

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 5, 2012

J. Dong
M. Chen
Huawei Technologies
K. Patel
Cisco Systems
July 4, 2011

BGP End-of-RIB Usage Extension and Negotiation
draft-dong-idr-end-of-rib-use-extension-00

Abstract

This document describes the use of BGP End-of-RIB marker in improving BGP routing convergence during initial route exchange. A mechanism to negotiate the extension of End-of-RIB usage is also specified.

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 January 5, 2012.

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.

Table of Contents

1.	Introduction	3
2.	BGP Convergence Problem in Initial Route Exchange	3
3.	End-of-RIB Usage Extension	3
4.	BGP End-of-RIB Capability	4
5.	IANA Considerations	5
6.	Security Considerations	5
7.	Acknowledgements	5
8.	References	5
8.1.	Normative References	5
8.2.	Informative References	5
	Authors' Addresses	6

1. Introduction

BGP Graceful Restart (GR) [[RFC4724](#)] defines an End-of-RIB marker to convey routing convergence information during BGP restart. It is also specified that the generation of such a marker upon completion of the initial update would be useful for routing convergence in general.

Currently most of BGP routers neither generate End-of-RIB marker upon completion of initial route advertisement nor anticipate the arrival of End-of-RIB from peers during initial route exchange. In addition, for some BGP implementations receiving End-of-RIB marker in scenarios other than BGP GR may be treated as an error.

This document describes the benefit of using BGP End-of-RIB marker to inform completion of initial BGP route exchange. A mechanism to negotiate the extension of End-of-RIB usage is also specified.

2. BGP Convergence Problem in Initial Route Exchange

When a BGP speaker establishes BGP sessions with multiple peers, the initial route exchange begins. Normally whenever a route is received from one of the peers, BGP path calculation would be executed and new UPDATE message will be advertised to peers immediately. Since during initial route exchange the BGP speaker may consecutively receive different routes to the same prefixes from different peers, this normal procedure may cause the BGP speaker execute the path calculation for some prefixes for multiple times, and would further result in advertising non-optimal routes before routing convergence.

Apparently this is a waste of processing resource and also impacts routing convergence and stability of the network. Some optimization has been proposed. One typical approach is to set a timer for initial route exchange, and BGP speaker will not execute path calculation and advertise routes to peers until that timer expires.

The disadvantage of this approach is value of the timer would be critical for BGP convergence performance, and since the timer value would be fixed once configured, it cannot guarantee best performance and convergence time for different cases.

[3.](#) End-of-RIB Usage Extension

[RFC4724] defines the use of End-of-RIB marker in BGP Graceful Restart scenario, and it also specifies that "generation of such a marker upon completion of the initial update would be useful for routing convergence in general, and thus the practice is

Dong, et al.

Expires January 5, 2012

[Page 3]

Internet-Draft

End-of-RIB Usage Extension

July 2011

recommended".

Actually End-of-RIB marker should be used as an individual feature independent of whether BGP GR is enabled or not. One example is [\[RFC4684\]](#) specifies the use of End-of-RIB for RT Membership information advertisement.

Similar to the use in BGP Graceful Restart, End-of-RIB marker could also be used to inform completion of initial route exchange. Thus route calculation and further advertisement would be suspended until End-of-RIB marker is received from all or a predefined portion of BGP peers. In addition, a relative large timer could be used as a backup trigger to ensure path calculation and advertisement would always be executed within a predefined time range.

Although it is easy to understand the use of End-of-RIB in improving initial routing convergence, such benefit may not be obtained directly, as currently most of BGP routers neither generate End-of-RIB marker upon completion of initial route advertisement nor anticipate the arrival of End-of-RIB from peers during initial route exchange. In addition, for some BGP implementations receiving End-of-RIB marker in scenarios other than BGP GR may be treated as an error. Thus to use End-of-RIB for initial route exchange scenario, some negotiation between the sending and receiving BGP speaker would be necessary.

[4.](#) BGP End-of-RIB Capability

A new BGP capability called End-of-RIB Capability is defined. The Capability code for this capability is to be assigned. The Capability length field is zero.

By advertising this capability to a peer, a BGP speaker conveys to the peer that the speaker support advertising and receiving End-of-RIB marker and the related procedures described in this document.

After capability negotiation, if both the peer speaker and local speaker support this capability, then End-of-RIB marker MUST be sent to peer after finishing initial route advertisement, and both speakers MUST use the End-of-RIB marker received from peer as notification of initial exchange completion and trigger of local route processing and further advertisement. If any one of the peering speakers does not support this extension, End-of-RIB MUST not be used in initial route exchange scenario.

When End-of-RIB is used for initial exchange, a timer MAY also be used to control the maximum initial delay of route processing and

UPDATEs advertisement. The timer value SHOULD be configurable.

[5.](#) IANA Considerations

A new BGP capability - End-of-RIB Capability is defined in this document. The Capability code needs to be assigned by the IANA.

[6.](#) Security Considerations

This document does not change the security properties of BGP.

[7.](#) Acknowledgements

The authors would like to thank John Scudder, Shunwan Zhuang, Qing Zeng for their valuable comments and discussions to this document.

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

8.2. Informative References

- [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.

Dong, et al. Expires January 5, 2012 [Page 5]

Internet-Draft End-of-RIB Usage Extension July 2011

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Building, No.3 Xinxu Rd
Beijing, 100085
China

Email: jie.dong@huawei.com

Mach Chen
Huawei Technologies
Huawei Building, No.3 Xinxu Rd
Beijing, 100085

China

Email: mach.chen@huawei.com

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

Email: keyupate@cisco.com