Internet Engineering Task Force
Internet-Draft
Updates: 7606 (if approved)
Intended status: Standards Track
Expires: 23 April 2026

B. Decraene
Orange
J. G. Scudder
HPE

20 October 2025

The Key List BGP Attribute for NLRI Error handling
draft-decraene-idr-nlri-error-handling-01

Abstract

RFC 7606 partially revises the error handling for BGP UPDATE messages. It reduces the cases of BGP session reset by defining and using less impactful error handling approaches, such as attribute discard and treat-as-withdraw when applicable. The treat-as-withdraw approach requires that the entire NLRI field of the MP_REACH_NLRI attribute be successfully parsed. This typically means parsing errors in MP_REACH_NLRI cannot be handled by any means short of session reset. This is exacerbated by the use of non-key data within NLRI, which introduces parsing complexity and additional error cases.

This specification defines a non-transitive BGP attribute, the "NLRI_KEY_LIST attribute", to encode NLRIs as per the format of MP_UNREACH_NLRI. This attribute is used to allow the treat-as-withdraw error-handling approach to be used in case an error in the MP_REACH_NLRI attribute prevents the parsing of its NLRIs.

This document updates RFC 7606 by mandating that the NLRI_KEY_LIST attribute appear before the MP_REACH_NLRI (or any other) attribute in an UPDATE message.

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 https://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 23 April 2026.

Table of Contents

1. Introduction 2
1.1. Requirements Language 3
2. NLRI_KEY_LIST Capability 4
3. NLRI_KEY_LIST Attribute 4
3.1. NLRI_KEY_LIST Format 4
3.2. Sending the NLRI_KEY_LIST attribute 4
3.3. Receiving the NLRI_KEY_LIST attribute 5
3.4. NLRI_KEY_LIST attribute Error Handling 5
4. Operational Considerations 6
5. IANA Considerations 6
6. Security Considerations 7
7. Acknowledgements 7
8. References 7
8.1. Normative References 7
8.2. Informative References 8
Authors' Addresses 8

1. Introduction

According to the base BGP specification [RFC4271], a BGP speaker that receives an UPDATE message containing a malformed attribute is required to reset the session over which the offending attribute was received. This behavior is undesirable because a session reset impacts not only routes with the offending attribute but also other valid routes exchanged over the session.

[RFC7606] revises BGP error handling, with the goal of minimizing the impact on routing of a malformed UPDATE message, while maintaining protocol correctness to the extent possible. For most BGP attributes, a malformed attribute may be handled using attribute discard or treat-as-withdraw. Both approaches preserve the routing of all the NLRIs not advertised in the affected BGP UPDATE message.

However, as indicated in Section 3 of [RFC7606], treat-as-withdraw can only be used if the entire NLRI field of the MP_REACH_NLRI attribute is successfully parsed. This typically means parsing errors in MP_REACH_NLRI cannot be handled by any means short of session reset.

[RFC4760] allows the Border Gateway Protocol (BGP) to advertise general routing information in the Network Layer Reachability Information (NLRI) field of the UPDATE message. Some specifications, such as [RFC8277], [I-D.ietf-idr-bgp-car], and [I-D.ietf-idr-bgp-ct] carry both a key field and a non-key field in the NLRI. The key field is typically the real NLRI. The non-key field carries extra data that is NLRI-specific and hence not located in the BGP path attributes for packing optimization purposes. For example, [RFC8277] carries the Prefix in the key field and one label (stack) in the non-key field. As another example, [I-D.ietf-idr-bgp-car] defines a BGP CAR SAFI explicitly carrying Key Fields and Non-Key Fields as a list of TLVs. In case of a BGP withdraw, the key is indicated in the MP_UNREACH_NLRI attribute to withdraw the unfeasible routes, while the non-key data is typically not encoded.

This specification defines a new BGP non-transitive attribute, the "NLRI_KEY_LIST attribute", to carry the NLRIs using the simple and existing format of MP_UNREACH_NLRI. Its most important use is for AFI/SAFI whose NLRI are considered to be at elevated risk of malformation. An example are AFI/SAFI that encode both a key field and a non-key field in the NLRIs of the MP_REACH_NLRI attribute, while encodes NLRI in a simpler way in the MP_UNREACH_NLRI attrbiute, e.g., with only the key field. For such AFI/SAFI, in case of an error in the MP_UNREACH_NLRI attribute preventing the identification of all NLRIs key, the parsing of the NLRI_KEY_LIST attribute is more likely allow the identification of all NLRIs key. This attribute is used to allow the treat-as-withdraw error-handling approach to be used when there is an error in the MP_REACH_NLRI attribute that prevents the parsing of its NLRIs.

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. NLRI_KEY_LIST Capability

To avoid the overhead of sending and receiving an attribute which is not understood by the BGP speaker receiving it, this document defines a new BGP Capability [RFC5492] "NLRI_KEY_LIST", of type TBD2, and of length zero. A BGP speaker that supports the reception of the NLRI_KEY_LIST attribute {#receiving} SHOULD advertise the NLRI_KEY_LIST Capability Advertisements. A BGP speaker SHOULD NOT send the NLRI_KEY_LIST attribute unless its peer has advertised the NLRI_KEY_LIST Capability. Note however that if the attribute is sent, it will cause no harm (an incapable implementation will disregard the attribute, per the base BGP specification). A potential reason to send the attribute to a peer that has not advertised support is to avoid fragmenting a peer group.

3. NLRI_KEY_LIST Attribute

3.1. NLRI_KEY_LIST Format

The NLRI_KEY_LIST attribute is an optional, non-transitive BGP path attribute with type code TBD1. The format of the NLRI_KEY_LIST attribute is the same as the format of the MP_UNREACH_NLRI as defined in Section 4 of [RFC4760] and the relevant specification for the AFI/SAFI in question.

3.2. Sending the NLRI_KEY_LIST attribute

The NLRI_KEY_LIST attribute may be sent in a BGP UPDATE message carrying the MP_REACH_NLRI attribute. It MUST NOT be sent in an UPDATE message not carrying the MP_REACH_NLRI attribute. To facilitate the determination of the NLRI key list in an UPDATE message with a malformed attribute, the NLRI_KEY_LIST SHALL be encoded as the very first path attribute in an UPDATE message, followed by the MP_REACH_NLRI attribute. (This represents an update to Section 5.1 [RFC7606], which mandated that the MP_REACH_NLRI come first.) The ordering of NLRIs within the NLRI_KEY_LIST MUST be the same as their ordering within the corresponding MP_REACH_NLRI.

If the AFI/SAFI specification allows for different NLRI encodings in the MP_UNREACH_NLRI, the sender MUST use the simplest encoding. The receiver MUST accept any valid encoding. For example, [RFC3107] allows the use of either the MPLS label stack originally sent or the static 0x800000 value. The latter is simpler in that the size is smaller, fixed, and the number of labels to parse is minimized.

The NLRI_KEY_LIST attribute is generally useful as its encoding is simpler than the encoding of the MP_REACH_NLRI, hence it maximizes the chances of handling an error in the MP_REACH_NLRI attribute using

the treat-as-withdraw approach. In particular the NLRI_KEY_LIST attribute does not carry the variable length "Network Address of Next Hop" field nor the "Length of Next Hop Network Address" which, if erroneous, trigger a BGP session reset as per [RFC7606]. It is notably, although not exclusively, useful for AFI/SAFI carrying non-key data in the NLRI such as [RFC8277], [I-D.ietf-idr-bgp-car], and [I-D.ietf-idr-bgp-ct] as these NLRI are longer and more complex, hence have a higher probability of error. In addition, in case of error, they have a lower probability of being able to parse the full list of NLRIs. It is less useful when the NLRI encoding is the same for MP_REACH_NLRI and MP_UNREACH_NLRI.

3.3. Receiving the NLRI_KEY_LIST attribute

An UPDATE message with a malformed MP_REACH_NLRI attribute and a correctly formed NLRI_KEY_LIST attribute SHALL be handled using the approach of "treat-as-withdraw". The UPDATE message SHALL be handled as if received with only the NLRI_KEY_LIST attribute - all other attributes being ignored - and the NLRI_KEY_LIST attribute handled as an MP_UNREACH_NLRI attribute.

In the case of an UPDATE message with a correctly formed MP_REACH_NLRI attribute, the NLRI_KEY_LIST attribute SHOULD be parsed and its list of NLRI compared to the list of NLRI present in the MP_REACH_NLRI attribute. In case of difference, the NLRI_KEY_LIST attribute SHALL be ignored. However, because this reveals an error in either the NLRI_KEY_LIST attribute or the MP_REACH_NLRI attribute, a BGP speaker must provide debugging facilities to permit issues caused by a malformed attribute to be diagnosed. At a minimum, such facilities must include logging an error listing the NLRI involved and containing the entire malformed UPDATE message when such an attribute is detected. The malformed UPDATE message should be analyzed, and the root cause should be investigated.

3.4. NLRI_KEY_LIST attribute Error Handling

The NLRI_KEY_LIST attribute has the same format as the MP_UNREACH_NLRI and hence has the same conditions under which it is considered malformed. As per Section 3.3, an UPDATE message with a malformed NLRI_KEY_LIST attribute is handled using the approach of "attribute discard".

4. Operational Considerations

The choice of whether or not to send the NLRI_KEY_LIST attribute is up to the implementor and the operator. As is discussed elsewhere in this document, the attribute is considered more likely to be valuable for AFI/SAFI with more complex NLRI encodings, and less likely to be valuable in the case where the encodings used for the MP_REACH_NLRI and MP_UNREACH_NLRI are the same.

Drawbacks of sending the attribute include space overhead in the UPDATE message, as well as time overhead to form the attribute on the sender side and to validate it on the receiver side.

The primary advantage of sending the attribute is avoidance of session reset with concomitant service disruption, but one may also observe the potential to detect non-fatal errors which would otherwise be invisible, when the NLRI_KEY_LIST attribute is compared to the MP_REACH_NLRI attribute. The latter might motivate an operator to configure support even for less complex AFI/SAFI.

An implementation SHOULD provide a configuration option allowing the operator to send, or not send, the attribute with any AFI/SAFI. The default may differ for different AFI/SAFI; this specification does not, in any case, mandate a default.

5. IANA Considerations

IANA is requested to allocate a new optional, non-transitive attribute called "NLRI_KEY_LIST" from the BGP Path Attributes registry of the Border Gateway Protocol (BGP) Parameters group.

Value	Code	Reference
TBD1	NLRI_KEY_LIST	(this doc)

Table 1

IANA is requested to allocate a new capability called "NLRI_KEY_LIST" from the Capability Codes registry.

Value	Description	Reference
TBD2	NLRI_KEY_LIST	(this doc)

Table 2

6. Security Considerations

The NLRI_KEY_LIST attribute does not change BGP security considerations. An attacker having the ability to send or modify a BGP message has the ability to withdraw any NLRI, with or without the NLRI_KEY_LIST attribute.

7. Acknowledgements

The authors of this specification thank Jeffrey Haas and Robert Raszuk for their review and comments.

8. References

8.1. 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/rfc/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/rfc/rfc4271>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/rfc/rfc4760>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/rfc/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/rfc/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/rfc/rfc8174>.

8.2. Informative References

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, DOI 10.17487/RFC3107, May 2001, <https://www.rfc-editor.org/rfc/rfc3107>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/rfc/rfc8277>.

[I-D.ietf-idr-bgp-car]
Rao, D. and S. Agrawal, "BGP Color-Aware Routing (CAR)", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-car-16, 20 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-car-16>.

[I-D.ietf-idr-bgp-ct]
Vairavakkalai, K. and N. Venkataraman, "BGP Classful Transport Planes", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-ct-39, 28 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ct-39>.

Authors' Addresses

Bruno Decraene
Orange
Email: bruno.decraene@orange.com

John G. Scudder
HPE
Email: jgs@juniper.net