        Enhanced Route Refresh Capability for BGP-4
        draft-ietf-idr-bgp-enhanced-route-refresh-06.txt

Abstract

   In this document we enhance the existing BGP route refresh mechanisms
   to provide for the demarcation of the beginning and the ending of a
   route refresh. The enhancement can be used to facilitate correction
   of BGP RIB inconsistencies in a non-disruptive manner.

Table of Contents

1. Introduction

It is sometimes necessary to perform routing consistency validations
such as checking for possible missing withdraws between BGP speakers
[RFC4271]. Currently such validations typically involve off-line,
manual operations which can be tedious and time consuming.

In this document we enhance the existing BGP route refresh mechanisms
[RFC2918] to provide for the demarcation of the beginning and the
ending of a route refresh (which refers to the complete re-
advertisement of the Adj-RIB-Out to a peer, subject to routing
policies). The enhancement can be used to facilitate on-line, non-
disruptive consistency validation of BGP routing updates.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to
be interpreted as described in [RFC2119] only when they appear in all
upper case. They may also appear in lower or mixed case as English
words, without any normative meaning.

3. Protocol Extensions

The BGP protocol extensions introduced in this document include the
definition of a new BGP capability, named "Enhanced Route Refresh
Capability", and the specification of the message subtypes for the
ROUTE-REFRESH message.

3.1. Enhanced Route Refresh Capability

   The "Enhanced Route Refresh Capability" is a new BGP capability
   [RFC5492].   IANA has assigned a Capability Code of 70 for this
   capability . The Capability Length field of this capability is zero.

   By advertising this capability to a peer, a BGP speaker conveys to
   the peer that the speaker supports the message subtypes for the
   ROUTE-REFRESH message and the related procedures described in this
   document.

3.2. Subtypes for ROUTE-REFRESH Message

   The "Reserved" field of the ROUTE-REFRESH message specified in
   [RFC2918] is re-defined as the "Message Subtype" with the following
   values:

          0 - Normal route refresh request [RFC2918]
                  with/without ORF [RFC5291]
          1 - Demarcation of the beginning of a route refresh operation.
              Also known as a "BoRR message" or just a "BoRR".
          2 - Demarcation of the ending of a route refresh operation.
              Also known as a "EoRR message" or just a "EoRR".

   The remaining values of the message subtypes are reserved for future
   use. The use of the new message subtypes is described in the
   Operations section.

4. Operation

   A BGP speaker that supports the message subtypes for the ROUTE-
   REFRESH message and the related procedures SHOULD advertise the
   "Enhanced Route Refresh Capability".

   The following procedures are applicable only if a BGP speaker has
   received the "Enhanced Route Refresh Capability" from a peer.

   Before the speaker starts a route refresh that is either initiated
   locally, or in response to a "normal route refresh request" from the
   peer, the speaker MUST send a BoRR message. After the speaker
   completes the re-advertisement of the entire Adj-RIB-Out to the peer,
   it MUST send an EoRR message.

   Conceptually the "entire Adj-RIB-Out" for a peer in this section
   refers to all the route entries in the "Adj-RIB-Out" for the peer at
   the start of the route refresh operation. These route entries
   comprise of both, the reachability as well as unreachability

information. When a route entry in the "ADJ-RIB-Out" changes, only
the modified route entry needs to be advertised.

In processing a ROUTE-REFRESH message from a peer, the BGP speaker
MUST examine the "message subtype" field of the message and take the
appropriate actions. The message processing rules for ROUTE-REFRESH
message with subtype of 0 are described in [RFC2918] and [RFC5291].
A BGP speaker can receive a BoRR message from a peer at anytime,
either as a result of a peer responding to a ROUTE-REFESH message, or
as a result of a peer unilaterally initiating a route refresh. When
a BGP speaker receives a BoRR message from a peer, it MUST mark all
the routes with the given <AFI, SAFI> from that peer as stale. As it
receives routes from its peer's subsequent Adj-RIB-Out re-
advertisement, these replace any corresponding stale routes. When a
BGP speaker receives an EoRR message from a peer, it MUST immediately
remove any routes from the peer that are still marked as stale for
that <AFI, SAFI>. Such purged routes MAY be logged for future
analysis.

An implementation MAY impose a locally configurable upper bound on
how long it would retain any stale routes. Once the upper bound is
reached, the implementation MAY remove any routes from the peer that
are still marked as stale for that <AFI, SAFI> without waiting for an
EoRR message.

The following procedures are specified in order to simplify the
interaction with the BGP Graceful Restart [RFC4724]. For a BGP
speaker that supports the BGP Graceful Restart, it MUST NOT send a
BoRR for an AFI/SAFI to a neighbor before it sends the EOR for the
AFI/SAFI to the neighbor. A BGP speaker that has received the
Graceful Restart Capability from its neighbor, MUST ignore any BoRRs
for an AFI/SAFI from the neighbor before the speaker receives the EoR
for the given AFI/SAFI from the neighbor. The BGP speaker SHOULD log
an error of the condition for further analysis.

5. Error Handling

This document defines a new NOTIFICATION error code:

      Error Code   Symbolic Name

         TBD   ROUTE-REFRESH Message Error

The following error subcodes are defined as well:

         Subcode   Symbolic Name

           1   Invalid Message Length

The error handling specified in this section is applicable only when
a BGP speaker has received the "Enhanced Route Refresh Capability"
from a peer.

If the length, excluding the fixed-size message header, of the
received ROUTE-REFRESH message with Message Subtype 1 and 2 is not 4,
then the BGP speaker MUST send a NOTIFICATION message with the Error
Code of "ROUTE-REFRESH Message Error" and the subcode of "Invalid
Message Length". The Data field of the NOTIFICATION message MUST
contain the complete ROUTE-REFRESH message.

When the BGP speaker receives a ROUTE-REFRESH message with a "Message
Subtype" field other than 0, 1 or 2, it MUST ignore the received
ROUTE-REFRESH message. It SHOULD log an error for further analysis.

6. IANA Considerations

This document defines the Enhanced Route Refresh Capability for BGP.
The Capability Code 70 has been assigned by the IANA. This document
also defines two new subcodes for the Route Refresh message. They
need to be registered with the IANA. We request IANA to create a new
registry for the Route Refresh message subcodes as follows:

    Under "Border Gateway Protocol (BGP) Parameters":
    Registry: "BGP Route Refresh Subcodes"
    Reference: [draft-ietf-idr-bgp-enhanced-refresh-06.txt]
    Registration Procedure(s): Values 0-127 Standards Action, values
    128-254 First Come, First Served, Value 255 reserved

    Value   Code   Reference
    0   Route-Refresh   [RFC2918], [RFC5291]
    1   BoRR   [draft-ietf-idr-bgp-enhanced-refresh-06.txt]
    2   EoRR   [draft-ietf-idr-bgp-enhanced-refresh-06.txt]
    255   Reserved

In addition, this document defines an NOTIFICATION error code and
several error subcodes for the ROUTE-REFRESH message. The
NOTIFICATION error code need to be registered with the IANA. We
request IANA to create a new registry for the error subcodes as
follows:

```
        Under "BGP Error Subcodes":
        Registry: "BGP ROUTE-REFRESH Message Error subcodes"
        Reference: [draft-ietf-idr-bgp-enhanced-refresh-06.txt]
        Registration Procedure(s): Values 0-127 Standards Action, values
        128-255 First Come, First Served

        Value  Code                      Reference
        0      Reserved
        1      Invalid Message Length    [draft-ietf-idr-bgp-enhanced-
refresh-06.txt]
```

7. Security Considerations

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

8. Acknowledgements

   The authors would like to thank Pedro Marques, Pradosh Mohapatra,
   Robert Raszuk, Pranav Mehta, Shyam Sethuram, Bruno Decraene, Martin
   Djernaes, Jeff Haas, Ilya Varlashkin, Rob Shakir, Paul Jakma, Jie
   Dong, Qing Zeng, Albert Tian, Jakob Heitz and Chris Hall for their
   review and comments. The authors would like to thank John Scudder
   for the review and contribution to this document.

9. Normative References

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

   [RFC2918] Chen, E., "Route Refresh Capability for BGP-4", RFC 2918,
              September 2000.

   [RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway
              Protocol 4 (BGP-4)", RFC 4271, January 2006.

   [RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y.
              Rekhter, "Graceful Restart Mechanism for BGP", RFC 4724,
              January 2007.

   [RFC5291] Chen, E. and Y. Rekhter, "Outbound Route Filtering
              Capability for BGP-4", RFC 5291, August 2008.

   [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement
              with BGP-4", RFC 5492, February 2009.

Authors' Addresses

    Keyur Patel
    Cisco Systems
    170 W. Tasman Drive
    San Jose, CA 95124 95134
    USA

    Email: keyupate@cisco.com

    Enke Chen
    Cisco Systems
    170 W. Tasman Drive
    San Jose, CA 95124 95134
    USA

    Email: enkechen@cisco.com

    Balaji Venkatachalapathy

    Email: balaji_pv@hotmail.com