

BGP Over QUIC

[draft-chen-idr-bgp-over-quic-00](#)

Shuanglong Chen, Yongkang Zhang, Haibo Wang, **Zhenbin (Robin) Li**
Huawei Technologies

Motivation

- Currently, BGP has the following challenges:
 - ✓ **BGP session establishing delay is prolonged:** when establishing a BGP session, a three-way handshake is adopted to establish a TCP connection, and then TLS handshake authentication is also performed.
 - ✓ **Head-of-line Block:** Routing information of different address families is sent over the same TCP connection. If any packet is blocked, the sending of routes of all address families is affected.
 - ✓ **In encryption scenarios:** The BGP over TLS configuration is complex.

```
peer {group-name | ipv4-address} ssl-policy role {client | server}
peer {group-name | ipv4-address} ssl-policy name ssl-policy-name
peer {group-name | ipv4-address} ssl-server certificate
```
- With the development of open network operating systems, BGP is gradually integrated into the IT world. Using QUIC in this way is becoming a possible option. In addition, establishing connections from cloud devices to network devices has higher requirements on security and network adaptability.

What can QUIC bring to BGP?

- Similar to TCP, QUIC is a UDP-based, byte-stream-based reliable data transmission service. In addition, by integrating with TLS 1.3, QUIC also supports functions such as establishing connections with minimum latency and providing confidentiality and integrity protection for the transmitted data, and multi-stream multiplexing.

QUIC [RFC9000] [RFC9001] is a UDP-based transport protocol that provides the following functions:

1. **Reliable data transmission** service based on byte streams similar to TCP.
2. Support **low-latency connection** establishment.
3. Authentication of the server or client(Optional) is provided during connection establishment.
4. QUIC provides **confidentiality and integrity protection** for transport data and key fields in QUIC headers. QUIC also supports periodic key updates.
5. Supports **stream multiplexing**, including unidirectional and bidirectional streams.
6. Supports connection migration

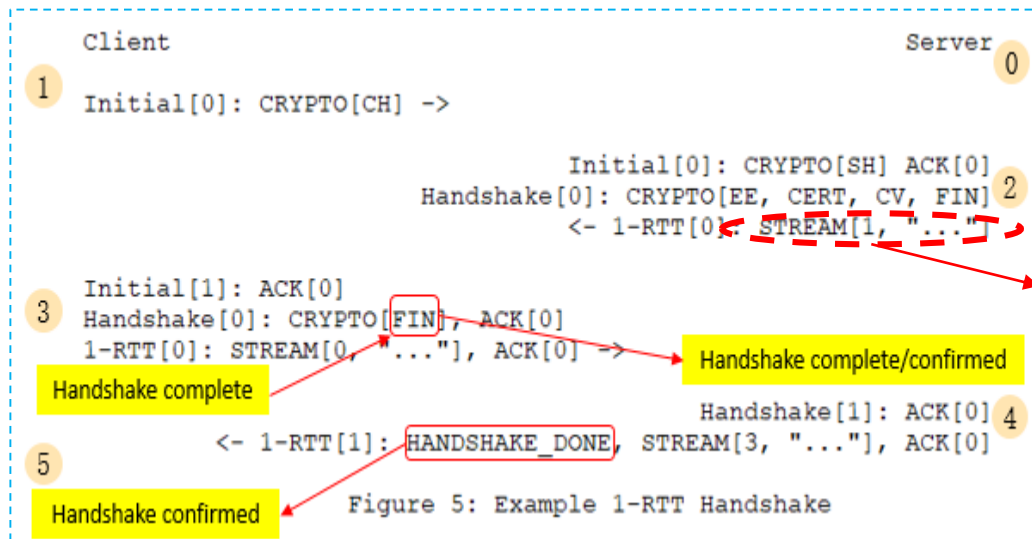
In addition, the configuration of BGP over TLS can be simplified. Peers can share the basic configuration of QUIC or be configured independently.

QUIC 1-RTT and 0-RTT

- When both communication parties initiate a communication connection, the first data packet may carry valid service data, which is referred to as 0-RTT, and vice versa, is referred to as 1-RTT.

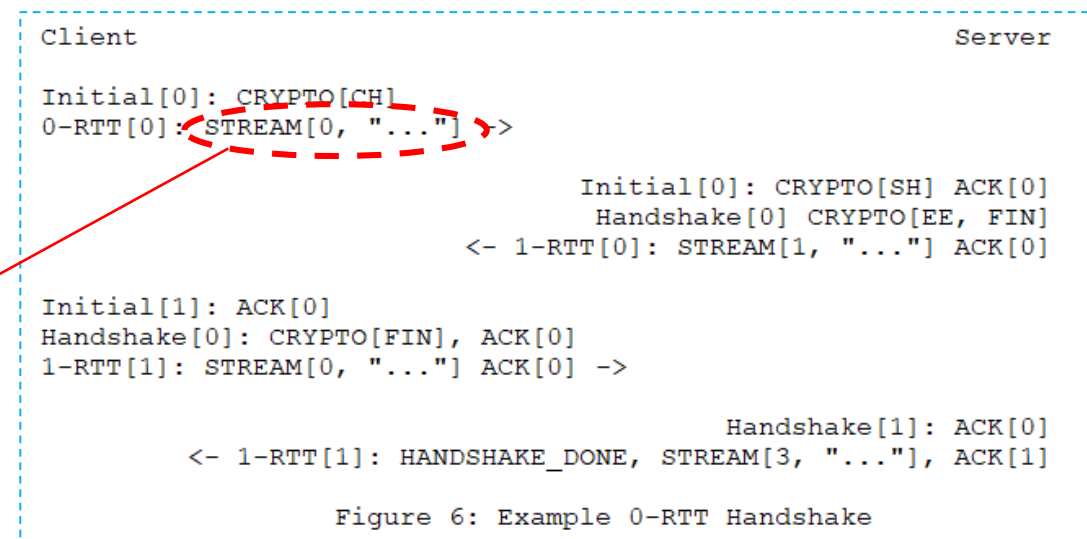
QUIC 1-RTT

- When the Quic client and Quic server initiate a communication connection for the first time, they need to exchange TLS parameters. Therefore, the first QUIC packet initiated by the client cannot carry APP data. As shown in the following figure, theoretically, the server can carry APP Data (Stream [1, "..."]) in response to the Initial packet.



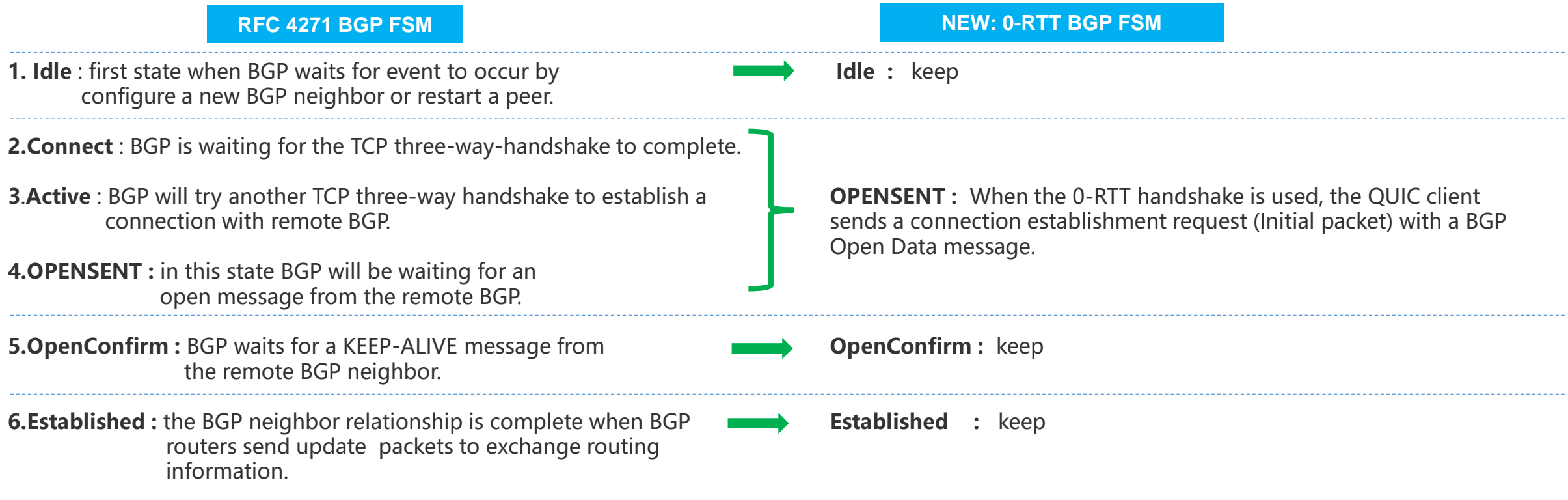
QUIC 0-RTT

- For a non-first connection, TLS parameters have been exchanged before. Therefore, when two communicating parties initiate a communication connection, the first data packet may carry valid service data.



Use Cases 1: BGP session establishment using 0-RTT FSM

- QUIC provides minimal connection setup delay. The BGP session setup delay is shortened from TLS 1.3(1 RTT) + TCP(3 RTT) to QUIC(1 RTT). If a BGP session is not established for the first time, the RTT can be set to 0 to shorten the BGP session setup delay. When the 0-RTT handshake is used, the QUIC client sends a connection establishment request (Initial packet) with a BGP Open Data message.
- 0-RTT FSM as follows:



Use Cases2: Stream mapping to solve Head-of-line Block

- QUIC also supports stream-level flow control (Stream multiplexing).
- BGP can take use of the stream multiplexing to solve the head-of-line issues. The following QUIC stream mapping modes can be selected:
 - Option 1, Mapping streams based on address families: One or more address family can be mapped to one stream.
 - Option 2, Mapping streams based on VRFs: One or more VRFs can be mapped to one stream.
 - Option 3, Mapping streams based on prefix: it can be combinations of prefixes.

Note that regardless of which mapping mode is selected, data of the same object MUST be received and transmitted using the same QUIC stream.

QUIC Events vs. TCP Events

- Both BGP-over-TCP and BGP-over-Quic have similar definitions and processing of transport-layer events.

QUIC Event	Event Name	TCP Event(RFC 4271)
Event 29	ManualStart_with_Passive Quic Establishment	Event 4: ManualStart_with_PassiveTcpEstablishment
Event 30	AutomaticStart_with_Passive Quic Establishment	Event 5: AutomaticStart_with_PassiveTcpEstablishment
Event 31	AutomaticStart_with_DampPeerOscillations_and_Passive Quic Establishment	Event 7: AutomaticStart_with_DampPeerOscillations_and_PassiveTcpEstablishment
Event 32	Quic Connection_Valid	Event 14: TcpConnection_Valid
Event 33	Quic _CR_Invalid	Event 15: Tcp_CR_Invalid
Event 34	Quic _CR_Acked	Event 16: Tcp_CR_Acked
Event 35	Quic ConnectionConfirmed	Event 17: TcpConnectionConfirmed
Event 36	Quic ConnectionFails	Event 18: TcpConnectionFails

Next-step

- Solicit comments and refine the draft
- Welcome new co-authors to co-work on the draft
- Possible open source work, implementation and verification