Workgroup: Interdomain Routing Working Group
Internet-Draft:
draft-wang-idr-bgp-error-enhance-00
Published: 24 October 2021
Intended Status: Standards Track
Expires: 27 April 2022

Authors: H. Wang                M. Shen
         Huawei Technologies   Huawei Technologies
         J. Dong
         Huawei Technologies

# Revised Error Handling for BGP Messages

## Abstract

This document supplements and revises RFC7606. According to RFC 7606, when an UPDATE packet received from a neighbor contains an attribute of incorrect format, the BGP session cannot be reset directly. Instead, the BGP session must be reset based on the specific problem. Error packets must minimize the impact on routes and do not affect the correctness of the protocol. Different error handling methods are used. The error handling methods include discarding attributes, withdrawing routes, disabling the address family, and resetting sessions.

RFC 7606 specifies the error handling methods of some existing attributes and provides guidance for error handling of new attributes.

This document supplements the error handling methods for common attributes that are not specified in RFC7606, and provides suggestions for revising the error handling methods for some attributes. The general principle remains unchanged: Maintain established BGP sessions and keep valid routes updated. However, discard or delete incorrect attributes or packets to minimize the impact on the current session.

## 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

## Copyright Notice

## Table of Contents

## 1. Introduction

According to RFC 4271, a BGP session that receives an UPDATE message containing a malformed attribute needs to reset the session that receives the malformed attribute.

According to our experience during maintenance, malformed packets may be incorrectly encapsulated due to software bugs or misunderstanding of standards in software development. Interrupting a neighbor causes neighbor flapping, which does not help solve the problem. The malformed packets may not be recognized by intermediate routers and cannot be incorrectly checked and propagated to other routers that establish sessions. When they reach the router that recognizes and checks the attribute, the neighbor flapping may also occured. Even because routes are propagated multiple times, a route containing malformed packets may be received from multiple sessions at a checkpoint, causing multiple sessions to be reset, and the harm is multiplied.

For the preceding reasons, RFC 7606 defines a new method for processing incorrect UPDATE packets. If the Update packet received from a neighbor contains incorrect attributes, the BGP session cannot be reset directly. Instead, the BGP session needs to be handled in a specific manner based on the principle that incorrect packets affect routes as little as possible and do not affect protocol correctness. The error handling methods include discarding attributes, withdrawing routes, disabling the address family, and resetting sessions.

However, the error handling methods of some common attributes are not provided in RFC7606 or are different from those of vendors. This document supplements the error handling methods of some common attributes and provides suggestions for modifying the error handling methods of some attributes.

## 2. Scenarios

```
+-----+    +-----+   +-----+
| RT1 |----| RR  |---| RT2 |
+-----+    +-.---+   +-----+
             |
             \
           +-----+
           | RT3 |
           +-----+
```

Figure 1 A simple network

Figure 1 shows a simple network. When RT3 has some software bugs or misunderstands the RFC, it may send a malformed packet. The RR receives the packet and considers it a malformed packet according to the error handling rules and resets the session. Later the session between RT3 and RR is re-established. RT3 will resend the packet to RR, and RR continues to repeat the previous action. This will happen continuously until the operator modifies the configuration, such as deleting the configuration about that new feature of the property, and sometimes they don't know how to correct the problem. During this process, RT3 services are unavailable. Frequent neighbor reestablishment and route updating also consumes more RR's system resources.

If the RR does not understand the attribute, the RR sends packets to RT1 and RT2. RT1 and RT2 may perform the same operation as RR. As a result, services between RT1 and RT2 are interrupted.

## 3.  Error-Handling Procedures Update for NLRI

### 3.1.  Prefix Length Error

According to [RFC7606], when a NLRI/UNLRI or MP_REACH_NLRI/MP_REACH_UNLRI with invalid length, eg, IPv4 Prefix length is more than 32, we must drop this Prefix and ignore the following Prefixes. We may keep the prefixes we have parsed correctly before.

Then we may also try to continue parse the next update packet if we can correctly find it.

The NLRI/UNLRI or MP_REACH_NLRI/MP_REACH_UNLRI with invalid length is malformed.

### 3.2.  Appears More Than Once

[RFC7606] described like this:

If the MP_REACH_NLRI attribute or the MP_UNREACH_NLRI [RFC4760] attribute appears more than once in the UPDATE message, then a NOTIFICATION message MUST be sent with the Error Subcode "Malformed Attribute List".

Revised suggestion:

If the MP_REACH_NLRI attribute or the MP_UNREACH_NLRI [RFC4760] attribute appears more than once in the UPDATE message, only the last MP_REACH_NLRI/MP_UNREACH_NLRI SHOULD be processed, the others would be ignore.

## 4.  Error-Handling Procedures Update for Existing Attributes

### 4.1. Nexthop

[RFC4271] define the IP address in the NEXT_HOP meet the following criteria to be considered semantically incorrect:

a) It is the IP address of receiving speaker.

b) The IP address is not EBGP directly neighbor's address or not share a common subnet with the receiving BGP speaker.

An update message with the case a) MAY be install to the RIB but treat as invalid.

Whether an update message with the case b) SHOULD be considered semantically incoorect depends on the user's configuration.

The following criteria also must to be considered semantically incorrect:

c) The IP address is all zero.

d) The IP address is all one.

e) The IP address is multicast address(Class D) or reserved address (Class E).

f) The IP address is not a invalid ip address.

An update message with the case c) to f) SHOULD be logged, and the route will be treat-as-withdraw.

### 4.2. MP_REACH_NLRI

[RFC7606] suggest to do "session reset" or "AFI/SAFI disable" approach. But this approach is too strict.

If the Length of Next Hop Network Address field of the MP_REACH attribute is inconsistent with that which was expected, the attribute is considered malformed. The whole MP_REACH attribute will be ignore and try to parse the next update packet. When it cannot correctly locate the next update packet, it will do the procedure suggested according to [RFC7606] . Otherwise, only the error SHOULD be logged and continued to do packet parsing.

An update message may both contained MP_REACH_NLRI and MP_REACH_UNLRI. If there are same Prefixes in both MP_REACH_NLRI and MP_REACH_UNLRI, the message SHOULD NOT be consider malformed. In this case, it should be firstly process the Prefixes in the MP_REACH_NLRI then process the Prefixes in the MP_REACH_UNLRI.

### 4.3. Prefix SID

According to [RFC8669], an update message containing a malformed or invalid BGP Prefix-SID attribute will be ignore and not advertise it to other BGP peers. But this procedure may lead to unexpected results.

The error handling is revised to be treat-as-withdraw.

### 4.4. AGGREGATE and AS4_AGGREGATOR

When the router-id in AGGREGATE or AS4_AGGREGATE attibute is zero, the attribute SHOULD be consider semantically incorrect, and the attribute SHOULD be logged and discard.

### 4.5. ORIGINATOR_ID

The error handling of [RFC4456] and [RFC7606] is revised as follows.

When the BGP Identifier in ORIGINATOR_ID attibute is zero, the attribute SHOULD be consider semantically incorrect, and the attribute SHOULD be logged and the UPDATE message SHALL be handled using the approach of "treat-as- withdraw".

### 4.6. Cluster-List

The error handling of [RFC4456] and [RFC7606] is revised as follows.

When the CLUSTER_ID value in ORIGINATOR_ID attibute is zero, the attribute SHOULD be consider semantically incorrect, and the attribute SHOULD be logged and the UPDATE message SHALL be handled using the approach of "treat-as- withdraw".

## 5. IANA Considerations

This document makes no request of IANA.

## 6. Security Considerations

This document helps reduce the impact of malformed packets on the network and devices.

## 7. References

### 7.1. 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>.

### 7.2. Informative 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, <https://www.rfc-editor.org/info/rfc4271>.

**[RFC4456]** Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

**[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>.

**[RFC8669]** Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <https://www.rfc-editor.org/info/rfc8669>.

### Authors' Addresses

Haibo Wang
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China

Email: rainsword.wang@huawei.com

Ming Shen
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China

Email: shenming2@huawei.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China

Email: jie.dong@huawei.com