Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 16 November 2026

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

15 May 2026

BMP Extension for Path Status TLV
draft-ietf-grow-bmp-path-marking-tlv-05

Abstract

The BGP Monitoring Protocol (BMP) provides an interface for obtaining BGP path information, which is conveyed through BMP Route Monitoring (RM) messages. This document specifies a BMP extension to convey the status of a path after being processed by the BGP process.

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

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Multiple paths with different path statuses (e.g., the "best path", "backup 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 not carried in the BGP UPDATE Message [RFC4271] or in the BMP Route Monitoring Message [RFC7854].

External systems can use the path status for various applications. For example, operators commonly use path status during troubleshooting. Having such status stored and tracked enables the development of tools that facilitate this process. Optimization systems can consider path status in their process, e.g., as a validation source (since it can compare the calculated state to the actual outcome of the network, such as primary and backup path). Moreover, path status information can complement other centralized sources of data. For example, flow collectors can leverage it to identify the exact forwarding paths, yielding more accurate traffic matrices than existing methods.

This document defines a Path Status TLV to convey the BGP path status to a BMP server. The BMP Path Status TLV is carried in the BMP Route Monitoring (RM) Message [RFC7854].

2. Path Status TLV encoding

The path status TLV follows the common header defined in [I-D.ietf-grow-bmp-tlv] and [I-D.ietf-grow-bmp-tlv-ebit].

```
 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 TLV Header (Variable bits)                  |
+---------------------------------------------------------------+
|                       Path Status (4 octets)                  |
+---------------------------------------------------------------+
| Reason Code (2 oct., opt.)    |
+-------------------------------+
```

Figure 2: Encoding of Path Status TLV

* The common TLV header that can encode IANA-registered TLV or Enterprise-specific markings using [I-D.ietf-grow-bmp-tlv-ebit].

* Path Status (4 Octets): indicates the path status of the BGP Update PDU encapsulated in the RM Message. The path status is encoded using a bitmap where each bit position encodes a specific role of the path. Multiple bits may be set when multiple path statuses apply to a path. All zeros are reserved for paths with no marking.

* 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 not included. If a reason code is carried, the reason code is indicated by a two-byte value.

3. IANA encoding of Path Status TLV

3.1. IANA path status codes

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

Table 1: IANA-Registered Path Status

Table 1 includes a list of IANA status codes. This list might be extended. An explanation of each of the types is included next:

* An invalid route is a route that does not enter the BGP decision process as indicated in Section 9.1.2 of RFC4271 [RFC4271].

* The best route is defined in Section 9.1 of RFC4271 [RFC4271].

* Nonselected routes are those that are not selected in the BGP decision process. Backup routes are considered nonselected, while the best and primary routes are not considered as nonselected.

* A primary route is a path used for traffic forwarding. See draft-ietf-rtgwg-bgp-pic [I-D.ietf-rtgwg-bgp-pic]. A prefix can have more than one primary path when multipath is configured draft-lapukhov-bgp-ecmp-considerations [I-D.lapukhov-bgp-ecmp-considerations]. The best path is also a primary path.

* A backup path is 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 primary path failures.

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

* The best external path is defined in draft-ietf-idr-best-external [I-D.ietf-idr-best-external].

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

* Filtered in inbound policy routes are those that are filtered in the Adj-RIB-In policy.

* Filtered in outbound policy routes are those that are filtered in the Adj-RIB-Out policy.

* Stale routes refer to paths which have been declared stale by the BGP Graceful Restart mechanism, as described in Section 4.1 of [RFC4724].

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

3.2. IANA reason codes

Table 2 includes a list of IANA reason codes. This list can be extended in future documents. This document includes a brief explanation of each code and the path status they are intended to explain. Please see Section 5.4 for notes on potentially inconsistencies in the path marking data.

* Invalid routes due to AS loop and unresolvable nexthop are defined in Section 9.1.2 of [RFC4271]. These codes target routes of type "Invalid".

* The reason codes starting with 'not preferred' are aimed at paths not selected as best, and describe the reason they were ranked lower in the decision process. AIGP is explained in RFC7311 [RFC7311]. The rest of the codes are described in Section 9.1.2.2 of [RFC4271].

Value	Reason Code
0x0001	Invalid due to AS loop
0x0002	Invalid due to unresolvable nexthop
0x0003	Not preferred for 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

4. Path Marking Statistics

This section defines statistics to account and summarize Path Status TLV in the BGP Local-RIB. The statistics specified in this section are exported by BMP speakers using the Stats Reports message type defined in Section 4.8 of [RFC7854].

Unless otherwise stated, all counters defined in this section are unsigned 64-bit integers. Counters are maintained per BMP session and per monitored peer. The reset behavior of these counters (for example, upon BMP session restart or operator-triggered reset) follows the general semantics of BMP Stats Reports as defined in RFC7854 [RFC7854].

4.1. Per-Status Counters

For each IANA-registered Path Status bit defined in Table 1, an implementation SHOULD maintain a counter that reflects the number of paths currently marked with that status. These counters are conceptually similar to the per-status gauges defined for primary, backup, stale, and suppressed routes in the BMP RIB statistics document [I-D.ietf-grow-bmp-bgp-rib-stats].

The following counters are defined:

* Type = TBD1: Current number of paths for which the Path Status TLV has the Invalid status bit set.

* Type = TBD2: Current number of paths for which the Path Status TLV has the Best status bit set.

* Type = TBD3: Current number of paths for which the Path Status TLV has the Nonselected status bit set.

* Type = TBD4: Current number of paths for which the Path Status TLV has the Primary status bit set.

* Type = TBD5: Current number of paths for which the Path Status TLV has the Backup status bit set.

* Type = TBD6: Current number of paths for which the Path Status TLV has the Non-installed status bit set.

* Type = TBD7: Current number of paths for which the Path Status TLV has the Best-external status bit set.

* Type = TBD8: Current number of paths for which the Path Status TLV has the Add-Path status bit set.

* Type = TBD9: Current number of paths for which the Path Status TLV has the Filtered in inbound policy status bit set.

* Type = TBD10: Current number of paths for which the Path Status TLV has the Filtered in outbound policy status bit set.

* Type = TBD11: Current number of paths for which the Path Status TLV has the Stale status bit set.

* Type = TBD12: Current number of paths for which the Path Status TLV has the Suppressed status bit set.

Since the Path Status field is encoded as a bitmap, multiple status bits can be set for a single path. The corresponding per-status counters are incremented independently; that is, a single path may contribute to more than one per-status counter.

4.2. Generic Path Status Counters

In addition to the per-status counters defined above, the following generic counters MAY be maintained by a BMP implementation:

* Type = TBD13: Current number of paths that carry a Path Status TLV with at least one status bit set (that is, paths that are explicitly marked).

* Type = TBD14: Current number of paths that either do not carry a Path Status TLV or carry a Path Status TLV with all bits set to zero.

5. 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 statuses defined in Table 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 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.

The remainder of this section covers additional points related to the implementation of the BMP Path Marking TLV.

5.1. Configuration of BMP Path Marking

Implementations supporting the BMP Path Marking TLV should provide an option for enabling/disabling the Path Marking TLV over BMP monitoring sessions. Furthermore, the configuration options for this TLV should provide the means to enable/disable the transmission of reason codes, if the reason codes are supported by the implementation.

5.2. Scalability and churn

The Path Marking TLV introduces metadata on the routes, which could increase the churn (Section 8.1.6 of RFC4098 [RFC4098]) of paths within the BMP session. For instance, if BMP Path Marking is configured, and a non-installed path changes status to a backup route, the device should send an update about this path with the new markings, even if its BGP attributes remain unchanged. Enabling reason codes could additionally increase the churn. Churn could be more pronounced during the start of a BGP session, where the device is processing all available routes.

If churn is undesired, an implementation could make use of "state compression" to hide state until paths converge (Section 5 of [RFC7854]). It could also initially send BMP routes without the path marking TLV, even if it were configured, and then add them once the implementation considers the path to be stable enough. This document does not provide a definitive solution for churn since it depends on the capabilities of an implementation and the requirements of an operator.

5.3. Paths with no markings

Some BGP routes might not require any type of status or reasons. For example, a path in Adj-RIB-In where the BGP best path decision has not been applied yet, falls under this category, since there is really nothing to mark for that path. This document suggests applying an explicit marking of this route, by attaching a BMP path marking TLV with no bits set. This will help BMP monitor stations to differentiate this case from those in which markings are not configured, or not yet attached by the device.

5.4. Path markings applicability and consistency

The status and reason codes from Table 1 and Table 2 are included based on use cases from network operators and defined following the most relevant protocol references available. While implementations are strongly encouraged to align with these code definitions, this document does not enforce strict validity rules for code combinations to accommodate the diversity of BGP implementations.

The experience during testing of this TLV revealed scenarios where implementations might combine codes differently than originally anticipated. For example, one test implementation marked routes with both 'Invalid' and 'Best' status bits set, which is contradictory from the point of view of [RFC4271], but made sense for their specific implementation.

Operators should apply their own validation checks on the data from TLVs and discuss potential inconsistencies with their vendors, and raise bugs if applicable.

5.4.1. Significance of status and origin RIBs

This document refrains from imposing on any implementation the requirement to mark specific status from specific RIBs. Some implementations 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 filtered routes when obtained from the Adj-RIB-In pre-policy, while others could do it only from the Adj-RIB-In post-policy. To remove ambiguities in implementations, it is recommended that the meaning of status (and reason codes) does not depend on the origin RIB of a route.

5.5. Multiple TLVs assigned to the same route.

We advocate for the use of TLV grouping wherever feasible (Section 5.2.1 of [I-D.ietf-grow-bmp-tlv]). The inclusion of all marking information within a single message is recommended. In situations where multiple TLVs are associated with a single route, all markings and reasons will be applicable to that route.

5.6. Enterprise-specific status

Implementations introducing their own status and reason codes are advised to adhere to [I-D.ietf-grow-bmp-tlv-ebit] and use the enterprise-bit (ebit) for vendor-specific status and reasons.

For scenarios where a path 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.

5.7. Multiple reason codes

The path marking TLV was not designed to optimally hold more than one reason code per path. However, if needed by a specific use case, the implementation can use two or more path markings TLVs for the same path listing the multiple reasons that apply to it.

6. Acknowledgments

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

7. IANA Considerations

This document requests that IANA assign the following new TLV type to the BMP Route Monitoring TLVs.

Type = 5: indicates that this is the IANA-registered Path marking TLV. The value field is defined in Section 2.

RFC Editor and IANA registry note: The registry is created with Section 10 of [I-D.ietf-grow-bmp-tlv] and populated with initial values 1-4. This document adds value 5 to the registry. Please remove this sentence before publishing the document as RFC.

In addition, this document requests IANA to assign new statistic types in the BMP Statistics Types (https://www.iana.org/assignments/bmp-parameters/bmp-parameters.xhtml#statistics-types) registry, part of the BMP parameters (https://www.iana.org/assignments/bmp-parameters/bmp-parameters.xhtml) registry group, for the Path Marking Statistics defined in Section 4.

Unless otherwise stated, all statistics defined below are 64-bit unsigned integer gauges whose Stat Data value field and Stat Length follow the rules defined in Section 4.8 of [RFC7854].

This document requests IANA to assign the following new BMP Statistics Types and to update the registry to reference the RFC number of this document:

* Type TBD1: Current number of paths for which the Path Status TLV has the Invalid status bit set.

* Type TBD2: Current number of paths for which the Path Status TLV has the Best status bit set.

* Type TBD3: Current number of paths for which the Path Status TLV has the Nonselected status bit set.

* Type TBD4: Current number of paths for which the Path Status TLV has the Primary status bit set.

* Type TBD5: Current number of paths for which the Path Status TLV has the Backup status bit set.

* Type TBD6: Current number of paths for which the Path Status TLV has the Non-installed status bit set.

* Type TBD7: Current number of paths for which the Path Status TLV has the Best-external status bit set.

* Type TBD8: Current number of paths for which the Path Status TLV has the Add-Path status bit set.

* Type TBD9: Current number of paths for which the Path Status TLV has the Filtered in inbound policy status bit set.

* Type TBD10: Current number of paths for which the Path Status TLV has the Filtered in outbound policy status bit set.

* Type TBD11: Current number of paths for which the Path Status TLV has the Stale status bit set.

* Type TBD12: Current number of paths for which the Path Status TLV has the Suppressed status bit set.

* Type TBD13: Current number of paths that carry a Path Status TLV with at least one status bit set (paths that are explicitly marked).

* Type TBD14: Current number of paths that either do not carry a Path Status TLV or carry a Path Status TLV with all bits set to zero.

8. Security Considerations

Using the path status information may affect other applications which rely on this information for operational decisions. Operators should secure BMP sessions and control access to TLV data to mitigate these risks.

9. References

9.1. Normative References

[I-D.ietf-grow-bmp-bgp-rib-stats]
Srivastava, M., Liu, Y., Lin, C., and J. Li, "Advanced BGP Monitoring Protocol (BMP) Statistics Types", Work in Progress, Internet-Draft, draft-ietf-grow-bmp-bgp-rib-stats-17, 3 December 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-bgp-rib-stats-17>.

[I-D.ietf-grow-bmp-tlv]
Lucente, P., Gu, Y., Younsi, M., and P. Francois, "BMP v4: Extended TLV Support for BGP Monitoring Protocol (BMP)", Work in Progress, Internet-Draft, draft-ietf-grow-bmp-tlv-20, 2 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-tlv-20>.

[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-06, 17 January 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-tlv-ebit-06>.

[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 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>.

[RFC7311] Mohapatra, P., Fernando, R., Rosen, E., and J. Uttaro, "The Accumulated IGP Metric Attribute for BGP", RFC 7311, DOI 10.17487/RFC7311, August 2014, <https://www.rfc-editor.org/info/rfc7311>.

[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>.

9.2. Informative References

[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., Mohapatra, P., and Y. Qu, "BGP Prefix Independent Convergence", Work in Progress, Internet-Draft, draft-ietf-rtgwg-bgp-pic-23, 15 February 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-bgp-pic-23>.

[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-13, 6 January 2025, <https://datatracker.ietf.org/doc/html/draft-lapukhov-bgp-ecmp-considerations-13>.

[RFC4098] Berkowitz, H., Davies, E., Ed., Hares, S., Krishnaswamy, P., and M. Lepp, "Terminology for Benchmarking BGP Device Convergence in the Control Plane", RFC 4098, DOI 10.17487/RFC4098, June 2005, <https://www.rfc-editor.org/info/rfc4098>.

Appendix A. Implementation Status

Note to the Editor: Please remove this section before publication.

This section records the status of known implementations of the BMP Path Marking TLV specified in this document. The information is based on interoperability and testbed activities, including the Swisscom IETF Daisy testbed, and is intended to help reviewers and operators understand current implementation and deployment status. This section is non-normative.

A.1. Huawei VRP NE8000

Huawei implemented the BMP Path Marking TLV, including Path Status and Path Reason, on NE8000 platforms running VRP R025C00SPC305T and subsequent R025C10 test images.

The implementation exports Path Marking information in BMP Route Monitoring messages for IPv4 and IPv6 unicast, as well as VPNv4 and VPNv6 Local-RIB. Support includes ADD-PATH and VPN peer distinguishers.

In the test environment, BMP sessions were established from NE8000 PEs to a collector and configured to monitor Adj-RIB-In and Adj-RIB-Out, both pre-policy and post-policy, as well as VPNv4 and VPNv6 Local-RIB across multiple VRF instances.

The implementation was exercised against a set of path selection scenarios, including best path, backup path, non-installed paths, and policy-filtered paths. The exported Path Status bits, such as Best, Nonselected, Primary, Backup, and Non-installed, and the corresponding Path Reason codes, for example not preferred due to local preference or router ID, were verified against the corresponding VRF RIB and BGP routing state.

A.2. pmacct BMP Collector

The BMP collector used in the implementation is pmacct with support for the Path Marking TLVs as defined in this document.

The implementation includes updates to align TLV type and value handling with the current draft, as well as changes to the TLV parsing logic to correctly process Path Status and Path Reason information in BMP Route Monitoring messages.

In the implementation tests, the collector is configured to receive BMP sessions and export decoded data as JSON logs, including Path Marking TLVs where available. This allows comparison between the BGP path state observed on the NE8000, for example best, backup, non-installed, or policy-filtered, and the corresponding BMP-exported Path Status and Reason values.

A.3. NetGauze and Wireshark Decoders

NetGauze is used as an independent BMP decoder to validate the encoding of Path Marking TLVs. Updates were made to improve BMPv4 support, including handling of ADD-PATH for IPv4/IPv6 MPLS unicast, Loc-RIB Peer Up messages, and general BMP message parsing. These changes help decode and inspect Route Monitoring messages carrying Path Status and Path Reason TLVs.

Wireshark nightly builds are used to inspect BMP Route Monitoring messages and verify the presence and decoding of Path Marking TLVs, including VPN peer distinguishers and SAFI-specific updates. Improvements in BMP TLV and Group TLV decoding help inspect VRF context and associated path-marking information.

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