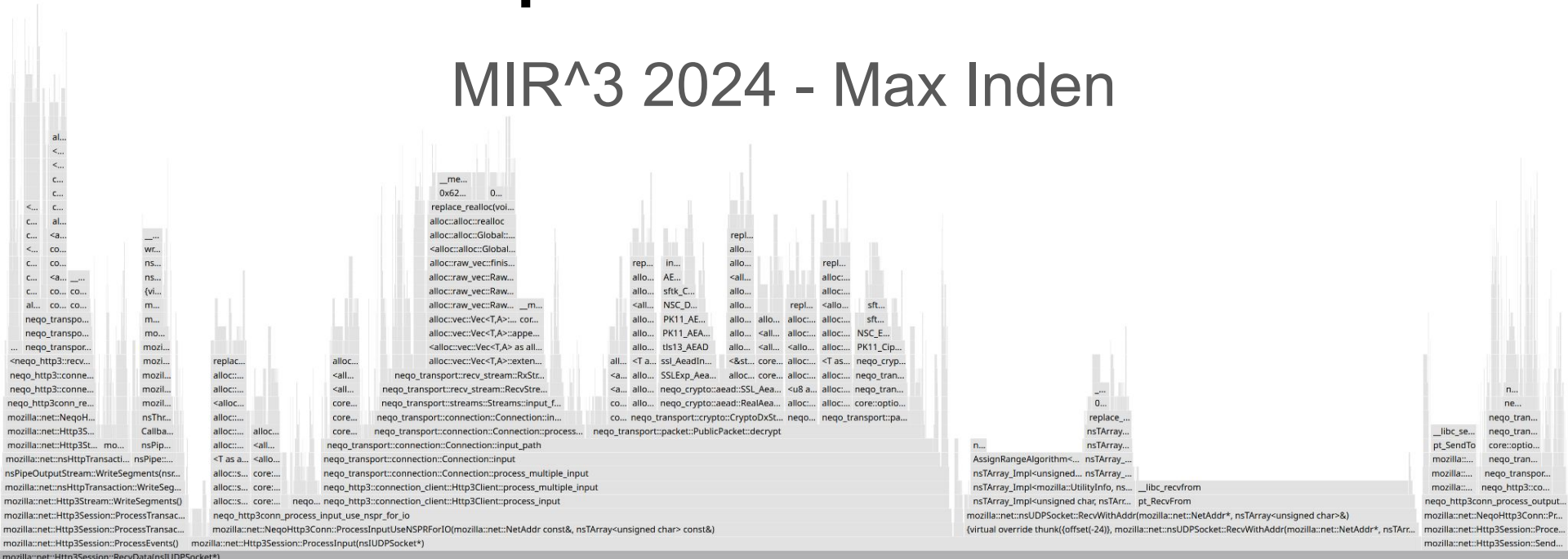
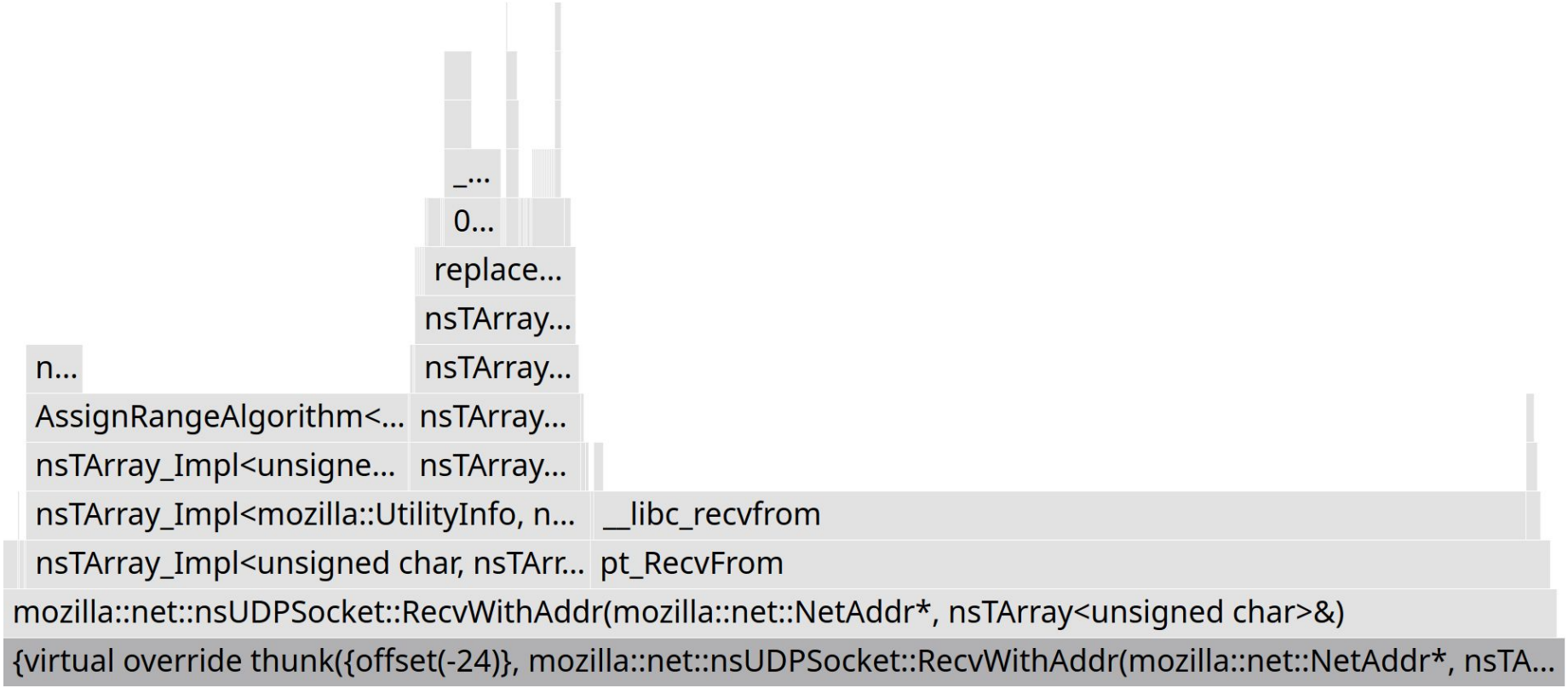
Fast UDP I/O for Firefox's QUIC implementation in Rust

MIR^3 2024 - Max Inden



Firefox socket thread CPU profile

Firefox socket thread CPU profile



TCP vs QUIC

TCP in the kernel is heavily optimized, e.g. with TCP segmentation offload.

“In short, a standard TCP configuration sees a nearly 5x speed-up over UDP”

De Bruijn, Willem, and Eric Dumazet. "Optimizing UDP for content delivery: GSO, pacing and zerocopy." Linux Plumbers Conference. 2018.

In 2018 Linux got UDP GSO (Google) and GRO (RedHat).

Firefox current network stack - NSPR

- Single datagram
  - e.g. sendto recvfrom on Linux
- No segmentation
- Using Netscape Portable Runtime

Modern system calls

- Multi-message (e.g. sendmmsg, recvmmsg)
- Segmentation (e.g. GSO, GRO)
  - same 5 tuple
  - same length, but last
- Using Rust UDP library quinn-udp

Platform support

	Linux / Android	Windows	Apple
Send			
segmentation	Generic Segmentation Offload (GSO)	UDP Segmentation Offload (USO)	
Single msg	sendmsg	WSASendMsg	sendmsg
Multi msg	sendmmsg		sendmsg_x
Receive			
segmentation	Generic Receive Offload (GRO)	UDP Receive Segment Coalescing Offload (URO)	
Single msg	recvmsg	WSARecvMsg	recvmsg
Multi msg	recvmmsg		recvmsg_x

Experiments

- in Firefox Nightly (or `network.http.http3.use_nspr_for_io`)
  - GRO (Linux, Android, Windows)
    - 75th of read syscalls read 2 or more packets, 95th read 10 or more packets.
    - 75th of read syscalls read 1.2 KiB total, 95th read 2.4kbyte total.
- proof of concept
  - GSO (Linux, Android, Windows)
  - sendmmsg (MacOS)
  - recvmmsg (Linux, Android, Windows, MacOS)
- Getting close to 8 Gbit/s (localhost)

Additional wins

- Explicit Congestion Notification (ECN)
  - With modern syscalls, we can now read auxiliary IP metadata, e.g. ECN.
  - On Firefox Nightly 75th percentile of QUIC connections see >= 0.6% CE marks on receive path.
- Optimized memory management
  - Using Rust's borrow checker, we can use long-lived send & recv buffers.

Open questions

- Pacing vs GSO
- packet trains on the internet
- send & receive buffer size
- measure power consumption

Further reading

- Cloudflare: Accelerating UDP packet transmission for QUIC

- De Bruijn, Willem, and Eric Dumazet. "Optimizing UDP for content delivery: GSO, pacing and zerocopy." Linux Plumbers Conference. 2018.

- LWN: Generic receive offload

- Mozilla: Fast UDP for Firefox