Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 26, 2021

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

February 22, 2021

BMP (BGP Monitoring Protocol) Seamless Session
draft-tppy-bmp-seamless-session-00

Abstract

This document describes an optional BMP session lifecycle extension to prevent data duplication of previously exported messages when TCP session is re-established. It prevents loss of messages between TCP session re-establishments and increase overall BMP scalability.

Table of Contents

1. Introduction

With the constant increase of BGP paths, the increase of BMP BGP RIB coverage from RFC8671 [RFC8671] and draft-ietf-grow-bmp-local-rib [I-D.ietf-grow-bmp-local-rib], the addition of new TLVs such as draft-cppy-grow-bmp-path-marking-tlv [I-D.cppy-grow-bmp-path-marking-tlv] and draft-xu-grow-bmp-route-policy-attr-trace [I-D.xu-grow-bmp-route-policy-attr-trace], more BMP messages and BGP contexts, such as peering, route-policy or RIB, are exported from BMP client to server.

With each BMP session re-establishment, clients export the initial BGP RIB via BMP route-monitoring messages as described in section 5 of RFC7854 [RFC7854]. Regardless if the same messages were already exported in a previous BMP session or not. This leads to data duplication and unnecessary strain of the BMP client and server.

In a network most times BMP sessions are re-established within a short period of time due to connectivity interruption between BMP client and server or restart of the BMP server due to maintenance. Even though most BMP client implementations support a BMP buffering mechanism, messages are not buffered across BMP session re-establishment, thus leading to a loss of messages.

Therefore, the proposed BMP session lifecycle improvement covers

- Brief loss of connectivity between BMP client and server
- Seamless Maintenance of BMP server

It is based on RFC7413, TCP Fast Open [RFC7413], which allows previously established TCP transport sessions to be re-established more efficiently.

This draft describes how the BMP application MUST behave during TCP transport re-establishment period in order to prevent metric loss.

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

3. Definitions

Brief loss of connectivity between BMP client and server: Describes a period of time, in seconds, starting from the point in time in which the BMP client detects loss of connectivity to the BMP server and tries to re-establish the TCP session.

Maintenance of BMP server: Describes a period of time, in seconds, starting from when the BMP server daemon is restarted for maintenance purposes and the BMP client tries to re-establish the TCP session.

4. BMP Client and Server Capability

To support brief loss of connectivity between BMP client and server, the BMP client and server MUST support TCP Fast Open as described in RFC7413 [RFC7413].

To support seamless maintenance of a BMP server, the BMP client and server MUST support TCP Fast Open as described in RFC7413 [RFC7413] and the restart of the BMP server MUST distinguish between normal and seamless restart, wherever TCP Fast Open cookies are preserved or not.

5. Updated BMP Session Lifecycle

Section 3 of RFC7413 [RFC7413] describes the TCP Fast Open extension in the initial TCP SYN packet and the cookie handling during initial and subsequent re-establishment of the TCP transport session.

Section 3.3 of RFC7854 [RFC7854] describes that the BMP session closes with the TCP session. This behavior is extended with a configurable BMP session timeout.

The BMP session timeout starts counting down under the following conditions:

- Configured value is bigger than 0
- Current TCP session was established with Fast Open extension and cookie has been saved
- BMP buffer is not full
- TCP session is going to be terminated

The default BMP session timeout is 60 seconds.

While the time is counting down, all the BMP messages, regardless of message type, MUST be buffered. At this stage, the BMP session is still considered to be alive.

When a TCP session is re-established with TCP Fast Open extension and the cookie is identical to the previous TCP session with the same BMP peer, the BMP session remains alive, BMP buffer is exported and normal operation continues.

When a TCP session is re-established without TCP Fast Open extension or with TCP Fast Open extension but the cookie is not identical to the previous TCP session with the same BMP peer, the BMP session is considered terminated and starts with a new BMP Initiation message.

When a TCP session is not re-established within the configured timeout, then the BMP buffer is discarded and the BMP session is considered terminated.

When the BMP buffer is full before the TCP session is re-established, then the BMP buffer is discarded and the BMP session is considered terminated.

6. Security Considerations

The same security considerations apply as for TCP Fast Open RFC7413 [RFC7413].

7. Operational Considerations

From the perspective of the BMP server, the TCP Fast Open mechanism is rather transparent since it is entirely handled by the operating system kernel: this also means a BMP Server application can't determine if the TCP session was established with SYN Cookies or without them.

Upon terminating the existing BMP session(s), the BMP server should dump to persistent storage the BGP RIBs currently in memory. In terms of encoding, MRT format could be used for the task (ie. draft-petrie-grow-mrt-bmp)

At restart, the BMP server should first restore the content of BGP RIBs from persistent storage before accepting any incoming connection from BMP clients. Only once this process is finished, connections can then be accepted again so that messages buffered by BMP clients are applied to the last known BGP RIBs upon termination.

8. References

8.1. Normative References

[RFC7413] Cheng, Y., Chu, J., Radhakrishnan, S., and A. Jain, "TCP Fast Open", RFC 7413, DOI 10.17487/RFC7413, December 2014, <https://www.rfc-editor.org/info/rfc7413>.

[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>.

8.2. Informative References

[I-D.cppy-grow-bmp-path-marking-tlv] Cardona, C., Lucente, P., Francois, P., Gu, Y., and T. Graf, "BMP Extension for Path Status TLV", draft-cppy-grow-bmp-path-marking-tlv-07 (work in progress), October 2020.

[I-D.ietf-grow-bmp-local-rib] Evens, T., Bayraktar, S., Bhardwaj, M., and P. Lucente, "Support for Local RIB in BGP Monitoring Protocol (BMP)", draft-ietf-grow-bmp-local-rib-09 (work in progress), January 2021.

[I-D.xu-grow-bmp-route-policy-attr-trace] Xu, F., Graf, T., Gu, Y., Zhuang, S., and Z. Li, "BGP Route Policy and Attribute Trace Using BMP", draft-xu-grow-bmp-route-policy-attr-trace-05 (work in progress), July 2020.

[RFC8671] Evens, T., Bayraktar, S., Lucente, P., Mi, P., and S. Zhuang, "Support for Adj-RIB-Out in the BGP Monitoring Protocol (BMP)", RFC 8671, DOI 10.17487/RFC8671, November 2019, <https://www.rfc-editor.org/info/rfc8671>.

Authors' Addresses

Thomas Graf
Swisscom
Binzring 17
Zurich 8045
Switzerland

Email: thomas.graf@swisscom.com

Paolo Lucente
NTT
Siriusdreef 70-72
Hoofddorp, WT 2132
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