Network Working Group
Request for Comments: 5291
Category: Standards Track

E. Chen
Cisco Systems
Y. Rekhter
Juniper Networks

August 2008

Outbound Route Filtering Capability for BGP-4

Status of This Memo

Abstract

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) that the peer would use to constrain/filter its outbound routing updates to the speaker.

1. Introduction

Currently, it is not uncommon for a BGP speaker [BGP-4] 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.

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 RFC 2119 [RFC2119].

3. Outbound Route Filter (ORF)

This document uses the terms "Address Family Identifier (AFI)" and "Subsequent Address Family Identifier (SAFI)". In the context of this document, the meaning of these terms is the same as in [BGP-MP].

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 Network Layer Reachability Information (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 to support matching granularity on a per ORF entry basis. It can be either 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.

When an ORF is defined, an ORF-specific matching rule MUST be specified so that there is no ambiguity regarding which ORF entry is considered as the matching entry in the ORF when a route is passed through the ORF.

4. 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 MAY carry multiple ORF entries in one or more ORFs, 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, as shown in Figures 1 and 2.

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 either 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 ORF entries" component is a two-octet field that contains the total length (in octets) of the ORF entries that follows for the specified ORF type.

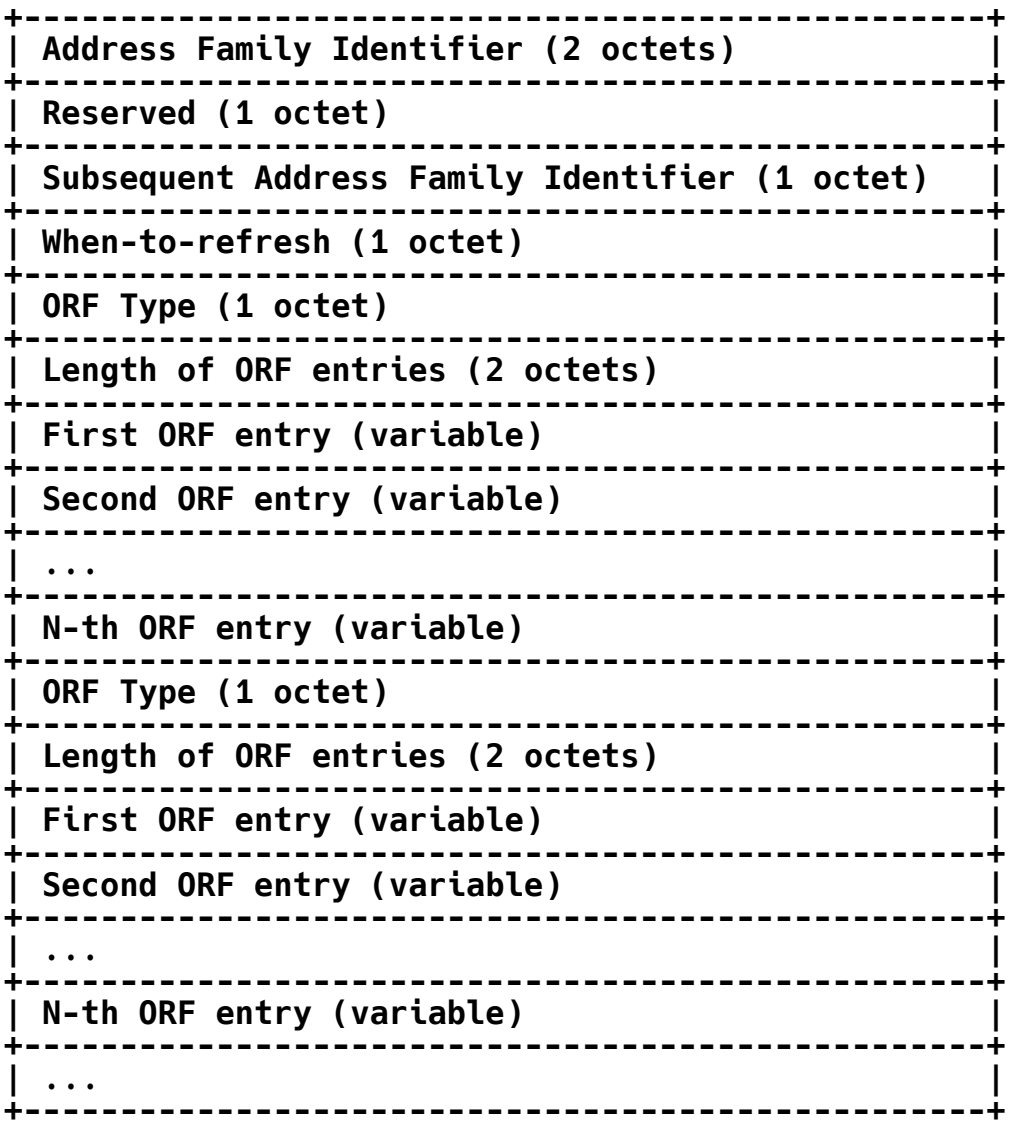

Figure 1: Carrying ORF Entries in the ROUTE-REFRESH Message

The rest of the components in the common part are encoded in the first octet of each ORF-entry (from the most significant to the least significant bit) as shown in Figure 2:

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 receipt.

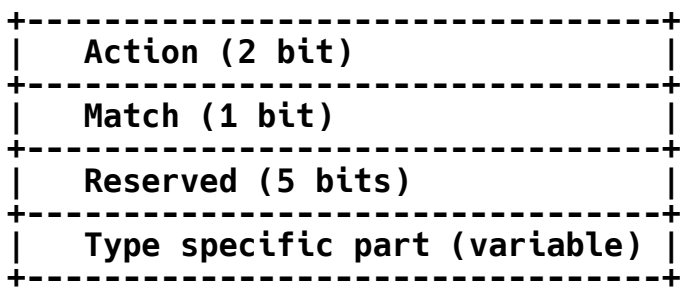

Figure 2: ORF Entry Encoding

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

5. Outbound 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 Outbound Route Filtering Capability, as described below.

The Outbound 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 entries as shown in Figure 3.

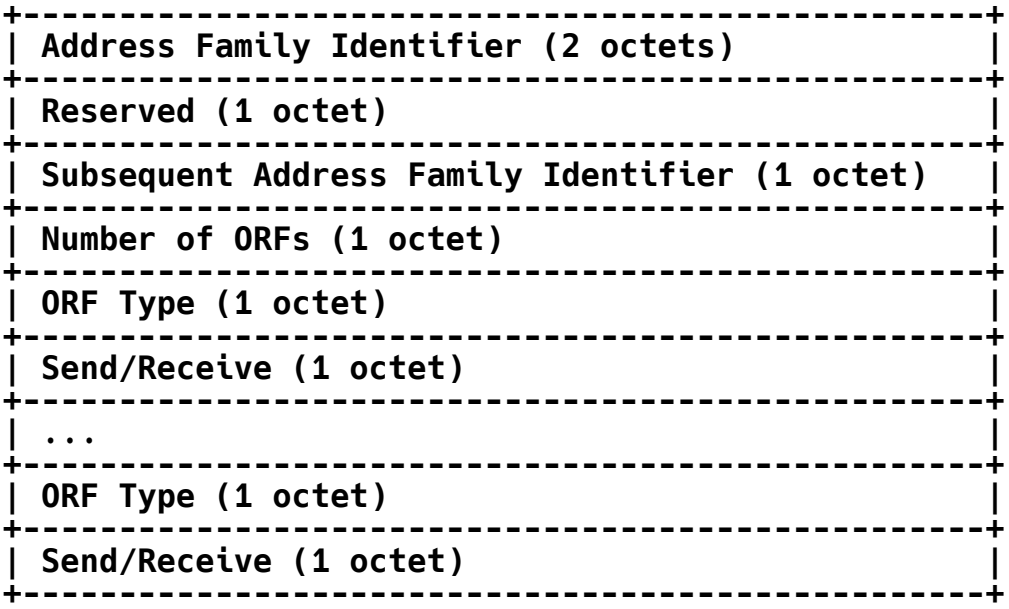

Figure 3: Outbound Route Filtering Capability Encoding

The use and meaning of these fields are as follows:

Address Family Identifier (AFI):

This field is the same as the one used in [BGP-MP].

Subsequent Address Family Identifier (SAFI):

This field is the same as the one used in [BGP-MP].

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.

6. Operation

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 Outbound Route Filtering Capability to the peer using BGP Capabilities advertisement [BGP-CAP].

A BGP speaker that implements the Outbound Route Filtering Capability MUST support the BGP ROUTE-REFRESH message, as defined in [BGP-RR]. A BGP speaker that advertises the Outbound Route Filtering Capability to a peer using BGP Capabilities advertisement [BGP-CAP] does not have to advertise the BGP Route Refresh Capability to that peer.

Consider a BGP speaker that advertises the Outbound Route Filtering Capability indicating its willingness to receive a particular set of <AFI/SAFI, ORF-Type> from its peer, and that receives the Outbound 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 is non-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 the 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 MUST 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 Outbound Route Filtering Capability indicating its willingness to receive ORF entries from the speaker, and the speaker advertises to the peer the Outbound 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 Outbound 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 an <AFI/SAFI, ORF-type> carried by the message does not match <AFI/SAFI, ORF-type> that the speaker is willing to receive from the peer (as advertised to the peer in the Outbound Route Filtering Capability), the specified ORF entries in the message are ignored. Otherwise, the speaker modifies the specified ORF previously received, according to the ORF entries carried in the message. If any of the fields of an ORF entry in the message contains an unrecognized value, the whole specified ORF previously received is removed.

If the Action component of an ORF entry is REMOVE, but the ORF previously received does not contain the specified entry, the ORF entry in the message is ignored.

ORF entries with either REMOVE or REMOVE-ALL cannot 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 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.

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 ORFs (if any) previously 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.

After a BGP speaker makes changes to the ORF entries previously sent to a peer, the speaker MUST send to the peer the updated ORF entries with either (a) When-to-refresh set to IMMEDIATE, or (b) When-to-refresh set to DEFER followed by a plain ROUTE-REFRESH message. The latter MUST be used by the speaker when there are other policy changes (in addition to the ORF entries) that require the peer to re-advertise all the routes.

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 removed (either via REMOVE-ALL, or via a sequence of REMOVE).

If a particular route maintained by a BGP speaker does not 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 combining the results (combining of PERMIT and DENY results in DENY).

7. IANA Considerations

This document defines a new BGP Capability - Outbound Route Filtering Capability. The Capability Code for the Outbound Route Filtering Capability is 3.

As specified in this document, an ORF entry contains the ORF-Type field for which IANA has created and now maintains a registry entitled "BGP Outbound Route Filtering (ORF) Types".

IANA maintains and registers values for ORF-Type field as follows:

- ORF-Type value 0 is reserved.

- ORF-Type values 1 through 63 are to be assigned by IANA using either the Standards Action process defined in RFC 5226 [RFC5226], or the Early IANA Allocation process defined in RFC 4020 [RFC4020].

- ORF-Type values 64 through 127 are to be assigned by IANA, using the "First Come First Served" policy defined in RFC 5226 [RFC5226].

- ORF-Type values 128 through 255 are vendor-specific, and values in this range are not to be assigned by IANA.

8. Manageability Considerations

The management objects for BGP ORFs will be defined separately, outside this document. However, it is suggested that the following management objects be defined:

The ORF Capability object, which describes the ORF Capability exchanged over a BGP session, should include the ORF types and the Send/Receive values advertised and received for a BGP peer.

The ORF entry object should contain the ORF entries of each ORF sent and received for a BGP peer.

9. Security Considerations

This extension to BGP does not change the underlying security issues [BGP-4].

10. Acknowledgments

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

11. Normative References

[BGP-4] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[BGP-MP] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[BGP-CAP] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", RFC 3392, November 2002.

[BGP-RR] Chen, E., "Route Refresh Capability for BGP-4", RFC 2918, September 2000.

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

[RFC4020] Kompella, K. and A. Zinin, "Early IANA Allocation of Standards Track Code Points", BCP 100, RFC 4020, February 2005.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

Authors' Addresses

Enke Chen
Cisco Systems, Inc.
170 W. Tasman Dr.
San Jose, CA 95134

Email: enkechen@cisco.com

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Ave
Sunnyvale, CA 94089

Email: yakov@juniper.net

Full Copyright Statement

Intellectual Property