

Global Routing Operations
Internet-Draft
Updates: 7854 (if approved)
Intended status: Standards Track
Expires: 19 September 2024

P. Lucente
NTT
Y. Gu
Huawei
18 March 2024

BMP v4: TLV support for BMP Route Monitoring and Peer Down Messages
draft-ietf-grow-bmp-tlv-14

Abstract

Most of the message types defined by the BGP Monitoring Protocol (BMP) make provision for data in TLV format. However, Route Monitoring messages (which provide a snapshot of the monitored Routing Information Base) and Peer Down messages (which indicate that a peering session was terminated) do not. Supporting (optional) data in TLV format across all BMP message types allows for a homogeneous and extensible surface that would be useful for the most different use-cases that need to convey additional data to a BMP station. While it is not intended for this document to cover any specific utilization scenario, it defines a simple way to support TLV data in all message types.

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 19 September 2024.

Copyright Notice

Copyright (c) 2024 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. Terminology	3
3. TLV encoding	3
4. BMP Message Format	5
4.1. Common Header	5
4.2. TLV data in Route Monitoring	5
4.2.1. Group TLV	5
4.2.2. VRF/Table Name TLV	6
4.2.3. Stateless parsing TLV	6
4.2.4. Wire-format example	7
4.3. TLV data in Peer Down	9
4.4. TLV data in other BMP messages	9
5. Error handling	9
6. Security Considerations	9
7. Operational Considerations	9
8. IANA Considerations	9
9. Normative References	10
Acknowledgements	11
Authors' Addresses	11

1. Introduction

The BGP Monitoring Protocol (BMP) is defined in RFC 7854 [RFC7854].

The Route Monitoring message consists of:

- * Common Header
- * Per-Peer Header
- * BGP Update PDU

The Peer Down Notification message consists of:

- * Common Header
- * Per-Peer Header

- * Reason
- * Data (only if Reason code is 1, 2 or 3)
- * TLV (only if Reason code is 6)

This means that both Route Monitoring and Peer Down messages have a non-extensible format (except for the specific case of Peer Down Reason Code 6 as defined by RFC 9069 [RFC9069]). In the Route Monitoring case, this prevents the transmission of characteristics of transported NLRIs (e.g. to help with stateless parsing) or of vendor-specific data. In the Peer Down case, this prevents matching with TLVs previously sent with the Peer Up message. The proposal of this document is to:

- * Bump the BMP version, for backward compatibility
- * Allow all defined BMP message types to make provision for TLV data
- * Change the structure of Route Monitoring so that the BGP Update PDU is enclosed in a TLV

2. Terminology

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 RFC 2119 [RFC2119] RFC 8174 [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. TLV encoding

The TLV data type is already defined in Section 4.4 of [RFC7854] for the Initiation and Peer Up message types. A TLV consists of:

- * 2 octets of TLV Type,
- * 2 octets of TLV Length,
- * 0 or more octets of TLV Value.

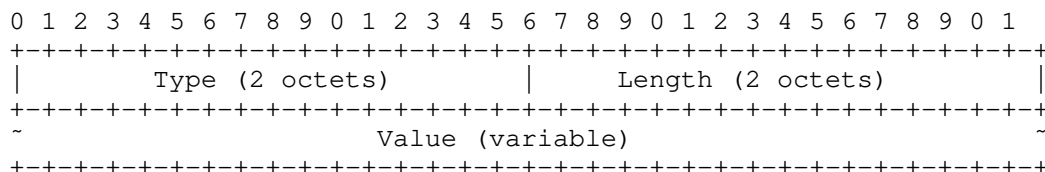


Figure 1

TLVs SHOULD be sorted by the sender by their code point. Multiple TLVs of the same type can be repeated as part of the same message, and it is left to the specific use-cases whether all, any, the first or the last TLV should be considered as well as whether ordering matters.

Route Monitoring messages may require per-NLRI TLVs, that is, there may be a need to map TLVs to NLRIs contained in the BGP Update message, for example, to express additional characteristics of a specific NLRI. For this purpose specifically, TLVs in Route Monitoring messages MUST be indexed, with the index starting at one (1) to refer to the first NLRI. Index zero (0) specifies that a TLV does apply to all NLRIs contained in the BGP Update message. The Index field is 2 bytes long of which the top-most bit of the Index field, G-bit, is reserved to flag a Group Index (more in Section 4.2.1). Multiple TLVs of the same type and with the same index can be repeated as part of the same message. Indexed TLVs are encoded as in the following figure:

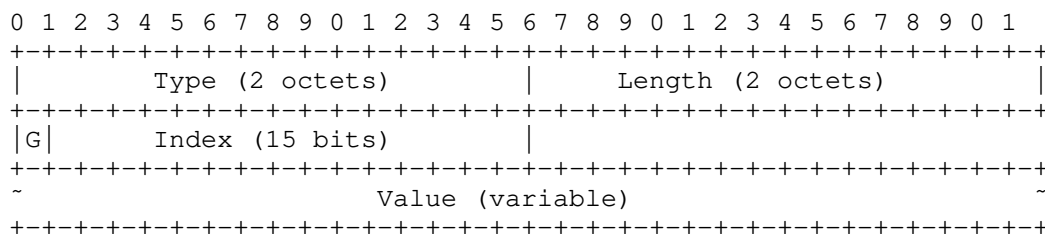


Figure 2

Indexed TLVs SHOULD be sorted by the sender by their code point and index value. Also in indexed TLVs, the reported length refers to the total encoded TLV value (ie. it does exclude the length of the index field).

Of the BMP message types defined so far, indexed TLVs apply only to Route Monitoring messages and, for example, they do not apply to Route Mirroring messages because the sender may not be aware of the payload of the transported BGP Update message.

4. BMP Message Format

4.1. Common Header

Section 4.1 of [RFC7854] defines the Common Header. While the structure remains unaltered, the following two definitions are changed:

- * Version: Indicates the BMP version. This is set to '4' for all messages.
- * Message Length: Total length of the message in bytes (including headers, encapsulated BGP message and optional data)

4.2. TLV data in Route Monitoring

The Route Monitoring message type is defined in Section 4.6 of [RFC7854]. The consistency model selected to extend this message type with TLVs is with the Route Mirroring type where the Per-peer header is being followed by TLVs.

The BGP Update PDU Section 4.3 of [RFC4271] is encoded itself as part of a BGP Message TLV with code point TBD1 and index set to zero. A Route Monitoring message MUST contain one BGP Message TLV which may be preceded and/or followed by other optional TLV data.

Corollary, in BMPv4 the BGP Update PDU is not just encoded as part of the message as it was the case for BMPv3 but it is rather enclosed in a TLV.

4.2.1. Group TLV

In a Route Monitoring message where the BGP Update PDU carries N NLRIs, indexed TLVs do allow to handle the cases of 1:1 and N:1 relationship among NLRIs and TLVs. The case of N:M relationship does instead require a form of grouping and hence this document defines a Group TLV to achieve this.

The TLV MUST contain:

- * A 2 bytes Group Index where the top-most bit, G-bit or Group Bit MUST be set to one (1). The full 2 bytes value, that is including the G-bit, MUST be unique to the message

- * Two or more 2 bytes indexes whose values MUST be less or equal to the amount of NLRIs packed in the BGP Update PDU.

A NLRI index can be listed as part of multiple Group TLVs within the same message. NLRI indexes within a Group TLV SHOULD be sorted by the sender. A Group Index can not reference an NLRI index 0. Finally, a Group TLV MUST NOT include its own or another Group Index. Multiple TLVs can point to the same Group Index, i.e. a group can be reused within the same Route Monitoring message.

The Group TLV code point is TBD2. It is recommended that this TLV is encoded first in order to ease parsing of the Route Monitoring message at the BMP station side.

4.2.2. VRF/Table Name TLV

The Information field contains a UTF-8 string whose value MUST be equal to the value of the VRF or table name (ie. RD instance name) being conveyed. The string size MUST be within the range of 1 to 255 bytes.

The VRF/Table Name TLV code point is TBD3

4.2.3. Stateless parsing TLV

Stateless parsing helps scaling the amount of Route Monitoring messages that can be processed at collection time, avoiding to have to correlate them to BGP capabilities received as part of the Peer Up message, for example.

Some BGP capabilities are not per AFI/SAFI, like 4-bytes ASN RFC 6793 [RFC6793], and hence these can be part of the BMP Peer flags section of a Route Monitoring message. Those that are, instead, per AFI/SAFI require finer granularity and hence the use of an indexed TLV.

The Stateless Parsing TLV code point is TBD4 and is organized as a 2-bytes registry. The following code points are defined for the Stateless Parsing registry:

- * Type = TBD5: the BGP Update PDU is encoded with the ADD-PATH capability RFC 7911 [RFC7911], value MUST be 0 for false and 1 for true.
- * Type = TBD6: the BGP Update PDU is encoded with the Multiple Labels capability RFC 8277 [RFC8277], value MUST be 0 for false and 1 for true.

If a Stateless parsing TLV is not present in a Route Monitoring message, the receiver MUST fall back to use capabilities present in the BGP Open PDU contained in the relevant BMP Peer Up message in order to properly parse BGP Update PDUs.

It is recommended that the Stateless Parsing TLV is encoded preceeding the BGP Message TLV in order to ease parsing of the Route Monitoring message at the BMP station side.

4.2.4. Wire-format example

The diagram in Figure 3 shows an example of a Route Monitoring message carrying a BGP UPDATE containing 10 NLRIIs. The TLVs are comprised of:

1. a Group TLV with index 0x000b, pointing to NLRI 1, 2, 3 and 10
2. a Group TLV with index 0x000c, pointing to NLRI 4, 5 and 6
3. a Stateless Parsing TLV with Index 0x800b (Group TLV bit set to 1, Index 0x000b)
4. a TLV pertaining to NLRI 7
5. a TLV pertaining to the NLRIIs listed in the Group TLV defined in 1
6. a TLV pertaining to the NLRIIs listed in the Group TLV defined in 2

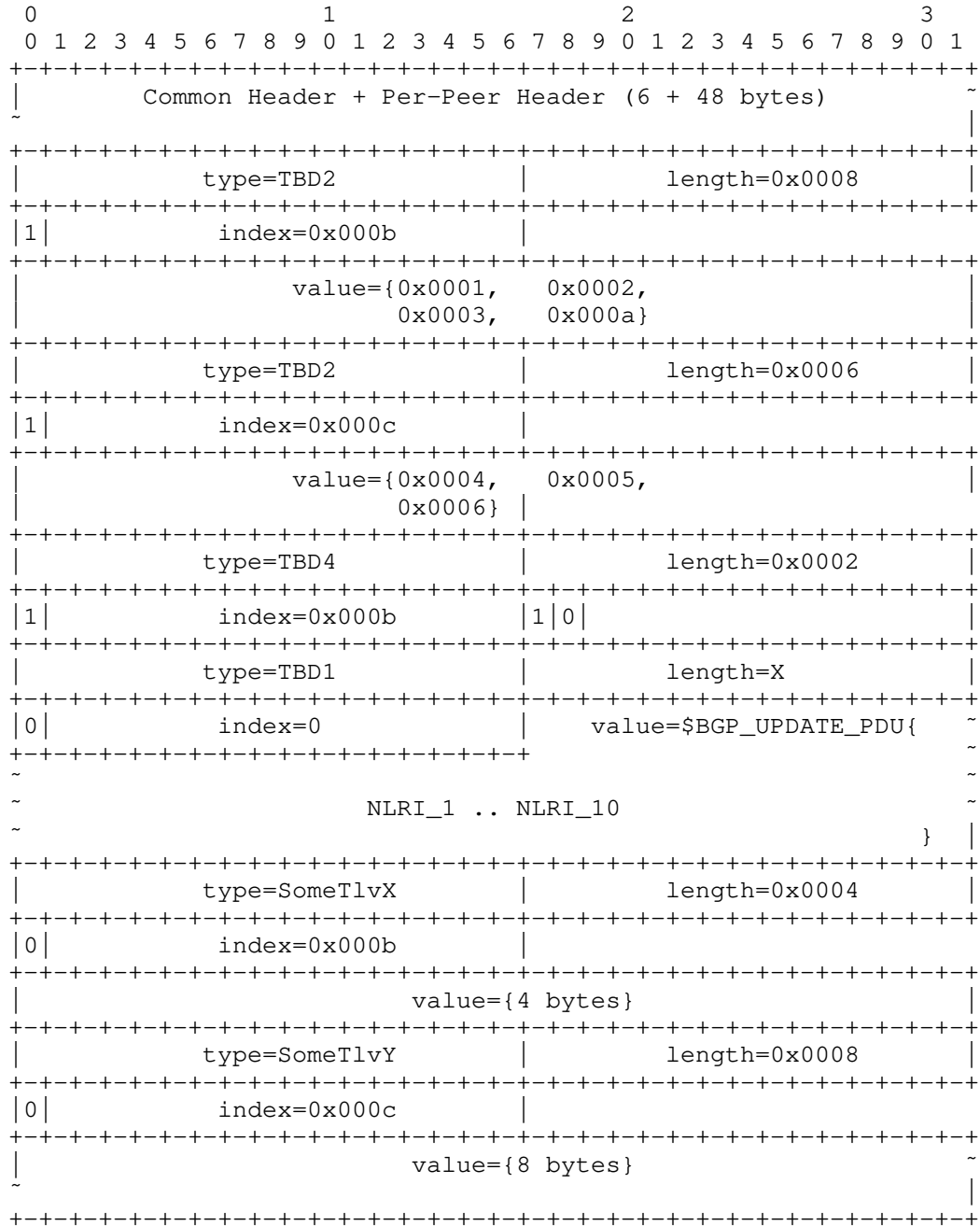


Figure 3

4.3. TLV data in Peer Down

The Peer Down Notification message type is defined in Section 4.9 of [RFC7854]. The consistency model selected to extend this message type with TLVs is with the Peer Up type where optional TLVs are placed at the end of the message.

This means for Reason codes 1 or 3, a BGP Notification PDU follows; the PDU MAY be followed by TLV data. For Reason code 2, a 2-byte field to give additional FSM info follows; this field MAY be followed by TLV data. For all other Reason codes, TLV data MAY follow the Reason field.

4.4. TLV data in other BMP messages

All other message types defined in RFC7854 [RFC7854] do already provision for TLV data. It is RECOMMENDED that all future BMP message types will also provide for TLV data following a consistency model with existing message types.

5. Error handling

It is worth nothing that RFC8654 [RFC8654] permits BGP Update and other messages to grow to a length of 65535 octets. This may cause a BMP PDU that attempts to encapsulate such long messages to overflow.

6. Security Considerations

It is not believed that this document adds any additional security considerations.

7. Operational Considerations

In Route Monitoring messages, the number of TLVs can be bound to the amount of NLRIs carried in the BGP Update message. This may degrade the packing of information in such messages and have specific impacts on the memory and CPU used in a BMP implementation. As a result of that it should always be possible to disable such features to mitigate their impact.

8. IANA Considerations

This document requests the renaming of the "Peer Up TLVs" registry defined by BMP Peer Up Message Namespace [I-D.ietf-grow-bmp-peer-up] into "Peer Up and Peer Down TLVs" and the definition of one new registry "BMP Route Monitoring TLVs". As part of the "BMP Route Monitoring TLVs" registry, the following new TLV types are defined (Section 4.2):

- * Type = TBD1: Support for BGP Message TLV. The value field is defined in Section 4.2
- * Type = TBD2: Support for grouping of TLVs. The value field is defined in Section 4.2.1. The recommended value for this TLV is 0.
- * Type = TBD3: Support for VRF/Table Name TLV. The value field is defined in Section 4.2.2
- * Type = TBD4: Support for Stateless Parsing TLV. The value field is defined in Section 4.2.3. The recommended value for this TLV is 1.

This document also requests the definition of a "Stateless Parsing TLV" registry seeded as follows:

- * Type = TBD5: ADD-PATH capability flag. Set to 1 if the BGP Update PDU enclosed in the Route Monitoring message was encoded according to the capability. The flag is defined in Section 4.2.3. The recommended value for this flag is 0.
- * Type = TBD6: Multiple Labels capability flag. The Set to 1 if the BGP Update PDU enclosed in the Route Monitoring message was encoded according to the capability. The flag is defined in Section 4.2.3. The recommended value for this flag is 1.

9. Normative References

- [I-D.ietf-grow-bmp-peer-up]
Scudder, J. and P. Lucente, "BMP Peer Up Message Namespace", Work in Progress, Internet-Draft, draft-ietf-grow-bmp-peer-up-03, 6 February 2024,
<<https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-peer-up-03>>.
- [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>>.

- [RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, DOI 10.17487/RFC6793, December 2012, <<https://www.rfc-editor.org/info/rfc6793>>.
- [RFC7854] Scudder, J., Ed., Fernando, R., and S. Stuart, "BGP Monitoring Protocol (BMP)", RFC 7854, DOI 10.17487/RFC7854, June 2016, <<https://www.rfc-editor.org/info/rfc7854>>.
- [RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", RFC 7911, DOI 10.17487/RFC7911, July 2016, <<https://www.rfc-editor.org/info/rfc7911>>.
- [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>>.
- [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/info/rfc8277>>.
- [RFC8654] Bush, R., Patel, K., and D. Ward, "Extended Message Support for BGP", RFC 8654, DOI 10.17487/RFC8654, October 2019, <<https://www.rfc-editor.org/info/rfc8654>>.
- [RFC9069] Evens, T., Bayraktar, S., Bhardwaj, M., and P. Lucente, "Support for Local RIB in the BGP Monitoring Protocol (BMP)", RFC 9069, DOI 10.17487/RFC9069, February 2022, <<https://www.rfc-editor.org/info/rfc9069>>.

Acknowledgements

The authors would like to thank Jeff Haas, Camilo Cardona, Thomas Graf, Pierre Francois, Ben Maddison, Tim Evens, Luuk Hendriks, Maxence Younsi, Ahmed Elhassany, Colin Petrie, Dhananjay Pakti and Shunwan Zhuang for their valuable input. The authors would also like to thank Greg Skinner and Zongpeng Du for their review.

Authors' Addresses

Paolo Lucente
NTT
Veemweg 23
3771 Barneveld
Netherlands
Email: paolo@ntt.net

Yunan Gu
Huawei
Huawei Bld., No.156 Beiqing Rd.
Beijing
100095
China
Email: guyunan@huawei.com