ï»¿

                     Extended Message Support for BGP

Abstract

   The BGP specification (RFC 4271) mandates a maximum BGP message size
   of 4,096 octets. As BGP is extended to support new Address Family
   Identifiers (AFIs), Subsequent AFIs (SAFIs), and other features,
   there is a need to extend the maximum message size beyond 4,096
   octets. This document updates the BGP specification by extending the
   maximum message size from 4,096 octets to 65,535 octets for all
   messages except for OPEN and KEEPALIVE messages.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8654.

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 new AFIs, SAFIs, and
    other capabilities (e.g., BGPsec [RFC8205] and BGP - Link State (BGP-
    LS) [RFC7752]), there is a need to extend the maximum message size
    beyond 4,096 octets. This document provides an extension to BGP to
    extend the message size limit from 4,096 octets to 65,535 octets for
    all messages except for OPEN and KEEPALIVE messages.

1.1. 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.

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 is a BGP KEEPALIVE, which
    consists of 19 octets.

3. BGP Extended Message Capability

    The BGP Extended Message Capability is a new BGP capability [RFC5492]
    defined with Capability Code 6 and Capability Length 0.

    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 receive and properly handle BGP
    Extended Messages (see Section 4).

    Peers that wish to use the BGP Extended Message Capability MUST
    support error handling for BGP UPDATE messages per [RFC7606].

4. Operation

    The BGP Extended Message Capability applies to all messages except
    for OPEN and KEEPALIVE messages. These exceptions reduce the
    complexity of providing backward compatibility.

    A BGP speaker that is capable of receiving BGP Extended Messages
    SHOULD advertise the BGP Extended Message Capability to its peers
    using BGP Capabilities Advertisement [RFC5492]. A BGP speaker MAY
    send BGP Extended Messages to a peer only if the BGP Extended Message
    Capability was received from that peer.

    An implementation that advertises the BGP Extended Message Capability
    MUST be capable of receiving a message with a length up to and
    including 65,535 octets.

    Applications generating information that 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 greater than 4,096 octets is received
    by a BGP listener who has not advertised the BGP Extended Message
    Capability, the listener will generate a NOTIFICATION with the Error
    Subcode set to Bad Message Length ([RFC4271], Section 6.1).

    A BGP UPDATE will (if allowed by policy, best path, etc.) typically
    propagate throughout the BGP-speaking Internet and hence to BGP
    speakers that may not support BGP Extended Messages. Therefore, an
    announcement in a BGP Extended Message where the size of the
    attribute set plus the NLRI is larger than 4,096 octets may cause
    lack of reachability.

A BGP speaker that has advertised the BGP Extended Message Capability
to its peers may receive an UPDATE from one of its peers that
produces an ongoing announcement that is larger than 4,096 octets.
When propagating that UPDATE onward to a neighbor that has not
advertised the BGP Extended Message Capability, the speaker SHOULD
try to reduce the outgoing message size by removing attributes
eligible under the "attribute discard" approach of [RFC7606]. If the
message is still too big, then it must not be sent to the neighbor
([RFC4271], Section 9.2). Additionally, if the NLRI was previously
advertised to that peer, it must be withdrawn from service
([RFC4271], Section 9.1.3).

If an Autonomous System (AS) has multiple internal BGP speakers and
also has multiple external BGP neighbors, care must be taken to
ensure a consistent view within the AS in order to present a
consistent external view. In the context of BGP Extended Messages, a
consistent view can only be guaranteed if all the Internal BGP (iBGP)
speakers advertise the BGP Extended Message Capability. If that is
not the case, then the operator should consider whether or not the
BGP Extended Message Capability should be advertised to external
peers.

During the incremental deployment of BGP Extended Messages and use of
the "attribute discard" approach of [RFC7606] in an iBGP mesh or with
External BGP (eBGP) peers, the operator should monitor any routes
dropped and any discarded attributes.

5. Error Handling

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

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

Error handling for UPDATE messages, as specified in Section 6.3 of
[RFC4271], is unchanged. However, if a NOTIFICATION is to be sent to
a BGP speaker that has not advertised the BGP Extended Message
Capability, the size of the message MUST NOT exceed 4,096 octets.

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

6. Changes to RFC 4271

[RFC4271] states "The value of the Length field MUST always be at
least 19 and no greater than 4096." This document changes the latter
number to 65,535 for all messages except for OPEN and KEEPALIVE
messages.

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

7. IANA Considerations

IANA has made the following allocation in the "Capability Codes"
registry:

   +-------+--------------------+----------+

```
          | Value | Description         | Reference |
          +=======+=====================+===========+
          | 6     | BGP Extended Message | RFC 8654  |
          +-------+---------------------+-----------+
```

          Table 1: Addition to "Capability Codes"
                        Registry

8. Security Considerations

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

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

   If a remote speaker 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 that support BGP Extended Messages may:

   * act to reduce the outgoing message (see Section 4) and, in doing
      so, cause an attack by discarding attributes one or more of its
      peers may be expecting. The attributes eligible under the
      "attribute discard" approach must have no effect on route
      selection or installation [RFC7606].

   * act to reduce the outgoing message (see Section 4) and, in doing
      so, allow a downgrade attack. This would only affect the
      attacker's message, where 'downgrade' has questionable meaning.

   * incur resource load (processing, message resizing, etc.) when
      reformatting the large messages.

9. References

9.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://www.rfc-editor.org/info/rfc2119>.

   [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,
              <https://www.rfc-editor.org/info/rfc4271>.

   [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement
              with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February
              2009, <https://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,
              <https://www.rfc-editor.org/info/rfc7606>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

9.2. Informative References

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

   [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,

            <https://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
   Arrcus & IIJ
   5147 Crystal Springs
   Bainbridge Island, WA 98110
   United States of America

   Email: randy@psg.com

   Keyur Patel
   Arrcus, Inc.

   Email: keyur@arrcus.com

   Dave Ward
   Cisco Systems
   170 W. Tasman Drive
   San Jose, CA 95134
   United States of America

   Email: dward@cisco.com