Global Routing Operations
Internet Draft
Intended status: Standards Track
Expires: August 19, 2025

Y. Liu
China Mobile
C. Lin
New H3C Technologies
T. Graf
Swisscom
P. Lucente
NTT

February 19, 2025

Using BMP over QUIC connection
draft-liu-grow-bmp-over-quic-02

Abstract

The BGP Monitoring Protocol (BMP) provides a convenient interface for obtaining route views by monitoring BGP sessions. BMP operates over TCP and is unidirectional (from client to server). QUIC provides multiple simultaneous streams to carry data in one direction, enabling much better efficiency and performance for both peers, in particular unidirectional streams can provide reverse data protection for the sender. QUIC also provides shorter handshake and includes TLS. This document describes how to use BMP over the QUIC transport protocol, named BMPoQUIC.

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 August 19, 2025.

Table of Contents

1. Introduction

The BGP Monitoring Protocol (BMP) [RFC7854] defines a standard mechanisms for obtaining route views by monitoring BGP sessions. BMP operation uses TCP as its transport protocol to provide reliable communication. BMP establishes connection relationships between monitored router and monitoring station using a TCP session.

In BMP message communication, in order to simplify the implementation, only the monitored router reports messages to the monitoring station, and the station does not send messages to the router [RFC7854]. In other words, the BMP communication is actually

unidirectional (from router to station). As a consequence, the direction from the monitoring station to the monitored router may be used as an interface for malicious attacks on the router. As BMP supports more and more types of routes to be reported, the number of reported BMP messages is also increasing, which also brings huge challenges to TCP data transmission pressure.

QUIC [RFC9000] is a UDP-based multiplexed and secure transport protocol that provides connection-oriented and stateful interaction between a client and server. It can provide low latency and encrypted transport with resilient connections.

QUIC uses multiple simultaneous streams to carry data in one direction. Each stream is a separate unidirectional or bidirectional channel consisting of an ordered stream of bytes. In Addition, each stream has its own flow control, which limit bytes sent on a stream, together with flow control of the connection. Among them, the unidirectional stream is very consistent with the message transmission mechanism of BMP.

Therefore, QUIC is a proper transport protocol for the message transmission mechanism of BMP. This document specifies how to use QUIC as the secure transport protocol for BMP.

2. Terminology and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

In this document, the terms "client" and "server" are used to refer to the two ends of the QUIC connection. The client actively initiates the QUIC connection. The terms "monitored router" and "monitoring station" are used to refer to the two ends of the BMP session. The router sends BMP messages to the station, but the station does not respond to the router.

* Client: The endpoint that initiates a QUIC connection, the BMP monitored router.

* Server: The endpoint that accepts a QUIC connection, the BMP monitoring station.

3. Connection Management

3.1. Connection Establishment

QUIC connection establishment is described in [RFC9000]. During establishing connection, BMPoQUIC support is indicated by selecting the Application-Layer Protocol Negotiation (ALPN) [RFC7301] token as listed in the IANA sectionSection 7 in the TLS handshake.

The monitored router MUST also act as the client meanwhile the monitoring station must also act as the server.

The monitored router should be the initiator of the QUIC connection to the monitoring station meanwhile the monitoring station acts as a connection acceptor.

3.2. Connection Termination

3.2.1. QUIC Connection Termination Process

The typical QUIC connection termination process is described in [RFC9000].

3.2.2. BMPoQUIC Considerations for Connection Termination

When a BMP session is implemented based on a QUIC connection, the idle timeout should be disabled or the QUIC max_idle_timeout should be set appropriately in order to keep the QUIC connection persistent even if the BMP session is idle.

When a BMP monitoring station receives a termination message, it will graceful close the BMP session. The station SHOULD close the associated QUIC connection.

When a BMP monitored router is detecting the interruption of the QUIC connection, it SHOULD send a termination message to the BMP monitoring station.

4. Stream mapping and usage

There are seven kinds of BMP main message sent from monitored router to monitoring station, namely route monitoring message, statistics report message, peer down notification message, peer up notification message, initiation message, termination message and route mirroring message [RFC7854]. The seven kinds of BMP messages need to be mapped into QUIC streams.

QUIC [RFC9000] is a UDP-based multiplexed and secure transport protocol that provides connection-oriented and stateful interaction between a client and server. It can provide low latency and encrypted transport with resilient connections.

QUIC uses Stream ID to identify the stream. The least significant bit (0x1) of the stream ID identifies the initiator of the stream. The second least significant bit (0x2) of the stream ID distinguishes between bidirectional streams (with the bit set to 0) and unidirectional streams.

No BMP message is ever sent from the monitoring station to the monitored router. The monitored router MAY take steps to prevent the monitoring station from sending data or it MAY silently discard any data sent by the monitoring station. So BMP messages from monitored router SHOULD be mapped into unidirectional stream whose stream type is 0x2 according to the above.

4.1. Multi-stream Selection

4.1.1. Peer Stream

When a router has many peers and a large number of routes, if the related BMP messages are reported through an independent stream, the communication pressure of this stream will be very large and the efficiency will be very low. In order to reduce the communication pressure and improve the communication efficiency, multiple streams can be allocated according to the number of peers of the router, and each stream is used to transmit the BMP message of the specified peers, as shown in Figure 1. The number of streams can be configured as needed.

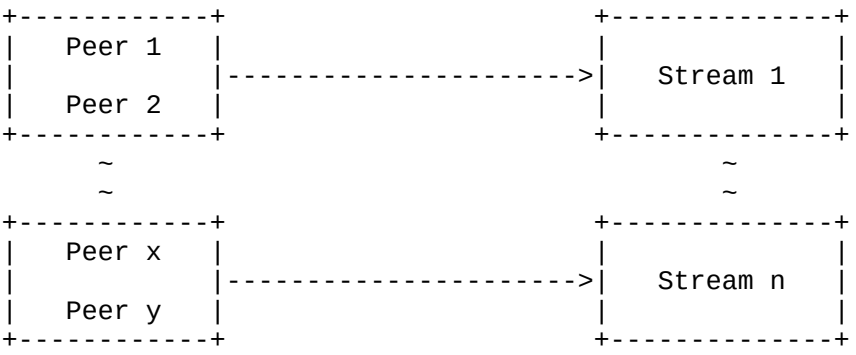

Figure 1: Peer Stream Structure

When multiple peer streams are used to transmit BMP messages, each stream needs to transmit peer-insensitive BMP messages (that is, BMP

messages that do not carry the per-peer header format) to ensure the order of BMP messages. Peer-insensitive BMP messages include Initiation Message and Termination Message [RFC7854].

4.1.2. Per-AFI/SAFI Stream without Control Stream

If multiple peer streams are used, each stream may carry different type (AFI/SAFI) routes from router, which may make the information seem a bit messy. So unidirectional QUIC streams can be created for Per-AFI/SAFI to carry the routes of the specific AFI/SAFI, as shown in the figure below.

Figure 2: Per AFI/SAFI Stream Structure

When per-AFI/SAFI streams are used to transmit BMP messages, it is necessary to distinguish which BMP messages can be categorized by per-AFI/SAFI and Determine how to send BMP messages for non-per-AFI/SAFI.

* For Initiation Message and Termination Message, they are all non-per-AFI/SAFI message, but they are messages necessary for session connection, so each Per-AFI/SAFI stream needs to send them.

* For Route Monitoring message, it is per-AFI/SAFI route message, and it can be carried over corresponding AFI/SAFI stream.

* For Route Mirroring message, if the message includes BGP PDU and the BGP PDU can distinguish AFI/SAFI information, it could be carried over corresponding AFI/SAFI stream. If not, it should be carried over any one of AFI/SAFI streams.

* For Statistics Report message, if the stat type is based on per-AFI/SAFI, it could be carried over corresponding AFI/SAFI stream. If not, it should be carried over any one of AFI/SAFI streams.

* For Peer up Notification message, it could include open message of one AFI/SAFI, and it should be carried over corresponding AFI/SAFI stream.

* For Peer down Notification message, as it does not include AFI/SAFI information, it should be carried over all corresponding AFI/SAFI streams which carry Peer up Notification message.

If BGP still uses TCP as the transport protocol, the Per AFI/SAFI Stream structure can be used selectively. If BGP uses QUIC as the transport protocol [I-D.draft-retana-idr-bgp-quic], the Per AFI/SAFI Stream structure MUST be used because of the implementation that per-AFI/SAFI streams (function channels) are used to carry routing information in one BGP over QUIC (BoQ) connection.

4.1.3. Per-AFI/SAFI Stream with Control Stream

Like the design of BGP over QUIC (BoQ) [I-D.draft-retana-idr-bgp-quic], the per-peer or per-AFI/SAFI streams (function channels) and the associated control mechanism (control channel) for the session are called "BMP channels". In one BMP over QUIC (BMPoQUIC) connection, one control channel and one or more function channels are used to carry BMP information, as shown in the figure below.

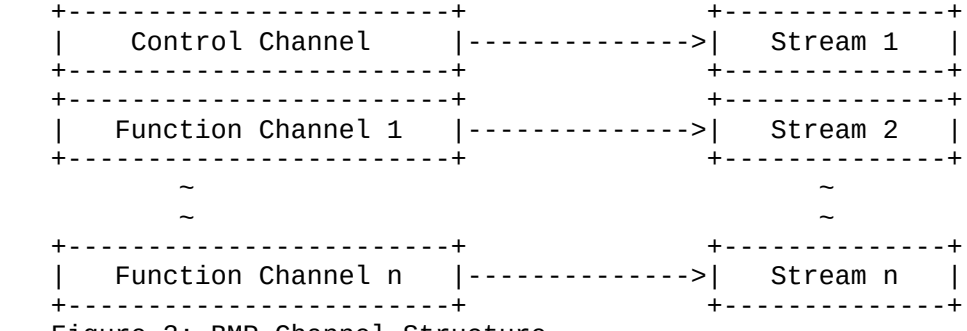

Figure 3: BMP Channel Structure

On a BMPoQUIC connection, the BMPoQUIC client first establishes a unidirectional stream for the "BMP control channel". The control channel is used to send the associated status control information, such as BMP session state control messages (Initiation Message and Termination Message) and BGP peer relationship state control messages (Peer up Notification message and Peer down Notification message).

As the Peer up Notification message is sent over the control channel, the BMPoQUIC client may create function channels for the peer by using unidirectional QUIC streams. These function channels are used to carry the relevant information for specific peer or AFI/SAFI.

These function channels include one Per-Peer Non-Per-AFI/SAFI function channel and several Per-Peer Per-AFI/SAFI function channels.

The Per-Peer Non-Per-AFI/SAFI function channel can be used to carry messages only related to peer and not to AFI/SAFI, such as Route Mirroring message without AFI/SAFI (if it cannot get AFI/SAFI) and Statistics Report message without AFI/SAFI (the Stat Type with not specifying Per-AFI/SAFI).

The Per-Peer Per-AFI/SAFI function channel can be used to transmit messages only related to AFI/SAFI for a specific peer, such as Route Monitoring message, Route Mirroring message with AFI/SAFI (if it can get AFI/SAFI from BGP Message) and Statistics Report message with AFI/SAFI (the Stat Type with specifying Per-AFI/SAFI).

As a Peer down Notification message is sent over the control channel, the router SHALL close all the related function channels. In the future, the Peer down Notification message may be expanded to include AFI/SAFI information, allowing the router to close the function channel of the specified AFI/SAFI for a peer.

According to [RFC7854], Initiation Message MUST be sent as the first message after the BMP session comes up. So it must be ensured that the Initiation message is the first message sent on each BMP channel, meaning that an Initiation message must be sent first for every BMP channel establishment.

5. Endpoint Authentication

BMPoQUIC uses QUIC which uses TLS version 1.3 or greater. Therefore, the TLS handshake process can be used for BMPoQUIC endpoint authentication. A third-party authentication mechanism can also be applied for BMPoQUIC endpoint authentication, such as a TLS client certificate.

6. Operational Considerations

The decision to use BMPoQUIC instead of the TCP-based mechanism in [RFC7854] is an operational decision, and an implementation MUST provide a configuration mechanism to enable BMPoQUIC on the BMP session.

Some connectivity problems (such as blocking UDP) could result in a failure to establish a QUIC connection. When this happens, monitored router SHOULD attempt to establish a TCP-based BMP session.

7. IANA Considerations

This document creates a new registration for the identification of BMPoQUIC in the "Application Layer Protocol Negotiation (ALPN) Protocol IDs registry established in [RFC7301].

The "BMPoQ" string identifies BMPoQUIC:

* Protocol: BMPoQUIC

* Identification Sequence: 0x42 0x4d 0x50 0x6f 0x51 ("BMPoQ")

* Specification: This document

8. Security Considerations

This document replaces the transport protocol layer of BMP from TCP to QUIC. The basic protocol specification of BMP is not modified, and therefore the new security risks are not introduced to the basic BMP protocol. BMPoQUIC enhances transport-layer security for BMP session according to [RFC9000].

This document does not require to support third-party authentication (e.g., backend Authentication) due to the fact that TLS does not specify this way of authentication. If third-party authentication is needed, TLS client certificates are recommended to be used here.

9. References

9.1. Normative References

[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/rfc/rfc2119>.

[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>.

[RFC9000] Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

9.2. Informative References

[RFC7301] Friedl, S., Popov, A., Langley, A., and E. Stephan, "Transport Layer Security (TLS) Application-Layer Protocol Negotiation Extension", RFC 7301, DOI 10.17487/RFC7301, July 2014, <https://www.rfc-editor.org/info/rfc7301>.

[I-D.draft-retana-idr-bgp-quic]

Retana, A., Qu, Y., Haas, J., Chen, S., and J. Tantsura, "BGP over QUIC", Work in Progress, Internet-Draft, draft-retana-idr-bgp-quic-05, 7 July 2024, <https://datatracker.ietf.org/doc/html/draft-retana-idr-bgp-quic-05>.

Authors' Addresses

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
Beijing
China

Email: linchangwang.04414@h3c.com

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

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