

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 5, 2011

K. Patel
R. Fernando
Cisco Systems
J. Scudder
J. Haas
Juniper Networks
February 2011

Graceful Restart Extensions for BGP
draft-keyupate-idr-bgp-gr-extension-00.txt

Abstract

The current BGP Graceful Restart mechanism limits the usage of BGP Graceful Restart to BGP protocol messages other than a BGP NOTIFICATION message. This document defines an extension to the BGP Graceful Restart that permits the Graceful Restart procedures to be performed when the BGP speaker receives a BGP NOTIFICATION Message. This document also defines a new BGP NOTIFICATION Cease Error subcode to prevent BGP speakers supporting the extension defined in this document from performing a Graceful Restart.

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 5, 2011.

Copyright Notice

Copyright (c) 2011 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

- 1. Introduction 4
 - 1.1. Requirements Language 4
- 2. Modifications to BGP Graceful Restart Capability 4
- 3. BGP Hard Reset Subcode 5
- 4. Operation 6
- 5. Acknowledgements 6
- 6. IANA Considerations 6
- 7. Security Considerations 6
- 8. References 7
 - 8.1. Normative References 7
 - 8.2. Informative References 7
- Authors' Addresses 7

1. Introduction

For many classes of errors, the BGP protocol must send a NOTIFICATION message and reset the peering session to handle the error condition. The BGP Graceful Restart extension defined in [RFC4724] requires that normal BGP procedures defined in [RFC4271] be followed when a NOTIFICATION message is sent or received. This document defines an extension to BGP Graceful Restart that permits the Graceful Restart procedures to be performed when the BGP speaker receives a NOTIFICATION message. This permits the BGP speaker to avoid flapping reachability and continue forwarding while the BGP speaker restarts the session to handle errors detected in the BGP protocol.

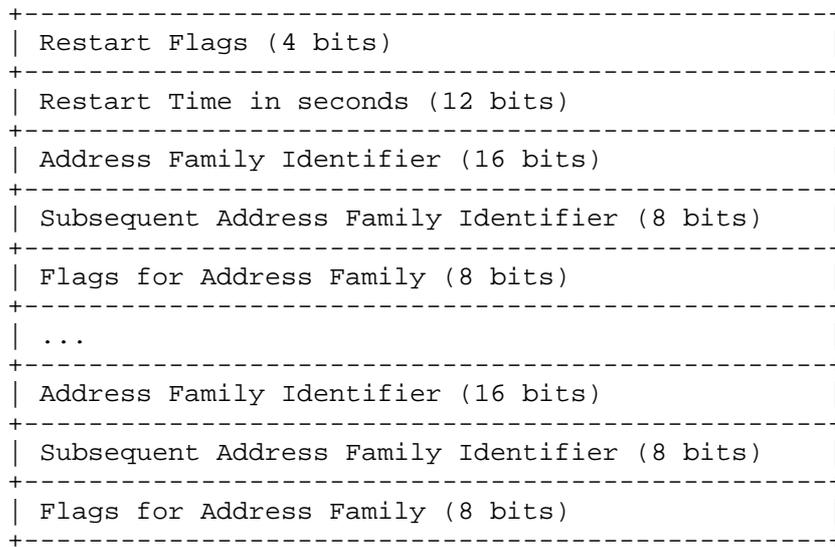
This document defines a BGP NOTIFICATION cease Error subcode for the Cease Error code to prevent BGP speakers supporting the extension defined in this document from performing a Graceful Restart.

1.1. 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].

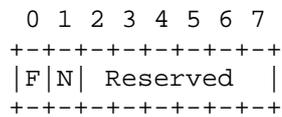
2. Modifications to BGP Graceful Restart Capability

The BGP Graceful Restart Capability is augmented to signal the Graceful Restart support for BGP NOTIFICATION messages. In particular, the flags field for Address Family is augmented as follows:



Flags for Address Family:

This field contains bit flags relating to routes that were advertised with the given AFI and SAFI.



The second most significant bit "N" is defined as a BGP Graceful Notification bit, which is used to indicate the Graceful Restart support for BGP NOTIFICATION messages. BGP speaker indicates the Graceful Restart support for BGP NOTIFICATION messages and its ability to handle the new BGP NOTIFICATION Cease message subcode and the format for a BGP NOTIFICATION Cease message defined in [RFC4486] when the Graceful NOTIFICATION bit is set (value 1).

3. BGP Hard Reset Subcode

A new BGP Cease message subcode is defined known as BGP Hard Reset Subcode. The value of this subcode is 9.

Whenever a BGP speaker receives a NOTIFICATION message with the Cease

Error code and Hard Reset Error subcode, the speaker MUST terminate the BGP session following the standard procedures in [RFC4271].

4. Operation

A BGP speaker that is willing to receive and send BGP NOTIFICATION messages in Graceful mode should advertise the BGP Graceful Notification Flag "N" using the Graceful Restart Capability as defined in [RFC4724].

When a BGP Speaker receives a BGP NOTIFICATION message, it SHOULD follow the standard rules of the receiving speaker mentioned in [RFC4724] for all AFI/SAFIs for which it has announced the BGP Graceful Notification flag. The BGP speaker generating a BGP NOTIFICATION message SHOULD follow the standard rules of the receiving Speaker in [RFC4724] for all AFI/SAFIs that were announced with the BGP Graceful Notification flag.

Once the session is re-established, both BGP speakers MUST set their "Forwarding State" bit to 1 if they want to apply planned graceful restart. The handling of the "Forwarding State" bit should be done as specified by the procedures of the Receiving speaker in [RFC4724] are applied.

As part of this extension, possible consecutive restarts SHOULD NOT delete a route (from the peer) previously marked as stale, until required by rules mentioned in [RFC4724].

5. Acknowledgements

The authors would like to thank Robert Raszuk for the review and comments.

6. IANA Considerations

This document defines a new BGP Cease message subcode known as BGP Hard Reset Subcode. IANA maintains the list of existing BGP Cease message subcodes. This document proposes defining a new BGP Cease message subcode known as BGP Hard Reset Subcode with the value 9.

7. Security Considerations

This extension to BGP does not change the underlying security issues inherent in the existing [RFC4724] and [RFC4271]

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2842] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", RFC 2842, May 2000.
- [RFC3392] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4", RFC 3392, November 2002.
- [RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.
- [RFC4486] Chen, E. and V. Gillet, "Subcodes for BGP Cease Notification Message", RFC 4486, April 2006.
- [RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y. Rekhter, "Graceful Restart Mechanism for BGP", RFC 4724, January 2007.

8.2. Informative References

- [RFC2858] Bates, T., Rekhter, Y., Chandra, R., and D. Katz, "Multiprotocol Extensions for BGP-4", RFC 2858, June 2000.

Authors' Addresses

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

Email: keyupate@cisco.com

Rex Fernando
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: rex@cisco.com

John Scudder
Juniper Networks
1194 N. Mathilda Ave
Sunnyvale, CA 94089
USA

Email: jgs@juniper.net

Jeff Haas
Juniper Networks
1194 N. Mathilda Ave
Sunnyvale, CA 94089
USA

Email: jhaas@juniper.net

