Global Routing Operations
Internet Draft
Intended status: Standards Track
Expires: July 07, 2026

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

January 6, 2026

Using BMP over QUIC connection
draft-liu-grow-bmp-over-quic-04

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 July 07, 2026.

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, typically a BMP monitored router.

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

3. Connection Management

3.1. Connection Establishment

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

Early Data (also known as 0-RTT) [RFC9001] provides weaker security than standard TLS, lacking forward secrecy and replay attack protection, while BMP information contains sensitive routing information. Therefore, BMPoQUIC implementations MUST NOT support Early Data. Servers MUST reject "early data" extensions in ClientHello messages, and clients MUST NOT attempt to send 0-RTT data.

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 (0x01) of the stream ID identifies the initiator of the stream (client-initiated with the bit set to 0). The second least significant bit (0x02) 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 unrecognized data sent by the monitoring station. So BMP messages from monitored router (as a client) SHOULD be mapped into unidirectional stream whose stream type is 0x2, or mapped into bidirectional stream whose stream type is 0x0, according to the above.

4.1. Multi-stream Selection

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 to carry BMP messages according to the number of peers or Address Family Identifier/Subsequent Address Family Identifier (AFI/SAFI).

This document introduces three multi-stream selections (section 4.2, section 4.3 and section 4.4), which can be freely chosen by network operator or user according to the level of complexity involved in configuring BGP. And the extent of complexity in BGP configuration is also decided by the network operator.

According to [RFC7854], Initiation Message MUST be sent as the first message after the BMP session comes up, meaning that the monitoring station MUST firstly receive the Initiation Message after the BMP session comes up. So when using peer stream or per-AFI/SAFI stream to send BMP messages, it must be ensured that the Initiation message is the first message sent on each BMP stream, meaning that an Initiation message must be sent first for every BMP stream establishment. When using function streams with control stream, it also must guarantee that message ordering by design implementation.

In the future, the information TLV of the Initiation message may be expanded to carry the multi-stream type (peer stream, per-AFI/SAFI stream, function streams with control stream, etc.), allowing the server (collector) to know, based on the multi-stream type, what kinds of BMP messages will be conveyed in the current stream. In addition, a termination message can be sent on any stream, which terminates the entire BMP session and closes the QUIC connection along with all of its opened streams.

4.2. Peer Stream

In order to reduce the communication pressure and improve the communication efficiency, multiple unidirectional streams can be allocated by router according to the number of BGP peers of the router, and each stream is used to transmit the BMP message of the specified peers, as shown in Figure 1.

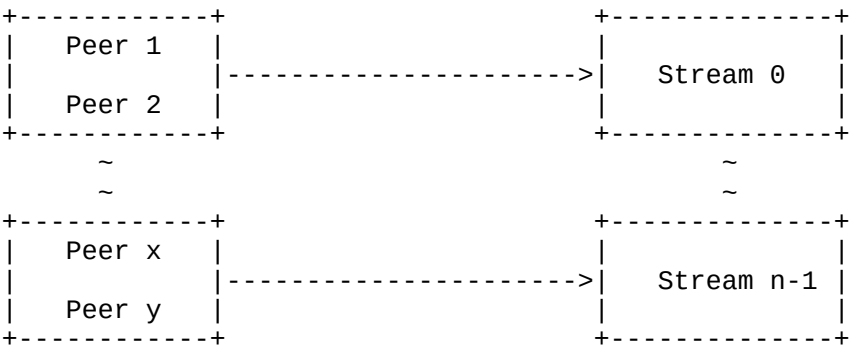


Figure 1: Peer Stream Structure

The number of streams can be configured as needed, and the corresponding relationship between peer and stream can be matched through configuration. For example, if there are five peers, when three streams are created, the stream 0 could carry the bmp messages of the peer 1 and peer 2, and the stream 1 could convey the bmp messages of the peer 3 and peer 4, and the stream 2 could send the

bmp messages of the peer 5. The default number of streams is 1, meaning that by default, stream 0 is created for all peers.

4.3. Per-AFI/SAFI Stream

When 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 a unidirectional QUIC stream can be created for Per-AFI/SAFI to carry the routes of the specific AFI/SAFI, as shown in the figure below. The unidirectional stream 0 is always created as a common stream to carry BMP messages which may or may not include AFI/SAFI information. If the common stream is used to carry the peer-related BMP messages, the Peer up Notification message MUST be sent prior to the other peer-related BMP messages.

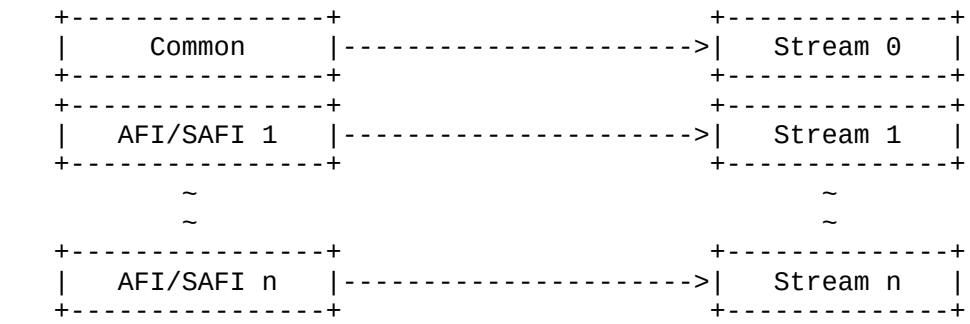


Figure 2: Per AFI/SAFI Stream Structure

In order to minimize resource occupation, it is recommended to create a stream for the AFI/SAFI when any BGP peer of the AFI/SAFI is in the up state. Of course, after creating the AFI/SAFI stream or common stream, an Initiation Message MUST be sent first before sending the Peer up Notification message.

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 Route Monitoring message, it is per-AFI/SAFI route message, and it SHOULD 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 includes the AFI/SAFI; otherwise, it does not include the AFI/SAFI. So it SHOULD be carried over the common stream (stream 0).

* For Statistics Report message, if the stat type is based on per-AFI/SAFI, it includes the AFI/SAFI. If not, it does not include the AFI/SAFI. So it SHOULD be carried over the common stream (stream 0).

* For Peer up Notification message, it could include open message containing multiple AFI/SAFIs, and it SHOULD be carried over all corresponding AFI/SAFI streams, and MUST be sent before sending other BMP messages which include peer information.

* For Peer down Notification message, as it does not include AFI/SAFI information, it SHOULD be carried over all corresponding AFI/SAFI streams or common stream 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], it is recommended that the Per AFI/SAFI Stream structure should be used in BMPoQUIC connection because of the implementation that per-AFI/SAFI streams (function channels) are also used to carry routing information in one BGP over QUIC (BoQ) connection.

4.4. Function Streams with Control Stream

In order to ensure the timing of BMP messages, the above two solutions (section 4.2 and section 4.3) must send an Initiation message or Peer up/down message on each stream, which may also increase the network load. Therefore, the solution with control stream is designed. 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 stream (control channel) for the BMP session are called "BMP channels". In one BMPoQUIC connection, one control channel and one or more function channels are used to carry BMP information, as shown in the figure below.

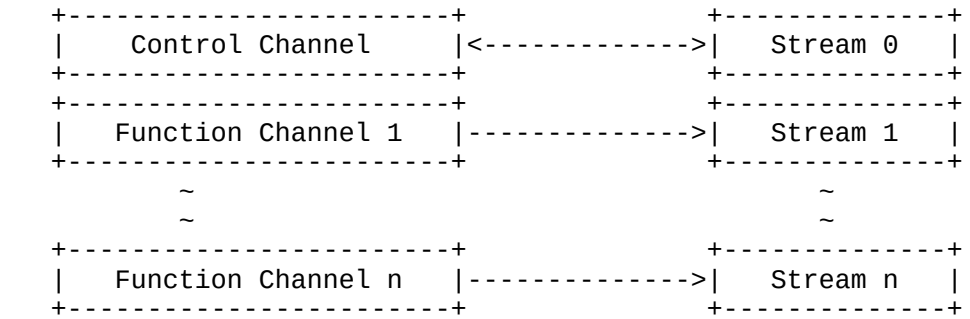


Figure 3: BMP Channel Structure

On a BMPoQUIC connection, the BMPoQUIC client first establishes a bidirectional 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). The control channel is also used to carry the associated control response information for keeping the order of BMP messages. In addition, the control channel is used to carry the periodic and low-load message, such as Statistics Report message.

When the BGP route association information with high load is needed to send, the function channel is created to deliver the key non-status BMP messages by monitored router, such as Route Monitoring message, and Route Mirroring message.

The Control channel always uses QUIC stream 0, which is a client-initiated bidirectional stream. The Function channels, which are initiated by a monitored router, are unidirectional streams.

4.4.1. Framing Layer

In QUIC layer, BMPoQUIC message are carried by QUIC STREAM frames. In BMPoQUIC layer, the two BMPoQUIC Frame types are defined, namely Data and Control Data, according to the BMP channels.

Data frames have following format:

```
BMPoQUIC Data Frame Format {
    Type (8) = 0,
    Length (24),
    Frame Payload (...)
}
```

Control Data frames have following format:

```
BMPoQUIC Control Data Frame Format {
    Type (8) = 1,
    Length (24),
    Sequence Number (32),
    Frame Payload (...)
}
```

Type: one octet, identifying the frame type.

Length: a 24-bit unsigned integer that describes the length in bytes of the frame payload.

Sequence Number: a 32-bit unsigned integer that indicating the sequence number of sending the control data by monitored router. After the sequence number increases to 2^32-1 by monitored router, the next sequence number returns to 0.

Frame Payload: BMP messages.

When the monitoring station needs send response message in control channel, the Length is set to 0 and the Sequence Number is same with corresponding BMP message from the monitored router.

4.4.2. Interactive Process

After the control channel is created, the Initiation Message is first sent over control channel. Only when the response message corresponding to the initialization message is received by the monitored router, the other BMP messages are allowed to be sent by the monitored router.

When the Peer up Notification message is sent over the control channel, and its response message from the monitoring station MUST be received, the monitored router can create function channels by using unidirectional QUIC streams. Or before the monitored router needs send other peer-related BMP messages over function channels, the corresponding function channel must be created. These function channels are used to carry the relevant non-status BMP messages with high load for specific peer or AFI/SAFI. So these function channels include one Per-Peer common function channel and several Per-Peer Per-AFI/SAFI function channels according to the AFI/SAFIs of the OPEN information contained in the Peer up Notification message [RFC7854].

The Per-Peer common function channel can be used to carry messages related to peer and need to be sent on the same channel, such as Route Mirroring message. For Route Mirroring message (regardless of whether it can get AFI/SAFI or not), it SHOULD be carried over the Per-Peer common function channel to prevent impacting the transmission efficiency of Route Monitoring messages in the Per-Peer Per-AFI/SAFI function channel, because Route Mirroring message may be extremely verbose.

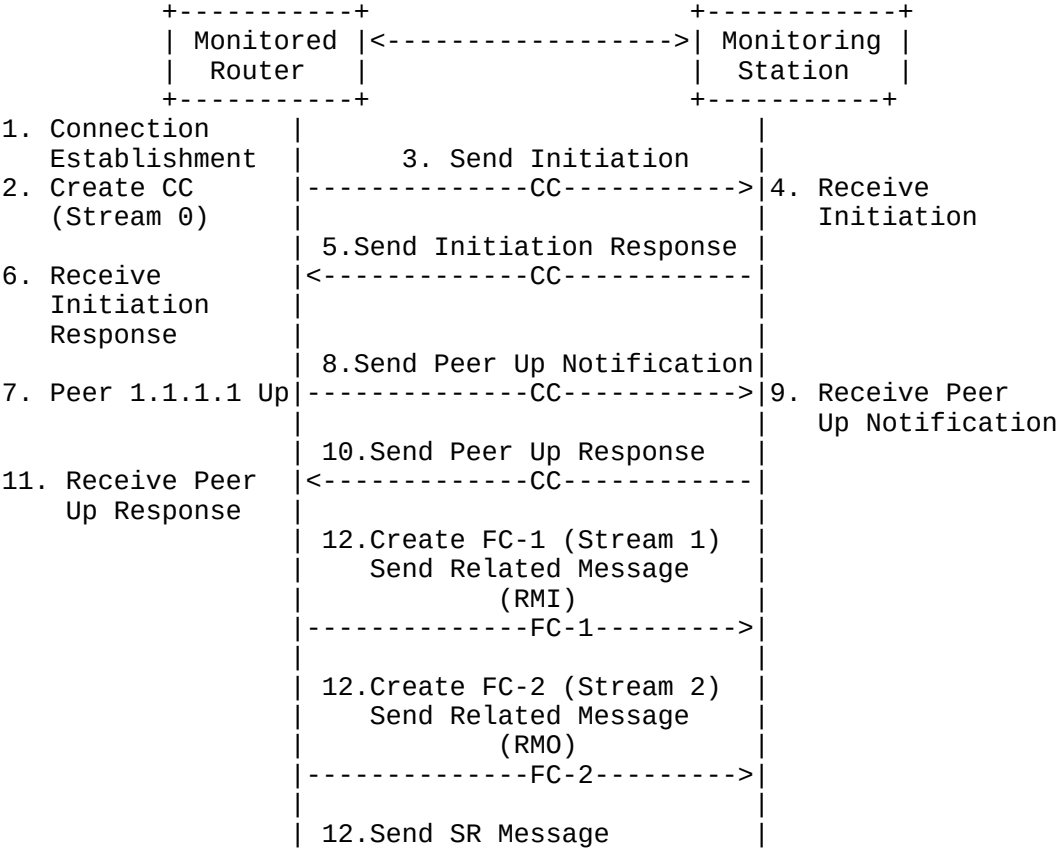
For Statistics Report message (regardless of whether it can get AFI/SAFI or not), since it may be a periodic message with low load, and different statistics may have some correlation. So it SHOULD always be carried over the control channel to maintain better atomicity of all statistics.

The Per-Peer Per-AFI/SAFI function channel SHOULD be used to transmit messages only related to AFI/SAFI for a specific peer, such as Route Monitoring message.

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

4.4.3. Use Case

In Monitored Router, the IPv4 unicast peer 1.1.1.1 is monitored. The Monitored Router and Monitoring Station establish a BMPoQUIC connection, than the relevant processing flow is shown in the figure 4.



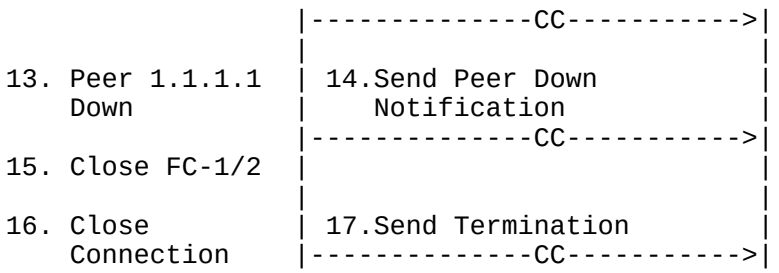


Figure 4: Typical Use Case

In figure 4, CC is the Control Channel, and FC is the Function Channel. RMI is the Route Mirroring Message, and RMO is Route Monitoring Message. SR is the Statistics Report Message.

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.

When using multiple stream, a configuration MAY be implemented to select to use which multi-stream selection.

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>.

[RFC9001] Thomson, M., Ed. and S. Turner, Ed., "Using TLS to Secure QUIC", RFC 9001, DOI 10.17487/RFC9001, May 2021, <https://www.rfc-editor.org/info/rfc9001>.

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