

BMP v4: TLV support

Support for BGP Monitoring Protocol (BMP) Route Monitoring and Peer Down Messages

draft-ietf-grow-bmp-tlv-15

Abstract

Most of ~~the message types defined by~~ the BGP Monitoring Protocol (BMP) message types make provision for data in Type, Length, Value (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 provides consistent~~ and extensible structures ~~among that would be useful for the most different use-cases that~~that the various cases where need to conveying additional data to a monitoringBMP station is required.

~~While it is not intended for this document to cover any specific utilization scenario, it~~ This document defines updates RFC 7854 a consistent and 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 21 July 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

Commenté [MB1]: Any mechanism to report failure due to version match?

Commenté [MB2]: Where BMP version negotiation is defined? How to ensure compatibility between peers? Some discussion is needed on these matters.

a mis en forme : Surlignage

a mis en forme : Surlignage

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. Message version 3

4. TLV encoding 3

5. BMP Message Format 5

5.1. Common Header 5

5.2. TLV data in Route Monitoring 5

5.2.1. Group TLV 6

5.2.2. VRF/Table Name TLV 6

5.2.3. Stateless parsing TLV 6

5.2.4. Wire-format example 7

5.3. TLV data in Peer Down 10

5.4. TLV data in other BMP messages 10

6. Error handling 10

7. Security Considerations 10

8. Operational Considerations 10

9. IANA Considerations 11

10. Normative References 11

Acknowledgements 13

Authors' Addresses 13

1. Introduction

The BGP Monitoring Protocol (BMP) version 3 is defined in ~~RFC 7854~~ [RFC7854].

Per Section 4.6 of [RFC7854], the BMP Route Monitoring message consists of:

- * Common ~~Headerheader~~ (Section 4.1 of [RFC7854])
- * Per-Peer ~~Headerheader~~
- * BGP Update PDU

Per Section 4.9 of [RFC7854], the Peer Down Notification message consists of:

- * Common ~~Headerheader~~
- * Per-Peer ~~Headerheader~~
- * Reason

Commenté [MB3]: De we really need to state this?

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

* TLV (only if Reason code is 6)

Commenté [MB4]: I think this can be deleted but references to the exact 7854 sections reminded in the next para.

~~This means that both~~ Both Route Monitoring and Peer Down BMP messages have a non-extensible format (except for the specific case of Peer Down Reason Code 6 as specified in Section 5.3 of [RFC9069]). In the Route Monitoring

case, this prevents the transmission of characteristics of transported NLRIs (e.g., to help with stateless parsing), status of a path after being processed by the BGP process 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 t~~This document ~~is to~~:

Commenté [MB5]: Can be elaborated further.

Commenté [MB6]: transmission of status of ...?

* Bumps the BMP version for all message types defined in ~~RFC 7854~~ [RFC7854] for backward compatibility.

* Changes the structure of Route Monitoring message type so that the BGP Update PDU is enclosed in a TLV. The BGP Message PDU TLV is Mandatory to be included.

* Allows all defined BMP message types to make provision for optional TLV data.

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.

~~xx~~

Commenté [MB7]: Add a sentence about BMP terms used in the doc.

~~xx~~

Commenté [BM18]: Add a section to list the main updates to 7854

3. Message ~~V~~ersion

For an exporter to flag a receiver that it does comply with this documentspecification, the Version field of the BMP Common ~~Headerheader~~, documented in Section 4.4 of [RFC7854], MUST be set to 4. This applies to every BMP message type.

Commenté [BM19]: How the monitoring/monitored entities agrees on the version(s) to use?

4. TLV ~~encoding~~Encoding

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

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

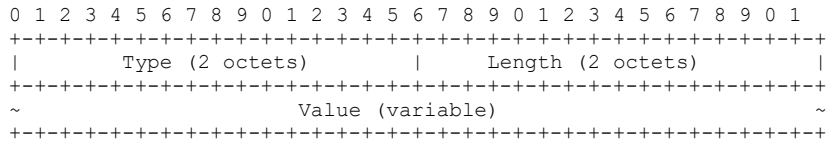


Figure 1: TLV Format

TLVs **SHOULD** be sorted by the sender by their **code-point** type. 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 and repeating is allowed.

- Commenté [MB10]:** What is the benefit? Shouldn't mandatory (if any) TLVs be listed first? What if this is not followed by a sender?
- Commenté [BMI11]:** To be consistent with the figure
- Commenté [MB12]:** That is? Do you meant the definition of a TLV? Or something else?

Route Monitoring messages may require per-NLRI TLVs. **That is**, there may be a need to map TLVs to NLRIs contained in **the a** BGP Update message, for example, to express additional characteristics of a specific NLRI. For this **purpose**, **specifically**, TLVs **enclosed** in **a** 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 **2-2-bytes** long of which the top-most bit, G-bit, is **reserved-used** to flag a Group Index (**more in** Section 5.2.1).

In general

TLVs of the same type and with **the same index** can be repeated as part of the same message, unless specified otherwise by the definition of the specific TLV. **Indexed TLVs** are encoded as **in the following figure** shown in Figure 2.

- Commenté [BMI13]:** With G-flag unset?
- Commenté [BMI14]:** How an impl knows it has to deal with a normal TLV or an indexed one?

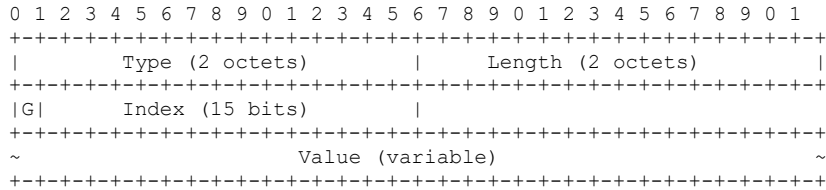


Figure 2: Indexing TLVs

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

A **decoder** can properly match indexed TLVs to the corresponding NLRI only if - or as long as - NLRIs are decoded successfully. In case of any parsing or error condition that prevents full decoding of **the a** BGP

- Commenté [MB15]:** Same comment as for TLVs.
- Commenté [BMI16]:** Any specific reason for this choice compared to enclosing the full length?
- Commenté [MB17]:** No used in RFC7854. Please a term that is already used there.

PDU, the decoder MUST ~~stop matching indexed TLVs to NLRIs~~.

~~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 a sender may not be aware of the payload of the transported BGP Update message.~~

Commenté [MB18]: Must ignore such received message?

Commenté [MB19]: Should that be tagged in registries such as <https://www.iana.org/assignments/bmp-parameters/bmp-parameters.xhtml#message-types> (as a new column)?

5. BMP Message Format

5.1. Common Header

~~Section 4.1 of [RFC7854] defines the Common Header.~~ While the ~~Structure~~structure of the Common header remains unaltered, the following two definitions are Changed compared to Section 4.1 of [RFC7854]:

* Version: Indicates the BMP version. This is set to '4' for all message types defined in RFC 7854 [RFC7854].

Commenté [MB20]: Already stated above

* Message Length: Total length of the message in bytes (including headers, encapsulated BGP Message PDU TLV and optional TLV data).

Commenté [MB21]: Isn't this already covered by the base spec:

5.2. TLV ~~D~~ata in Route Monitoring

~~The Route Monitoring message type is defined in Section 4.6 of [RFC7854]. For consistency with The consistency model selected by this document to extend encoding of such message type with TLVs is with the Route Mirroring type defined in (-Section 4.7 of [RFC7854]), this document extends the encoding of Route Monitoring message type with TLVs -where the Per-peer header is being followed by mandatory and optional 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 ~~preceeded~~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 [RFC7454] but it is rather enclosed in a TLV.

5.2.1. Group TLV

In a Route Monitoring message where ~~the a~~ BGP Update PDU carries N NLRIs, indexed TLVs do allow to handle the cases of 1:1 and N:1 relationship among TLVs and NLRIs (i.e., one TLV applies to one NLRI, N

TLVs apply to one same NLRI). The cases of 1:N and M:N relationships (i.e., one TLV applies to N NLRIs ~~and~~ M TLVs apply to N NLRIs) can benefit

by a form of grouping. ~~For that purpose, This is the context to define a Group TLV is defined to achieve this~~with the aim to limit both verbosity and repetitions.

The TLV value MUST contain:

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

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

An 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~~cannot reference an NLRI index 0. A Group TLV MUST NOT include its own or another Group Index. Multiple non-Group TLVs ~~can MAY~~ point to the same Group Index, ~~i.e.,~~ a group can be reused within the same Route Monitoring message.

The Group TLV ~~code-pointtype~~ 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.

5.2.2. 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 (~~i.e.,~~ RD instance name) being conveyed. The string size MUST be within the range of 1 to 255 bytes.

The VRF/Table Name TLV ~~code-pointtype~~ is TBD3

5.2.3. Stateless ~~parsing~~ 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-byte~~ 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 need to use an indexed TLV.

The Stateless Parsing TLV ~~code-pointtype~~ is TBD4 and is organized as a ~~5-bytes-byte~~ registry, one byte for the type, followed by AFI (2 bytes) and SAFI (1 byte) and finally one byte for the value. The type field encodes a ~~code-pointvalue~~ from the Stateless Parsing registry. The following ~~code-pointsvalues~~ are defined for this registry:

* Type = TBD5: the BGP Update PDU is encoded with the ADD-PATH capability ~~RFC-7911~~ [RFC7911], the value field 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], the value field MUST match

Commenté [MB22]: Group TLV?

a mis en forme : Surlignage

Commenté [MB23]: How?

Commenté [MB24]: MUST NOT?

Commenté [MB25]: Check of normative language is needed here

Commenté [MB26]: Cite a reference

Commenté [BMI27]: May be add a pointer to RFC9069#section 5.2.1?

Commenté [BMI28]: Add a reference

the Count for the relevant AFI/SAFI as it was encoded in the BGP Open PDU.

Only a single Stateless Parsing TLV is allowed in a message, ~~this TLV cannot be repeated~~. Also, the index of the TLV MUST be set to zero.

If the 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 ~~preceedingpreceeding~~ the BGP Message TLV in order to ease parsing of the Route Monitoring message at the BMP station side.

Commenté [BMI29]: RECOMMENDED?

5.2.4. Wire-format ~~example~~Example

~~The diagram in~~ Figure 3 shows an example of a Route Monitoring message carrying a BGP UPDATE containing 10 NLRI's. 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 NLRI's listed in the Group TLV defined in 1
6. a TLV pertaining to the NLRI's listed in the Group TLV defined in 2

```

0          1          2          3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Common Header + Per-Peer Header (6 + 48 bytes)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               type=TBD2                               |                               length=0x0008                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|                               index=0x000b                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               value={0x0001, 0x0002,                               |                               |
|                               0x0003, 0x000a}                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               type=TBD2                               |                               length=0x0006                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|                               index=0x000c                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               value={0x0004, 0x0005,                               |                               |
|                               0x0006} |                               |                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

```

|          type=TBD4          |          length=0x0002          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|1|          index=0x000b          |1|0|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          type=TBD1          |          length=X          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0|          index=0          |          value=$BGP_UPDATE_PDU{ ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~
~
~          NLRI_1 .. NLRI_10          ~
~
~          }          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          type=SomeTlvX          |          SomeTLVX          |
length=0x0004          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0|          index=0x000b          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          value={4 bytes}          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          type=SomeTlvY          |          SomeTLVY          |
length=0x0008          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0|          index=0x000c          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          value={8 bytes}          ~
~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Figure 3: An Example of XXX

Commenté [BMI30]: Move to an appendix

5.3. TLV ~~data-Data~~ in Peer Down

The Peer Down Notification message type ~~(-is defined in Section 4.9 of [RFC7854]) is extended following a consistent approach with . The consistency model selected by this document to extend encoding of such message type is with the Peer Up type (defined in Section 4.10 of [RFC7854]). That is, -where the message is extended so that 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 further followed by ~~optional~~ TLV data. For Reason code 2, a 2-byte field follows to provide additional ~~FSM~~ info; this field MAY be followed by ~~optional~~ TLV data. For all other Reason codes, ~~optional~~ TLV data MAY follow the Reason field.

Commenté [BMI31]: Redundant with «MAY»

Commenté [BMI32]: Please expand

Commenté [BMI33]: Redundant with MAY

5.4. TLV ~~data-Data~~ in ~~O~~ther BMP ~~messagesMessages~~

All other message types defined in ~~RFC7854~~ [RFC7854] do already provision for TLV data.

It is RECOMMENDED that all future defined BMP message types will also provide for optional TLV data following a consistency model for encoding with existing message types.

6. ~~Error handlingHandling~~

Commenté [BMI34]: The version handling is not discussed

~~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.

7. Security Considerations

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

8. 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.

9. IANA Considerations

This document requests ~~IANA to the renaming-rename 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 5.2):

- * Type = TBD1: Support for BGP Message TLV. The value field is defined in Section 5.2
- * Type = TBD2: Support for grouping of TLVs. The value field is defined in Section 5.2.1. The recommended value for this TLV is 0.
- * Type = TBD3: Support for VRF/Table Name TLV. The value field is defined in Section 5.2.2
- * Type = TBD4: Support for Stateless Parsing TLV. The value field is defined in Section 5.2.3. 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 5.2.3. 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 5.2.3. The recommended value for this flag is 1.

Commenté [MB35]: At least, a pointer to BMP is needed.

An obvious risk to cite is having a large number of TLVs/Indexed or not that will consume the resources when decoding, etc.

May consider if there is a need to have guards to limit the max number of TLVs present in a message, etc.

Commenté [BMI36]: Also consider adding some manageability considerations: version activation, for example.

a mis en forme : Surlignage

Commenté [BMI37]: This doc is in AUTH48, may be better to handle this now before that one is published as RFC.

Commenté [BMI38]: What is the policy for this new registry?

10. 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-05, 2 October 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-peer-up-05>>.
- [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>>.

Commenté [BMI39]: This one is informative

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: geyunan@huawei.com