Novinky v Knot DNS

• 4. 10. 2025 – Daniel Salzman

CZ.NIC | SPRÁVCE DOMÉNY CZ

O projektu

- Autoritativní DNS server
- Důraz na výkon
  - Odpovídání DNS dotazů
  - Zpracování mnoha nebo velkých zón
- Vyspělá podpora DNSSEC
- Pokročilé funkce
- Nástroje pro práci s DNS
- www.knot-dns.cz

Aktuálně podporované verze

- Knot DNS 3.5 (3.5.0)
  - Aktuální stabilní verze
  - Licence GPL-2+
  - EOL Q3 2027
- Knot DNS 3.4 (3.4.8)
  - Předchozí stabilní verze
  - Licence GPL-3+
  - EOL Q3 2026

Databázový backend – úvod

- Schopnost ukládat nebo načítat zónová data z databáze
  - Načítání a ukládání se konfiguruje nezávisle
- Nemá vliv na výkon zpracování DNS dotazů
  - Živá zóna je stále uložena a aktualizována v paměti Knot DNS serveru
- Možnost ukládát více variant jedné zóny – **instance**
  - Vhodné například pro oddělení podepsané a nepodepsané zóny
- Databáze může být umístěna odděleně od Knot DNS serveru
  - Možnost zabezpečení pomocí TLS včetně oboustranné autentizace
  - Přístup více Knot DNS serverů k jedné databázi

Databázový backend – požadavky

- Databáze Redis 7+ nebo Valkey
  - Možnost volby mezi perzistencí (primární databáze) a výkonem
  - Rozšiřitelnost funcionality pomocí modulů
  - Přirozená nabídka efektivní fronty událostí – Streams
- Načtený modul **knot** v databázi
  - Rozšiřuje databázi o příkazy a struktury pro práci se zónovými daty
  - Umožňuje správu zón přec CLI nebo API bez potřeby Knot DNS
  - Dostupný minimálně z projektových repozitářů pkg.labs.nic.cz

Databázový backend – konfigurace

- Lokální databáze

```
database:
  zone-db-listen: /run/redis/redis-server.sock
```

Vzdálená databáze přes TLS

```
database:
  zone-db-listen: 192.168.1.2
  zone-db-tls: on
  zone-db-cert-hostname: db.example.com
```

- Zóna načítána z instance 1, podepisována a ukládána jako instance 2

```
zone:
  - domain: example.com.
    dnssec-signing: on
    zone-db-input: 1
    zone-db-output: 2
```

Databázový backend – CLI – plnění zóny

```
$ redis-cli KNOT.ZONE.BEGIN example.com 1
(integer) 10

$ redis-cli KNOT.ZONE.STORE example.com 10 "@ SOA ns admin 1 86400 900 691200 3600"
OK
$ redis-cli KNOT.ZONE.STORE example.com 10 "@ NS ns"
OK
$ redis-cli KNOT.ZONE.STORE example.com 10 "ns 300 AAAA ::1"
OK

$ redis-cli KNOT.ZONE.COMMIT example.com 10
OK

$ redis-cli KNOT.ZONE.LOAD --compact example.com 1
1) "example.com. 600 SOA ns.example.com. admin.example.com. 1 86400 900 691200 3600"
2) "example.com. 600 NS ns.example.com."
3) "ns.example.com. 300 AAAA ::1"

$ redis-cli KNOT.ZONE.LIST --instances
1) "example.com.: 1"
```

Databázový backend – CLI – aktualizace zóny

```
$ redis-cli KNOT.UPD.BEGIN example.com 1
(integer) 10

$ redis-cli KNOT.UPD.ADD example.com 10 "test TXT \"Knot DNS\""
OK

$ redis-cli KNOT.UPD.DIFF example.com 10
1) 1) (empty array)
   2) 1) 1) "test.example.com."
         2) "600"
         3) "TXT"
         4) "\"Knot DNS\""

$ redis-cli KNOT.UPD.COMMIT example.com 10
OK

$ redis-cli KNOT.UPD.LOAD --compact example.com 1 1
1) 1) 1) 1) "example.com. 600 SOA ns.example.com. admin.example.com. 1 86400 900 691200 3600"
      2) 1) "example.com. 600 SOA ns.example.com. admin.example.com. 2 86400 900 691200 3600"
   2) 1) (empty array)
      2) 1) "test.example.com. 600 TXT \"Knot DNS\""
```

Externí validace zóny

- Možnost externího schvalování změn v zóně
- Lze kombinovat s **dnssec-validation** či **zonemd-verify**
- Server pracuje se současnou zónou, dokud se změna nepotvrdí nebo nedojde na **timeout**
- Čekání na validaci lze zjistit
  - **knotc zone-status +transaction <zone>** – hodnota **open-external**
  - D-Bus signalizace **dbus-event** – signál **external-verify**
- Textový výpis obsahu zóny může být jak v plné, tak i rozdílové podobě
- Publikaci aktualizované zóny je nutné potvrdit
  - **knotc zone-commit <zone>** – potvrzení změn
  - **knotc zone-abort <zone>** – odmítnutí změn

Externí validace zóny – příklad soubor

```
external:
  - id: all
    timeout: 60
    dump-new-zone: /tmp/%s.full
    dump-removals: /tmp/%s.diff
    dump-additions: /tmp/%s.diff

zone:
  - domain: example.com.
    external-validation: all
```

```
$ cat /tmp/example.com.diff
;; Removed records
example.com.            3600    SOA     dns1.example.com. hostmaster.example.com. 2010111201 10800 3600 1209600 7200
;; Added records
example.com.            3600    SOA     dns1.example.com. hostmaster.example.com. 2010111202 10800 3600 1209600 7200
horse.example.com.      3600    AAAA    1::1
tiger.example.com.      3600    AAAA    1::1
```

Externí validace zóny – příklad CLI

```
$ knotc zone-status +transaction example.com
[example.com.] transaction: open-external
```

```
$ knotc zone-diff example.com
[example.com.] -example.com. 3600 SOA dns1.example.com. hostmaster.example.com. 2010111201 10800 3600 1209600 7200
[example.com.] +example.com. 3600 SOA dns1.example.com. hostmaster.example.com. 2010111202 10800 3600 1209600 7200
[example.com.] +horse.example.com. 3600 AAAA 1::1
[example.com.] +tiger.example.com. 3600 AAAA 1::1
```

Optimalizace

- Efektivnější práce s mnoha zónami (100k+)
  - Dynamické přidávání nebo odebírání přes ovládací rozhraní
  - Zpracování změn v katalagové zóně
- Dřívější dostupnost ovládacího rozhraní po startu serveru
  - Dostupné bezprostředně po načtení konfigurace
- Podpora více ovládacích rozhraní
  - Možnost oddělení správy serveru od monitoringu
- Schopnost zpracovávat některé ovládací příkazy paralelně
  - Dynamické úpravy konfigurace vs. většina ostatních příkazů

DNS po TLS

- Podpora komunikace DNS po TLS 1.3 (DoT)
- Včetně XFR (XoT) v obou směrech (server vs. klient)
- Konfigurace obdobná jako u QUIC
- Přidána možnost autentizace protistrany ověřováním hostname v certifikátu (platí i pro QUIC)

DNS po TLS – ukázka konfigurace

Primární server

```
server:
    listen-tls: ::1
    ca-file: ca-cert.pem
    cert-file: primary-cert.pem
    key-file: primary-key.pem
    automatic-acl: on

remote:
  - id: secondary
    address: ::2
    tls: on
    cert-hostname: secondary.example.com

zone:
  - domain: example.com
    notify: secondary
```

Sekundární server

```
server:
    listen-tls: ::2
    ca-file: ca-cert.pem
    cert-file: secondary-cert.pem
    key-file: secondary-key.pem
    automatic-acl: on

remote:
  - id: primary
    address: ::1
    tls: on
    cert-hostname: primary.example.com

zone:
  - domain: example.com
    master: primary
```

Modul RRL – optimalizace

- Ochrana DNS klientů proti reflexním (amplifikačním) útokům po **UDP**
- Dříve
  - **R**esponse **R**ate **L**imiting – omezování rychlosti **odpovídání** na dotazy ze stejných síťových prefixů /24 (IPv4); /56 (IPv6)
  - Neefektivní implementace při vysokém provozu a hodně vláknech (XDP)
- Nově
  - Omezování rychlosti **dotazování** ze stejných síťových adres či prefixů /32, /24, /20, /18 (IPv4); /128, /64, /56, /48, /32 (IPv6)
  - Vysoce optimalizovaná implementace s využitím struktury KRU
- Parametr **rate-limit** – maximální počet dotazů za sekundu (prefixy s různými váhami)
- Parametr **instant-limit** – maximální počet dotazů v jednom časovém okamžiku
- Parametr **slip** – chování omezování (zahodit nebo odpovědět TC každý N-tý dotaz)

Modul RRL – nová funkce

- Částečná ochrana DNS serveru proti nadměrnému vytížení po **TCP**, **QUIC** a **TLS** ze stejných síťových adres či prefixů na základě spotřebovaného času
- Stejné prefixy a váhy jako u UDP
- Aplikuje se i na handshake
- Parametr **time-rate-limit** – maximální souhrnný počet mikrosekund v sekundovém okně
- Parametr **time-instant-limit** – maximální jednorázový počet mikrosekund

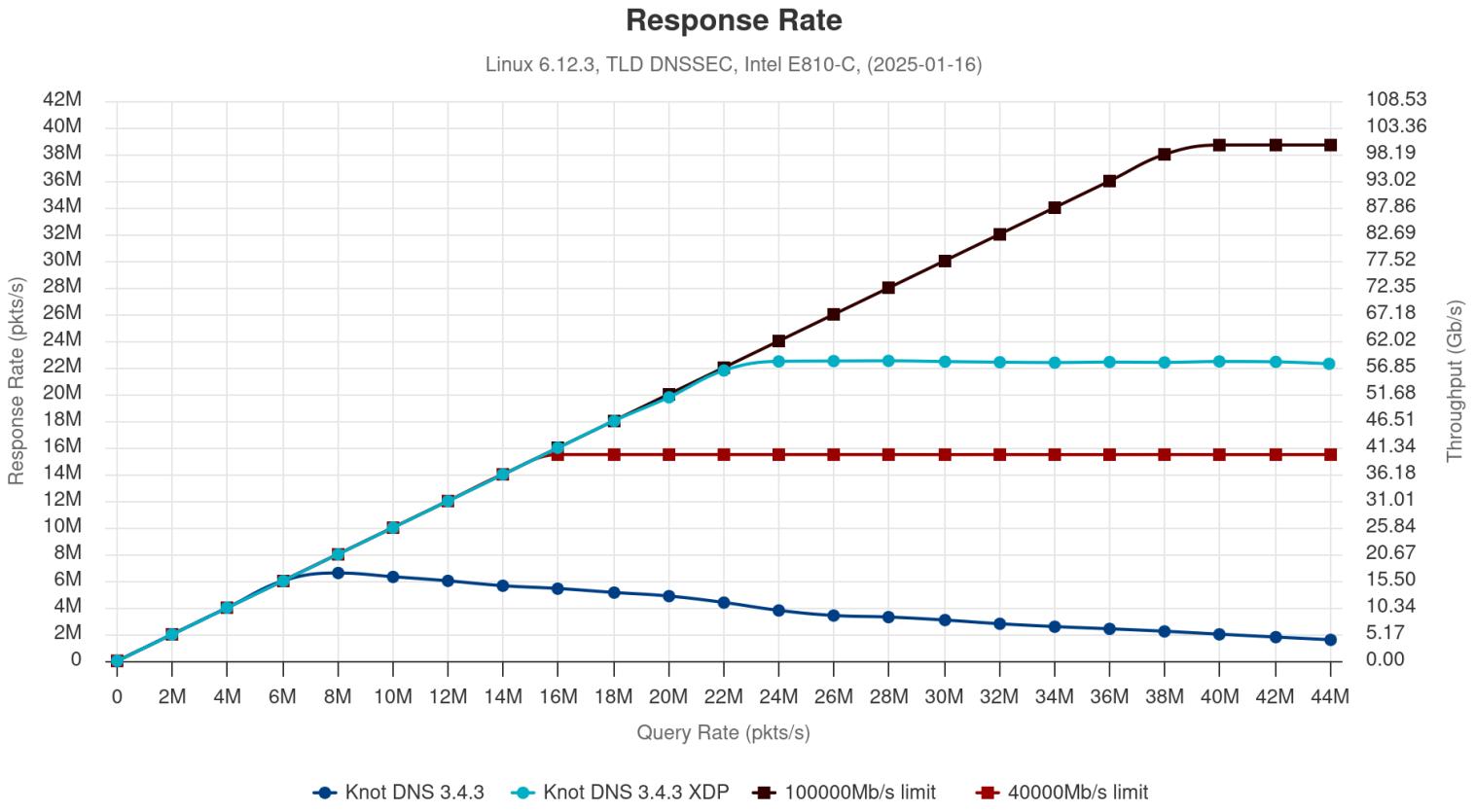
DNSSEC

- Podpora více úložišť klíčů pro zóny
  - Usnadnění migrace z jiného HSM
  - Možnost odděleného ukládání KSK a ZSK – bezpečnost vs. výkon
- Funkce pro předgenerování klíčů pro automatické použití
  - Možnost předběžné kontroly klíčů nebo jejich spolehlivější zálohování
- Možnost více aktivních podepisujících serverů pro jednu zónu
  - Schopnost automatické synchronizace nezávislých DNSKEY záznamů
- Volba **zonefile-skip**
  - Možnost ignorování zadaných typů záznamů při načítání zónového souboru
  - Využití např. při aktualizování DS přes DDNS v podepisované zóně z read-only souboru

Testování výkonu

- Úspěšné zprovoznění se 100GbE síťovými kartami v režimu XDP
- Intel E810 (ovladač **ice**)
  - Nestabilní do verze kernelu Linuxu 6.10.3
- NVIDIA Mellanox ConnectX-6 Dx (ovladač **mlx5**)
  - Maximální počet front 127
- Zapnutí SMT mělo pozitivní vliv v režimu XDP
- Velkými odpovědmi lze plně vytížit 100GbE linku i se starším HW
- https://www.knot-dns.cz/benchmark/

Testování výkonu – ukázka



Děkuji za pozornost!

• 4. 10. 2025 – Daniel Salzman

CZ.NIC | SPRÁVCE DOMÉNY CZ