Knot DNS with XDP

CcNSO Tech Day – ICANN 69

Ondřej Filip • ondrej.filip@nic.cz • 19 Oct 2020 • Prague

CZ DOMAIN REGISTRY

Knot DNS

- High-performance open-source authoritative DNS server (sister project Knot Resolver)
- Full featured
- Multi-threaded and mostly lock-free implementation
  - Scales well on multi-core systems
  - Non-stop operations even when adding or removing zones.

Knot DNS – some features

- Zone journal storage
- Persistent zone event timers
- YAML-based or database-based configuration
- Query processing modules with dynamic loading
- On-the-fly zone management and server reconfiguration

Knot DNS – some features

- Multithreaded DNSSEC signing
- Automatic DNSSEC key management
- Offline KSK operation
- PKCS #11 interface

Knot DNS – modules

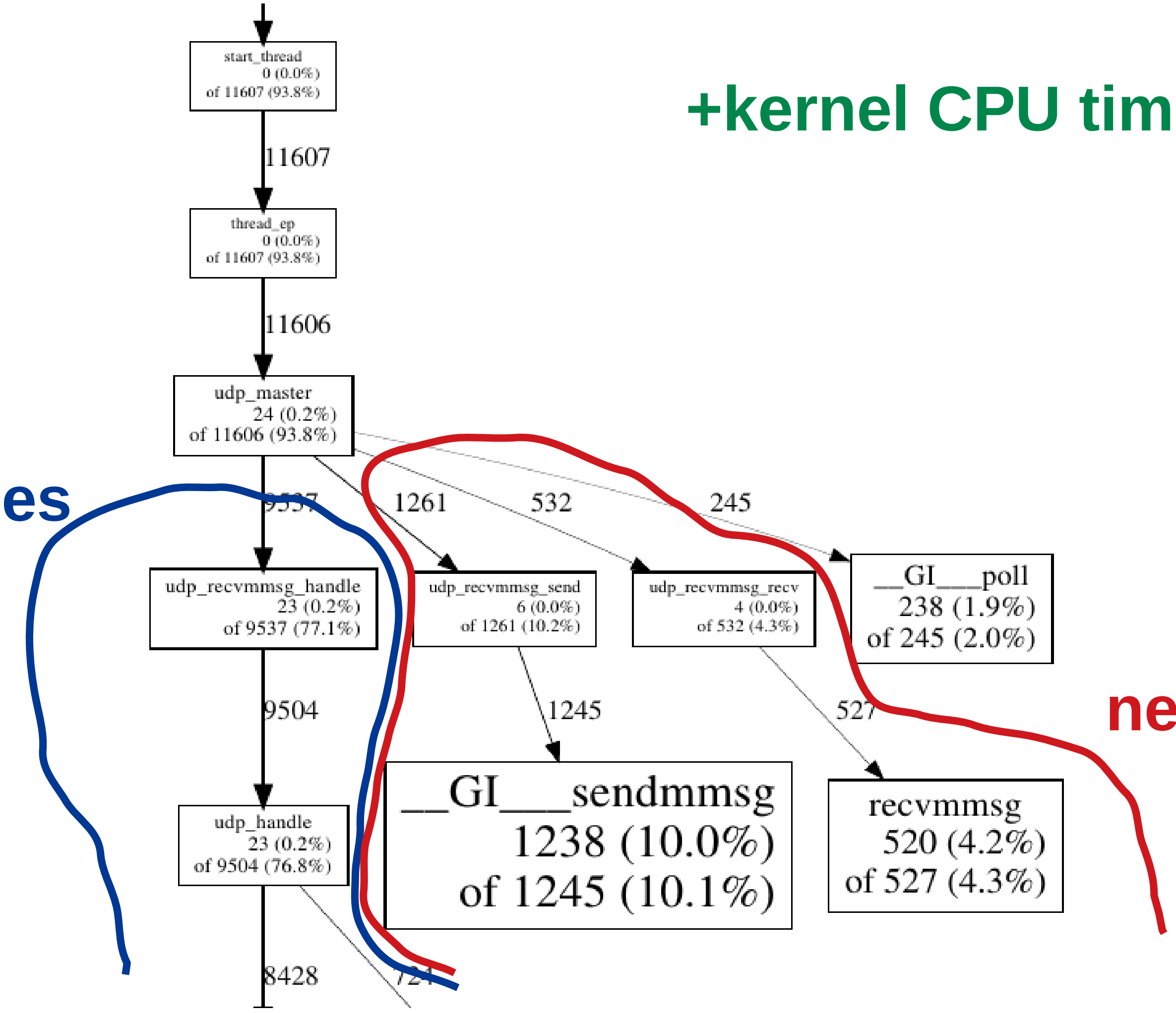
- Response rate limiting
- Forward and reverse records synthesis
- DNS request traffic statistics
- Dnstap traffic logging
- Online DNSSEC signing
- GeoIP response tailoring supporting ECS and DNSSEC

Implementation of XDP

NEW

- Brand new feature
- XDP (eXpress Data Path)
- Introduced in in version 3.0.0
- Currently 3.0.1

Authoritative DNS server profile

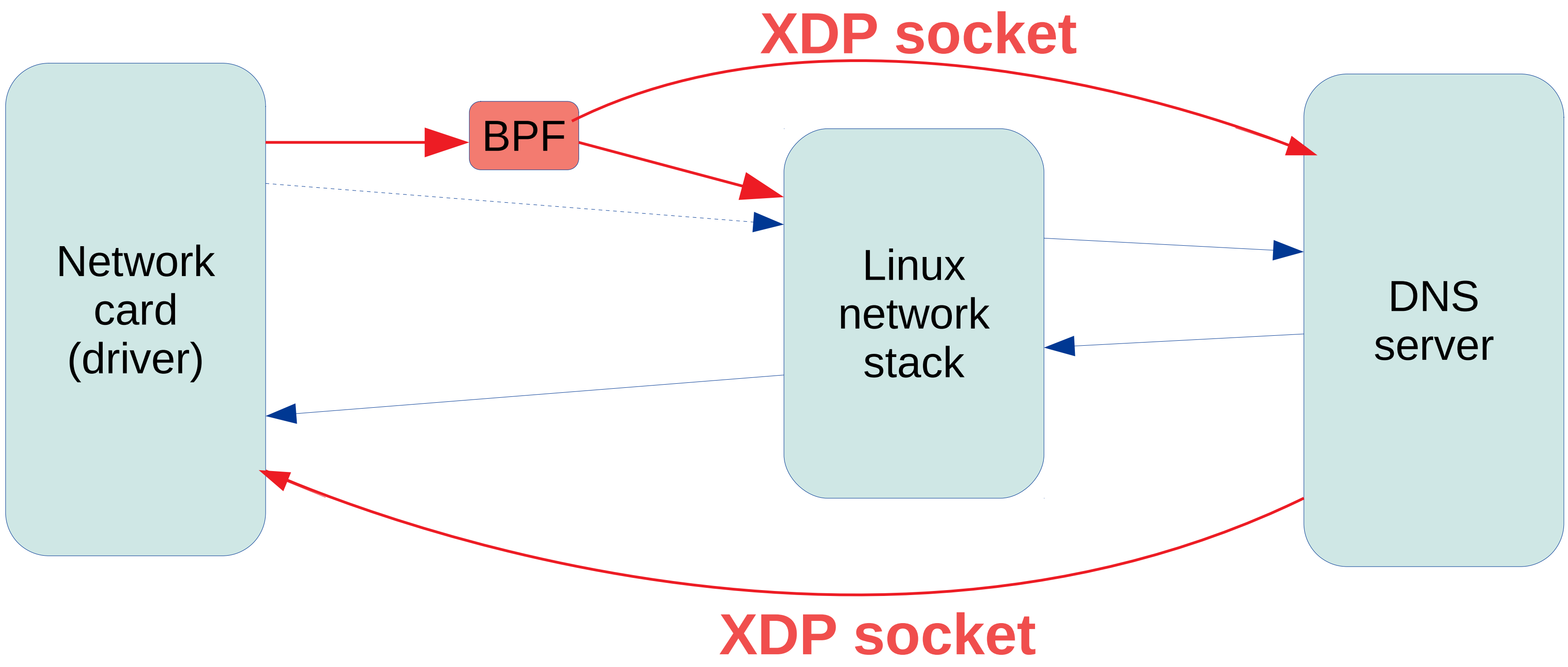


+kernel CPU time!

DNS routines

network I/O

Authoritative DNS server profile



BPF (Berkeley Packet Filter)

- Originally a firewall implementation
- “BPF program” instead of rules
  - written in C
  - compiled by Clang
  - verified by kernel upon load
  - limitations (size, no loops, …)

BPF (Berkeley Packet Filter)

- BPF program decides packet fate:
  - drop
  - hand-over to XDP socket
    - (DNS over UDP traffic)
  - pass to Linux stack
    - (TCP, other port, IPv6 extensions, IPSec, etc.)

XDP (eXpress Data Path)

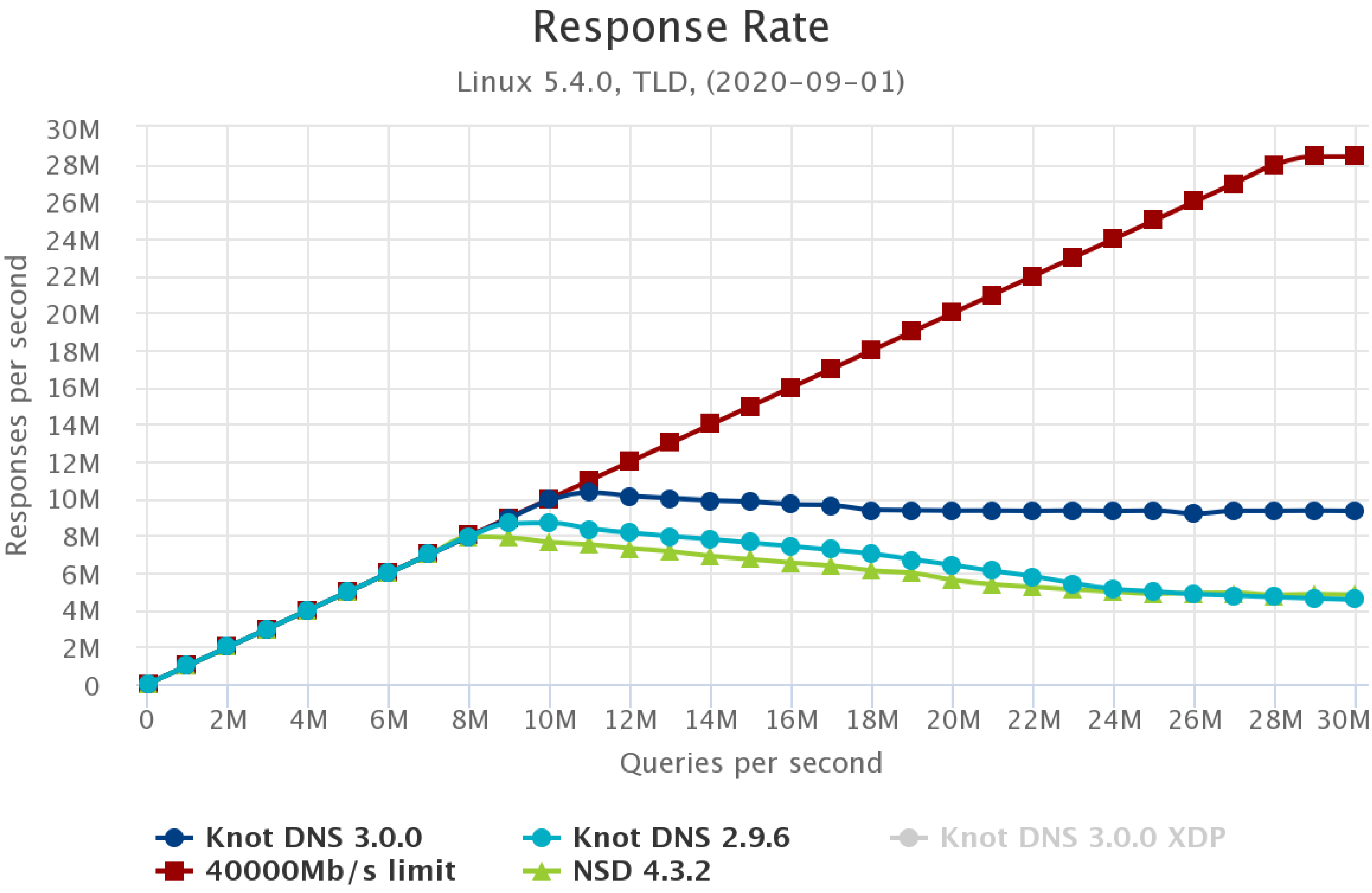
- Ethernet frames directly to userspace
- And back
- Zero-copy
- Need custom parsing (Ethernet + IP + UDP)
- Shared UMEM, care about buffer allocation

```
listen-xdp: ens2f0@53
```

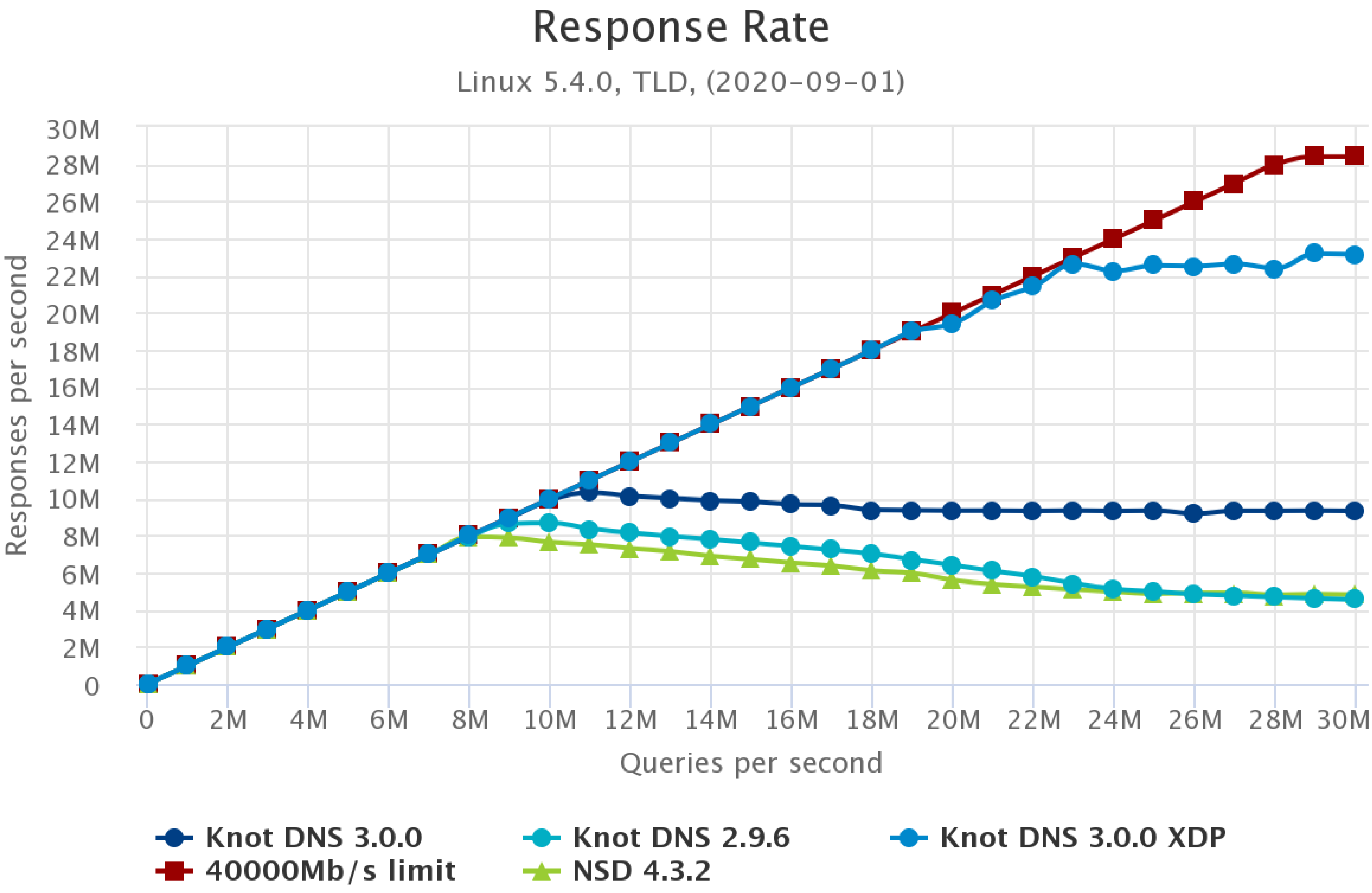
Requirements

- Linux kernel 4.18+ (5.x recommended)
- XDP-compatible network card to achieve speed-up
- CAP_SYS_ADMIN during server startup

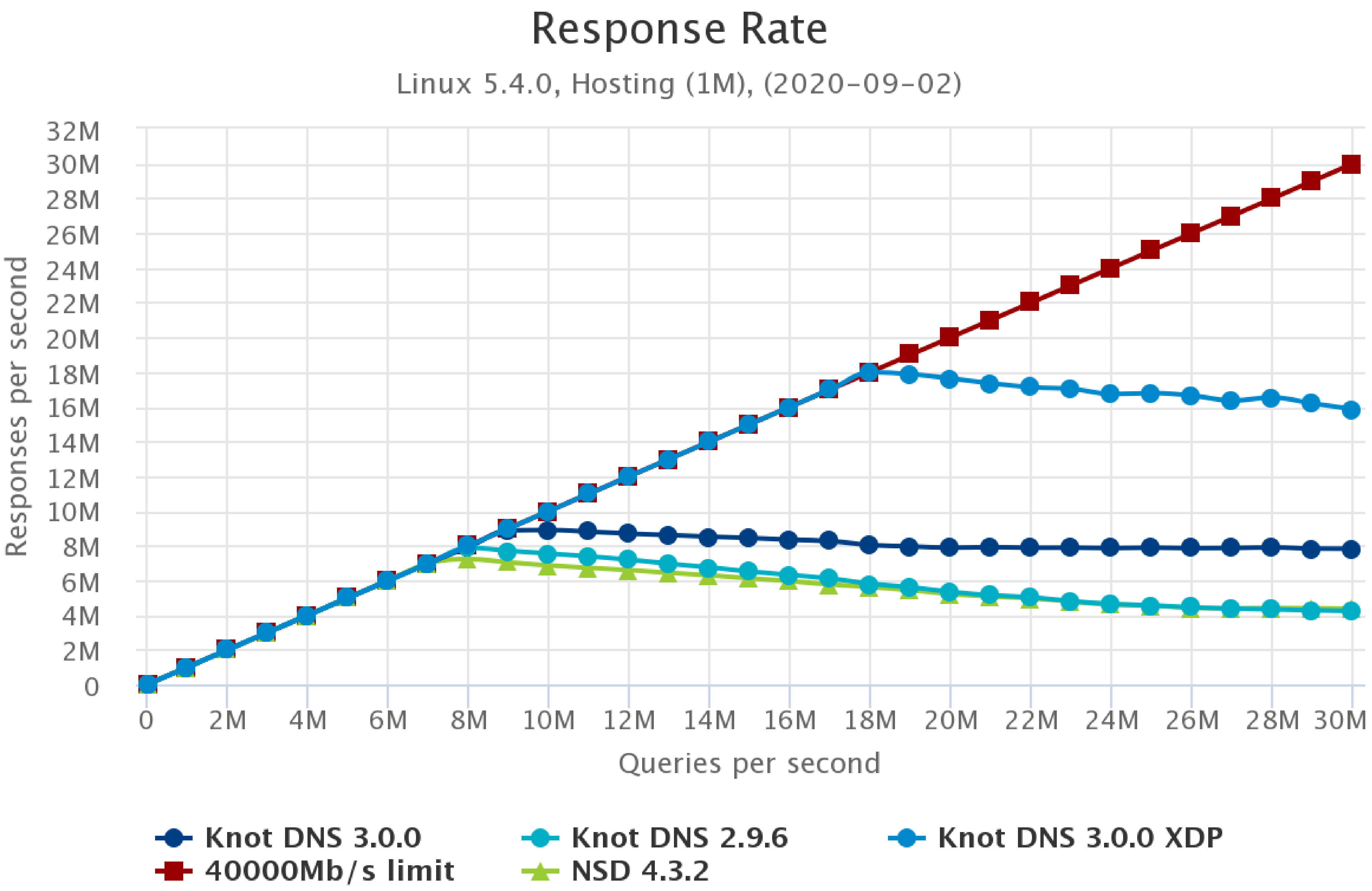
Knot 3.0.0 performance – TLD zone



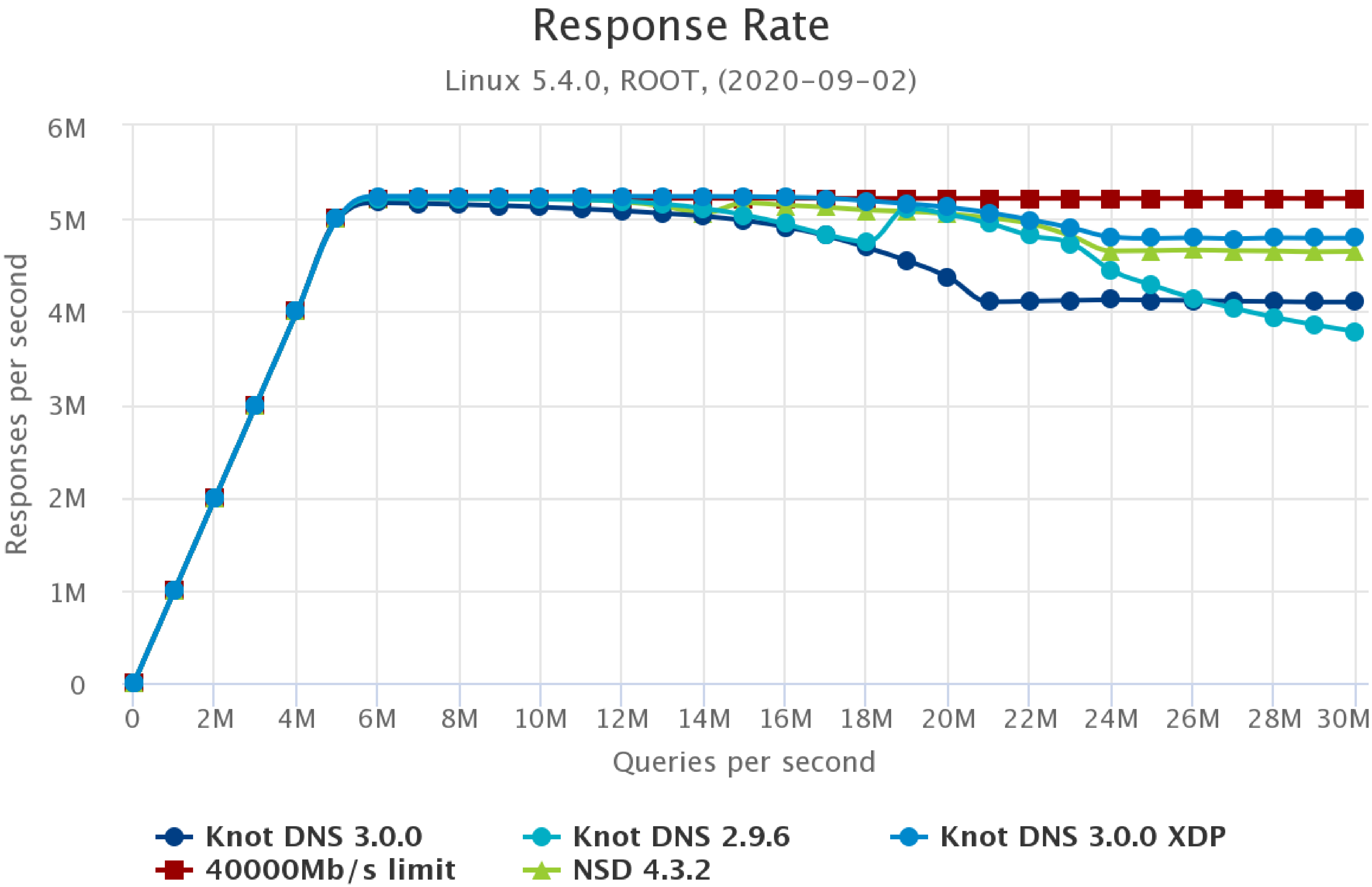
Knot 3.0.0 with XDP – TLD zone



Knot 3.0.0 with XDP – hosting

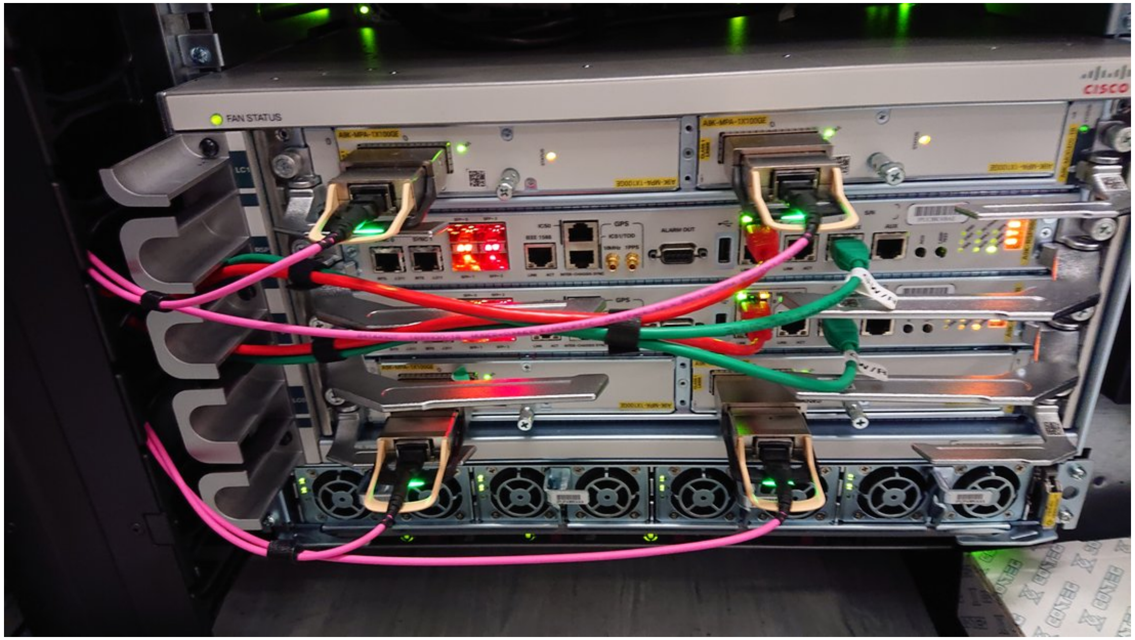


Knot 3.0.0 with XDP – root



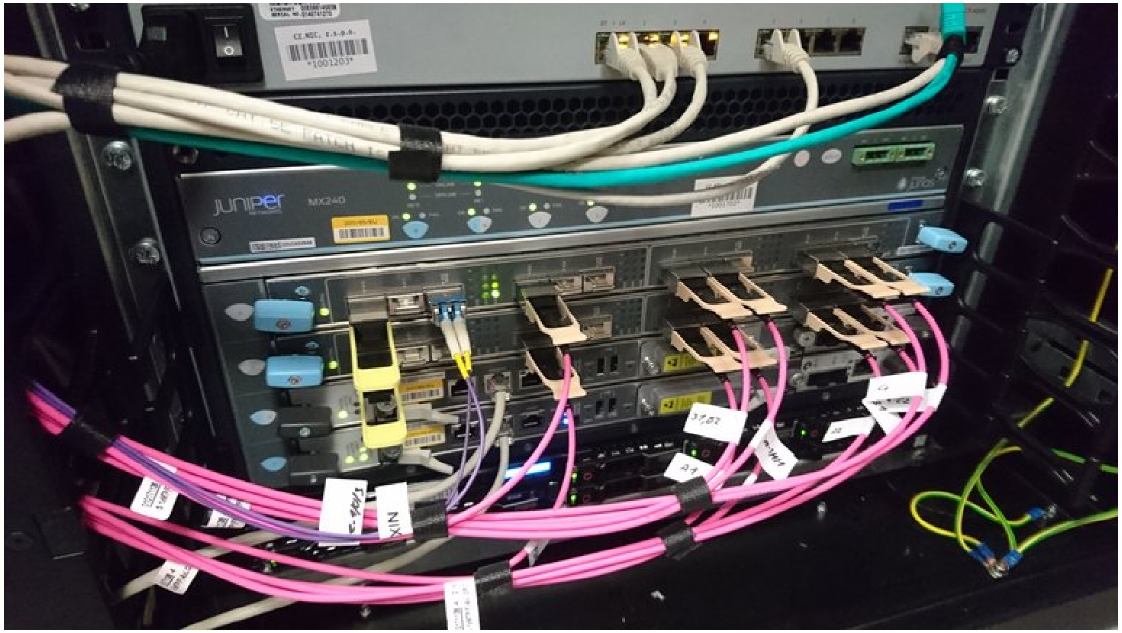
Impact

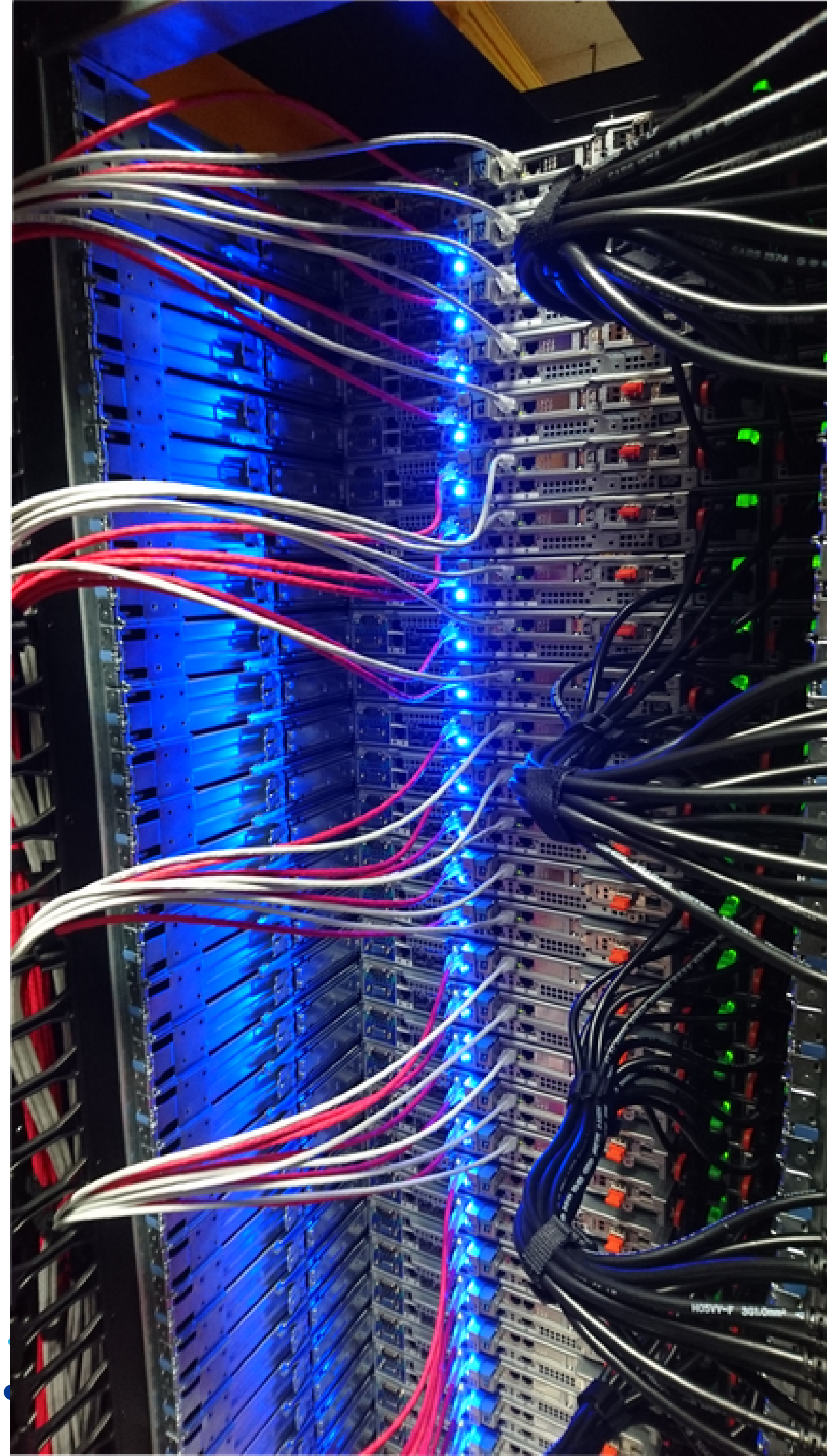
- 100Gbps stack - .cz +
- Now - 30 servers



Impact

- Another 100Gbps stack
- Now - 30 servers





Impact

- More than triple performance increase compared to 2.9.6 (and even more compared to older version)
- Currently 30 servers in our 100 Gbps stack
- Less than 10 servers in the future
- Significant cost reduction (!)
  - servers, rack space, power
- Installed in first location (in CZ) – no issues

XDP limitations

- Kernel IP stack bypassed
  - Routing decision
  - Statistics
  - Tcpdump
  - Filtering

- Workarounds

Thanks to

- Daniel Salzman
- Libor Peltan
- Zdeněk Brůna

Thank You!

KNOT DNS

Ondřej Filip • ondrej.filip@nic.cz • https://www.nic.cz • https://www.knot-dns.cz

CZ.NIC | CZ DOMAIN REGISTRY