the specified ORF, as specified in the ORF entries carried by the message. If any of the fields within an ORF entry contain an unrecognized value, the whole specified ORF is removed.

If the Action component of an ORF entry is REMOVE, but the ORF doesn't contain the specified entry, the entry is ignored.

ORF entries with either REMOVE or REMOVE-ALL can not remove locally configured outbound route filters.

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 received from the peer. However, the routes that have not be affected by the ORF entries carried in the message SHOULD NOT be re-advertised to the peer.

If the When-to-Refresh indicates DEFER, then after processing all the ORF entries carried in the message the speaker defers re-advertisement 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 received from the peer until the speaker receives a subsequent ROUTE-REFRESH message for the same AFI/SAFI either without any ORF entries, or with one or more ORF entries and When-to-refresh set to IMMEDIATE.

If the speaker receives from the peer a ROUTE-REFRESH message without any ORF entries, then the speaker sends to the peer all routes from the Adj-RIB-Out associated with the peer whose AFI/SAFI is the same as what is carried in the message and taking into account the ORF received from the peer.

The set of ORF entries that the speaker sends to the peer expresses the speaker's local preference, that the peer may or may not decide to honor.

During a single BGP session the speaker may pass multiple ORF entries to the peer.

The lifetime of an ORF is the duration of the BGP session during which the ORF is exchanged.

An ORF is removed when the last ORF entry is remove (either via REMOVE-ALL, or via a sequence of REMOVE).

If a particular route maintained by a BGP speaker doesn't match any of the ORF entries of any of the (non-empty) ORFs associated with a particular peer, then this route SHOULD NOT be advertised to the peer.

If a BGP speaker maintains multiple ORFs of different ORF-Types for a

particular peer, then the decision by the speaker to advertise a route to the peer is determined by passing the route through each such ORF, and and-ing the results (and-ing of PERMIT and DENY results in DENY).

9. IANA Considerations

As specified in this document, an ORF entry contains the ORF-Type field. ORF-Type value 0 is reserved. ORF-Type values 1 through 63 are to be assigned by IANA using the "IETF Consensus" policy defined in [RFC2434](#). ORF-Type values 64 through 127 are to be assigned by IANA, using the "First Come First Served" policy defined in [RFC2434](#). ORF-Type values 128 through 255 are vendor-specific, and values in this range are not to be assigned by IANA.

10. Security Considerations

This extension to BGP does not change the underlying security issues.

11. Acknowledgements

Some of the material in the document is "borrowed" from a proposal for selective updates by Yakov Rekhter, Kannan Varadhan, and Curtis Villamizar.

12. Normative References

[BGP-4] Rekhter, Y., and T. Li, "A Border Gateway Protocol 4 (BGP-4)", [RFC 1771](#), March 1995.

[BGP-MP] Bates, T., Chandra, R., Katz, D., and Rekhter, Y., "Multiprotocol Extensions for BGP-4", [RFC 2858](#), June 2000

[BGP-CAP] Chandra, R., Scudder, J., "Capabilities Advertisement with BGP-4", [RFC2842](#), May 2000

[BGP-COMMUNITIES] Chandra, R., Traina, P., and Li, T., "BGP Communities Attribute", [RFC1997](#), August 1996.

[BGP-EXT-COMMUNITIES] Ramachandra, S., Tappan, D., "BGP Extended Communities Attribute", [draft-ramachandra-bgp-ext-communities-02.txt](#)

[BGP-RR] Chen, E., "Route Refresh Capability for BGP-4", [RFC2918](#), September 2000

Chen & Rekhter

[Page 10]

Internet Draft [draft-ietf-idr-route-filter-09.txt](#)

August 2003

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

13. Author Information

Enke Chen
Redback Networks, Inc.
350 Holger Way
San Jose, CA 95134
e-mail: enke@redback.com

Yakov Rekhter
Juniper Networks
e-mail: yakov@juniper.net

