Ze života Knot DNS

Daniel Salzman • daniel.salzman@nic.cz • 10. 11. 2021

Knot DNS – stručný úvod

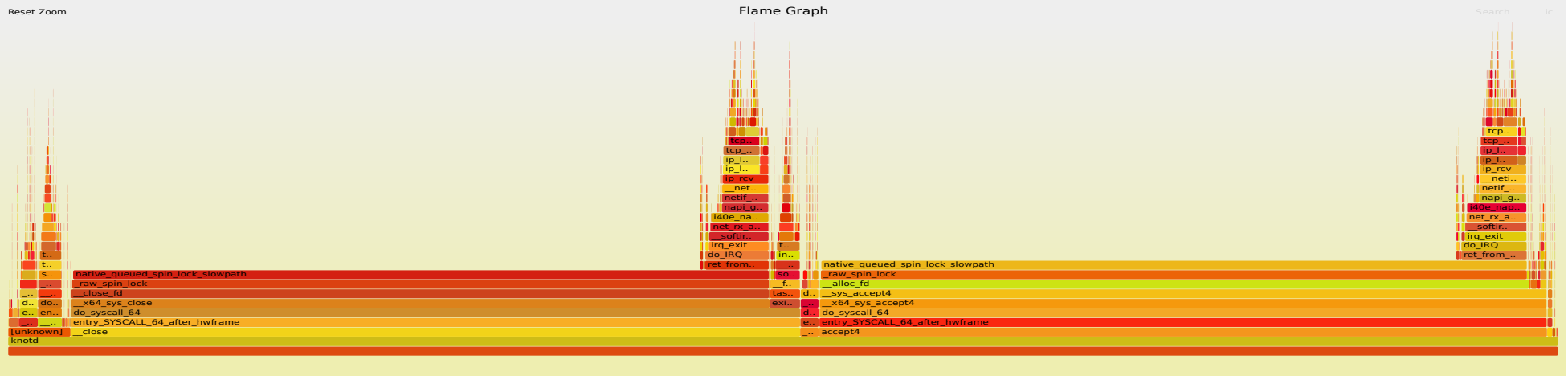
- Autoritativní DNS server
  - Primární nebo sekundární server
  - Vysoký výkon zpracování DNS dotazů
- Pokročilá podpora DNSSEC
  - Jednoduché použití
  - Výkonné podepisování
- A mnoho dalšího
  - www.knot-dns.cz

Výkon zpracování DNS dotazů

- Výkon nelze charakterizovat jedním číslem, protože závisí na spoustě faktorů
  - Velikost a struktura zón (TLD vs ROOT vs. jednoduchá zóna)
  - Počet zón (TLD vs. DNS hosting)
  - Zabezpeční DNSSEC (algoritmus DNSKEY, NSEC vs. NSEC3)
  - Typ DNS dotazu (pozitivní vs. negativní odpověď, DNSSEC)
  - Konfigurace démona (dodatečné moduly, počet adres,...)
- Stále narážíme na další scénáře nasazení, pro které je třeba optimalizovat implementaci
- Moderní hardware nabízí nové možnosti ale i výzvy
  - CPU se spoustou jader
  - Síťové karty s vysokou propustností
- Operační systém
  - Možnosti a nastavení síťového stacku (SO_REUSEPORT, recvmmsg/sendmmsg,…)
  - Přibývající bezpečnostní opravy v OS často zhoršují výkon aplikací

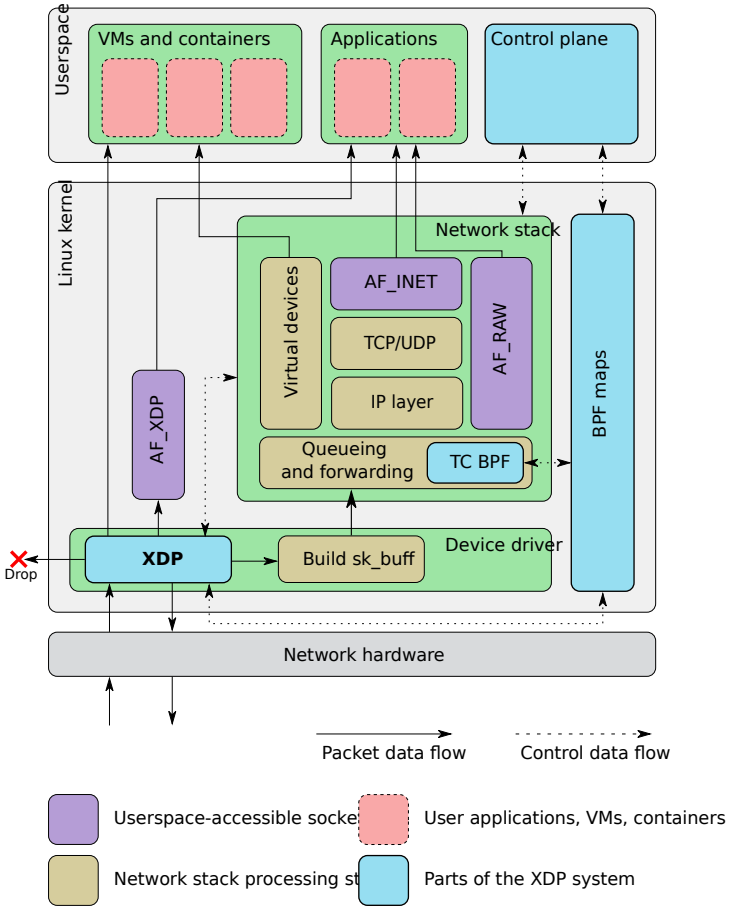
Problémy konvenčního síťového stacku

- Mnoho složitých vrstev zpracování
- Kopírování paketů mezi kernelem a aplikací
- Sdílená tabulka deskriptorů TCP spojení mezi vlákny procesu



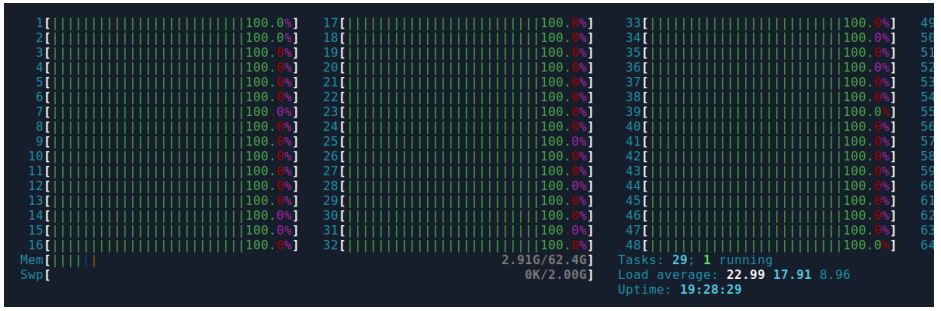
Režim XDP – popis

- Podpora pro UDP v Knot DNS 3.0 (pro TCP v 3.1)
- Klasifikace paketů v XDP (eBPF program)
  - Nativní režim – vyžaduje podporu ovladače
  - Emulovaný režim – nízký výkon, nevhodné do produkce
- Využití socketů AF_XDP (Linux 4.18+)
  - Předání paketů z XDP přímo do paměti aplikace v user space (Zero-Copy od Linux 5.0+)
- Nevyžaduje úpravu síťového ovladače
- Obcházení síťového stacku OS pro vybrané pakety
- Extra oprávnění procesu jsou třeba pouze při startu
- Ostatní provoz není dotčen a je zpracován normálně

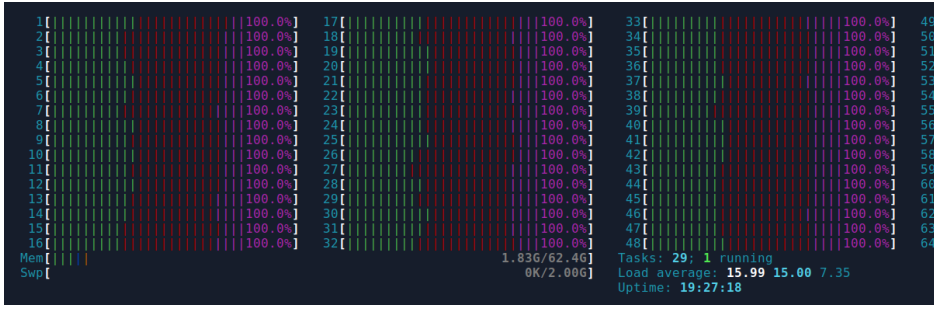


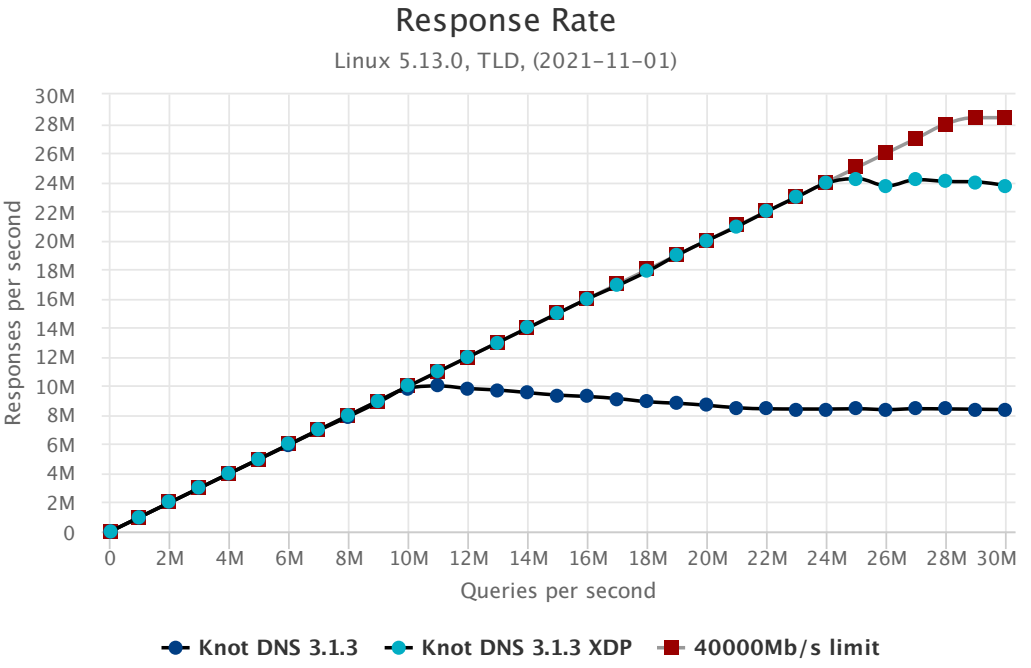
Výkon zpracování DNS dotazů po UDP

- htop: 64 cores, 30M qps, XDP



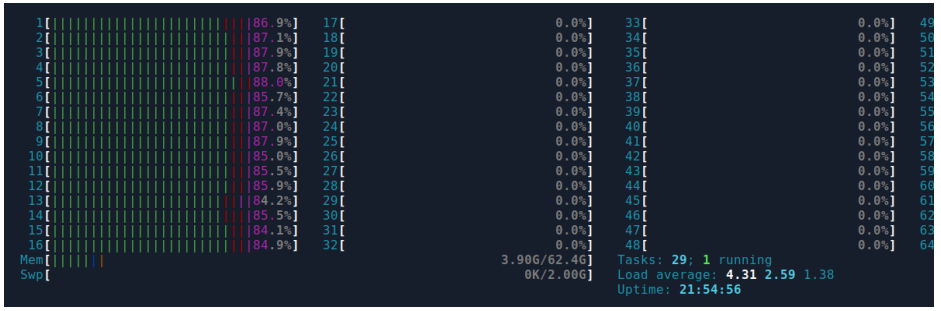
- htop: 64 cores, 30M qps



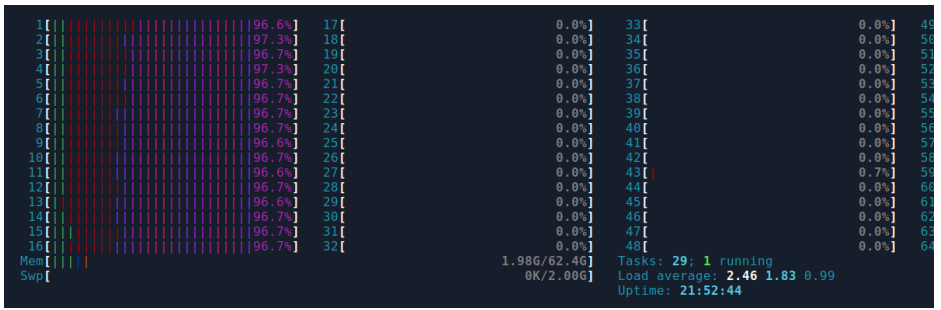


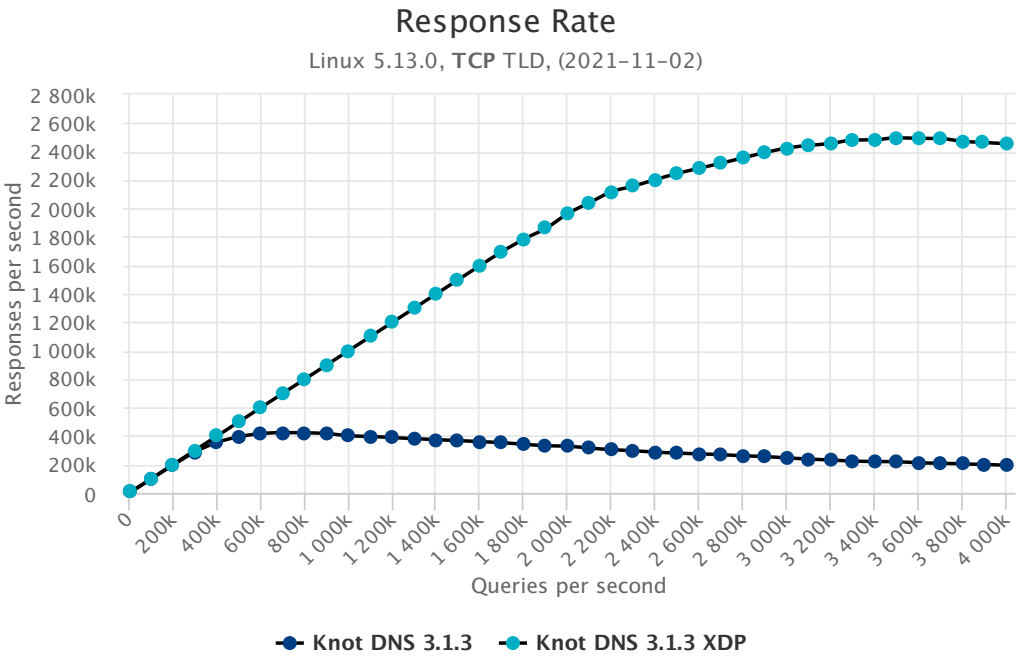
Výkon zpracování DNS dotazů po TCP

- htop: 16 cores, 4M qps (7Gbps), XDP



- htop: 16 cores, 4M qps





Měření výkonu zpracování DNS dotazů

- Dříve využívané nástroje již nevyhovují a špatně se používají
  - tcpreplay v režimu Netmap (UDP)
  - dpdk-tcp-generator (TCP)
- Absence vhodných alternativních nástrojů
- Potřeba vytvoření nového nástroje – **kxdpgun**
  - Součást Knot DNS – využití implementace režimu XDP
  - Vícevláknový generátor DNS provozu (UDP nebo TCP)
  - Vstupem je jednoduchý textový soubor s popisem dotazů (EDNS0, DO bit)
  - Na běžném serveru je možné generovat desítky miliónů dotazů za sekundu

Monitoring DNS provozu v režimu XDP

- Běžné nástroje (tcpdump, wireshark,...) nelze použít
- Modul statistik DNS provozu (**mod-stats**)
  - Rozšíření o nové čítače související s XDP
  - Přepracování modulu kvůli nežádoucí režii při vysokých rychlostech na procesorech s mnoha jádry
- Nový modul pro zjednodušený export DNS provozu (**mod-probe**)
  - Minimalistická a efektivní implementace bez zbytečných transformací
  - Předávání struktury parametrů (bez záznamů odpovědi) pro dvojici dotaz-odpověď
  - Možnost použití více paralelních instancí pro lepší rozložení zátěže
  - Možnost vzorkování od určité rychlosti
  - Rozhraní v C (integrace do dns-probe projektu ADAM) nebo Pythonu (jednoduché použití)

Režim XDP a routování

- Základní režim XDP je založen na symetrickém routování
  - Vstupní rozhraní dotazu = výstupní rozhraní odpovědi
  - Zdrojová MAC adresa dotazu = cílová MAC adresa dotazu
  - Pro správný chod vyžaduje předřazený router
- Knot DNS 3.1 přidává podporu „route-check“
  - Při zpracování dotazu v XDP se vyčte z OS routovací informace pro odpověď
  - Routovací tabulky může spravovat nezávislý démon (BIRD)
  - Nastavení správné cílové MAC adresy odpovědi
  - Pokud se výstupní rozhraní liší od vstupního, zpracuje se DNS dotaz konvenčně
  - Dotaz se zahodí pokud je routovací pravidlo blackhole/unreachable/prohibit

Plány do blízké budoucnosti

- Dokončení implementace DNS-over-TCP v režimu XDP
- Experimenty s DNS-over-QUIC v režimu XDP
- Testování se 100GbE síťovou kartou

Děkuji za pozornost

Daniel Salzman • daniel.salzman@nic.cz • knot-dns.cz