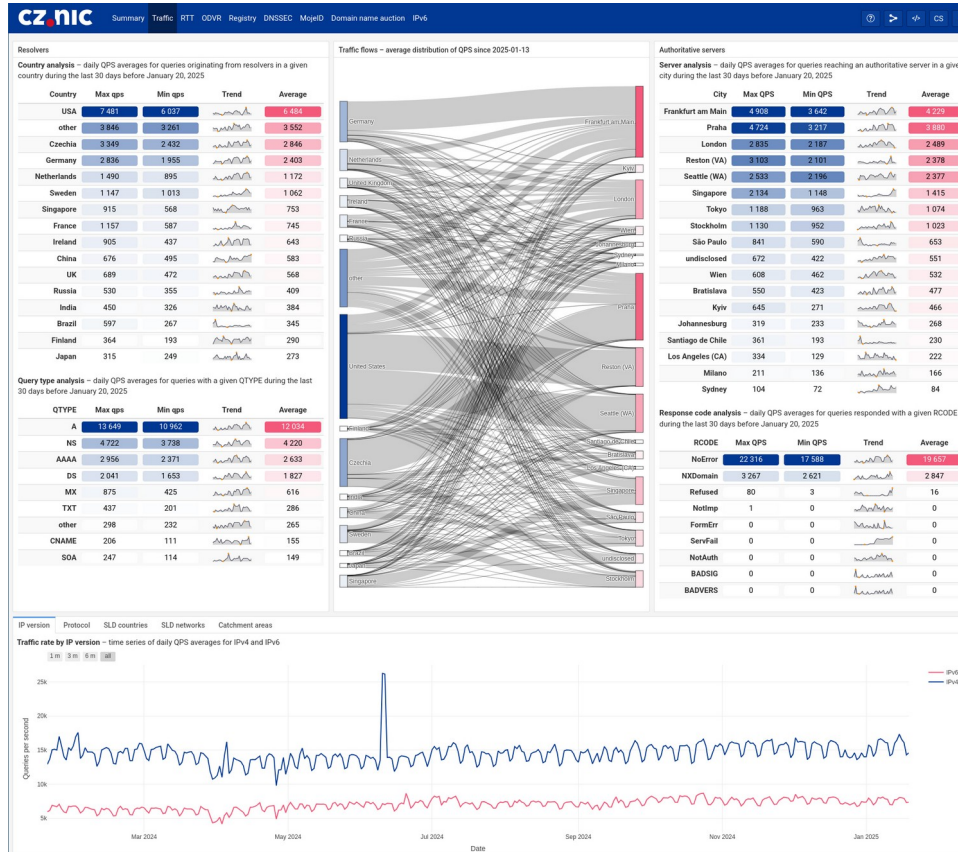


# Nasazení XDP v Knot DNS

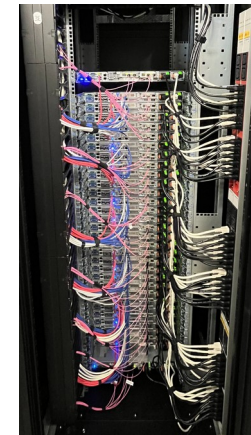
Lukáš Vacek • 21. 01. 2025

# DNS provoz

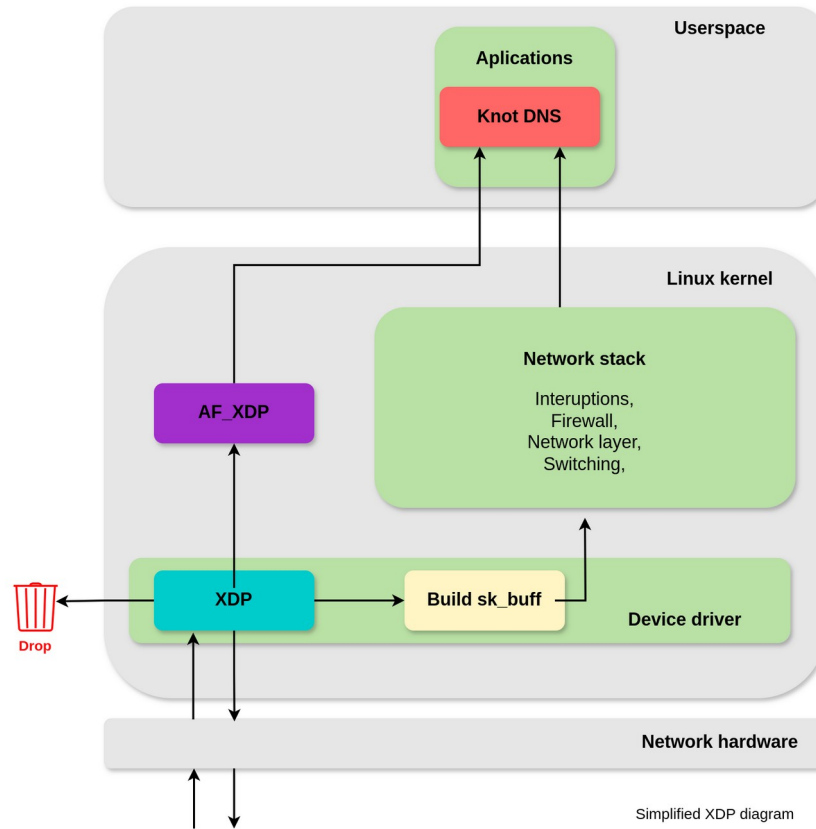


# Škálování výkonu DNS anycastu

- Přidávání nových uzlů (lokality)
- Navýšení kapacity linek v lokalitě
- Přidávání nových serverů v lokalitě
- Výkonější DNS démon



# Technologie XDP



# Prerekvizity HW a OS

- Jádro od verze. 4.18+ (doporučené 5.x+)
  
- Síťové karty (multiqueue)
  - Nvidia (Mellanox) ConnectX-6 Dx (driver mlx5\_core)
  - Intel series 700 (driver i40e)
  - Intel series E810 (driver ice)

# Nastavení Linuxu a optimalizace síťovky

- Systemd

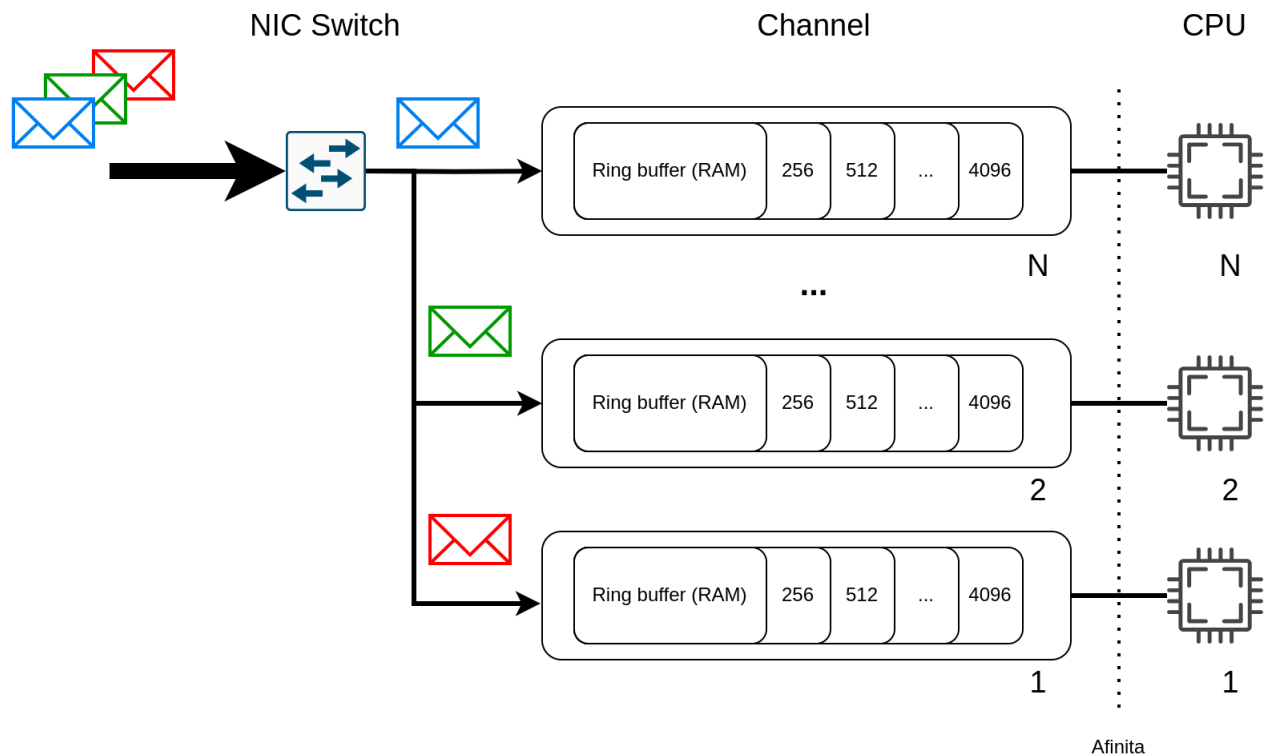
```
[Service]
CapabilityBoundingSet=CAP_NET_RAW CAP_NET_ADMIN CAP_SYS_ADMIN CAP_IPC_LOCK CAP_SYS_RESOURCE
AmbientCapabilities=CAP_NET_RAW CAP_NET_ADMIN CAP_SYS_ADMIN CAP_IPC_LOCK CAP_SYS_RESOURCE
```

- Konfigurace síťových karet

- Počet kanálů (channel) [#CPU nebo #CPU - 1]
- Velikost paměti (bufferu) [RX 256, TX 1024]
- Přerušení (IRQ)
- Afinity CPU

```
root@testbed-a-02:/home/lvacek#> ethtool -l enp196s0f1
Channel parameters for enp196s0f1:
Pre-set maximums:
RX:                n/a
TX:                n/a
Other:             1
Combined:          64
Current hardware settings:
RX:                n/a
TX:                n/a
Other:             1
Combined:          48
```

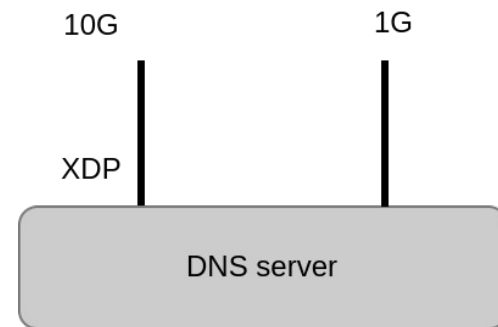
# Parametry síťovky



# Konfigurace Knota

- Module XDP
- Route-check
- Restart Knota

```
1 server:  
2   ...  
3  
4   listen: <ip1>  
5   listen: <ip2>  
6   # ...  
7   listen: <ipN>  
8  
9 xdp:  
10  listen: <interface name>  
11  route-check: <on/off>  
12
```





# Funguje to? Nevidím vůbec síťový provoz

- Nejsou aplikovány ACL pravidla pro část DNS provozu
- Tcpdump téměř mlčí
- Sběr provozu taky nefunguje jako před tím

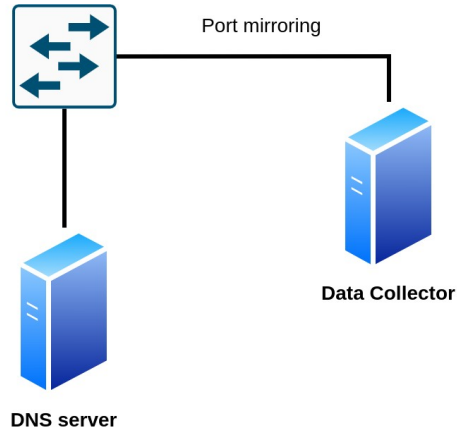


# ACL pravidla

- DNS provoz majoritně UDP → XDP
- Ostatní provoz jako předtím odbaven Linuxovým kernelem
- Lepší upočítat Knotem než nechat procházet linuxovým jádrem

# TCPDUMP a monitoring provozu

- Port mirroring



- `$> knotc -f stats`

- `$> ip link set dev <interface> xdp off`

```
9: vlan1301@enp196s0f1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 xdpgeneric/id:99 qdisc noqueue state UP group default qlen 1000
  link/ether 6c:fe:54:16:41:21 brd ff:ff:ff:ff:ff:ff
  inet 192.168.99.40/27 brd 192.168.99.63 scope global vlan1301
    valid_lft forever preferred_lft forever
  inet6 fd99:0:0:1301::40/64 scope global
    valid_lft forever preferred_lft forever
  inet6 fe80::6efe:54ff:fe16:4121/64 scope link
    valid_lft forever preferred_lft forever
```

```
server.zone-count = 397
xdp.rx_dropped[ix] = 0
xdp.rx_invalid[ix] = 0
xdp.tx_invalid[ix] = 0
xdp.rx_full[ix] = 0
xdp.fq_empty[ix] = 0
xdp.tx_empty[ix] = 0
xdp.tx_busy[ix] = 1
xdp.fq_fill[ix] = 1823
xdp.rx_fill[ix] = 0
xdp.tx_fill[ix] = 0
xdp.cq_fill[ix] = 1
mod-stats.request-protocol[udp4] = 344760391
mod-stats.request-protocol[tcp4] = 96784439
mod-stats.request-protocol[quic4] = 0
mod-stats.request-protocol[tls4] = 0
mod-stats.request-protocol[udp6] = 179379222
mod-stats.request-protocol[tcp6] = 94790613
mod-stats.request-protocol[quic6] = 0
mod-stats.request-protocol[tls6] = 0
mod-stats.request-protocol[udp4-xdp] = 947473225
mod-stats.request-protocol[tcp4-xdp] = 0
mod-stats.request-protocol[quic4-xdp] = 0
mod-stats.request-protocol[udp6-xdp] = 856310458
mod-stats.request-protocol[tcp6-xdp] = 0
mod-stats.request-protocol[quic6-xdp] = 0
mod-stats.server-operation[query] = 2519482052
mod-stats.server-operation[update] = 5
mod-stats.server-operation[notify] = 8742
mod-stats.server-operation[axfr] = 6546
mod-stats.server-operation[ixfr] = 380
mod-stats.server-operation[invalid] = 623
mod-stats.request-bytes[query] = 120086867957
mod-stats.request-bytes[update] = 2560
mod-stats.request-bytes[other] = 1912799
mod-stats.response-bytes[reply] = 1210826498386
mod-stats.response-bytes[transfer] = 229842
mod-stats.response-bytes[other] = 1090730
```

# Sběr dat

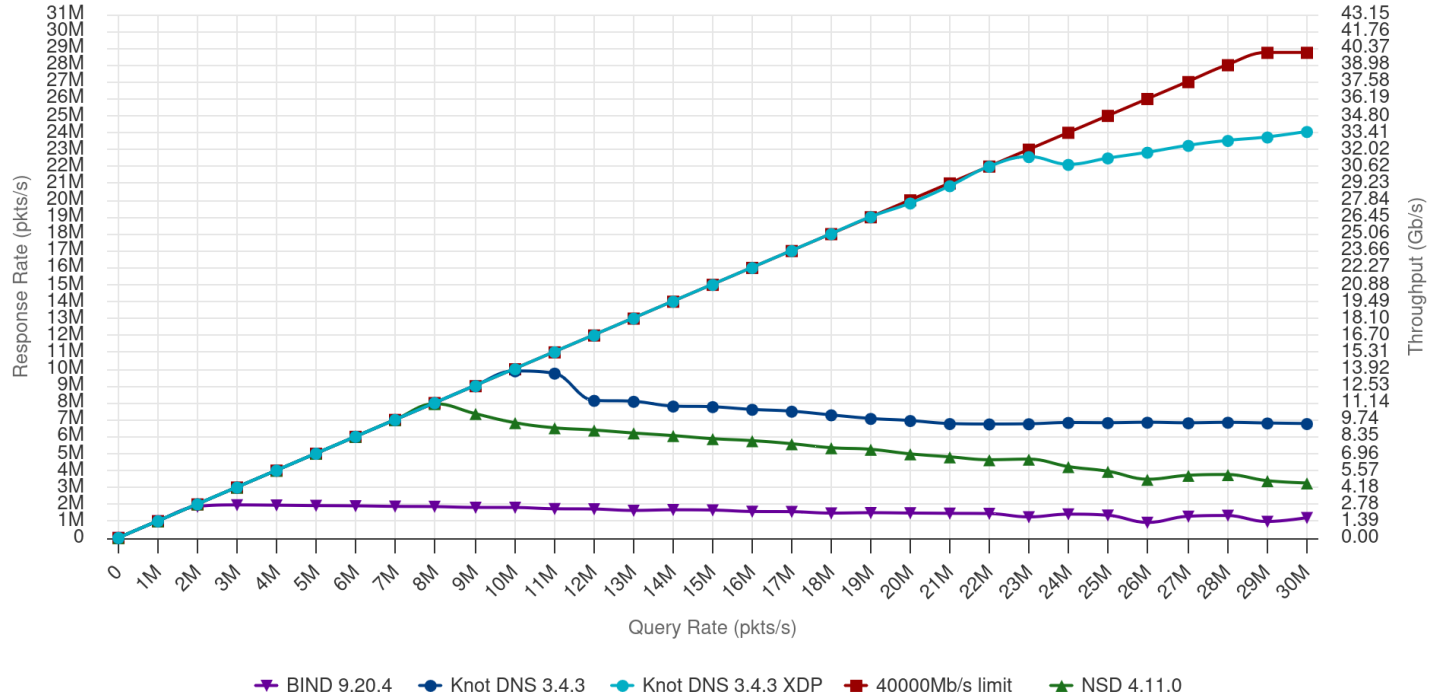
- Modul mod-probe
- Odesílání DNS provozu do UNIX socketu
- DNS Probe (Projekt ADAM)



# Benchmark 40G

## Response Rate

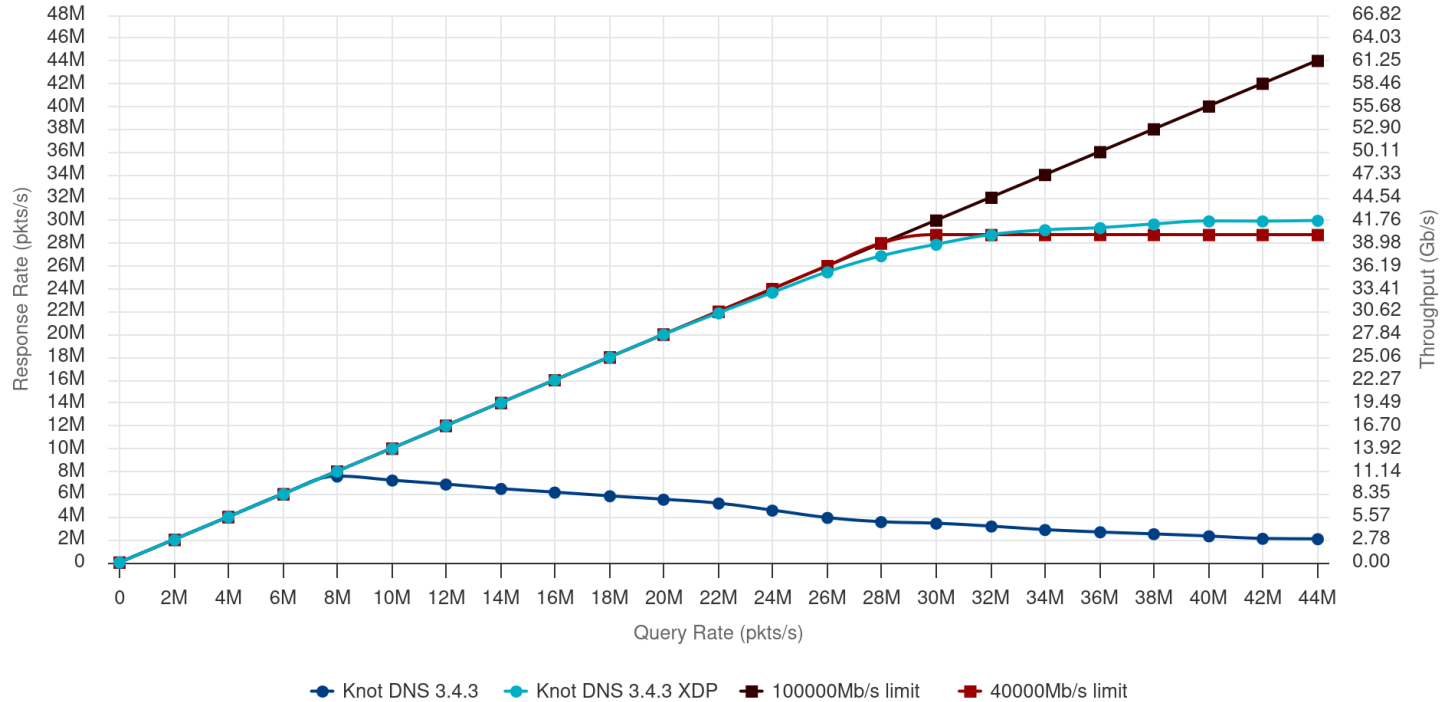
Linux 6.12.3, TLD, Intel XL710, (2025-01-06)



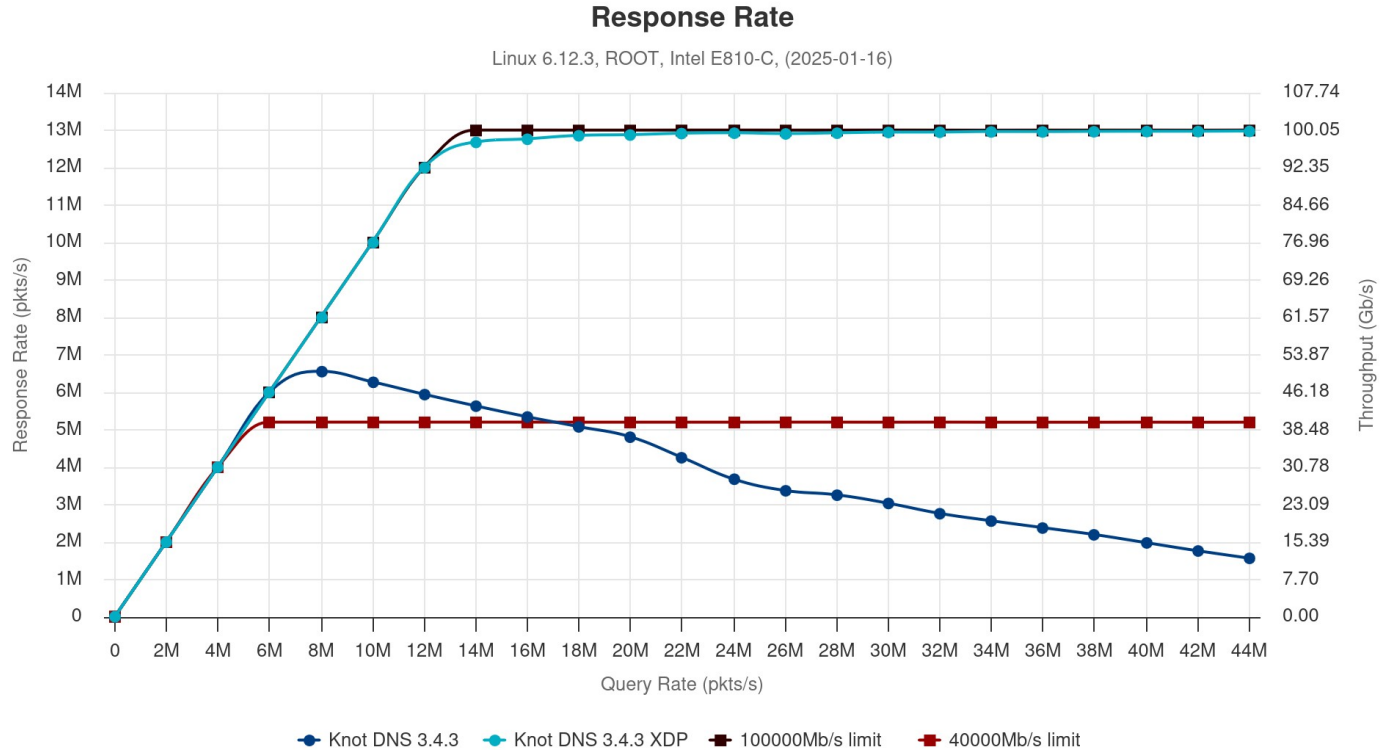
# Benchmark 100G

## Response Rate

Linux 6.12.3, TLD, Intel E810-C, (2025-01-16)



# Benchmark 100G – ROOT zone



# Děkuji vám za pozornost!

Lukáš Vacek



# Zdroje

- Mapa: <https://mapy.cz/>
- XDP diagramy:  
<https://github.com/tohojo/xdp-paper/blob/master/xdp-presentation.org>
- Konfigurace síťovky: <https://gitlab.nic.cz/knot/dns-benchmarking>
- Projekt ADAM: <https://adam.nic.cz/cs/>
- Knot DNS: <https://www.knot-dns.cz/>
- Obrázek na slidu 9 a 12 vytvořen ChatGPT