The sad reality of QUIC handshakes: How bad configuration leads to delay or DDoS.

Marcin Nawrocki, Nicholas Banks, Raphael Hiesgen,
Thomas C. Schmidt, Pouyan Fotouhi Tehrani, Matthias Wählisch

{marcin.nawrocki, m.waehlisch}@fu-berlin.de
nibanks@microsoft.com
pouyan.fotouhi.tehrani@fokus.fraunhofer.de
{raphael.hiesgen, t.schmidt}@haw-hamburg.de

Recap: RIPE 84 on Monday

Design goals of QUIC handshakes.

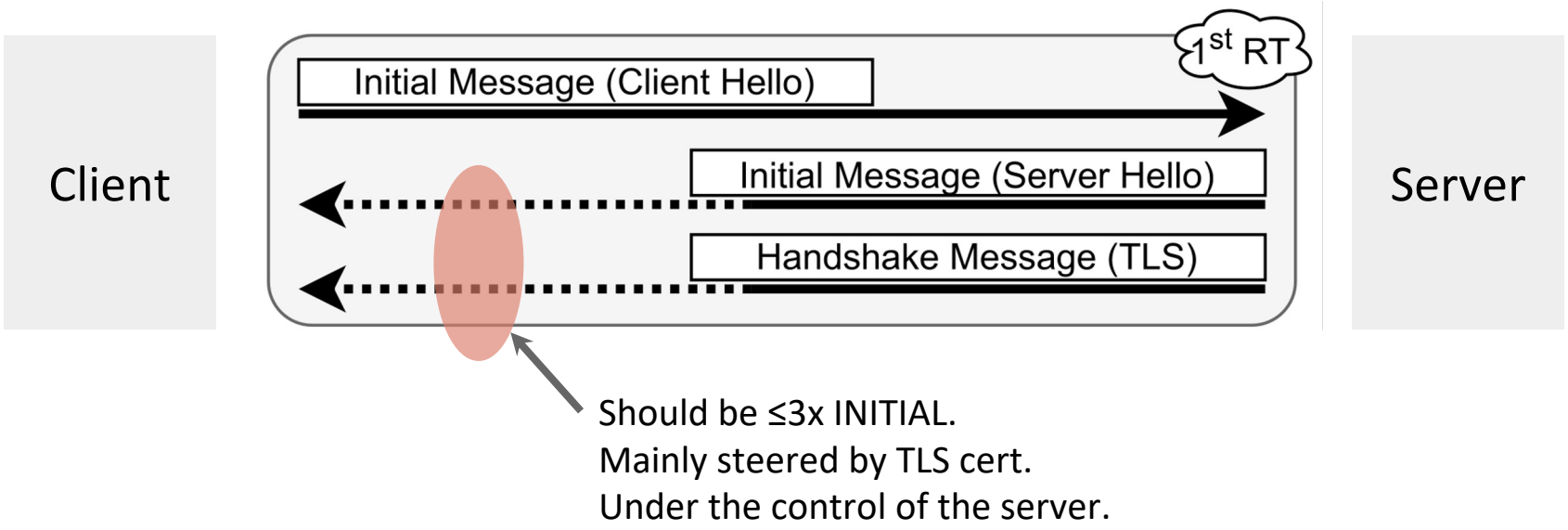
Reduce round trip times.

TCP/TLS/HTTP handshakes coalesced into 1RTT.

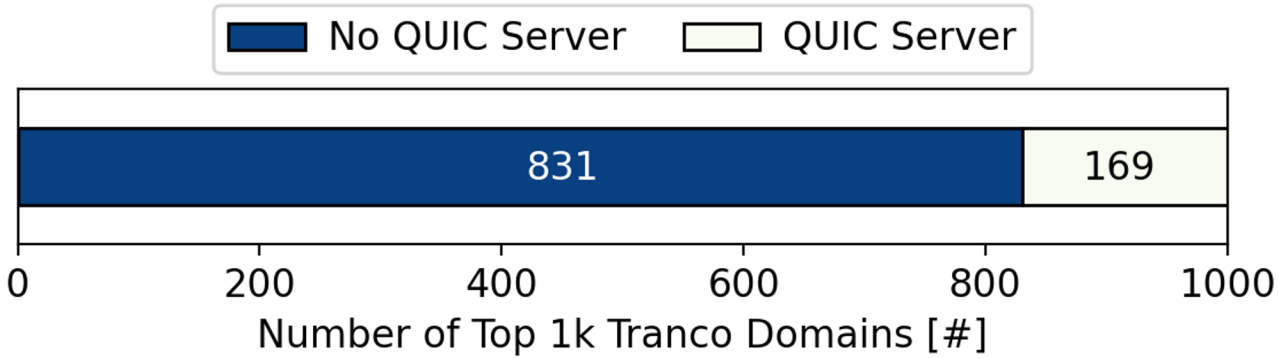
Prevent UDP amplification attacks.

RFC limits response size to 3x of an (unauthenticated) request.

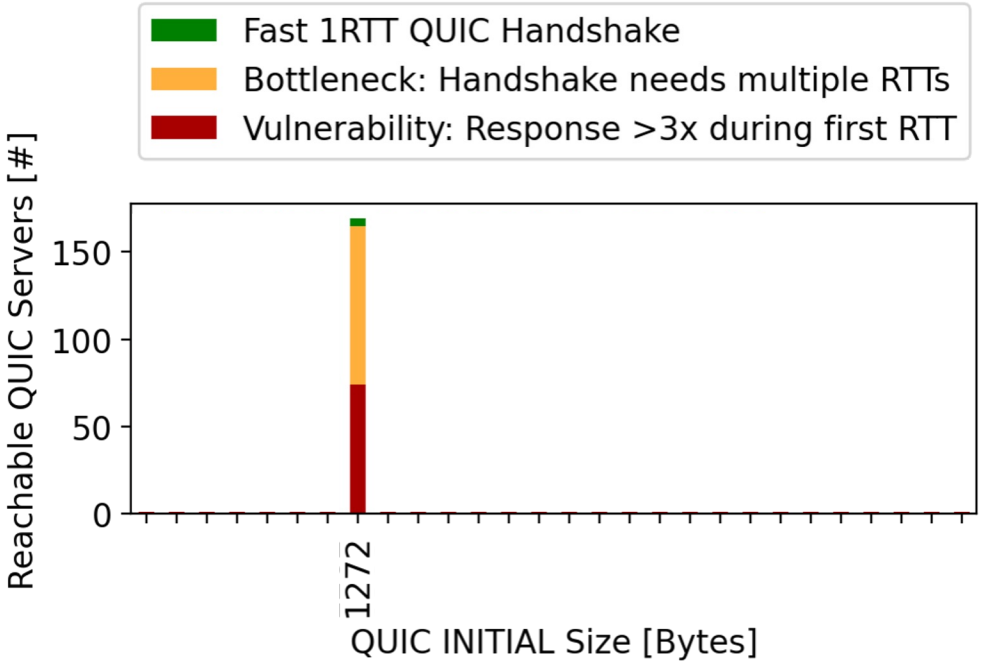
In practice.


Client
1st RT
Initial Message (Client Hello)
Initial Message (Server Hello)
Handshake Message (TLS)
Server
Should be ≤3x INITIAL.
Mainly steered by TLS cert.
Under the control of the server.

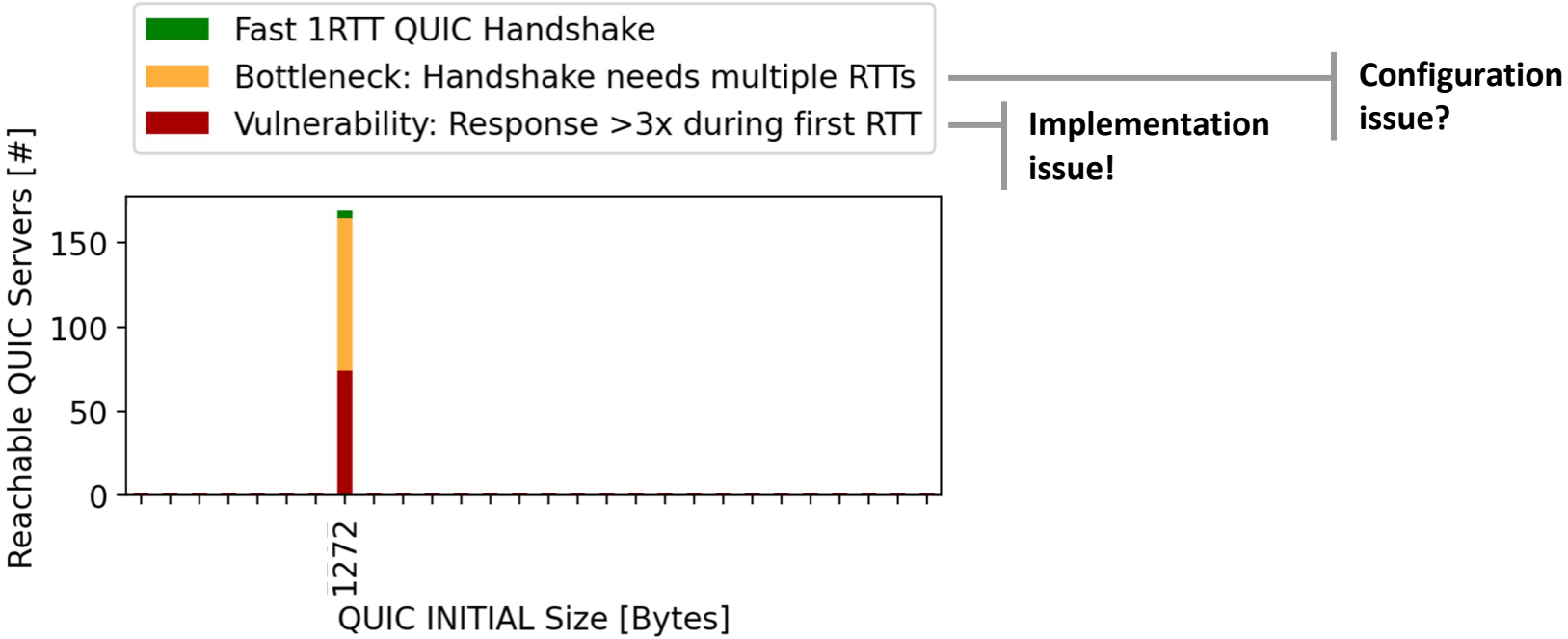
We actively scanned 1k top domains.



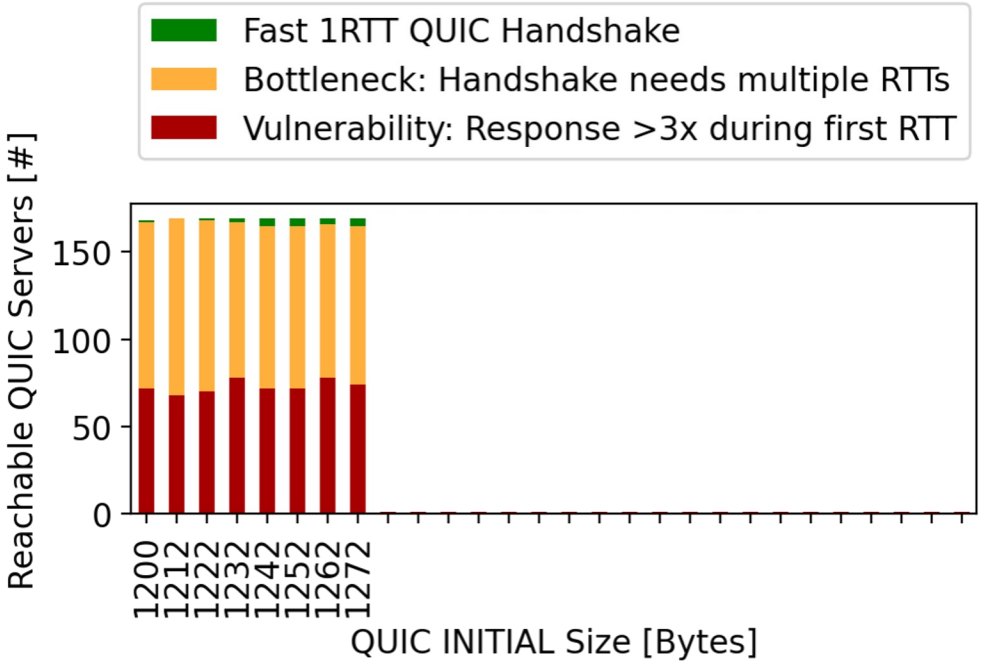
Have the design goals been met? No!



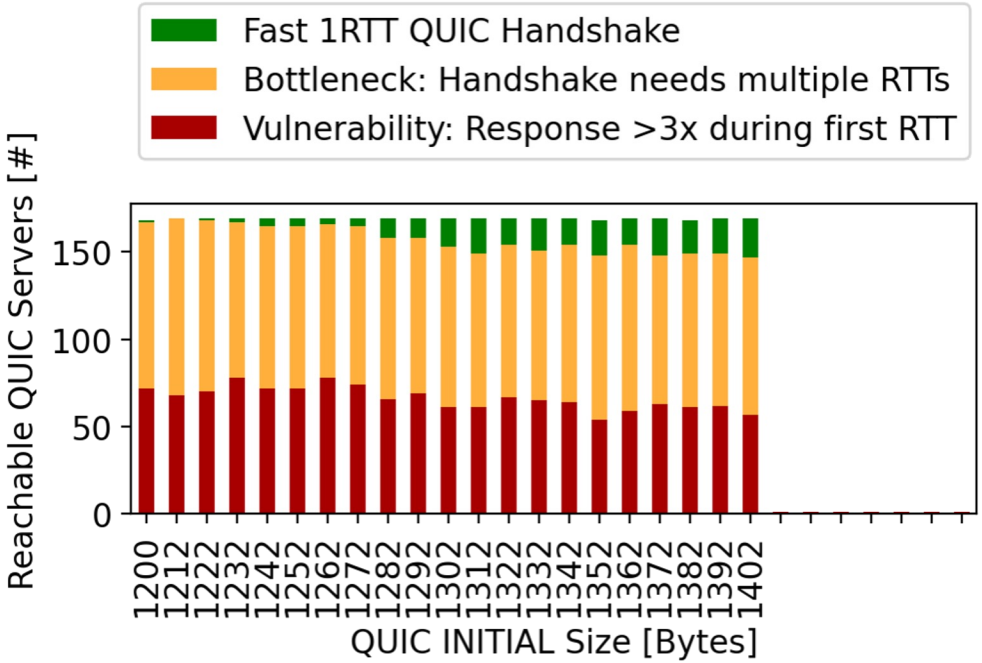
Have the design goals been met? No!



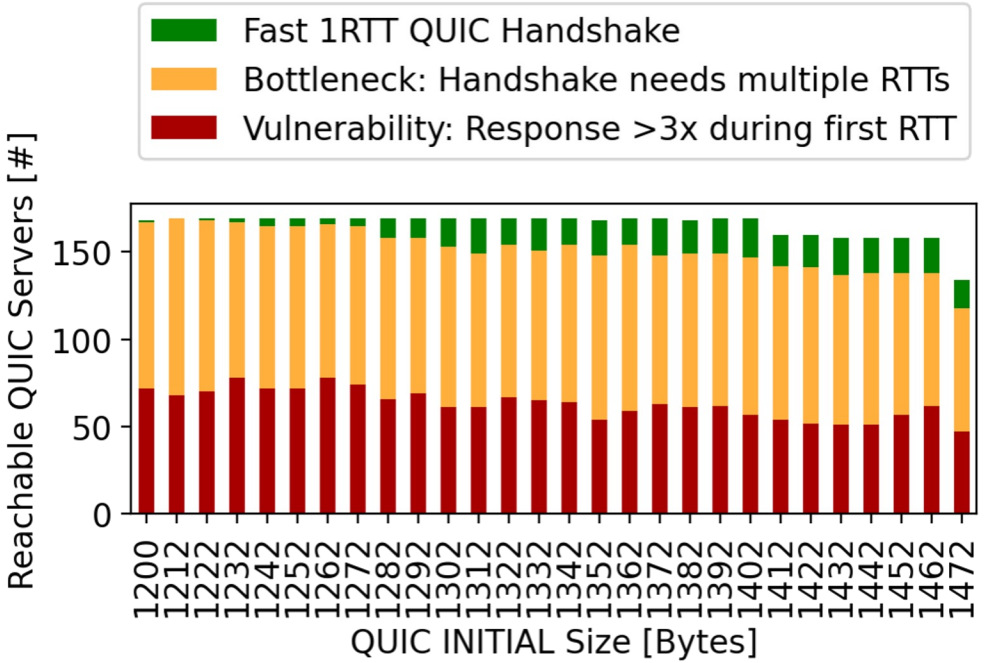
Small QUIC INITIALs lead to multiple RTTs.



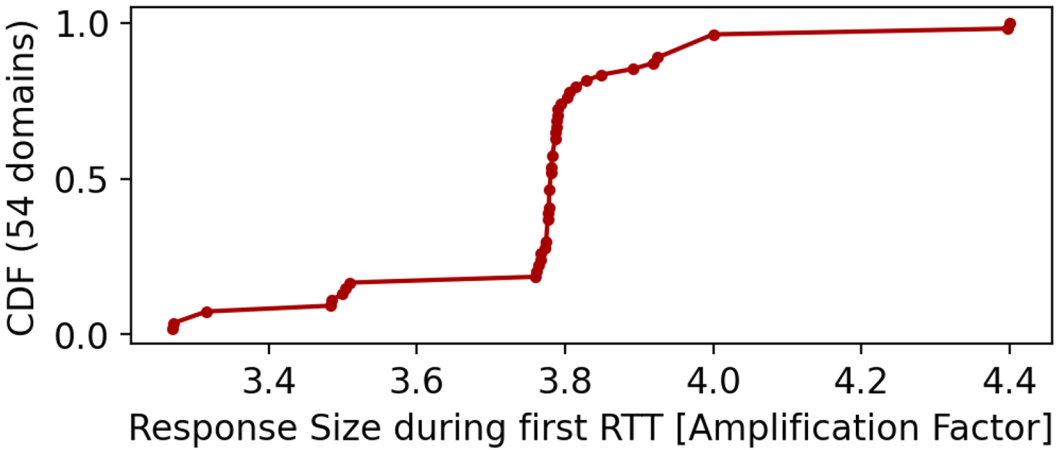
~1350 bytes are the sweet spot.



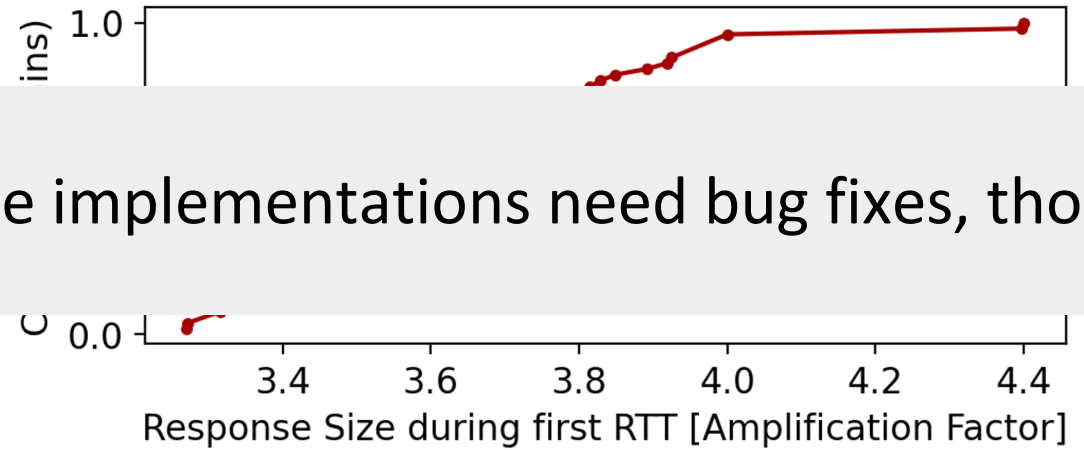
Larger QUIC INITIALs reduce reachability.



How bad is the amplification? Not bad.



How bad is the amplification? Not bad.



Some implementations need bug fixes, though!

What causes **multiple RTTs**?

RETRY tokens
(DDoS prevention)

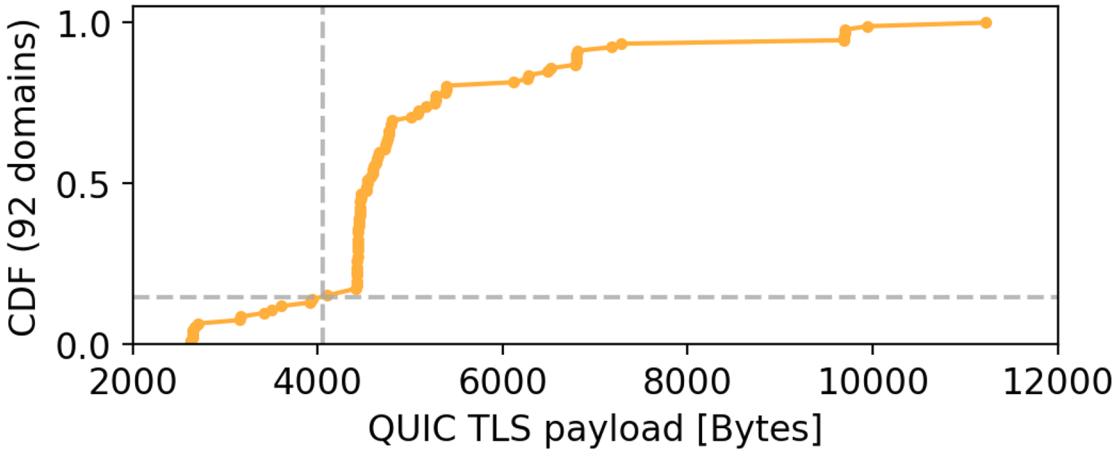
Only 2 domains.

Large TLS certificates
(in conflict with the 3x limit)

The majority!

85% of TLS data exceed the 3x threshold!

This causes multiple round trip times.



What is the best-performing INITIAL size to not induce an additional RTT given current TLS certs?

Currently, it is ~1350 bytes.

Trade-off:

Small INITIALs trigger multiple RTTs due to large TLS certificates.

Large INITIALs reduce reachability because QUIC forbids IP fragmentation!

Recommendations to operators

if you care about delay and DDoS prevention

1. Reduce the size of your TLS data!
   Be below 3x of common requests to fit into the 1RTT handshake.

2. Activate `RETRY` tokens!
   QUIC INITIAL DDoS floods are a rising threat vector! [ACM IMC'21]
   `RETRY` enables DDoS prevention for ZERO additional cost compared to current deployments (with multiple RTTs).

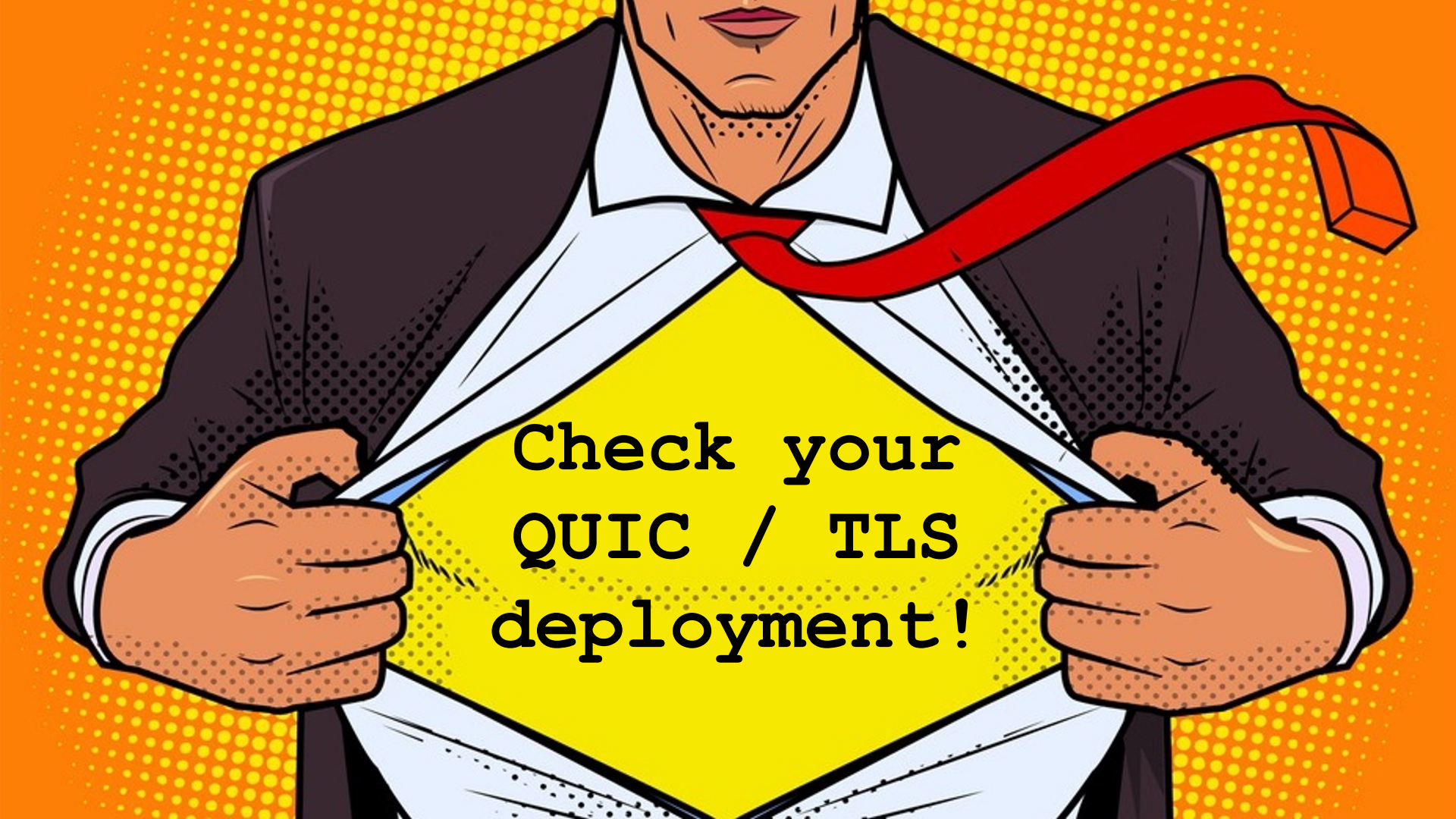
Test your QUIC server today! An open source tool.

```
> quicreach google.com
Success
```

```
> quicreach example.com
Failure
```

```
> quicreach '*' --stats
                        SERVER          RTT        TIME_I        TIME_H              SEND:RECV
              quic.aiortc.org    102.082 ms    106.934 ms    240.379 ms    4:5 2523:4900 (1.9x)
             ietf.akaquic.com     98.277 ms    100.906 ms    201.243 ms    3:5 2480:5869 (2.4x)
                quic.ogre.com
```

https://github.com/microsoft/quicreach


Check your
QUIC / TLS
deployment!

Backup

Are you saying we should increase the 3x limit?

No.

Let’s first find out what a space-efficient but secure TLS config looks like.

The 3x limit could encourage optimal configs.

How can you decrease your TLS data size?

1. Use algorithms with smaller signature footprint (elliptic instead of RSA).
2. We observe certificates with a long list of Subject Alt Names (SANs), reduce them!
3. This is a joint effort due to certificate chains.

Load balancers also add bloat!

We are scanning top domains with a very large user base. → Load balancers.

Packet encapsulation is used between L4 and L7 load balancers (+20 bytes).

RFC 9000 and the 3x threshold for unvalidated clients

8.1. Address Validation during Connection Establishment

Prior to validating the client address, servers MUST NOT send more than three times as many bytes as the number of bytes they have received.

17.2.5. Retry Packet

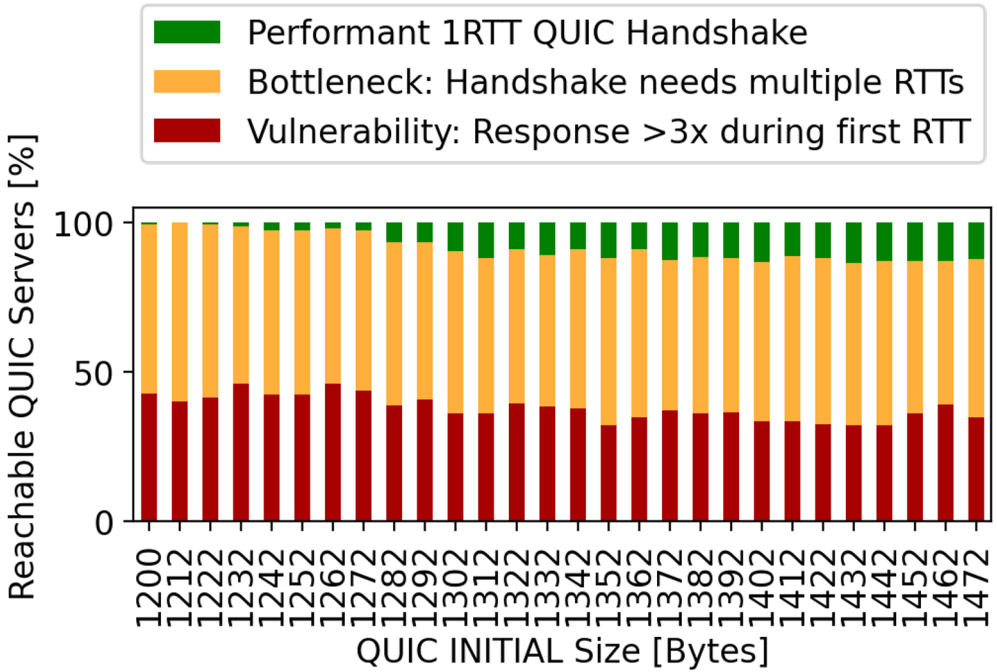
An opaque token that the server can use to validate the client's address.

RFC 9000 and IP fragmentation

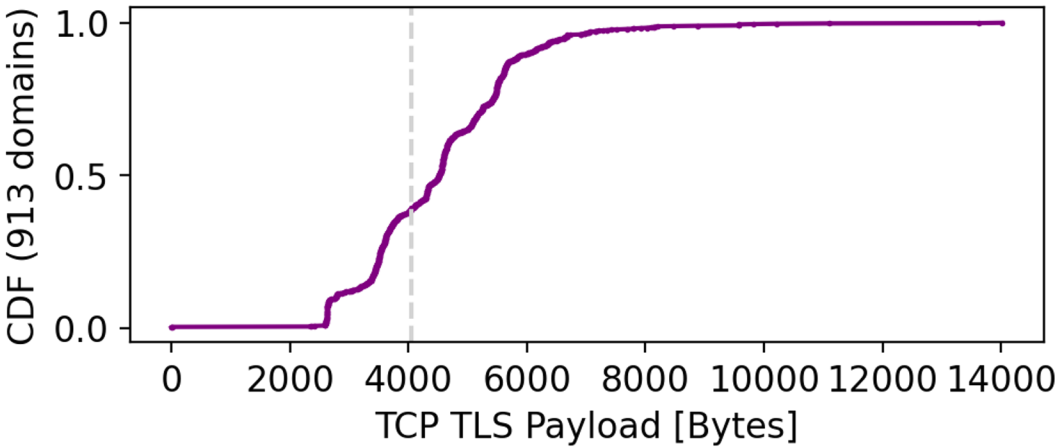
14. Datagram Size

UDP datagrams MUST NOT be fragmented at the IP layer. In IPv4 [IPv4], the Don't Fragment (DF) bit MUST be set if possible, to prevent fragmentation on the path.

QUIC INITIALS (relative).



>50% of 1k top domains have large TLS certificate data!



QUIC vs TCP certificate data.

TLS data received over QUIC or TCP has roughly the same size.

We observe ~100 bytes more with QUIC.