

QUIC
Internet-Draft
Intended status: Standards Track
Expires: November 11, 2019

M. Duke
F5 Networks, Inc.
May 10, 2019

QUIC-LB: Generating Routable QUIC Connection IDs
draft-duke-quic-load-balancers-04

Abstract

QUIC connection IDs allow continuation of connections across address/port 4-tuple changes, and can store routing information for stateless or low-state load balancers. They also can prevent linkability of connections across deliberate address migration through the use of protected communications between client and server. This creates issues for load-balancing intermediaries. This specification standardizes methods for encoding routing information and proposes an optional protocol called QUIC-LB to exchange the parameters of that encoding.

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 November 11, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
1.1.	Terminology	4
2.	Protocol Objectives	4
2.1.	Simplicity	4
2.2.	Security	5
2.3.	Robustness to Middleboxes	5
2.4.	Load Balancer Chains	5
3.	Routing Algorithms	6
3.1.	Plaintext CID Algorithm	6
3.1.1.	Load Balancer Actions	6
3.1.2.	Server Actions	6
3.2.	Obfuscated CID Algorithm	7
3.2.1.	Load Balancer Actions	7
3.2.2.	Server Actions	8
3.3.	Stream Cipher CID Algorithm	8
3.3.1.	Load Balancer Actions	9
3.3.2.	Server Actions	9
3.4.	Block Cipher CID Algorithm	10
3.4.1.	Load Balancer Actions	10
3.4.2.	Server Actions	11
4.	Protocol Description	11
4.1.	Out of band sharing	11
4.2.	QUIC-LB Message Exchange	12
4.3.	QUIC-LB Packet	12
4.4.	Message Types and Formats	13
4.4.1.	ACK_LB Message	13
4.4.2.	FAIL Message	13
4.4.3.	ROUTING_INFO Message	14
4.4.4.	STREAM_CID Message	14
4.4.5.	BLOCK_CID Message	15
4.4.6.	SERVER_ID Message	16
4.4.7.	MODULUS Message	16
4.4.8.	PLAINTEXT Message	16
5.	Config Rotation	17
5.1.	Configuration Failover	18
6.	Configuration Requirements	18
7.	Security Considerations	18
7.1.	Outside attackers	19
7.2.	Inside Attackers	19
8.	IANA Considerations	20
9.	References	20

Duke

Expires November 11, 2019

[Page 2]

9.1.	Normative References	20
9.2.	Informative References	20
Appendix A.	Acknowledgments	20
Appendix B.	Change Log	20
B.1.	Since draft-duke-quic-load-balancers-03	20
B.2.	Since draft-duke-quic-load-balancers-02	20
B.3.	Since draft-duke-quic-load-balancers-01	21
B.4.	Since draft-duke-quic-load-balancers-00	21
Author's Address	21

[1.](#) Introduction

QUIC packets usually contain a connection ID to allow endpoints to associate packets with different address/port 4-tuples to the same connection context. This feature makes connections robust in the event of NAT rebinding. QUIC endpoints designate the connection ID which peers use to address packets. Server-generated connection IDs create a potential need for out-of-band communication to support QUIC.

QUIC allows servers (or load balancers) to designate an initial connection ID to encode useful routing information for load balancers. It also encourages servers, in packets protected by cryptography, to provide additional connection IDs to the client. This allows clients that know they are going to change IP address or port to use a separate connection ID on the new path, thus reducing linkability as clients move through the world.

There is a tension between the requirements to provide routing information and mitigate linkability. Ultimately, because new connection IDs are in protected packets, they must be generated at the server if the load balancer does not have access to the connection keys. However, it is the load balancer that has the context necessary to generate a connection ID that encodes useful routing information. In the absence of any shared state between load balancer and server, the load balancer must maintain a relatively expensive table of server-generated connection IDs, and will not route packets correctly if they use a connection ID that was originally communicated in a protected NEW_CONNECTION_ID frame.

This specification provides a method of coordination between QUIC servers and low-state load balancers to support connection IDs that encode routing information. It describes desirable properties of a solution, and then specifies a protocol that provides those properties. This protocol supports multiple encoding schemes that increase in complexity as they address paths between load balancer and server with weaker trust dynamics.

Duke

Expires November 11, 2019

[Page 3]

1.1. Terminology

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, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying significance described in [RFC 2119](#).

In this document, "client" and "server" refer to the endpoints of a QUIC connection unless otherwise indicated. A "load balancer" is an intermediary for that connection that does not possess QUIC connection keys, but it may rewrite IP addresses or conduct other IP or UDP processing.

Note that stateful load balancers that act as proxies, by terminating a QUIC connection with the client and then retrieving data from the server using QUIC or another protocol, are treated as a server with respect to this specification.

When discussing security threats to QUIC-LB, we distinguish between "inside observers" and "outside observers." The former lie on the path between the load balancer and server, which often but not always lies inside the server's data center or cloud deployment. Outside observers are on the path between the load balancer and client. "Off-path" attackers, though not on any data path, may also be "inside" or "outside" depending on whether not they have network access to the server without intermediation by the load balancer and/or other security devices.

2. Protocol Objectives

2.1. Simplicity

QUIC is intended to provide unlinkability across connection migration, but servers are not required to provide additional connection IDs that effectively prevent linkability. If the coordination scheme is too difficult to implement, servers behind load balancers using connection IDs for routing will use trivially linkable connection IDs. Clients will therefore be forced choose between terminating the connection during migration or remaining linkable, subverting a design objective of QUIC.

The solution should be both simple to implement and require little additional infrastructure for cryptographic keys, etc.

2.2. Security

In the limit where there are very few connections to a pool of servers, no scheme can prevent the linking of two connection IDs with high probability. In the opposite limit, where all servers have many connections that start and end frequently, it will be difficult to associate two connection IDs even if they are known to map to the same server.

QUIC-LB is relevant in the region between these extremes: when the information that two connection IDs map to the same server is helpful to linking two connection IDs. Obviously, any scheme that transparently communicates this mapping to outside observers compromises QUIC's defenses against linkability.

However, concealing this mapping from inside observers is beyond the scope of QUIC-LB. By simply observing Link-Layer and/or Network-Layer addresses of packets containing distinct connection IDs, it is trivial to determine that they map to the same server, even if connection IDs are entirely random and do not encode routing information. Schemes that conceal these addresses (e.g., IPsec) can also conceal QUIC-LB messages.

Inside observers are generally able to mount Denial of Service (DoS) attacks on QUIC connections regardless of Connection ID schemes. However, QUIC-LB should protect against Denial of Service due to inside off-path attackers in cases where such attackers are possible.

Though not an explicit goal of the QUIC-LB design, concealing the server mapping also complicates attempts to focus attacks on a specific server in the pool.

2.3. Robustness to Middleboxes

The path between load balancer and server may pass through middleboxes that could drop the coordination messages in this protocol. It is therefore advantageous to make messages resemble QUIC traffic as much as possible, as any viable path must obviously admit QUIC traffic.

2.4. Load Balancer Chains

While it is possible to construct a scheme that supports multiple low-state load balancers in the path, by using different parts of the connection ID to encoding routing information for each load balancer, this use case is out of scope for QUIC-LB.

3. Routing Algorithms

In QUIC-LB, load balancers do not send individual connection IDs to servers. Instead, they communicate the parameters of an algorithm to generate routable connection IDs.

The algorithms differ in the complexity of configuration at both load balancer and server. Increasing complexity improves obfuscation of the server mapping.

The load balancer SHOULD route Initial and 0-RTT packets from the client using an alternate algorithm. Note that the SCID in these packets may not be long enough to represent all the routing bits. This algorithm SHOULD generate consistent results for Initial and 0RTT packets that arrive with the same source and destination connection ID. The load balancer algorithms below apply to all incoming Handshake and 1-RTT packets.

There are situations where a server pool might be operating two or more routing algorithms or parameter sets simultaneously. The load balancer uses the first two bits of the connection ID to multiplex incoming SCIDs over these schemes.

3.1. Plaintext CID Algorithm

3.1.1. Load Balancer Actions

The load balancer selects a number of bytes of the server connection ID (SCID) that it will use to route to a given server, called the "routing bytes". The number of bytes MUST have enough entropy to have a different code point for each server.

The load balancer shares this value with servers, as explained in [Section 4](#), along with the value that represents that server.

On each incoming packet, the load balancer extracts consecutive octets, beginning with the second byte. These bytes represent the server ID.

3.1.2. Server Actions

The server chooses a connection ID length. This MUST be at least one byte longer than the routing bytes.

When a server needs a new connection ID, it encodes its assigned server ID in consecutive octets beginning with the second. All other bits in the connection ID, except for the config rotation bits, MAY

be set to any other value. These other bits SHOULD appear random to observers.

The figure below clarifies the format. The first two bits are reserved for config rotation. The server can assign the next 6 bits to any value. The specified number of bytes encodes the server ID, and the server may decide how many trailing octets of information to include up to the QUIC limit of 18-octet CIDs.

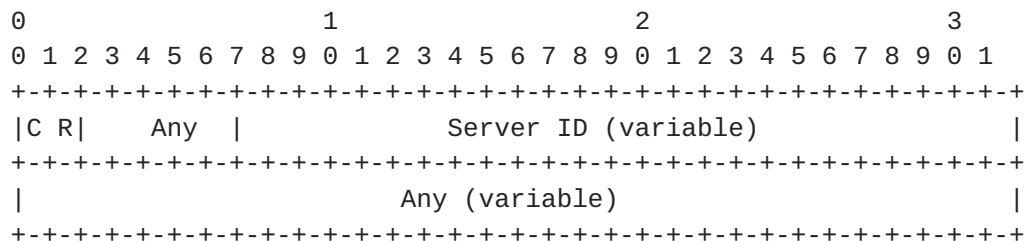


Figure 1: Plaintext CID Format

3.2. Obfuscated CID Algorithm

3.2.1. Load Balancer Actions

The load balancer selects an arbitrary set of bits of the server connection ID (SCID) that it will use to route to a given server, called the "routing bits". The number of bits MUST have enough entropy to have a different code point for each server, and SHOULD have enough entropy so that there are many codepoints for each server.

The load balancer MUST NOT select a routing mask that with more than 126 routing bits set to 1, which allows at least 2 bits for config rotation (see [Section 5](#)) and 16 for server purposes in a maximum-length connection ID.

The first two bits of an SCID MUST NOT be routing bits; these are reserved for config rotation.

The load balancer selects a divisor that MUST be larger than the number of servers. It SHOULD be large enough to accommodate reasonable increases in the number of servers. The divisor MUST be an odd integer so certain addition operations do not always produce an even number.

The load balancer also assigns each server a "modulus", an integer between 0 and the divisor minus 1. These MUST be unique for each server, and SHOULD be distributed across the entire number space between zero and the divisor.

The load balancer shares these three values with servers, as explained in [Section 4](#).

Upon receipt of a QUIC packet that is not of type Initial or 0-RTT, the load balancer extracts the selected bits of the SCID and expresses them as an unsigned integer of that length. The load balancer then divides the result by the chosen divisor. The modulus of this operation maps to the modulus for the destination server.

Note that any SCID that contains a server's modulus, plus an arbitrary integer multiple of the divisor, in the routing bits is routable to that server regardless of the contents of the non-routing bits. Outside observers that do not know the divisor or the routing bits will therefore have difficulty identifying that two SCIDs route to the same server.

Note also that not all Connection IDs are necessarily routable, as the computed modulus may not match one assigned to any server. Load balancers SHOULD drop these packets if not a QUIC Initial or 0-RTT packet.

[3.2.2](#). Server Actions

The server chooses a connection ID length. This MUST contain all of the routing bits and MUST be at least 8 octets to provide adequate entropy.

When a server needs a new connection ID, it adds an arbitrary nonnegative integer multiple of the divisor to its modulus, without exceeding the maximum integer value implied by the number of routing bits. The choice of multiple should appear random within these constraints.

The server encodes the result in the routing bits. It MAY put any other value into the non-routing bits except the config rotation bits. The non-routing bits SHOULD appear random to observers.

[3.3](#). Stream Cipher CID Algorithm

The Encrypted CID algorithm provides true cryptographic protection, rather than mere obfuscation, at the cost of additional per-packet processing at the load balancer to decrypt every incoming connection ID except for Initial and 0-RTT packets.

3.3.1. Load Balancer Actions

The load balancer assigns a server ID to every server in its pool, and determines a server ID length (in octets) sufficiently large to encode all server IDs, including potential future servers.

The load balancer also selects a nonce length and an 16-octet AES-ECB key to use for connection ID decryption. The nonce length **MUST** be at least eight octets and no more than 16 octets. The nonce length and server ID length **MUST** sum to 18 or fewer octets.

The load balancer shares these three values with servers, as explained in [Section 4](#).

Upon receipt of a QUIC packet that is not of type Initial or 0-RTT, the load balancer extracts as many of the earliest octets from the destination connection ID as necessary to match the nonce length. The server ID immediately follows.

The load balancer decrypts the server ID using 128-bit AES Electronic Codebook (ECB) mode, much like QUIC header protection. The nonce octets are padded to 16 octets using the as many of the first octets of the token as necessary. AES-ECB encrypts this nonce using its key to generate a mask which it applies to the encrypted server id.

```
server_id = encrypted_server_id ^ AES-ECB(key, padded-nonce)
```

For example, if the nonce length is 10 octets and the server ID length is 2 octets, the connection ID can be as small as 12 octets. The load balancer uses the first 10 octets (including the config rotation bits) of the connection ID for the nonce, pads it to 16 octets using the first 6 octets of the token, and uses this to decrypt the server ID in the eleventh and twelfth octet.

The output of the decryption is the server ID that the load balancer uses for routing.

3.3.2. Server Actions

When generating a routable connection ID, the server writes arbitrary bits into its nonce octets, and its provided server ID into the server ID octets. Servers **MAY** opt to have a longer connection ID beyond the nonce and server ID. The nonce and additional bits **MAY** encode additional information, but **SHOULD** appear essentially random to observers. The first two bits of the first octet are reserved for config rotation [Section 5](#), but form part of the nonce.

The server decrypts the server ID using 128-bit AES Electronic Codebook (ECB) mode, much like QUIC header protection. The nonce octets are padded to 16 octets using the as many of the first octets of the token as necessary. AES-ECB encrypts this nonce using its key to generate a mask which it applies to the server id.

`encrypted_server_id = server_id ^ AES-ECB(key, padded-nonce)`

3.4. Block Cipher CID Algorithm

The Block Cipher CID Algorithm, by using a full 16 octets of Plaintext and a 128-bit cipher, provides higher cryptographic protection and detection of spurious connection IDs. However, it also requires connection IDs of at least 17 octets, increasing overhead of client-to-server packets.

3.4.1. Load Balancer Actions

The load balancer assigns a server ID to every server in its pool, and determines a server ID length (in octets) sufficiently large to encode all server IDs, including potential future servers. The server ID will start in the second octet of the decrypted connection ID and occupy continuous octets beyond that.

The load balancer selects a zero-padding length. This SHOULD be at least four octets to allow detection of spurious connection IDs. The server ID and zero- padding length MUST sum to no more than 16 octets. They SHOULD sum to no more than 12 octets, to provide servers adequate space to encode their own opaque data.

The load balancer also selects an 16-octet AES-ECB key to use for connection ID decryption.

The load balancer shares these four values with servers, as explained in [Section 4](#).

Upon receipt of a QUIC packet that is not of type Initial or 0-RTT, the load balancer reads the first octet to obtain the config rotation bits. It then decrypts the subsequent 16 octets using AES-ECB decryption and the chosen key.

The decrypted plaintext contains the server id, zero padding, and opaque server data in that order. If the zero padding octets are not zero, the load balancer MUST drop the packet. The load balancer uses the server ID octets for routing.

3.4.2. Server Actions

When generating a routable connection ID, the server **MUST** choose a connection ID length of 17 or 18 octets. The server writes its provided server ID into the server ID octets, zeroes into the zero-padding octets, and arbitrary bits into the remaining bits. These arbitrary bits **MAY** encode additional information. Bits in the first and eighteenth octets **SHOULD** appear essentially random to observers. The first two bits of the first octet are reserved for config rotation [Section 5](#).

The server then encrypts the second through seventeenth octets using the 128-bit AES-ECB cipher.

4. Protocol Description

The fundamental protocol requirement is to share the choice of routing algorithm, and the relevant parameters for that algorithm, between load balancer and server.

For Obfuscated CID Routing, this consists of the Routing Bits, Divisor, and Modulus. The Modulus is unique to each server, but the others **MUST** be global.

For Stream Cipher CID Routing, this consists of the Server ID, Server ID Length, Key, and Nonce Length. The Server ID is unique to each server, but the others **MUST** be global. The authentication token **MUST** be distributed out of band for this algorithm to operate.

For Block Cipher CID Routing, this consists of the Server ID, Server ID Length, Key, and Zero-Padding Length. The Server ID is unique to each server, but the others **MUST** be global.

Each routing configuration also requires a unique two-bit config rotation codepoint (see [Section 5](#)) to identify it.

4.1. Out of band sharing

When there are concerns about the integrity of the path between load balancer and server, operators **MAY** share routing information using an out-of-band technique, which is out of the scope of this specification.

To simplify configuration, the global parameters can be shared out-of-band, while the load balancer sends the unique server IDs via the truncated message formats presented below.

4.2. QUIC-LB Message Exchange

QUIC-LB load balancers and servers exchange messages via the QUIC-LBv1 protocol, which uses the QUIC invariants with version number 0xF1000000. The QUIC-LB load balancers send the encoding parameters to servers and periodically retransmit until that server responds with an acknowledgement. Specifics of this retransmission are implementation-dependent.

4.3. QUIC-LB Packet

A QUIC-LB packet uses a long header. It carries configuration information from the load balancer and acknowledgements from the servers. They are sent when a load balancer boots up, detects a new server in the pool or needs to update the server configuration.



Figure 2: QUIC-LB Packet Format

The Version field allows QUIC-LB to use the Version Negotiation mechanism. All messages in this specification are specific to QUIC-LBv1. It should be set to 0xF1000000.

Load balancers MUST cease sending QUIC-LB packets of this version to a server when that server sends a Version Negotiation packet that does not advertise the version.

The length of the DCIL and SCIL fields are 0x00.

CR The 2-bit. CR field indicates the Config Rotation described in [Section 5](#).

Authentication Token The Authentication Token is an 8-byte field that both entities obtain at configuration time. It is used to verify that the sender is not an inside off-path attacker. Servers and load balancers SHOULD silently discard QUIC-LB packets with an incorrect token.

Message Type The Message Type indicates the type of message payload that follows the QUIC-LB header.

4.4. Message Types and Formats

As described in [Section 4.3](#), QUIC-LB packets contain a single message. This section describes the format and semantics of the QUIC-LB message types.

4.4.1. ACK_LB Message

A server uses the ACK_LB message (type=0x00) to acknowledge a QUIC-LB packet received from the load balancer. The ACK-LB message has no additional payload beyond the QUIC-LB packet header.

Load balancers SHOULD continue to retransmit a QUIC-LB packet until a valid ACK_LB message, FAIL message or Version Negotiation Packet is received from the server.

4.4.2. FAIL Message

A server uses the FAIL message (type=0x01) to indicate the configuration received from the load balancer is unsupported.

```

0                               1                               2                               3
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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Supp. Type | Supp. Type | ...
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

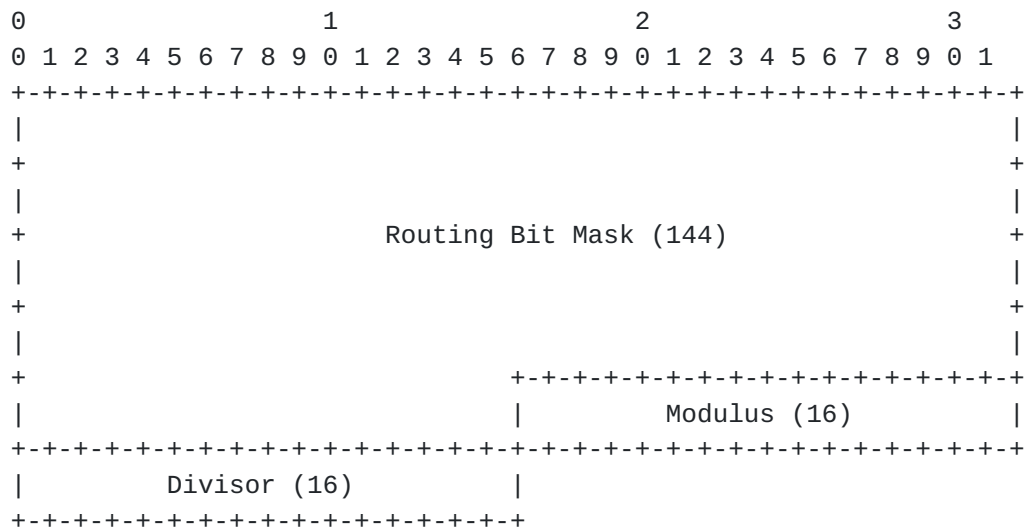
Servers MUST send a FAIL message upon receipt of a message type which they do not support, or if they do not possess all of the implied out-of-band configuration to support a particular message type.

The payload of the FAIL message consists of a list of all the message types supported by the server.

Upon receipt of a FAIL message, Load Balancers MUST either send a QUIC-LB message the server supports or remove the server from the server pool.

4.4.3. **ROUTING_INFO** Message

A load balancer uses the ROUTING_INFO message (type=0x02) to exchange all the parameters for the Obfuscated CID algorithm.



Routing Bit Mask The Routing Bit Mask encodes a '1' at every bit position in the server connection ID that will encode routing information.

These bits, along with the Modulus and Divisor, are chosen by the load balancer as described in [Section 3.2](#).

4.4.4. STREAM_CID Message

A load balancer uses the STREAM_CID message (type=0x03) to exchange all the parameters for using Stream Cipher CIDs.

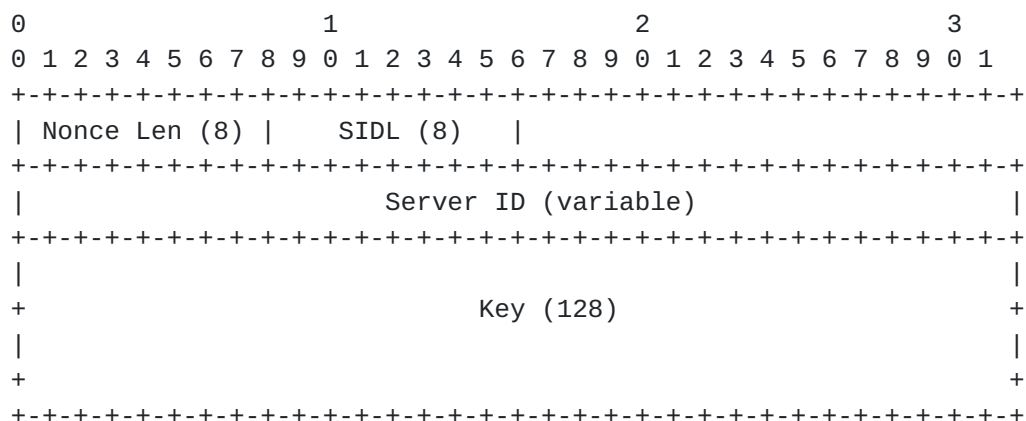


Figure 3: Stream CID Payload

Nonce Len The Nonce Len field is a one-octet unsigned integer that describes the nonce length necessary to use this routing algorithm, in octets.

SIDL The SIDL field is a one-octet unsigned integer that describes the server ID length necessary to use this routing algorithm, in octets.

Server ID The Server ID is the unique value assigned to the receiving server. Its length is determined by the SIDL field.

Key The Key is an 16-octet field that contains the key that the load balancer will use to decrypt server IDs on QUIC packets. See [Section 7](#) to understand why sending keys in plaintext may be a safe strategy.

4.4.5. BLOCK_CID Message

A load balancer uses the BLOCK_CID message (type=0x04) to exchange all the parameters for using Stream Cipher CIDs.

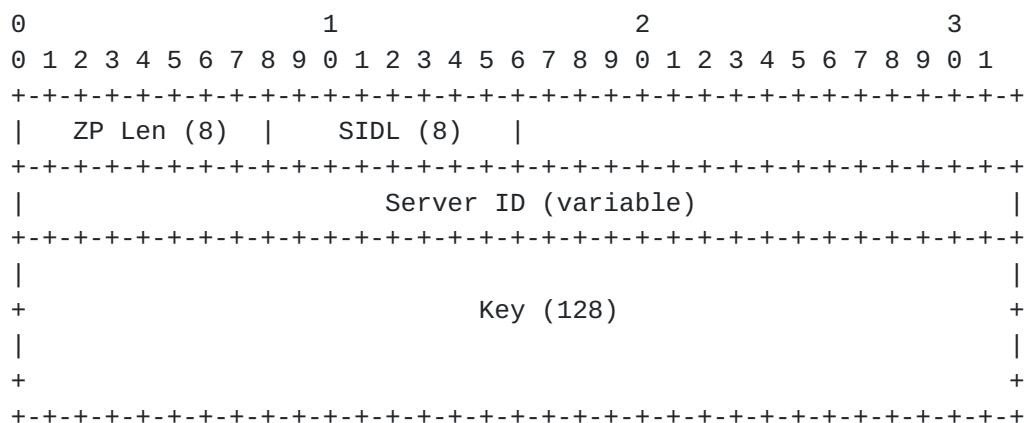


Figure 4: Block CID Payload

ZP Len The ZP Len field is a one-octet unsigned integer that describes the zero-padding length necessary to use this routing algorithm, in octets.

SIDL The SIDL field is a one-octet unsigned integer that describes the server ID length necessary to use this routing algorithm, in octets.

Server ID The Server ID is the unique value assigned to the receiving server. Its length is determined by the SIDL field.

Key The Key is an 16-octet field that contains the key that the load balancer will use to decrypt server IDs on QUIC packets. See [Section 7](#) to understand why sending keys in plaintext may be a safe strategy.

4.4.6. SERVER_ID Message

A load balancer uses the SERVER_ID message (type=0x05) to exchange explicit server IDs.

```

0                               1                               2                               3
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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   SIDL (8)   |   Server ID (variable)   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Load balancers send the SERVER_ID message when all global values for Stream or Block CIDs are sent out-of-band, so that only the server-unique values must be sent in-band. The fields are identical to their counterparts in the [Section 4.4.4](#) payload.

4.4.7. MODULUS Message

A load balancer uses the MODULUS message (type=0x06) to exchange just the modulus used in the Obfuscated CID algorithm.

```

0                               1                               2                               3
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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Modulus (16)           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
+                                     +
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Load balancers send the MODULUS when all global values for Obfuscated CIDs are sent out-of-band, so that only the server-unique values must be sent in-band. The Modulus field is identical to its counterpart in the ROUTING_INFO message.

4.4.8. PLAINTEXT Message

A load balancer uses the PLAINTEXT message (type=0x07) to exchange all parameters needed for the Plaintext CID algorithm.


```

0                               1                               2                               3
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
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|  SIDL (8)  |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|
+                               Server ID (variable)                               +
|
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

The SIDL field indicates the length of the server ID field. The Server ID field indicates the encoding that represents the destination server.

5. Config Rotation

The first two bits of any connection-ID MUST encode the configuration phase of that ID. QUIC-LB messages indicate the phase of the algorithm and parameters that they encode.

A new configuration may change one or more parameters of the old configuration, or change the algorithm used.

It is possible for servers to have mutually exclusive sets of supported algorithms, or for a transition from one algorithm to another to result in Fail Payloads. The four states encoded in these two bits allow two mutually exclusive server pools to coexist, and for each of them to transition to a new set of parameters.

When new configuration is distributed to servers, there will be a transition period when connection IDs reflecting old and new configuration coexist in the network. The rotation bits allow load balancers to apply the correct routing algorithm and parameters to incoming packets.

Servers MUST NOT generate new connection IDs using an old configuration when it has sent an Ack payload for a new configuration.

Load balancers SHOULD NOT use a codepoint to represent a new configuration until it takes precautions to make sure that all connections using IDs with an old configuration at that codepoint have closed or transitioned. They MAY drop connection IDs with the old configuration after a reasonable interval to accelerate this process.

5.1. Configuration Failover

If a server is configured to expect QUIC-LB messages, and it has not received these, it **MUST** generate connection IDs with the config rotation bits set to '0b11' and **MUST** use the "disable_migration" transport parameter in all new QUIC connections. It **MUST NOT** send NEW_CONNECTION_ID frames with new values.

A load balancer that sees a connection ID with config rotation bits set to '0b11' **MUST** revert to 5-tuple routing.

6. Configuration Requirements

QUIC-LB strives to minimize the configuration load to enable, as much as possible, a "plug-and-play" model. However, there are some configuration requirements based on algorithm and protocol choices above.

There are three levels of configuration that correspond to increasing levels of concern about the security of the load balancer-server path.

The complete information requirements are described in [Section 4](#). Load balancers **MUST** have configuration for all parameters of each routing algorithm they support.

If there is any in-band communication, servers **MUST** be explicitly configured with the token of the load balancer they expect to interface with. Endpoints that use Stream Cipher CIDs **MUST** have this token regardless of the configuration method.

Optionally, servers **MAY** be configured with the global parameters of supported routing algorithms. This allows load balancers to use Server ID and Modulus Payloads, limiting the information sent in-band.

Finally, servers **MAY** be directly configured with their unique server IDs or modulus, eliminating need for in-band messaging at all. In this case, servers and load balancers **MUST** enable only one routing algorithm, as there is no explicit message to agree on one or the other.

7. Security Considerations

QUIC-LB is intended to preserve routability and prevent linkability. Attacks on the protocol would compromise at least one of these objectives.

Note that the Plaintext CID algorithm makes no attempt to obscure the server mapping, and therefore does not address these concerns. It exists to allow consistent CID encoding for compatibility across a network infrastructure. Servers that are running the Plaintext CID algorithm SHOULD only use it to generate new CIDs for the Server Initial Packet, and SHOULD NOT send CIDs in QUIC NEW_CONNECTION_ID frames. Doing so might falsely suggest to the client that said CIDs were generated in a secure fashion.

A routability attack would inject QUIC-LB messages so that load balancers incorrectly route QUIC connections.

A linkability attack would find some means of determining that two connection IDs route to the same server. As described above, there is no scheme that strictly prevents linkability for all traffic patterns, and therefore efforts to frustrate any analysis of server ID encoding have diminishing returns.

7.1. Outside attackers

For an outside attacker to break routability, it must inject packets that correctly guess the 64-bit token, and servers must be reachable from these outside hosts. Load balancers SHOULD drop QUIC-LB packets that arrive on its external interface.

Off-path outside attackers cannot observe connection IDs, and will therefore struggle to link them.

On-path outside attackers might try to link connection IDs to the same QUIC connection. The Encrypted CID algorithm provides robust entropy to making any sort of linkage. The Obfuscated CID obscures the mapping and prevents trivial brute-force attacks to determine the routing parameters, but does not provide robust protection against sophisticated attacks.

7.2. Inside Attackers

As described above, on-path inside attackers are intrinsically able to map two connection IDs to the same server. The QUIC-LB algorithms do prevent the linkage of two connection IDs to the same individual connection if servers make reasonable selections when generating new IDs for that connection.

On-path inside attackers can break routability for new and migrating connections by copying the token from QUIC-LB messages. From this privileged position, however, there are many other attacks that can break QUIC connections to the server during the handshake.

Off-path inside attackers cannot observe connection IDs to link them. To successfully break routability, they must correctly guess the token.

8. IANA Considerations

There are no IANA requirements.

9. References

9.1. Normative References

[QUIC-TRANSPORT]

Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", [draft-ietf-quic-transport](#) (work in progress).

9.2. Informative 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/info/rfc2119>>.

Appendix A. Acknowledgments

Appendix B. Change Log

RFC Editor's Note: Please remove this section prior to publication of a final version of this document.

B.1. Since [draft-duke-quic-load-balancers-03](#)

- o Renamed Plaintext CID algorithm as Obfuscated CID
- o Added new Plaintext CID algorithm

B.2. Since [draft-duke-quic-load-balancers-02](#)

- o Added Config Rotation
- o Added failover mode
- o Tweaks to existing CID algorithms
- o Added Block Cipher CID algorithm
- o Reformatted QUIC-LB packets

B.3. Since [draft-duke-quic-load-balancers-01](#)

- o Complete rewrite
- o Supports multiple security levels
- o Lightweight messages

B.4. Since [draft-duke-quic-load-balancers-00](#)

- o Converted to markdown
- o Added variable length connection IDs

Author's Address

Martin Duke
F5 Networks, Inc.

Email: martin.h.duke@gmail.com

