Building modern TCP stack for high-performance DNS on top of XDP

Libor Peltan • libor.peltan@nic.cz • 2021-11-25

CZ DOMAIN REGISTRY

What is XDP

- Bypass kernel during packet processing
  - Only for relevant traffic
- Requires network stack in userspace
- Invisible for firewall and tcpdump
- Improves throughput & latency

DNS over UDP over XDP

- Knot DNS 3.0 (Sep 2020)
- Later improvements, routing helpers
- Performance several-times higher
  - See OARC36 next Tuesday

XDP-TCP: Motivation

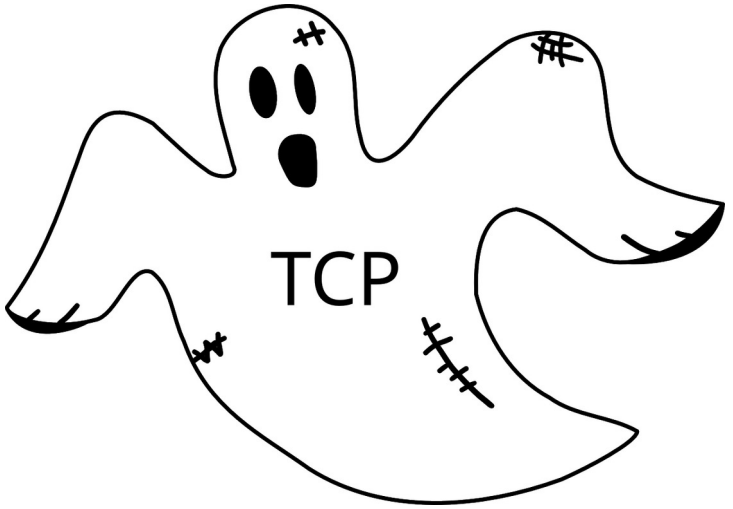
- Let’s try it
- Performance of conventional TCP
  - Worse in multi-thread process
- Slow-loris, DDoS
- Increasing use of DNS-over-TCP
  - DNSSEC, Flag Days
- DNS-over-TCP shall not be discouraged

XDP-TCP: Why custom stack

- “Halfway finished already”
- Designed with DNS in mind
- Existing stacks (like lwTCP) are for IoT
  - And no suitable API
- Diversity

Custom TCP stack implementation

- TCP has accumulated many ~~quirks~~ RFCs
- Balance between connection reliability and server security



XDP-TCP: Limitations

- No Congestion control
  - AXFR over XDP?
- No out-of-order
- No PMTUD
  - ICMP 

XDP-TCP: Results

- In Lab conditions
  (99% answer threshold, 1 query per conn)
  - ~0.1 Mqps → 1.4 Mqps
- SYN attack
  - 3.27 MSYNps → 3.5+ MSYNps
- Incomplete Query attack
  - XY kqps → 2.1 Mqps
  - XY differs highly with configuration

Kxdpgun with TCP

- Hi-perf debug and measure tool
- Using XDP & same routines
- Incl. TCP stack
- Options for SYN, DoS and memory depletion attacks

XDP-TCP: Future

- Finish it
- Keep on improving further
- Find users
- DNS over QUIC (DoQ) over XDP