

[qlog]

structured **event** logging
for (encrypted) **protocols**

Robin Marx robin.marx@kuleuven.be

What's in a name?

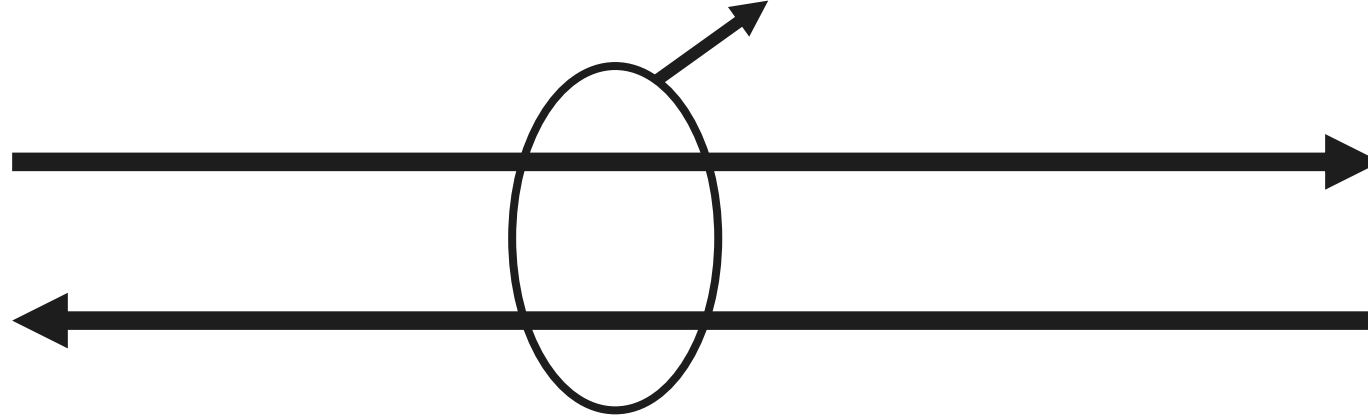
[qlog] = QUIC Logging

QUIC and HTTP/3 are complex

- Will need good debugging and analysis **tools**
- Tools need **data** to ingest

Typical network logging

get raw wire image
from one location

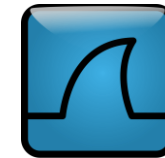


tv-netflix-problems-2011-07-06.pcap

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

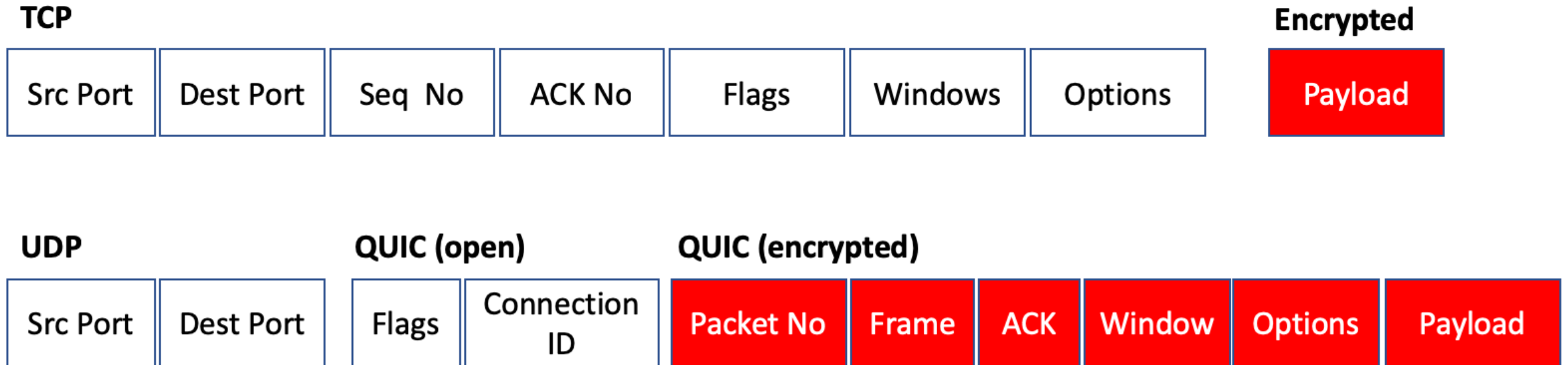
Apply a display filter ... <Ctrl-/> Expression... +

No.	Time	Source	Destination	Protocol	Length	Info
343	65.142415	192.168.0.21	174.129.249.228	TCP	66	40555 → 80 [ACK] Seq=1 Ack=1 Win=5888 Len=0 TSval=491519346 TSecr=551811827
344	65.142715	192.168.0.21	174.129.249.228	HTTP	253	GET /clients/netflix/flash/application.swf?flash_version=flash_lite_2.1&v=1.5&n
345	65.230738	174.129.249.228	192.168.0.21	TCP	66	80 → 40555 [ACK] Seq=1 Ack=188 Win=6864 Len=0 TSval=551811850 TSecr=491519347
346	65.240742	174.129.249.228	192.168.0.21	HTTP	828	HTTP/1.1 302 Moved Temporarily
347	65.241592	192.168.0.21	174.129.249.228	TCP	66	40555 → 80 [ACK] Seq=188 Ack=763 Win=7424 Len=0 TSval=491519446 TSecr=551811852
348	65.242532	192.168.0.21	192.168.0.1	DNS	77	Standard query 0x2188 A cdn-0.nflximg.com
349	65.276870	192.168.0.1	192.168.0.21	DNS	489	Standard query response 0x2188 A cdn-0.nflximg.com CNAME images.netflix.com.edge
350	65.277992	192.168.0.21	63.80.242.48	TCP	74	37063 → 80 [SYN] Seq=0 Win=5840 Len=0 MSS=1460 SACK_PERM=1 TSval=491519482 TSecr=
351	65.297757	63.80.242.48	192.168.0.21	TCP	74	80 → 37063 [SYN, ACK] Seq=0 Ack=1 Win=5792 Len=0 MSS=1460 SACK_PERM=1 TSval=3295
352	65.298396	192.168.0.21	63.80.242.48	TCP	66	37063 → 80 [ACK] Seq=1 Ack=1 Win=5888 Len=0 TSval=491519502 TSecr=3295534130
353	65.298687	192.168.0.21	63.80.242.48	HTTP	153	GET /us/nrd/clients/flash/814540.bun HTTP/1.1
354	65.318730	63.80.242.48	192.168.0.21	TCP	66	80 → 37063 [ACK] Seq=1 Ack=88 Win=5792 Len=0 TSval=3295534151 TSecr=491519503
355	65.321733	63.80.242.48	192.168.0.21	TCP	1514	[TCP segment of a reassembled PDU]



wireshark

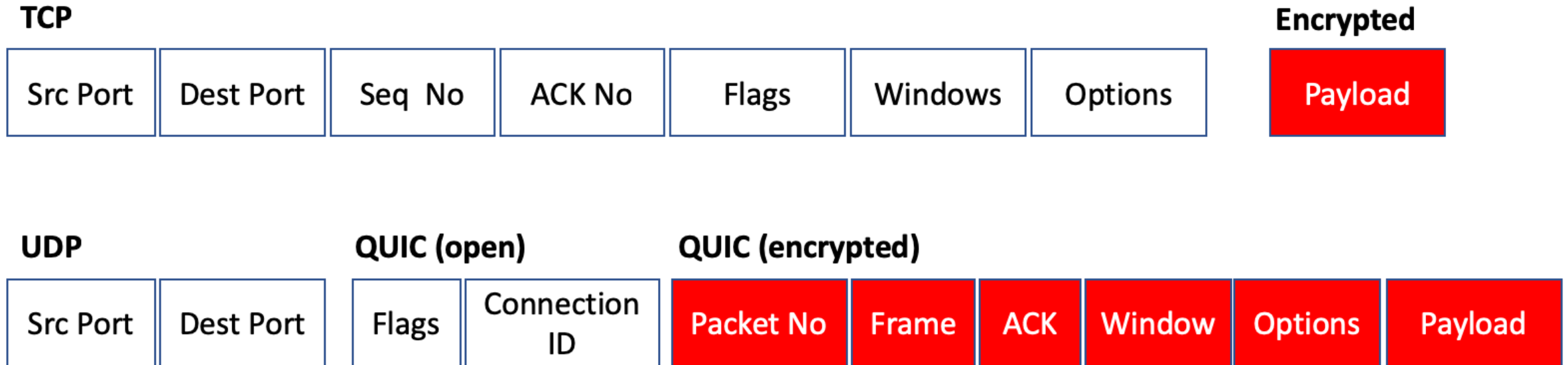
1. QUIC is almost entirely encrypted



Storing full packet captures and TLS secrets is bad for:

- scalability
- privacy

1. QUIC is almost entirely encrypted



2. not everything is sent on the wire

congestion control, decision making, internal errors, ...

[qlog] structured endpoint logging



Event examples

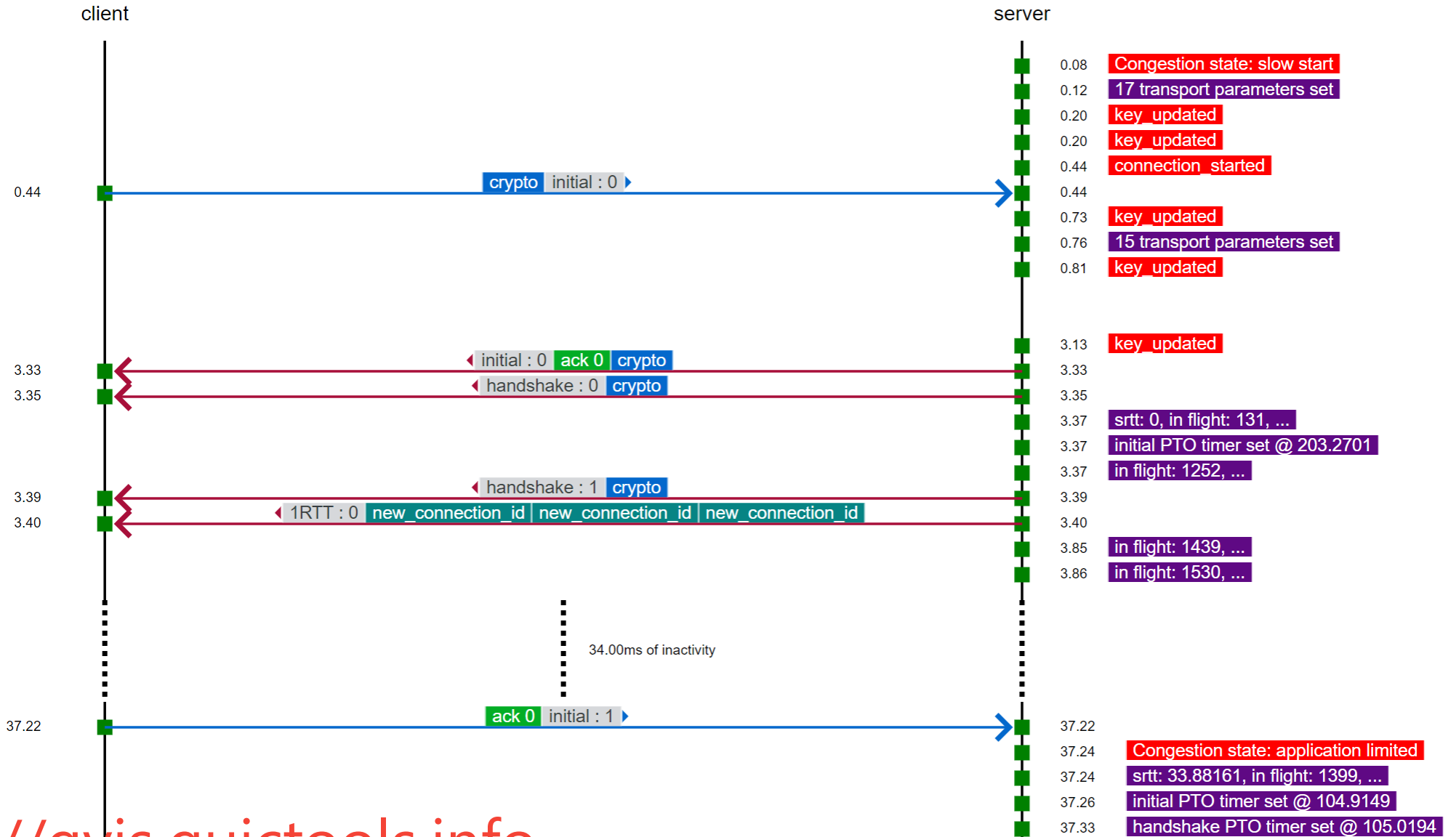
```
{
  "time": 15000,
  "name": "transport:packet_received",
  "data": {
    "header": {
      "packet_type": "1rtt",
      "packet_number": 25
    },
    "frames": [
      {
        "frame_type": "ack",
        "acked_ranges": [
          [10,15],
          [17,20]
        ]
      }
    ]
  }
}
```

```
{
  "time": 15001,
  "name": "recovery:metrics_updated",
  "data": {
    "min_rtt": 25,
    "smoothed_rtt": 30,
    "latest_rtt": 25,

    "congestion_window": 60,
    "bytes_in_flight": 77000,
  }
}
```



QUIC and HTTP/3 tools

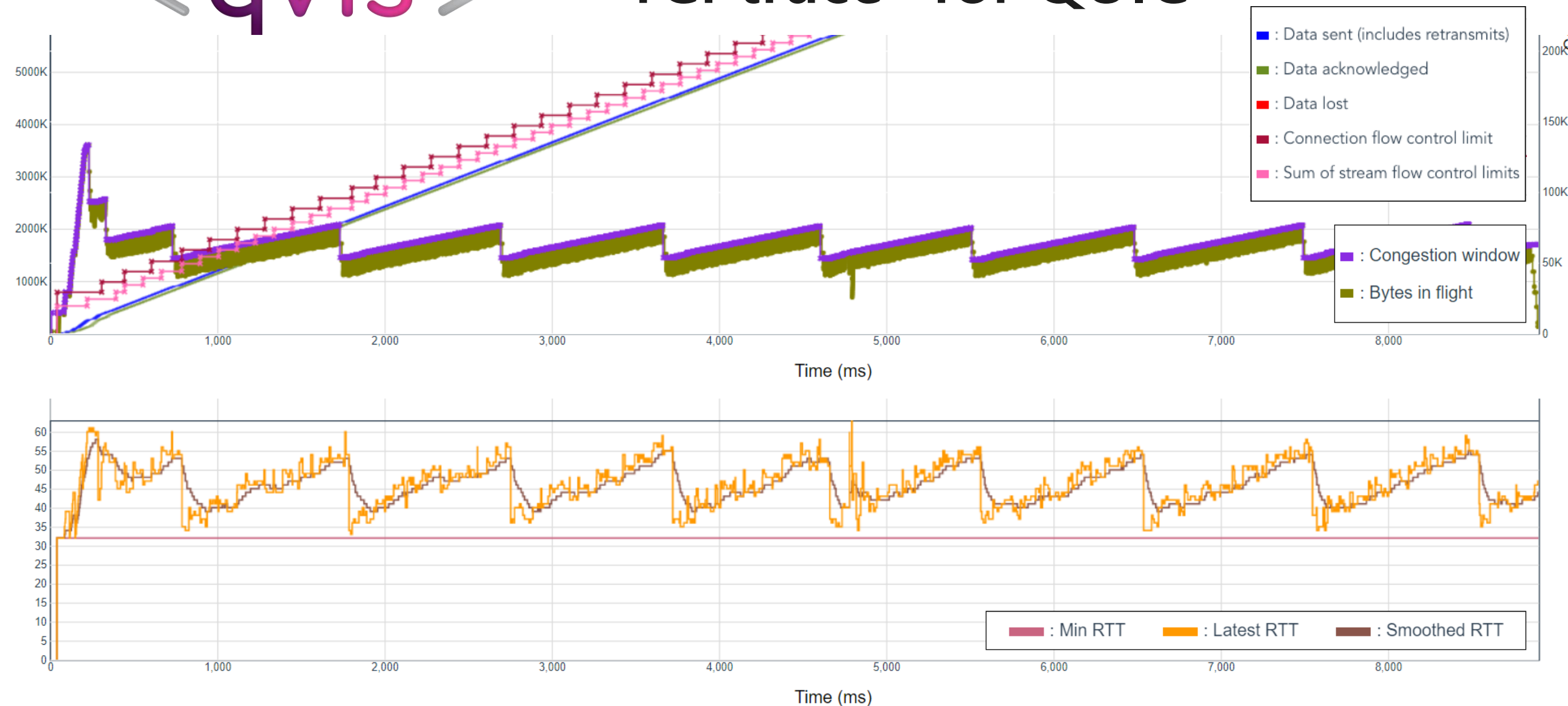


<https://qvis.quictools.info>

<https://github.com/quiclog/qvis>

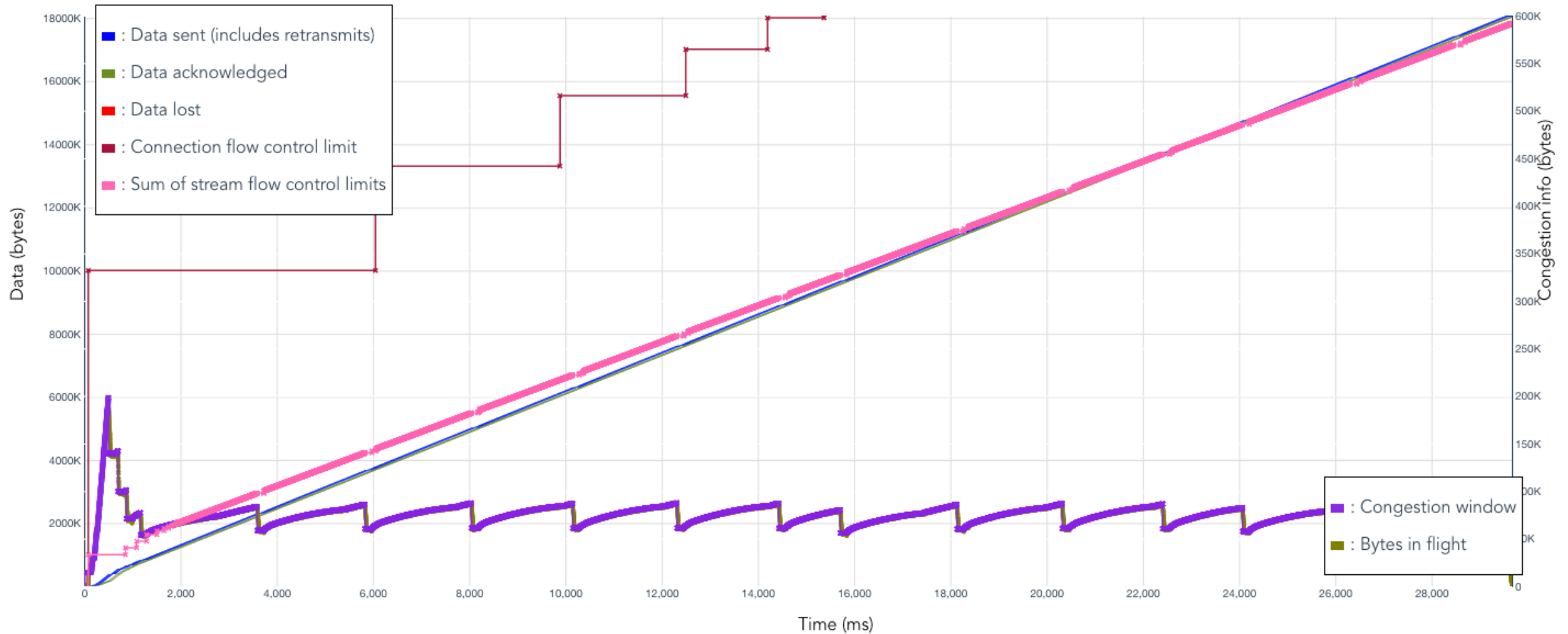


"TCPtrace" for QUIC





"TCPtrace" for QUIC



[qlog] CC deployment debugging

facebook

BBR

- 0-RTT bandwidth estimate too low, because no previously ACKed packets
- Cross-Atlantic growth too slow, due to bad RTT estimates
- Implicit ACKs during handshake caused huge RTT swings, preventing BW growth¹

Flow control

- FC for reset streams was not released, leading to FC-limited transfer

→ Tooling is crucial for non-experts, still useful for experts

¹ http://cs.brown.edu/~tab/papers/QUIC_WWW21.pdf

<https://github.com/facebookincubator/mvfst/commit/39d1a991450d02b2a618bb2fa915c5cf462ee245>

<https://github.com/facebookincubator/mvfst/commit/a5ed2117f36e875f47422ff103cd2026754f3545>

[qlog] support

> **75%** of QUIC/H3 stacks support direct qlog output:

- mvfst



- ngtcp2



- quiche



- quic-go

- aioquic

- quicly / H2O



- neqo



- picoquic

- ...



mjoras 10:35 PM

@rmarx we currently have qlog enabled in prod with similar amounts of events being recorded a day as I quoted before (dozens of billions).

[qlog] adoption

qlog draft adoption in QUIC wg

- Expected before or during IETF 111
- Part of recharter

Goals

- Flesh out schema's for QUIC and HTTP/3
- **Prepare qlog for broader use with other protocols / applications**
 - TCP
 - Multipath TCP and QUIC, MASQUE
 - TLS, HTTP, DNS, WebTransport
 - ...

<https://qlog.edm.uhasselt.be/anrw/>

<https://tools.ietf.org/html/draft-marx-qlog-main-schema-02>

<https://tools.ietf.org/html/draft-marx-qlog-event-definitions-quic-h3-02>

[qlog] for TCP

Use eBPF and kprobes to get info from kernel

- Combine with pcaps to get full info

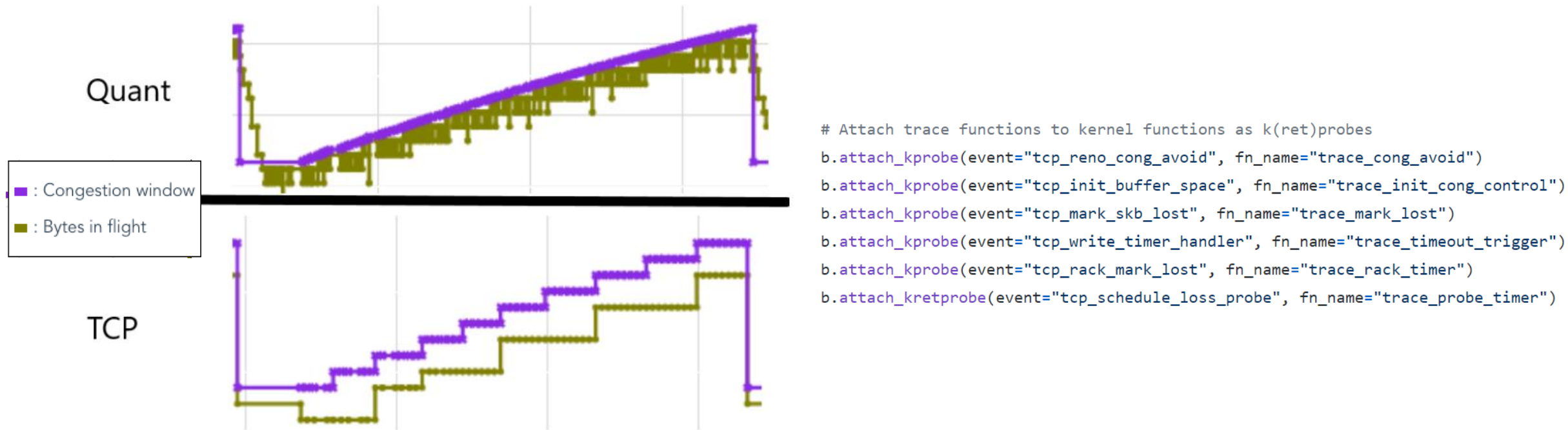
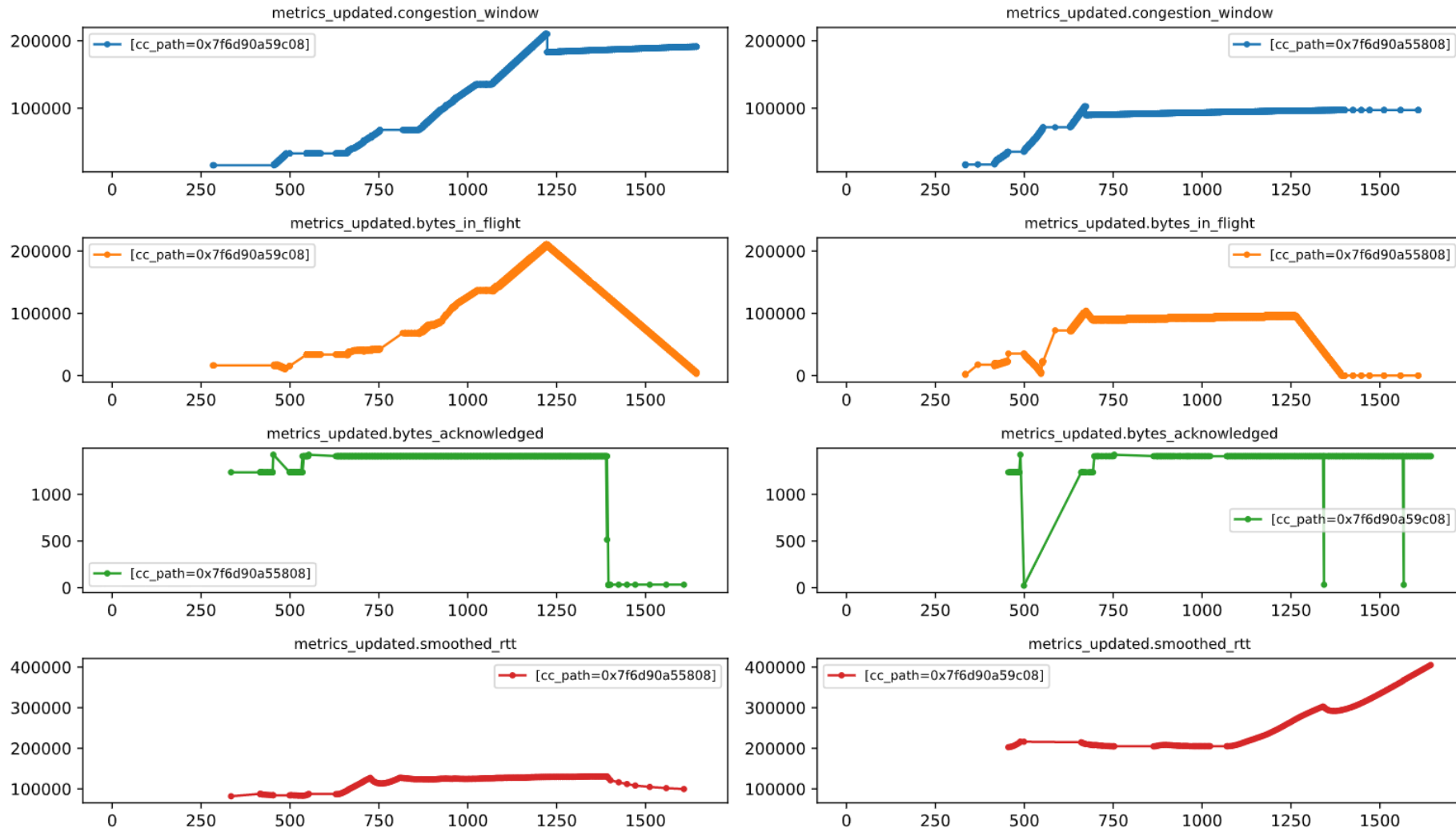


Figure 29: CWND increasing during one congestion avoidance phase in NS-3 (10ms, 5Mbps)

[qlog] for MPQUIC

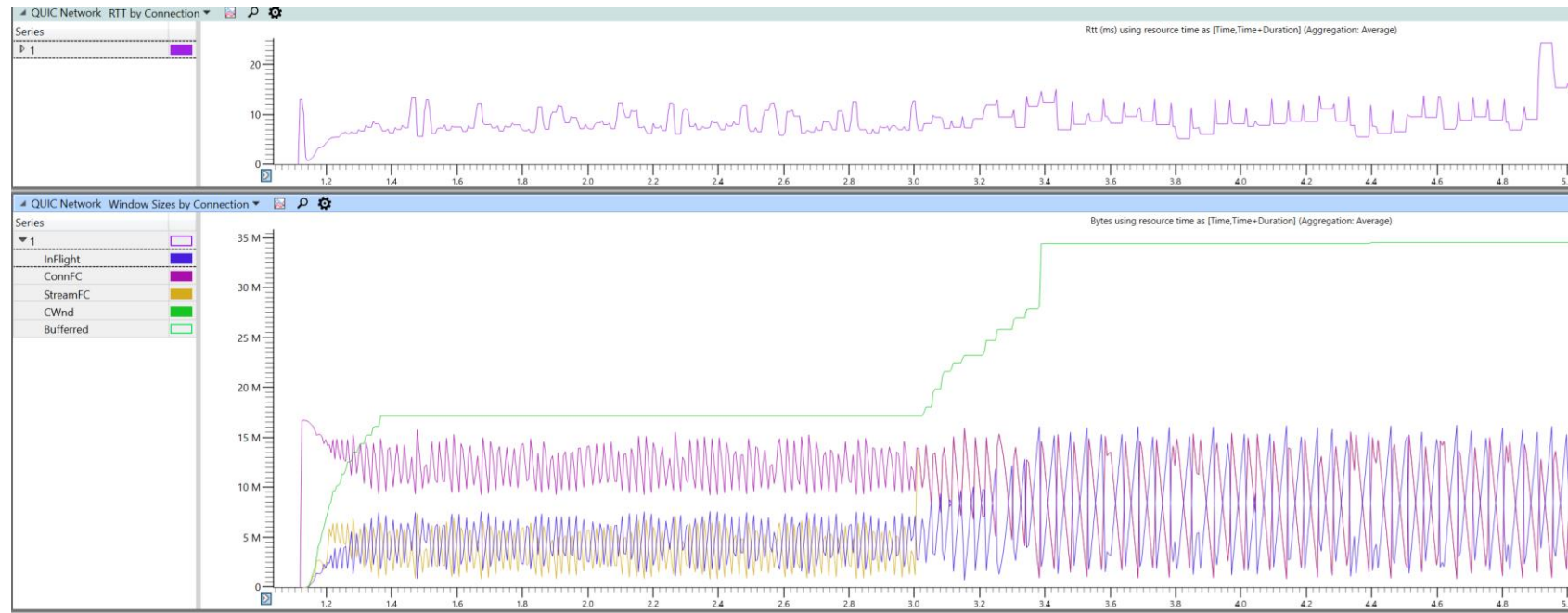
Log path ID with each event, split out later



[qlog] serialization format

qlog is currently **JSON**-based

- 500 MB transfer → 300 MB qlog
- With compression: 18 MB
- Tests on Gbps networks
 - difficult to scale
 - Binary formats or different tools can help there

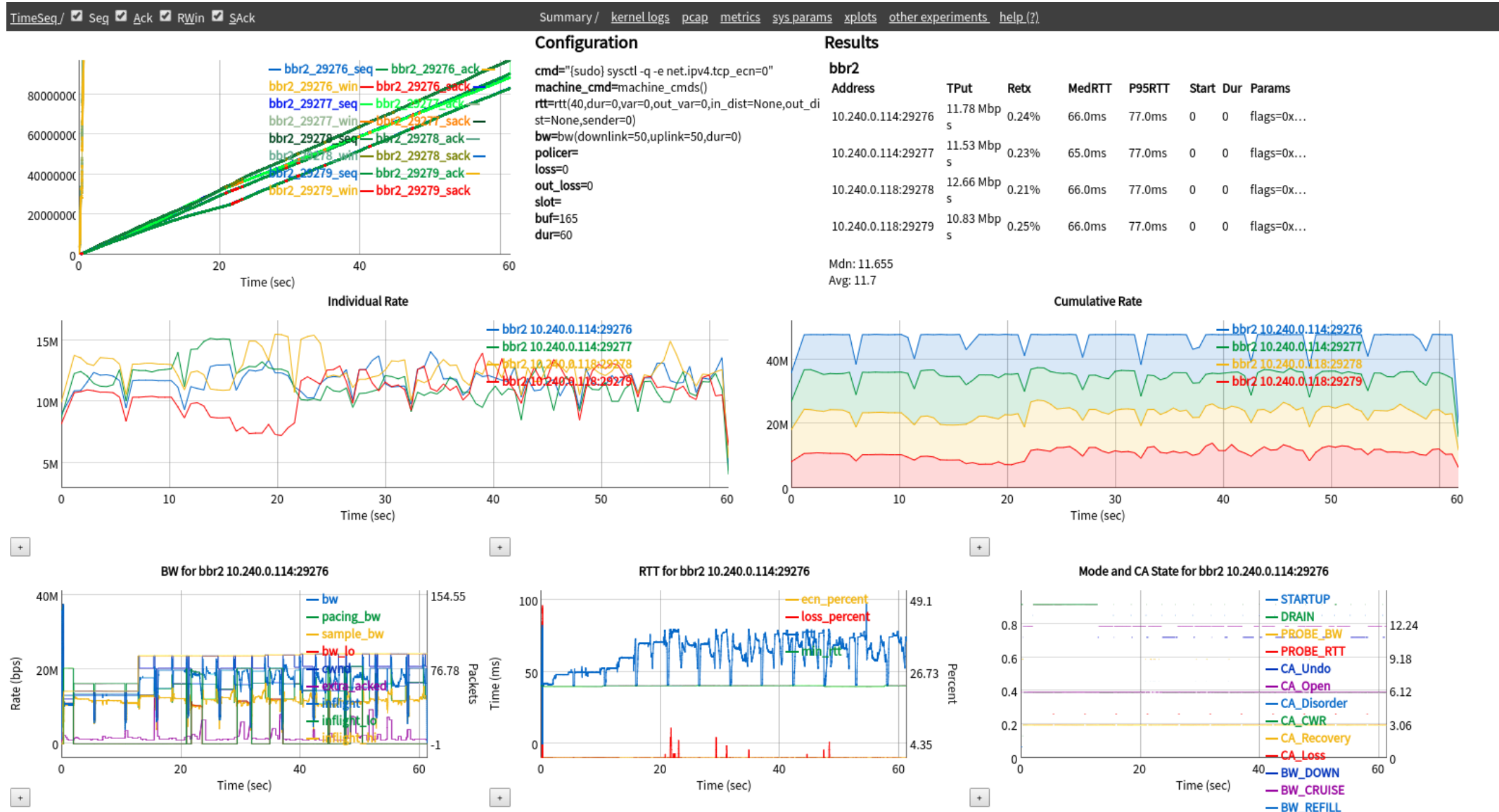


<https://github.com/nibanks/quic-perf>

<https://github.com/quiclog/internet-drafts/issues/30>

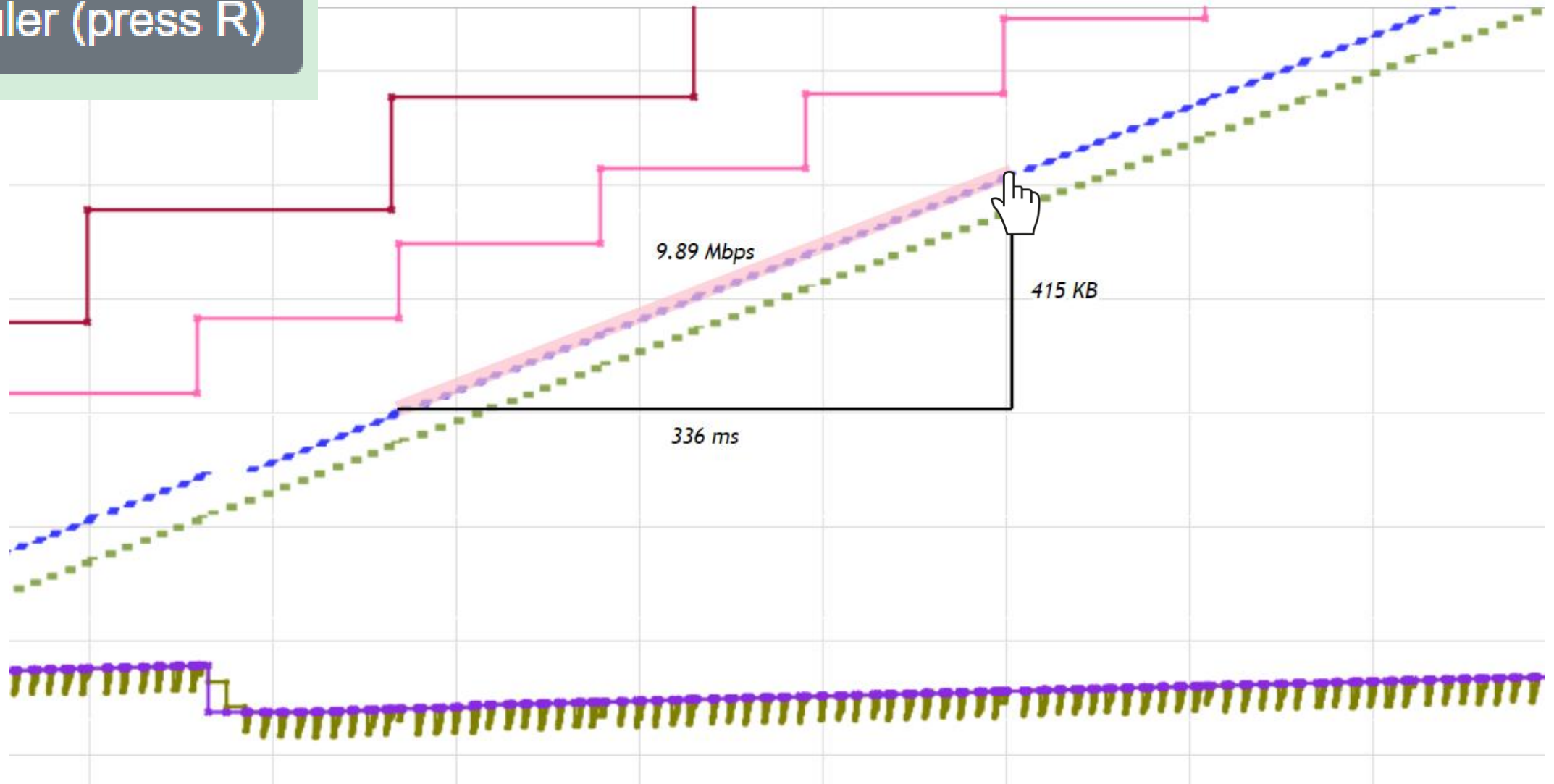
<https://github.com/microsoft/msquic/tree/main/src/plugins/wpa>

How do you actually work your magic?



How do you actually work your magic?

Ruler (press R)



Next steps

Not standardizing tools

- But hoping to make creation of re-usable tools and test suites easier
- Working backwards from tools to see what we need in qlog is a possible approach

Come join us soon!

- Drafts adoption in the QUIC wg (part of recharter)
- Expected before or during IETF 111

In the mean time

- Join us on github.com/quiclog/internet-drafts
- Join the qlog IETF mailing list ietf.org/mailman/listinfo/qlog

Give feedback now!