Network Working Group
Internet-Draft
Updates: 4271 (if approved)
Intended status: Standards Track
Expires: January 1, 2020

R. Bush
IIJ & Arrcus
K. Patel
Arrcus, Inc.
D. Ward
Cisco Systems

June 30, 2019

Extended Message support for BGP
draft-ietf-idr-bgp-extended-messages-31

Abstract

The BGP specification mandates a maximum BGP message size of 4,096 octets. As BGP is extended to support newer AFI/SAFIs and other features, there is a need to extend the maximum message size beyond 4,096 octets. This document updates the BGP specification RFC4271 by extending the maximum message size from 4,096 octets to 65,535 octets for all except the OPEN and KEEPALIVE messages.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The BGP specification [RFC4271] mandates a maximum BGP message size of 4,096 octets. As BGP is extended to support newer AFI/SAFIs and newer capabilities (e.g., BGPsec [RFC8205] and BGP-LS [RFC7752]), there is a need to extend the maximum message size beyond 4,096 octets. This draft provides an extension to BGP to extend its message size limit from 4,096 octets to 65,535 octets for all except the OPEN and KEEPALIVE messages.

2. BGP Extended Message

A BGP message over 4,096 octets in length is a BGP Extended Message.

BGP Extended Messages have a maximum message size of 65,535 octets. The smallest message that may be sent consists of a BGP KEEPALIVE which consists of 19 octets.

3. Extended Message Capability for BGP

To advertise the BGP Extended Message Capability to a peer, a BGP speaker uses BGP Capabilities Advertisement [RFC5492]. By advertising the BGP Extended Message Capability to a peer, a BGP speaker conveys that it is able to send, receive, and properly handle, see Section 4, BGP Extended Messages.

The BGP Extended Message Capability is a new BGP Capability [RFC5492] defined with Capability code 6 and Capability length 0.

A peer which does not advertise this capability MUST NOT send BGP Extended Messages, and BGP Extended Messages MUST NOT be sent to it.

Peers that wish to use the BGP Extended Message capability must support Error Handling for BGP UPDATE Messages per [RFC7606].

4. Operation

The Extended Message Capability applies to all messages except for the OPEN and KEEPALIVE messages. The former exception is to reduce the complexity of providing a backward compatibility

A BGP speaker that is capable of sending and receiving BGP Extended Messages SHOULD advertise the BGP Extended Message Capability to its peers using BGP Capabilities Advertisement [RFC5492]. A BGP speaker MAY send Extended Messages to its peer only if both peers have negotiated the Extended Message Capability with each other.

An implementation that advertises support for BGP Extended Messages MUST be capable of receiving a message with a Length up to and including 65,535 octets.

Applications generating information which might be encapsulated within BGP messages MUST limit the size of their payload to take the maximum message size into account.

If a BGP message with a Length lgreater than 4,096 octets is received by a BGP listener who has not advertised the Extended Message Capability, the listener MUST treat this as a malformed message, and MUST generate a NOTIFICATION with the Error Subcode set to Bad Message Length (see [RFC4271] Sec 6.1).

A BGP announcement will (policy, best path, etc., allowing) propagate throughout the BGP speaking Internet; and hence to BGP speakers which may not have the Extended Message capability. Therefore, an announcement in an Extended Message where the size of the attribute

set plus the NLRI can not be decomposed to 4,096 octets or less may cause lack of reachability.

A speaker capable of BGP Extended Messages having a mixture of peers some of which have not exchanged the BGP Extended Message capability, may receive an announcement from one of its capable peers that would (due to the new AS on the path, new added attributes, etc.) produce an ongoing announcement that would be over 4,096 octets. When propagating that update onward to a neighbor with which it has not negotiated the BGP Extended Message capability, the sender SHOULD try to reduce the outgoing message size by downgrading BGPsec to BGP4, decomposing a multi-NLRI update producing multiple updates with fewer NLRI per update, removing attributes eligible under the attribute discard approach of [RFC7606], etc. If the resulting message would still be over the 4,096 octet limit, the sender SHOULD treat-as-withdraw per [RFC7606].

In an iBGP mesh, all peers SHOULD support the BGP Extended Message Capability and [RFC7606]. Only then is it consistent to deploy with eBGP peers.

During the incremental deployment of BGP Extended Messages and [RFC7606] in an iBGP mesh, or with eBGP peers, the operator should monitor any routes dropped as "treat-as-withdraw".

It is RECOMMENDED that BGP protocol developers and implementers are conservative in their application and use of Extended Messages. Future protocol specifications will need to describe how to handle peers which can only accommodate 4,096 octet messages.

5. Error Handling

A BGP speaker that has the ability to use Extended Messages but has not advertised the BGP Extended Messages capability, presumably due to configuration, SHOULD NOT accept an Extended Message. A speaker SHOULD NOT implement a more liberal policy accepting BGP Extended Messages.

A BGP speaker that does not advertise the BGP Extended Messages capability might also genuinely not support Extended Messages. Such a speaker will follow the error handling procedures of [RFC4271] if it receives an Extended Message. Similarly, any speaker that treats an improper Extended Message as a fatal error, MUST treat it similarly.

The inconsistency between the local and remote BGP speakers MUST be flagged to the network operator through standard operational

interfaces. The information should include the NLRI and as much relevant information as reasonably possible.

6. Changes to RFC4271

[RFC4271] states "The value of the Length field MUST always be at least 19 and no greater than 4,096." This document changes the latter number to 65,535 for all except the OPEN and KEEPALIVE messages.

[RFC4271] Sec 6.1, specifies raising an error if the length of a message is over 4,096 octets. For all messages except the OPEN message, if the receiver has advertised the BGP Extended Messages Capability, this document raises that limit to 65,535.

7. IANA Considerations

The IANA has made an early allocation for this new BGP Extended Message Capability referring to this document.

Registry: BGP Capability Code

Value	Description	Document
6	BGP-Extended Message	[this draft]

8. Security Considerations

This extension to BGP does not change BGP's underlying security issues; see [RFC4272].

Section 5 allows a receiver to accept an Extended Message even though it had not advertised the capability. This slippery slope could lead to sloppy implementations sending Extended Messages when the receiver is not prepared to deal with them, e.g. to peer groups. At best, this will result in errors; at worst, buffer overflows.

Due to increased memory requirements for buffering, there may be increased exposure to resource exhaustion, intentional or unintentional.

As this draft requires support for [RFC7606] update error handling, it inherits the security considerations of [RFC7606]. BGP peers may avoid such issues by using Authenticated Encryption with additional Data (AEAD) ciphers [RFC5116] and discard messages that do not verify.

If a remote attacker is able to craft a large BGP Extended Message to send on a path where one or more peers do not support BGP Extended Messages, peers which support BGP Extended Messages may act to reduce the outgoing message, see Section 4, and in doing so produce a downgrade attack, e.g. convert BGPsec to BGP4.

If a remote attacker is able to craft a large BGP Extended Message to send on a path where one or more peers do not support BGP Extended Messages, peers which support BGP Extended Messages may incur resource load (processing, message resizing, etc.) reformatting the large messages. Worse, ([RFC7606] "treat-as-withdraw" may consistently withdraw announcements causing inconsistent routing.

BGP routes are filtered by policies set by the operators. Implementations may provide policies to filter routes that would cause the "treat-as-withdraw" from being passed by an extended message speaker.

9. Acknowledgments

The authors thank Alvaro Retana for an amazing review, Enke Chen, Susan Hares, John Scudder, John Levine, and Job Snijders for their input; and Oliver Borchert and Kyehwan Lee for their implementations and testing.

10. References

10.1. Normative References

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <http://www.rfc-editor.org/info/rfc4272>.

[RFC5116] McGrew, D., "An Interface and Algorithms for Authenticated Encryption", RFC 5116, DOI 10.17487/RFC5116, January 2008, <http://www.rfc-editor.org/info/rfc5116>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <http://www.rfc-editor.org/info/rfc5492>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <http://www.rfc-editor.org/info/rfc7606>.

10.2. Informative References

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <http://www.rfc-editor.org/info/rfc7752>.

[RFC8205] Lepinski, M., Ed. and K. Sriram, Ed., "BGPsec Protocol Specification", RFC 8205, DOI 10.17487/RFC8205, September 2017, <https://www.rfc-editor.org/info/rfc8205>.

Authors' Addresses

Randy Bush
IIJ & Arrcus
5147 Crystal Springs
Bainbridge Island, Washington 98110
US

Email: randy@psg.com

Keyur Patel
Arrcus, Inc.

Email: keyur@arrcus.com

Dave Ward
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
US

Email: dward@cisco.com