QUICsand: Quantifying QUIC Reconnaissance Scans and DoS Flooding Events

Marcin Nawrocki, Raphael Hiesgen, Thomas C. Schmidt, Matthias Wählisch

{marcin.nawrocki, m.waehlisch}@fu-berlin.de
{raphael.hiesgen, t.schmidt}@haw-hamburg.de

Freie Universität Berlin

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

In a nutshell

Is QUIC used for DoS attacks?

Yes.

Network telescopes allow us to observe these attacks.

QUIC: New protocol, well-known foundations.

QUIC

UDP

By implementation, based on UDP.

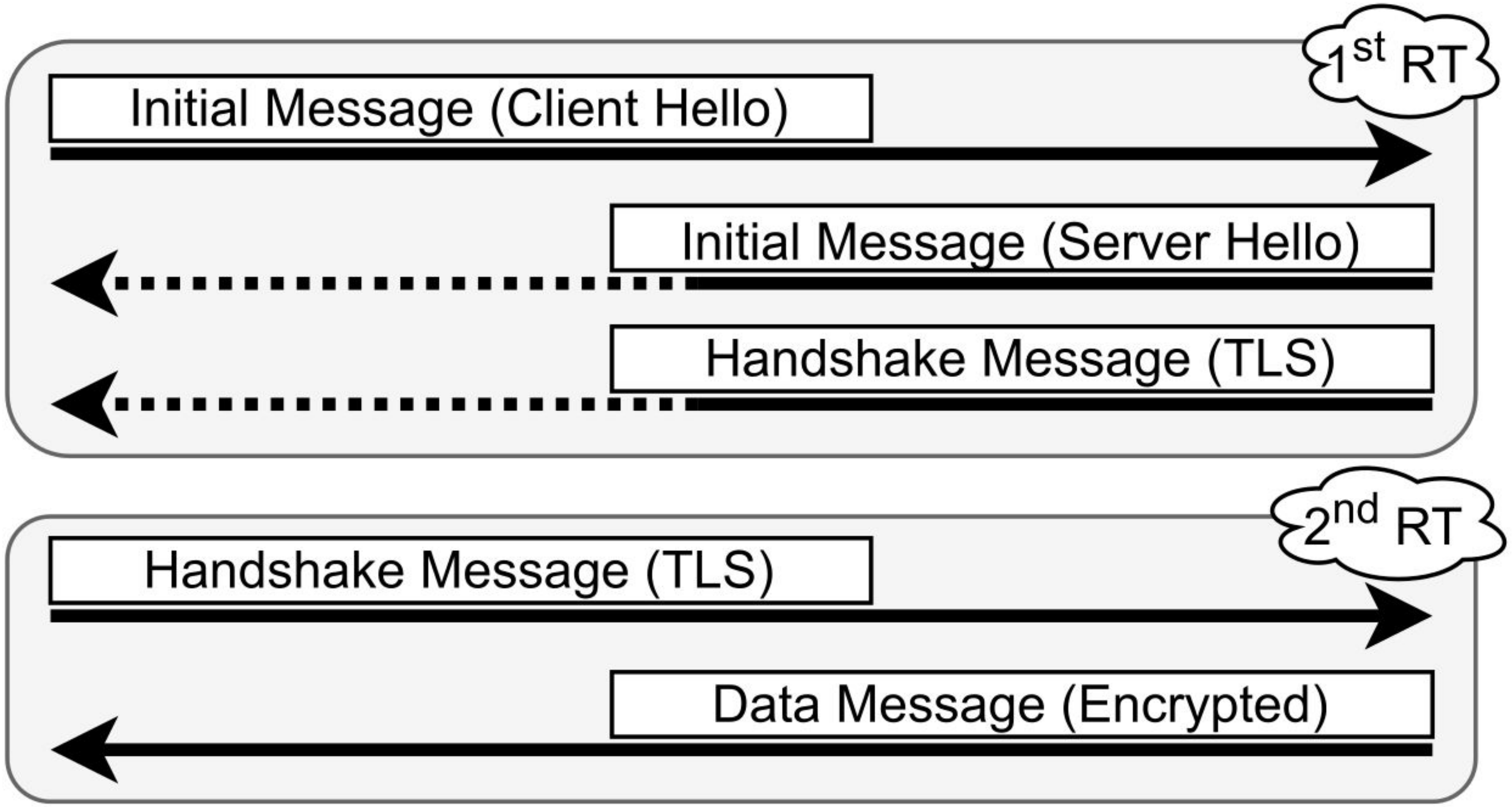
Prevents ossification by middleboxes.

TCP

By design, akin to TCP.

Connection-oriented, base for HTTP/3.

A *typical* QUIC handshake (1-RTT)

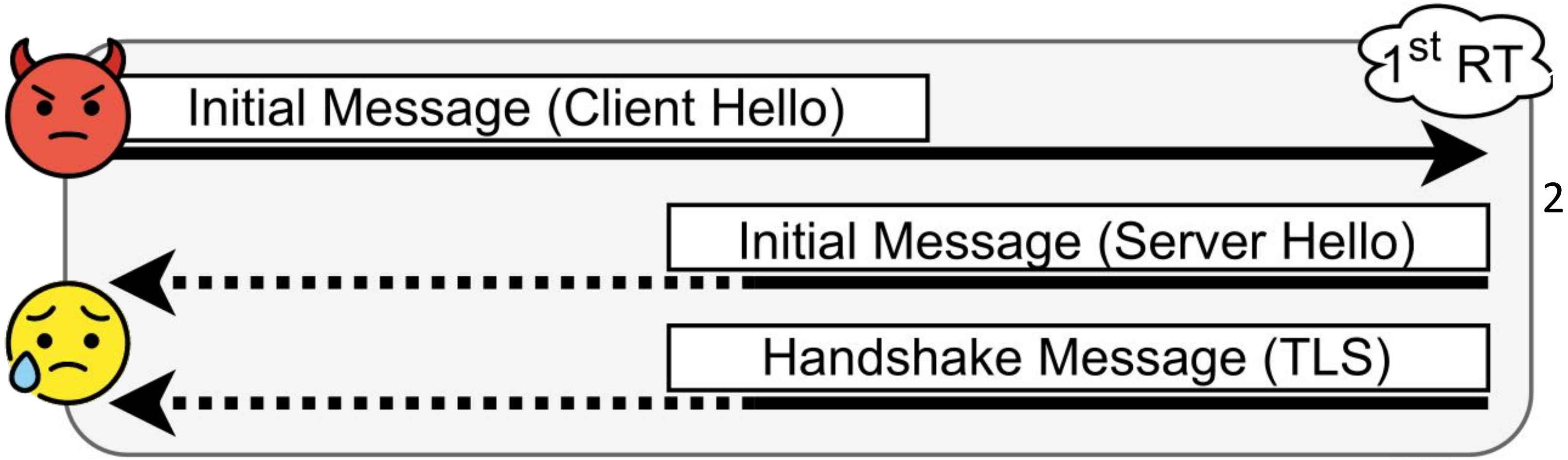


Problem?

During the first round-trip, the server responds to an **unverified** source.

Reflective amplification attacks?

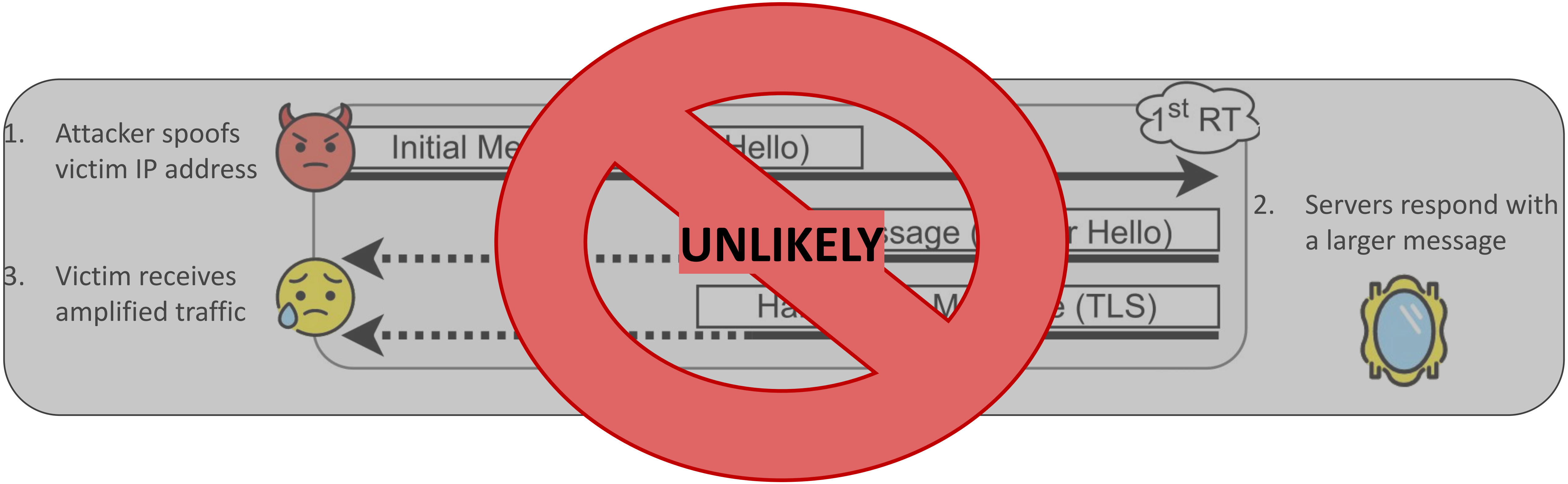
1. Attacker spoofs victim IP address



2. Servers respond with a larger message

3. Victim receives amplified traffic

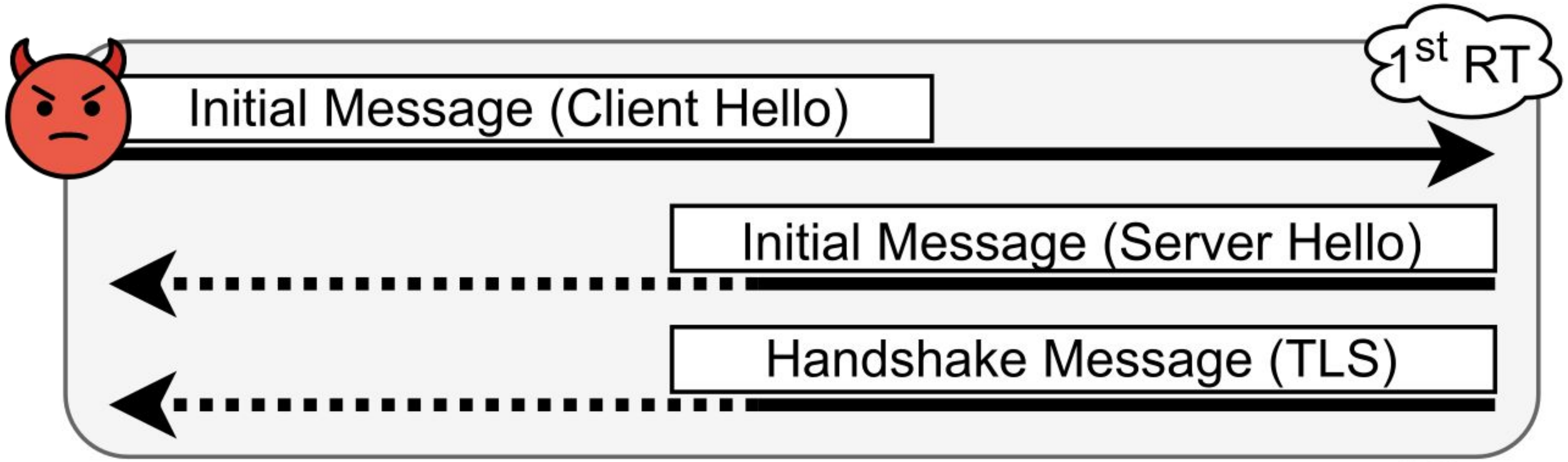
Reflective amplification attacks?



QUIC RFC forbids responses to unverified clients larger than 3x request.
Many UDP-based protocols exist with a higher amplification factor.

Randomly spoofed QUIC INITIAL floods

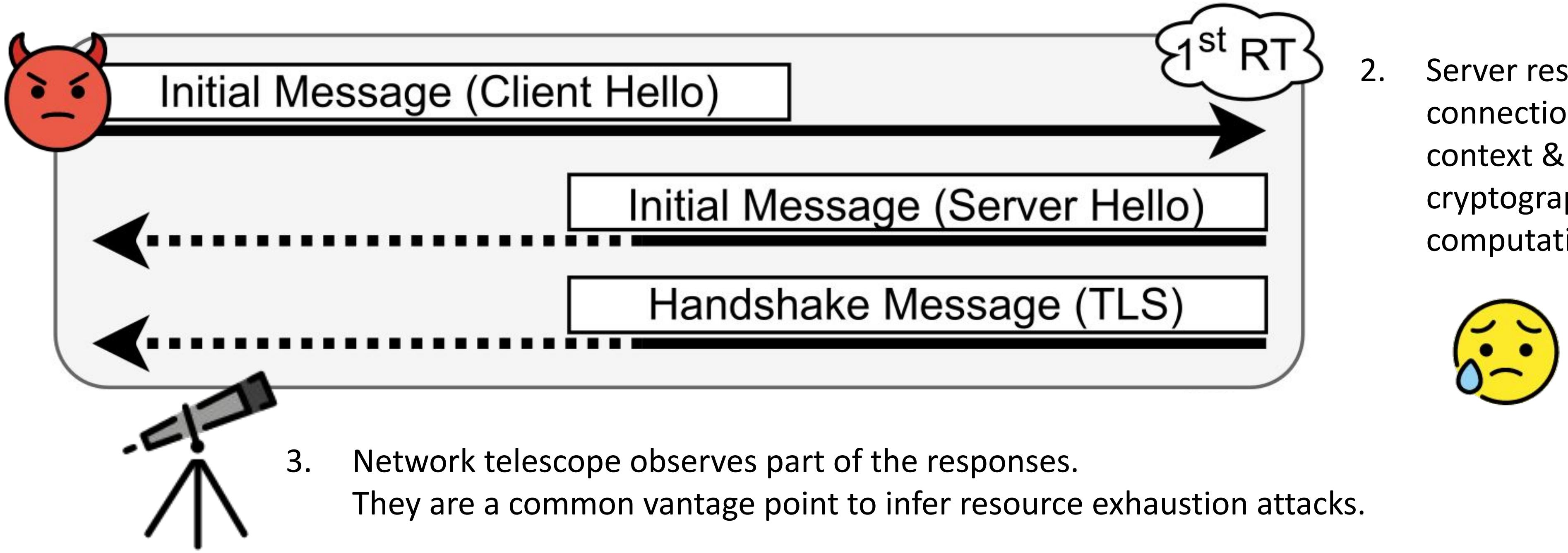
1. Attacker randomly spoofs IP addresses



2. Server reserves connection context & cryptographic computations

Randomly spoofed QUIC INITIAL floods

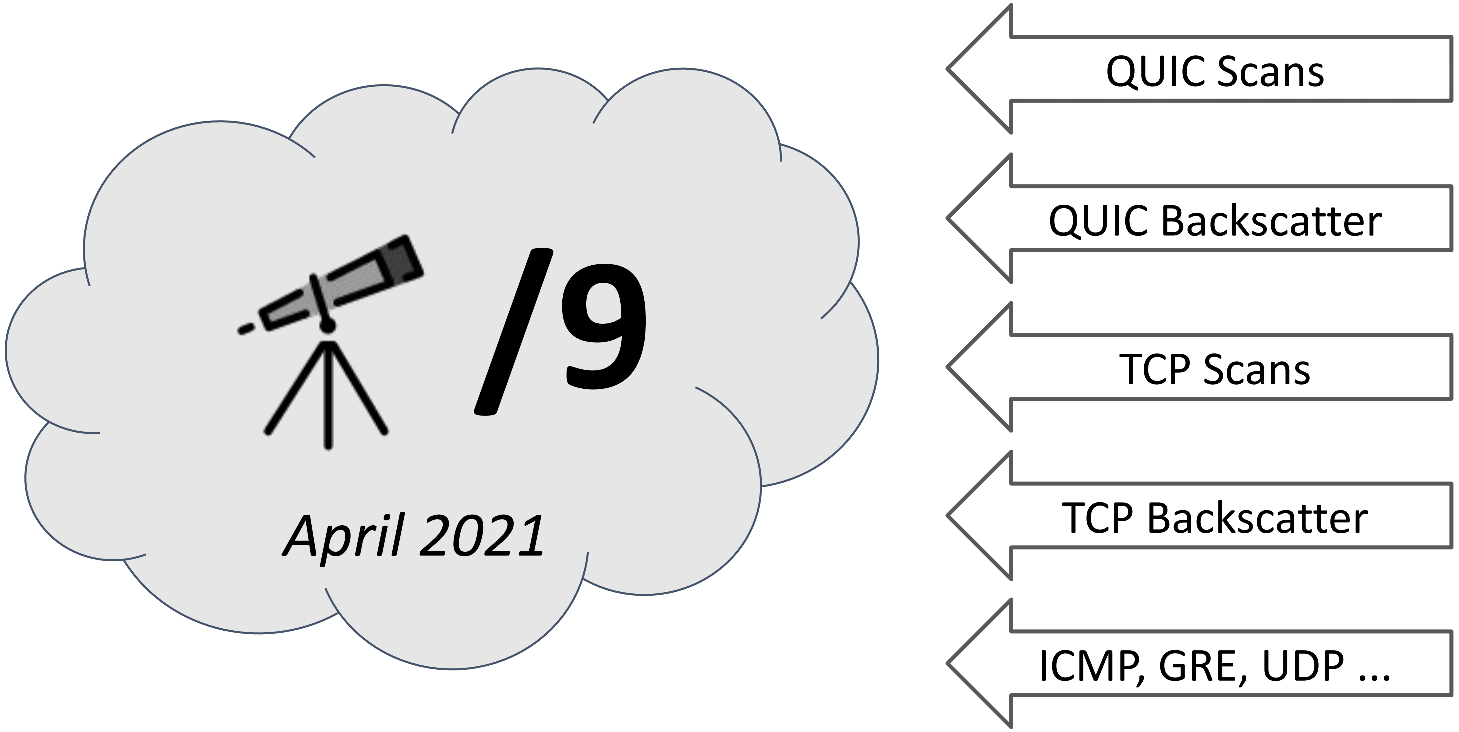
1. Attacker randomly spoofs IP addresses



2. Server reserves connection context & cryptographic computations

3. Network telescope observes part of the responses.
   They are a common vantage point to infer resource exhaustion attacks.

Setup: Passive traffic capture@UCSD telescope.



How to detect QUIC backscatter@telescope?

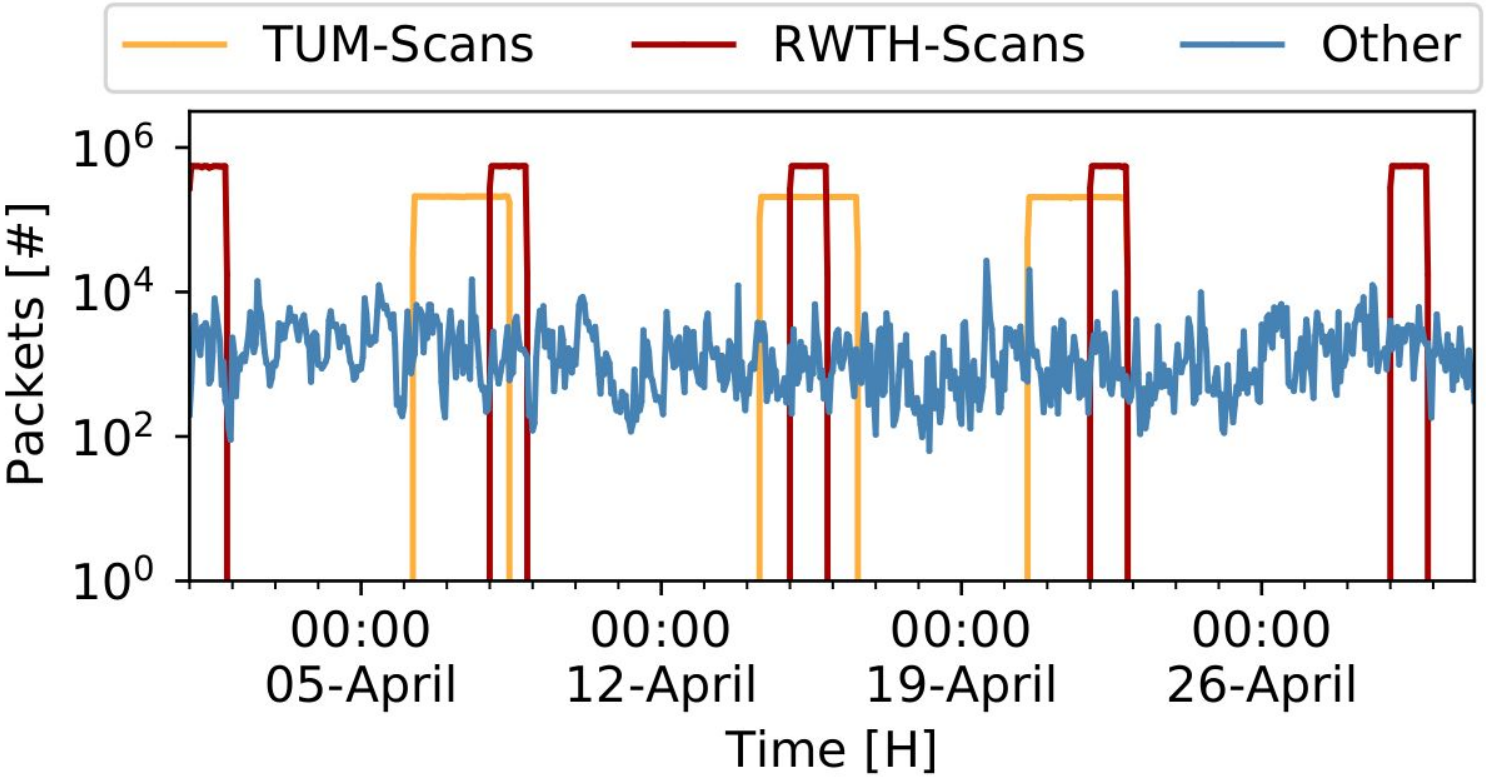
We use Wireshark to detect QUIC traffic based on the payload (DPI), *not only by ports*.

We detected 92M QUIC packets.

Then, we identify *scans* and *backscatter*:

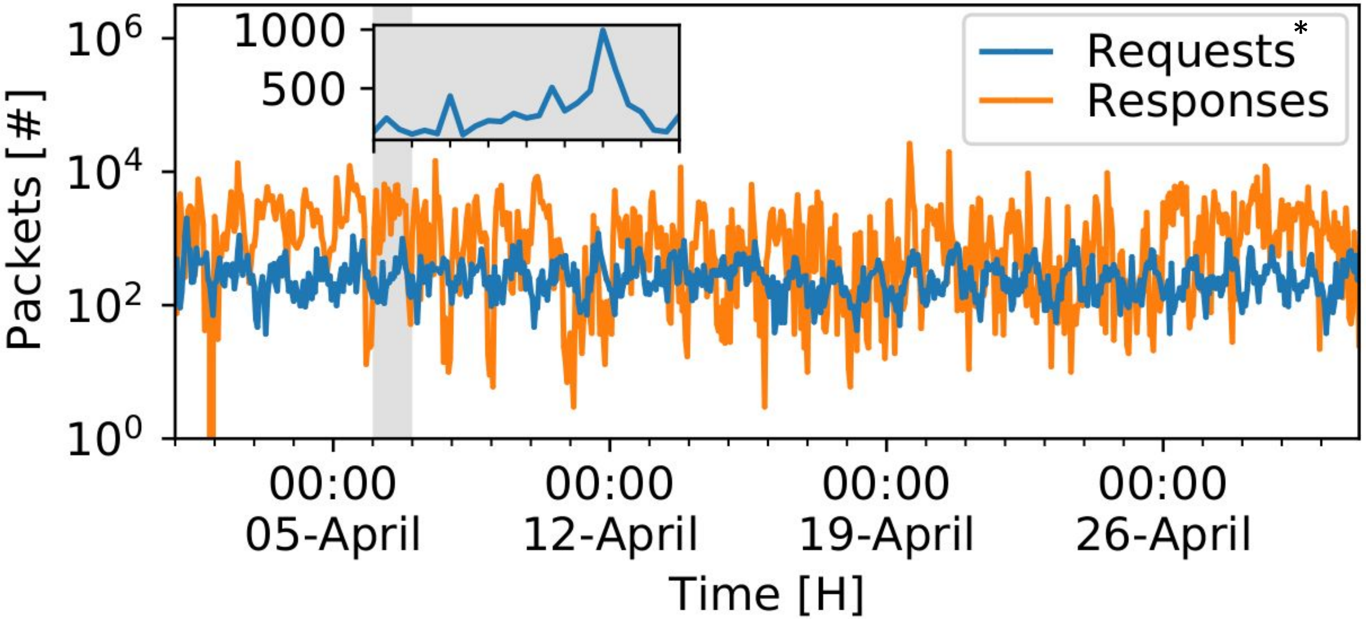
a) QUIC requests are part of scanning activities.
b) QUIC responses are backscatter due to QUIC floods.

Was data sanitization necessary? Yes: Research scanners dominate QUIC IBR



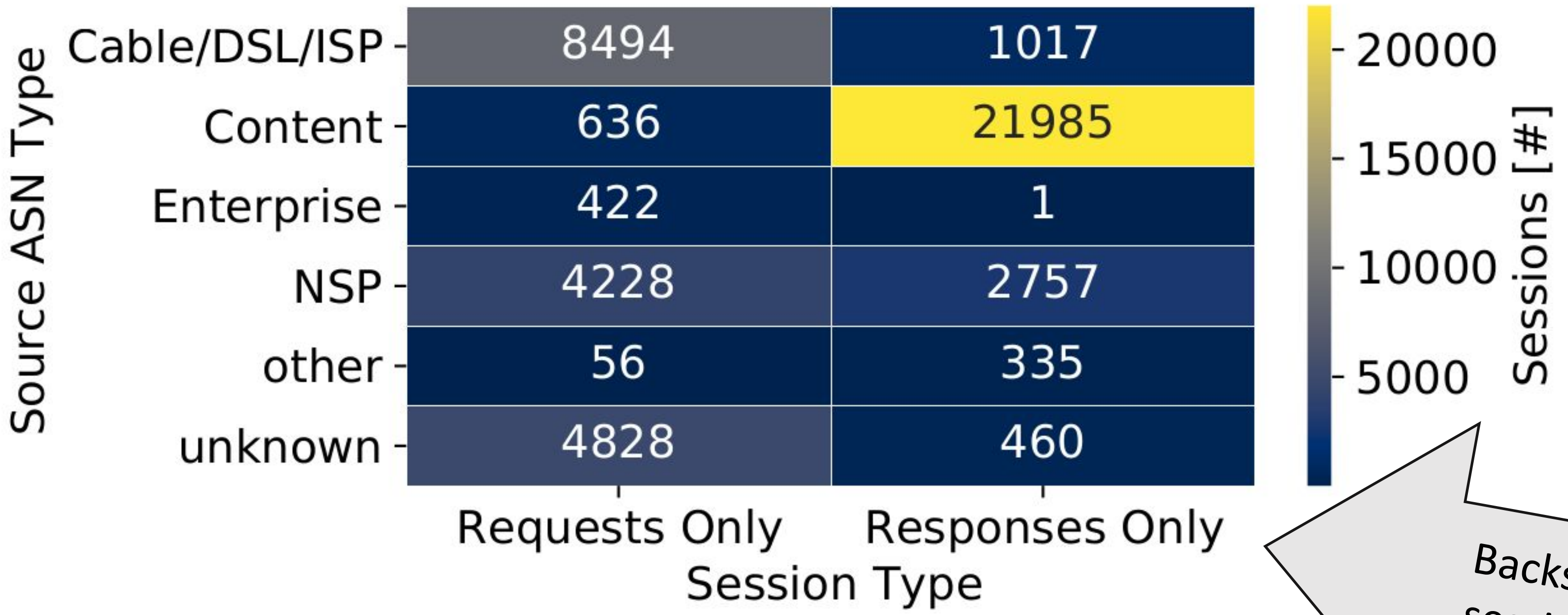
In 2022, we see also Censys scans.

Erratic response traffic hints at DoS events



*sanitized

Sources of QUIC traffic in the telescope



Backscatter, but are response sessions really DDoS events?

How to infer DoS attacks?

We apply a common* method and thresholds to identify attacks.

1. Group packets from the same source into sessions:
idle timeout == 5 minutes

1. Response (backscatter) sessions are an attack if:
> 60 seconds, > 25 packets, and maximum PPS > 0.5

* Moore, David, et al. "Inferring internet denial-of-service activity." *ACM Transactions on Computer Systems (TOCS)* 24.2 (2006): 115-139.

How many attacks did you find?

2905

QUIC floods in April 2021.

How many attacks did you find?

2905

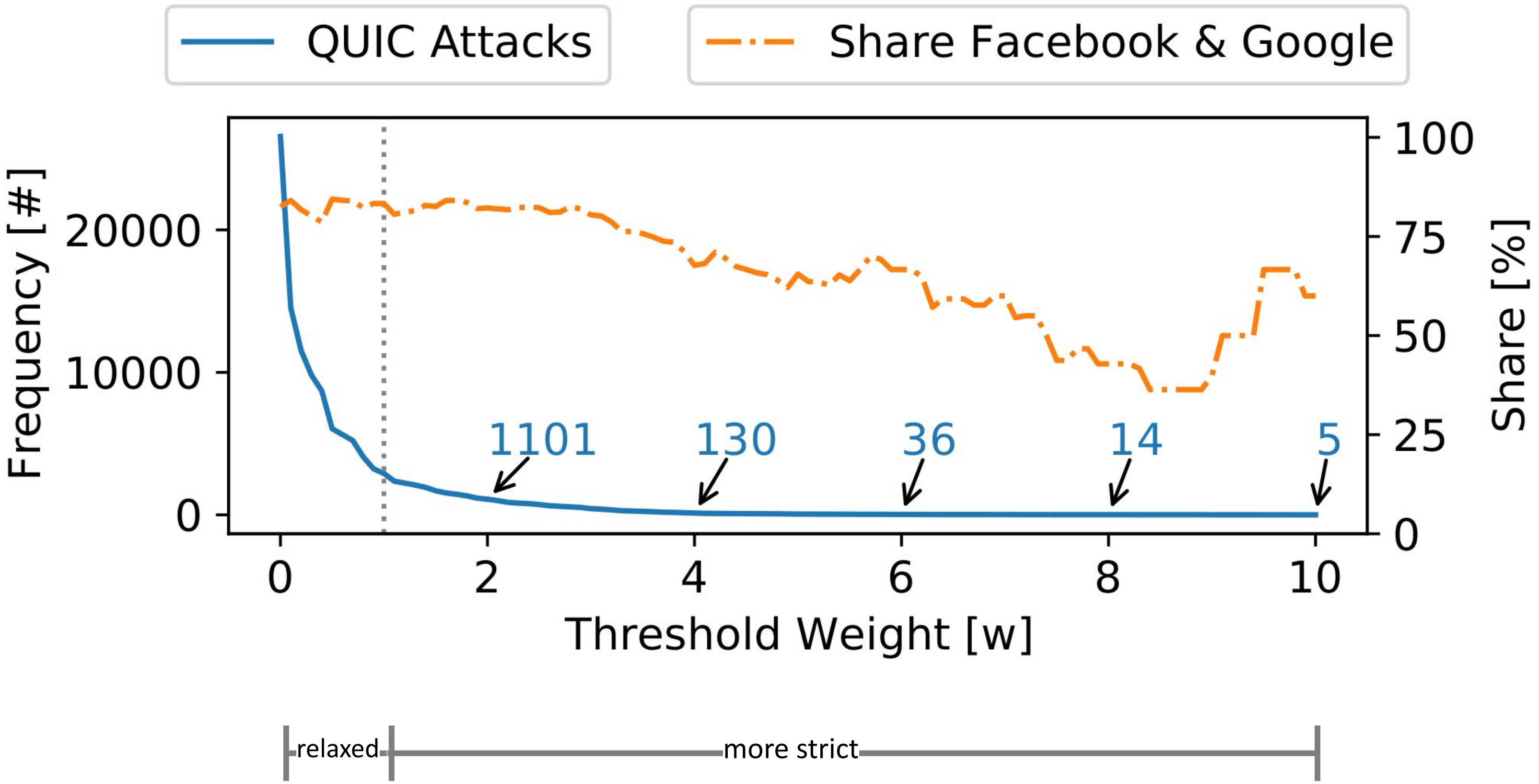
QUIC floods in April 2021.

Google
58%

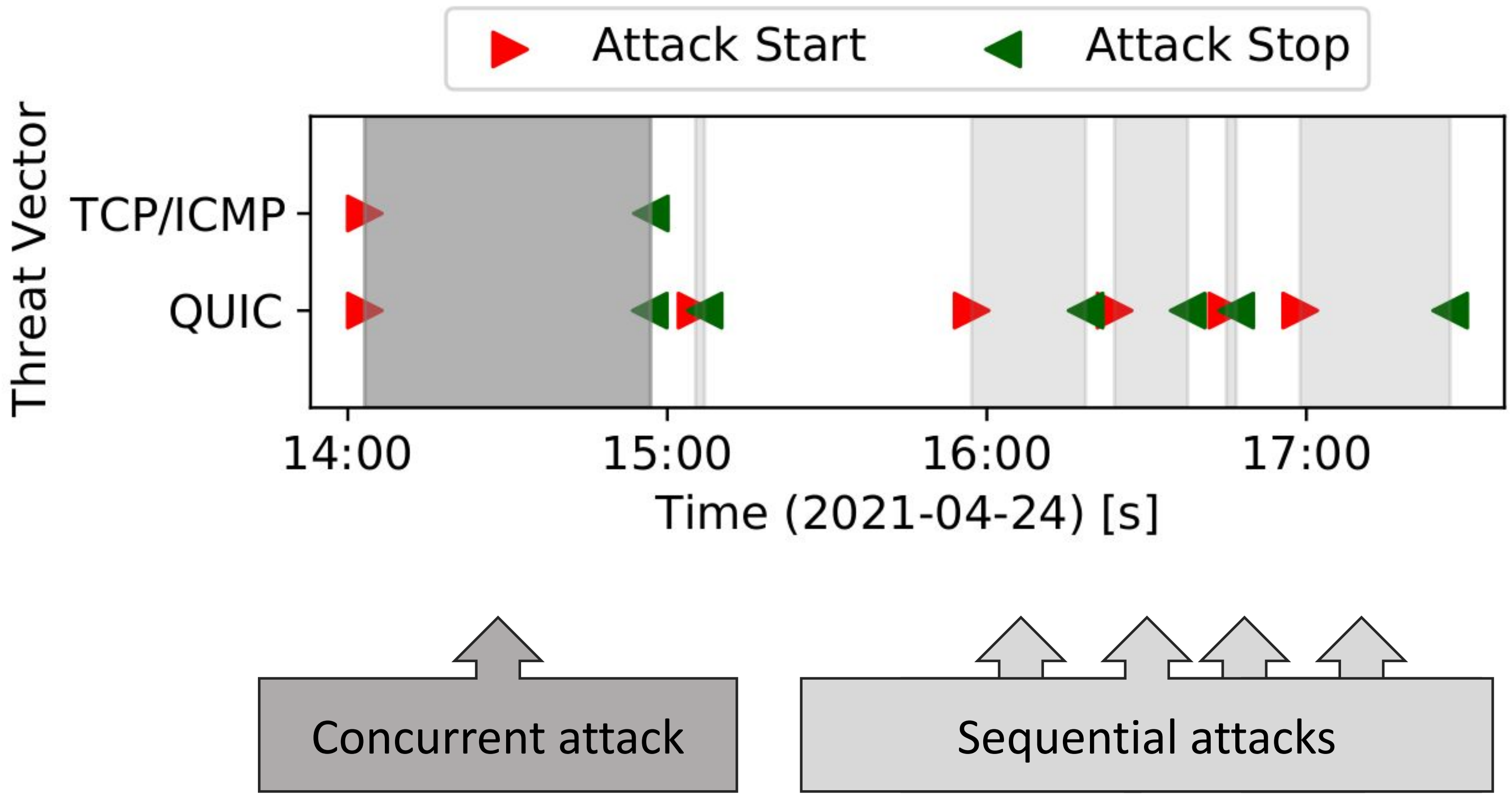
facebook
25%

Victims

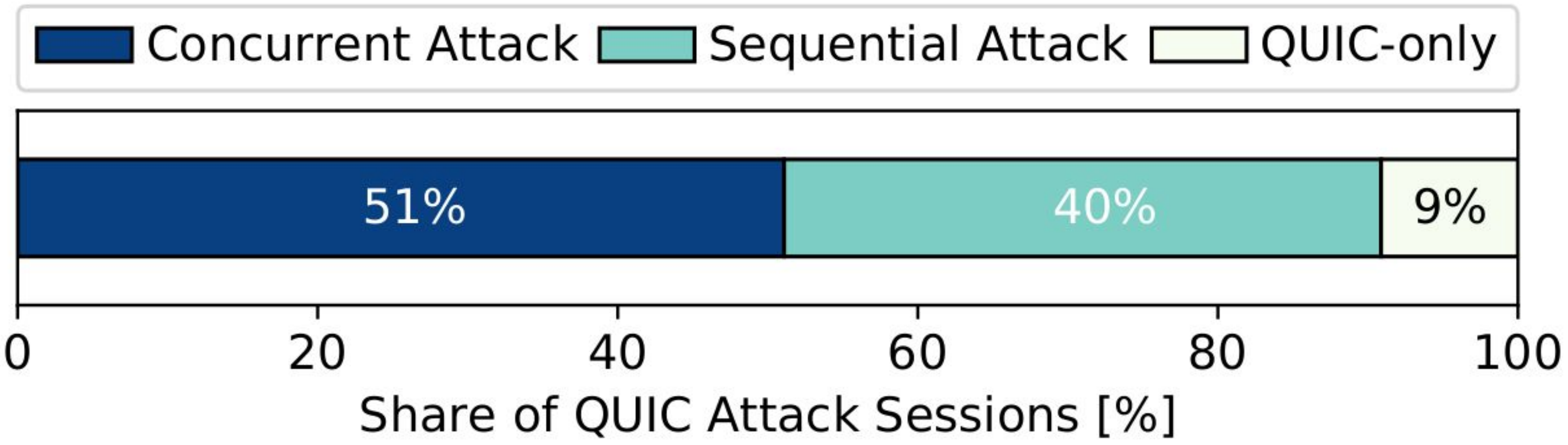
This trend remains even if we apply 10x stricter DoS detection thresholds.



A closer look at a single victim



Multi-vector attacks are common: QUIC INITIAL and TCP SYN floods co-occur



A mitigation option: QUIC RETRY.

Similar to TCP SYN cookies, RETRY messages force the client to return with a unique token.

This proves its authenticity but adds a full round-trip to the connection setup.

https://www.br.de/nachrichten/bayern/anstehen-statt-einloggen-warum-behoerden-noch-kaum-digital-sind

Do QUIC floods really work? Yes, NGINX is vulnerable without RETRY.

Attack	NGINX Config		Results			
Volume [pps]	QUIC Retry	Workers [#]	Client [# Req]	Server [# Resp]	Service Available	Extra RTT
10	✘	4	3,001	12,004	100%	✘
100	✘	4	30,001	81,520	68%	✘
1,000	✘	4	300,001	81,520	7%	✘

Do QUIC floods really work? Yes, NGINX is vulnerable without RETRY.

More CPUs just **delay** the problem

Attack	NGINX Config		Results			
Volume [pps]	QUIC Retry	Workers [#]	Client [# Req]	Server [# Resp]	Service Available	Extra RTT
10	✘	4	3,001	12,004	100%	✘
100	✘	4	30,001	81,520	68%	✘
1,000	✘	4	300,001	81,520	7%	✘
1,000	✘	auto=128	300,001	1,200,004	100%	✘
10,000	✘	auto=128	499,798	521,728	26%	✘
100,000	✘	auto=128	498,505	320,222	26%	✘

Do QUIC floods really work? Yes, NGINX is vulnerable without RETRY.

More CPUs just **delay** the problem

Attack	NGINX Config		Results			
Volume [pps]	QUIC Retry	Workers [#]	Client [# Req]	Server [# Resp]	Service Available	Extra RTT
10	✘	4	3,001	12,004	100%	✘
100	✘	4	30,001	81,520	68%	✘
1,000	✘	4	300,001	81,520	7%	✘
1,000	✘	auto=128	300,001	1,200,004	100%	✘
10,000	✘	auto=128	499,798	521,728	26%	✘
100,000	✘	auto=128	498,505	320,222	26%	✘
1,000	✓	4	300,001	300,001	100%	✓
10,000	✓	4	499,798	499,798	100%	✓
100,000	✓	4	499,798	499,798	100%	✓

Enabling RETRY **prevents** the DoS

Do we want the QUIC RETRY option?

This is **not** about NGINX (or any other implementation).
This is a fundamental QUIC design challenge.

In 2021, no RETRY packets in the DoS backscatter.
RETRY is not used by the large content providers under attack.

Update: April 2021 vs January 2022

Number of QUIC INITIAL floods **doubled**. We now identify off-net servers, which reveals even **more attacks** on Google and Facebook.

First attacks on Cloudflare visible.

Two cases of DoS events mitigated by RETRY packets :).

Conclusion & Outlook

QUIC INITIAL floods are an actively misused (multi-)attack vector.

We detected and quantified QUIC DoS attacks using a network telescope.

Can we fine-tune the DoS thresholds?

Is the deployment of RETRY worth the cost?

More details?

Full paper, ACM IMC 2021

https://doi.org/10.1145/3487552.3487840
https://arxiv.org/pdf/2109.01106.pdf

Artifacts available

https://zenodo.org/record/5504169

QUICsand: Quantifying QUIC Reconnaissance Scans and DoS Flooding Events

Marcin Nawrocki
marcin.nawrocki@fu-berlin.de
Freie Universität Berlin
Germany

Raphael Hiesgen
Raphael.Hiesgen@haw-hamburg.de
HAW Hamburg
Germany

Thomas C. Schmidt
t.schmidt@haw-hamburg.de
HAW Hamburg
Germany

Matthias Wählisch
m.waehlisch@fu-berlin.de
Freie Universität Berlin
Germany

ABSTRACT

In this paper, we present first measurements of Internet background radiation originating from the emerging transport protocol QUIC. Our analysis is based on the UCSD network telescope, correlated with active measurements. We find that research projects dominate the QUIC scanning ecosystem but also discover traffic from non-benign sources. We argue that although QUIC has been carefully designed to restrict reflective amplification attacks, the QUIC handshake is prone to resource exhaustion attacks, similar to TCP SYN floods. We confirm this conjecture by showing how this attack vector is already exploited in multi-vector attacks: On average, the Internet is exposed to four QUIC floods per hour and half of these attacks occur concurrently with other common attack types such as TCP/ICMP floods.

CCS CONCEPTS

• **Security and privacy** → **Denial-of-service attacks**; • **Networks** → **Transport protocols**; *Public Internet*.

ACM Reference Format:
Marcin Nawrocki, Raphael Hiesgen, Thomas C. Schmidt, and Matthias Wählisch. 2021. QUICsand: Quantifying QUIC Reconnaissance Scans and DoS Flooding Events. In *ACM Internet Measurement Conference (IMC '21), November 2–4, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3487552.3487840

1 INTRODUCTION

QUIC is a secure transport protocol originally developed by Google and tested in Chrome browsers since 2013 [26]. It has been recently standardized by the IETF as RFC 9000 [22] and at the same time enjoys rapidly growing deployment by major Web operators and browsers. In 2017, Google estimated that QUIC accounted for 7% of Internet traffic [26] and, by the end of 2020, Facebook announced that 75% of its Internet traffic is QUIC [14]. Despite its recent standardization, QUIC has already many implementations [37] and concurrently supported QUIC versions [40]. In 2021, scans of the complete IPv4 address space detected around 2 million QUIC servers [40].

Key design objectives in QUIC were privacy and security. QUIC was built to reduce the attack surface on the transport layer, which includes attacks such as reflective amplifications [41] and resource exhaustions. Security considerations in the QUIC RFC [22] span 18 pages and discuss properties against active and passive attackers.

In this paper, we report about early observations that indicate regularly ongoing attacks based on QUIC. We argue that the strong security model in QUIC does not preclude misuse and measure clear signals of DDoS attacks in Internet background radiation. We confirm these results by correlating our observations with several complementary data sources. Our findings indicate that QUIC servers are indeed prone to resource exhaustion attacks and these flaws are currently exploited in multi-vector attacks. We believe that it will be crucial to monitor such attack attempts early in the QUIC deployment phase before they enfold their full potential.

The main contributions of our paper summarize as follows:

(1) We present the first study on QUIC Internet background radiation as seen by a large network telescope.
(2) We show a significant bias by research scanners but also detect scanning activity from non-benign sources.
(3) Surprisingly, we find high-volume backscatter events suggesting that QUIC is used in multi-vector resource exhaustion attacks, targeting well-known companies.
(4) We benchmark a popular web server implementation to test its DoS resiliency and validate our observations.
(5) We show the efficacy of QUIC's built in defense mechanism with RETRY messages, which remains unused in the wild based on our measurements.

In the remainder of this paper, we present background and related work in Section 2. We outline the QUIC attack scenarios that we base on in Section 3, and introduce our measuement method and data sources in Section 4. We analyse QUIC scanning and backscatter events in Section 5. Finally, we discuss our findings in Section 6.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
IMC '21, November 2–4, 2021, Virtual Event, USA
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-9129-0/21/11...$15.00
https://doi.org/10.1145/3487552.3487840