Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 20 March 2025

C. Cardona
P. Lucente
NTT
P. Francois
INSA-Lyon
Y. Gu
Huawei
T. Graf
Swisscom

16 September 2024

BGP Monitoring Protocol (BMP) Extension for Path Status TLV

draft-ietf-grow-bmp-path-marking-tlv-02

Abstract

The BGP Monitoring Protocol (BMP) provides an interface for obtaining BGP ~~Path~~ path information. ~~BGP Path Information~~Such information is conveyed within BMP Route Monitoring ~~(RM)~~ messages. This document ~~proposes an~~ specifies a BMP extension ~~to BMP~~ to convey the status of a path after being processed by the BGP process. This extension makes use of the TLV ~~mechanisms~~mechanisms described in draft-ietf-grow-bmp-tlv ~~[I-D.ietf-grow-bmp-tlv]~~ and draft-ietf-grow-bmp-tlv-ebit ~~[I-D.ietf-grow-bmp-tlv-ebit]~~.I

Commenté [MB1]: This can be deleted.

Anyway, citations should be avoided in abstract.

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

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 20 March 2025.

Table of Contents

1. Introduction

~~For a given prefix, m~~Multiple paths with different path status ~~(~~, e.g., the "best-path", "back-up path", "invalid", and so on~~),~~ may co-exist for a given prefix in the BGP RIBs after being processed by the BGP decision process.

The path status information is ~~currently~~ not carried in the BGP Update Message ~~RFC4271~~ (Section 4.3 of [RFC4271]) or in the BMP Update Message ~~RFC7854~~ (Section 4.6 of [RFC7854]).

External systems can use the path status for various applications. For example, ~~The~~ the path status is commonly checked by operators when performing troubleshooting. Having such status stored and tracked in a centralized system can enable the development of tools that facilitate this process. ~~Optimisation~~Optimization systems can ~~include~~ take into account the path status ~~in their process,~~, e.g., ~~and also use the status~~ as a validation source (since it can compare the calculated state to the actual outcome of the network, such as primary and backup path). ~~As a final example~~Also, path status information can complement data from other ~~centralized~~ sources (e.g., ~~of data, for example,~~ flow collectors).

Commenté [MB2]: Which one specifically from defined BMP message types?

Commenté [MB3]: Is this mention really needed?

Commenté [MB4]: May elaborate further on the use. You may indicate this can be used for correlation, cross-checking, etc.

This document defines a ~~so-called~~ Path Status TLV to convey the BGP path status to ~~the~~ a BMP server (Section 2.1). The BMP Path Status TLV is carried in the BMP Route Monitoring (RM) ~~Message~~message (Section 4.6 of [RFC7854].

2. Path Status TLV

~~This document defines two types of Path Status TLVs: one is the IANA-registered Path Status TLV, and the other is the Enterprise-specific Path Status TLV.~~

2.1. IANA-registered Path Status TLV

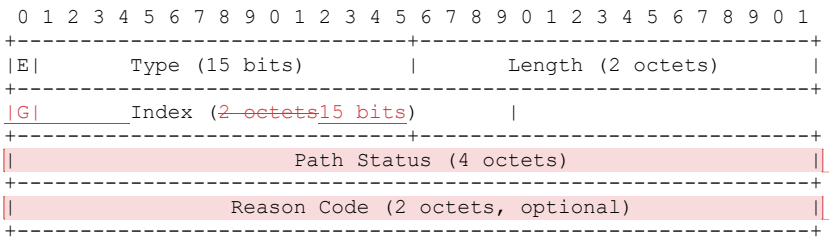

Figure 2: Encoding of IANA-Registered Path Status TLV

* E bit: ~~For an IANA-registered TLV, the E~~This bit MUST be set to 0 [I-D.ietf-grow-bmp-tlv-ebit].

* Type (15 Bits) = TBD2 ~~(15 Bits): indicates that it is the IANA-registered Path Status TLV.~~

* Length (2 Octets): indicates the length of the value field of the Path Status TLV. The value field ~~further~~ consists of the Path-Status field and Reason Code field.

* G-bit and Index (2 Octets): indicates the prefix that this TLV is describing. ~~Please see~~Refer [I-D.ietf-grow-bmp-tlv] for details of the use of the index field to associate the path marking content with one or more NLRIs.

* Path Status (4 Octets): indicates the path status of the BGP Update PDU encapsulated in ~~the~~ an RM Message. ~~Currently 10 types of path status are defined, as shown~~ Refer to ~~in~~ Table 1 for the allowed values. All zeros are reserved and MUST NOT be used.

* Reason Code (2 Octets, optional): indicates the reason of the path status indicated in the Path Status field. The reason code field is optional. If no reason code is carried, this field is empty. If a reason code is ~~carried~~present, the reason code is indicated by a 2-byte value, which is ~~defined~~ listed in Table 2.

Commenté [MB5]: The first sentence is stated right before the title.

Commenté [MB6]: Do we really need 4 octets for encoding the status?

Commenté [MB7]: The indicated length is inconsistent with the drawing! Please fix

Commenté [MB8]: No sure what is meant here

Commenté [MB9]: Better to point to the IANA registry

Commenté [MB10]: I guess this is not included, not empty.

Commenté [MB11]: Better to point to the IANA registry

Value	Path Ttype
0x00000001	Invalid
0x00000002	Best
0x00000004	Non-selected
0x00000008	Primary
0x00000010	Backup
0x00000020	Non-installed
0x00000040	Best-external
0x00000080	Add-Path
0x00000100	Filtered in inbound policy
0x00000200	Filtered in outbound policy
0x00000400	Invalid ROV
0x00000800	Stale
0x00001000	Suppressed

Table 1: IANA-Registered Path Types

~~Figure 1~~

The Path Status field contains a bitmap where each bit encodes a specific role of the path. Multiple bits may be set when multiple path status apply to a path.

* The ~~best-path~~best route is defined in ~~RFC4271~~ [RFC4271] and the best-external path is defined in ~~draft-ietf-idr-best-external~~ [I-D.ietf-idr-best-external].

* An invalid path is a route that does not enter the BGP decision process.

* A non-selected ~~path~~ route is a route that is not selected in the BGP decision process. Back-up routes are considered non-selected, while the best and ECMP routes are not considered as non-selected.

* A primary ~~path~~ route is a recursive or non-recursive ~~path~~ route whose next-hop resolution ends with an adjacency ~~draft-ietf-rtgwg-bgp-pic~~ [I-D.ietf-rtgwg-bgp-pic]. A prefix can have more than one primary path if multipath is configured ~~draft-lapukhov-bgp-ecmp-considerations~~ [I-D.lapukhov-bgp-ecmp-considerations]. A best~~-path~~ route is also considered as a primary ~~path~~route.

* A backup ~~path~~ route is also installed in the RIB, but it is not used until some or all primary paths become unreachable. ~~Backup paths are used for fast convergence in the event of failures.~~

* A non-installed ~~path~~ route refers to the route that is not installed into the IP routing table.

Commenté [MB12]: This is about the bit offset. Right?

Commenté [MB13]: Not described in the narrative text

Commenté [MB14]: Why not define this in a registry?

Commenté [MB15]: To be consistent with 4271

There other occurrences where the terminology is not aligned with 4271. Please check and align.

Commenté [MB16]: Do you need this mention? This I-D was expired since 2012!

Commenté [MB17]: You may order the description to match the order provided in Table 1

Commenté [MB18]: I initially thought this applies for Adj-RIBs-In, not Loc-RIB

Commenté [MB19]: How is this useful? Other than putting complexity in setting the corresponding bits

Commenté [MB20]: That is?

Commenté [MB21]: Why not simply refer to rfc7911?

* For the advertisement of multiple paths for the same address prefix without the new paths implicitly replacing any previous ones, the add-path status is applied [RFC7911].

* Stale refers to a path which has been declared stale by the BGP Graceful Restart mechanism as described in Section 4.1 of [RFC4724].

* Suppressed refers to a path which has been declared suppressed by the BGP Route Flap Damping mechanism as described in Section 2.2 of [RFC2439].

The Ppath ~~status~~ Status TLV does not force a BMP client to send any of these paths. It just provides a method to mark the paths that are available with their status.

Value	Reason code
[0x0001]	invalid for AS loop
[0x0002]	invalid for unresolvable next-hop
[0x0003]	not preferred for LOCAL PREF~~Local preference~~
[0x0004]	not preferred for AS Path Length
[0x0005]	not preferred for origin
[0x0006]	not preferred for MED
[0x0007]	not preferred for peer type
[0x0008]	not preferred for IGP cost
[0x0009]	not preferred for router ID
[0x000A]	not preferred for peer address
[0x000B]	not preferred for AIGP

Table 2: IANA-Registered Reason Codes

~~Figure 2~~

2.2. Enterprise-specific Path Status TLV

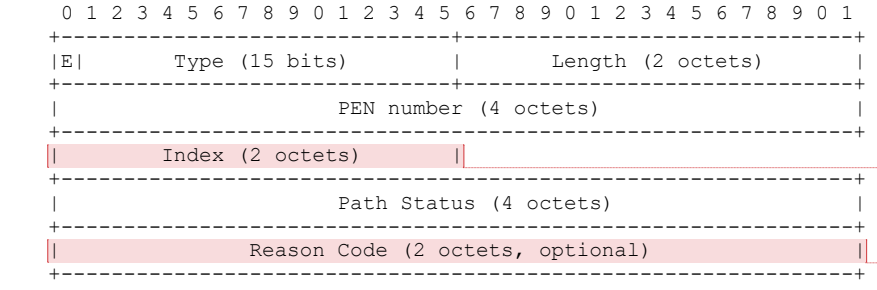

Figure 3: Encoding of Enterprise-specific Path Status TLV

* E bit: ~~For an Enterprise-specific TLV, the E bit~~ MUST be set to 1 [I-D.ietf-grow-bmp-tlv-ebit].

Commenté [MB22]: May indicate the status types for which a reason is valid?

Commenté [MB23]: Idem as for Table 1.

Commenté [MB24]: Why is this defined in the document?

Commenté [MB25]: Idem as for figure 2

Commenté [MB26]: Idem as for figure 2

* Type = 1 (15 Bits): indicates that ~~it's~~ it is the Enterprise-specific
  Path Status TLV.

* Length (2 Octets): indicates the length of the value field of the Path Status TLV. The value field further consists of the Path-Status field and Reason Code field.

* Index (2 Octets): indicates the prefix that this TLV is describing. The index is the encapsulation order, starting from 0, of the prefix in the BGP Update PDU.

* PEN Number (4 octets): indicates the IANA enterprise number IANA-PEN.

* Path Status (4 Octets): indicates the enterprise-specific path status. The format is to be determined w.r.t. each PEN number.

* Reason Code (2 octets, optional): indicates the reasons/ explanations of the path status indicated in the Path Status field. The format is to be determined w.r.t. each PEN number.

3. Implementation notes

The BMP path marking TLV remains optional within BMP implementations.

An implementation of the BMP path marking TLV may not fully support marking of all status defined in ~~table~~ Table ~~Figure~~ 1 or any future extensions. Similarly, an implementation may choose to support the inclusion of the reason code (for which support is also optional), without necessarily incorporating any of the reason codes defined in ~~table~~ Table ~~Figure~~ 2 or future extensions.

This document refrains from defining mechanisms for signaling the status or reason codes an implementation supports. This could be established through external means (e.g., documentation) or potentially addressed in a subsequent document.

Commenté [MB27]: Do we really need to say this?

The remainder of this section encompasses additional points related to the implementation of the BMP Path marking TLV.

3.1. Configuration of BMP Ppath ~~marking~~Marking

Implementations supporting the BMP Ppath ~~marking~~ Status TLV SHOULD provide ~~an~~ ~~Optiona~~ configuration parameter for ~~enabling or disabling~~controlling the Path ~~Marking~~ Status TLV over BMP sessions. Furthermore, the configuration options for this TLV SHOULD provide a configuration parameter to control ~~the means to enable and disable~~ the transmission of reason codes, if the reason code are supported by the implementation.

Commenté [MB28]: This is local to the implem. Not sure the normative language is needed.

Commenté [MB29]: Indicate a default value

Commenté [MB30]: Idem as previous comment

Commenté [MB31]: Indicate a default value

3.2. Paths with no ~~markings~~Markings

Some BGP routes might not require any type of status or reasons. For example, an unfiltered path obtained via the Adj-RIB-In ~~Adj-RIB-IN~~ may fall under this category since there is ~~really~~ nothing to mark for that

path. ~~We~~ This document suggests ~~a couple of approaches~~the following for signaling that a path has no markings: (1) An implicit form of marking, achieved by abstaining from appending any BMP marking TLV pointing toward the route. (2) Alternatively, an explicit marking of the packet through a TLV containing no marked status and no associated reason code.

Commenté [MB32]: For the sake of simplicity, the implicit approach seems reasonable.

I suggest to delete this text (given the statement in the sentence right before + there is no actual code for the explicit mode) :-))

3.3. Significance of status and origin RIBs

This document refrains from imposing any implementation to mark specific status from specific RIBs. ~~We recognize~~There is ~~the~~ diversity among implementations; some might be able to mark some status over one RIB while others do it on others. For instance, some might be able to mark Adj-RIB-In~~Adj-RIB-in~~ filtered routes when obtained from the Adj-RIB-In~~Adj-RIB-IN~~ pre, while other could do it only from the Adj-RIB-In ~~Adj-RIB-IN~~ post. To remove ambiguities in implementations, we recommend the meaning of status (and reason codes) to not depend on the origin RIB of a route.

Commenté [MB33]: I'm afraid this is not within the scope of the document.

As a general comment: the document should simply expose/export whatever seen/set by a BGP speaker.

3.4. Enterprise-specific status and reasons

Implementations introducing their own status and reason codes are advised to adhere to [I-D.ietf-grow-bmp-tlv-ebit] and use ~~ebit~~ E-bit and vendor specific status and reasons. Additionally, we recommend all implementations to provide comprehensive documentation for these codes.

Commenté [MB34]: I'm afraid this smells like discouraging registering new values

For scenarios where a ~~path~~ route state combines a standard status with an enterprise-specific reason code (or vice versa), the following alternatives are presented:

* Replication of the standard definitions within the enterprise-specific space, thus permitting direct marking within the same packet using the ebit.
* Assigning two TLVs to the same path(s): one containing the standard part and another housing the vendor-specific part.

Commenté [MB35]: I don't think this belongs to this doc.

3.5. Multiple TLVs ~~assigned~~ Assigned to the same route.

~~We advocate for the employment of TLV grouping wherever feasible.~~ The inclusion of all marking information within a single message is recommended, except on the case described in ~~section~~ Section 3.4. In situations where multiple TLVs are associated with a single route, all markings will be applicable to that route.

Commenté [BMI36]: This can be added to the description of the TLV itself

4. Acknowledgments

We would like to thank Jeff Haas and Maxence Younsi for their valuable comments.

5. IANA Considerations

This document requests that IANA assign the following new ~~parameters~~type from the IANA registry at XXX

Commenté [BMI37]: Please add where to find the registry

~~to the BMP parameters name space.~~

Type = TBD1 ~~(15 Bits):~~ indicates that it is ~~the IANA-registered~~ Path Status TLV.

6. Security Considerations

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

7. Normative References

[I-D.ietf-grow-bmp-tlv]
Lucente, P. and Y. Gu, "BMP v4: TLV support for BMP Route Monitoring and Peer Down Messages", Work in Progress, Internet-Draft, draft-ietf-grow-bmp-tlv-14, 18 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-tlv-14>.

[I-D.ietf-grow-bmp-tlv-ebit]
Lucente, P. and Y. Gu, "Support for Enterprise-specific TLVs in the BGP Monitoring Protocol", Work in Progress, Internet-Draft, draft-ietf-grow-bmp-tlv-ebit-05, 18 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-tlv-ebit-05>.

[I-D.ietf-idr-best-external]
Marques, P., Fernando, R., Chen, E., Mohapatra, P., and H. Gredler, "Advertisement of the best external route in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-best-external-05, 3 January 2012, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-best-external-05>.

[I-D.ietf-rtgwg-bgp-pic]
Bashandy, A., Filsfils, C., and P. Mohapatra, "BGP Prefix Independent Convergence", Work in Progress, Internet-Draft, draft-ietf-rtgwg-bgp-pic-21, 7 July 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-bgp-pic-21>.

[I-D.lapukhov-bgp-ecmp-considerations]
Lapukhov, P. and J. Tantsura, "Equal-Cost Multipath Considerations for BGP", Work in Progress, Internet-Draft, draft-lapukhov-bgp-ecmp-considerations-12, 28 December 2023, <https://datatracker.ietf.org/doc/html/draft-lapukhov-bgp-ecmp-considerations-12>.

[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>.

[RFC2439] Villamizar, C., Chandra, R., and R. Govindan, "BGP Route Flap Damping", RFC 2439, DOI 10.17487/RFC2439, November 1998, <https://www.rfc-editor.org/info/rfc2439>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A

Commenté [BMI38]: Relying on the status may have implications on other app mentioned in the intro and may influence them. Some caution need to be in place.

Commenté [BMI39]: I'd remove this

Commenté [BMI40]: Idem

Border Gateway Protocol 4 (BGP-4)", RFC 4271,
DOI 10.17487/RFC4271, January 2006,
<https://www.rfc-editor.org/info/rfc4271>.

[RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y. Rekhter, "Graceful Restart Mechanism for BGP", RFC 4724, DOI 10.17487/RFC4724, January 2007, <https://www.rfc-editor.org/info/rfc4724>.

[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>.

Authors' Addresses

Camilo Cardona
NTT
164-168, Carrer de Numancia
08029 Barcelona
Spain
Email: camilo@ntt.net

Paolo Lucente
NTT
Siriusdreef 70-72
2132 Hoofddorp
Netherlands
Email: paolo@ntt.net

Pierre Francois
INSA-Lyon
Lyon
France
Email: Pierre.Francois@insa-lyon.fr

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

Thomas Graf
Swisscom

Binzring 17
CH-8045 Zurich
Switzerland
Email: thomas.graf@swisscom.com