DNS pro začátečníky

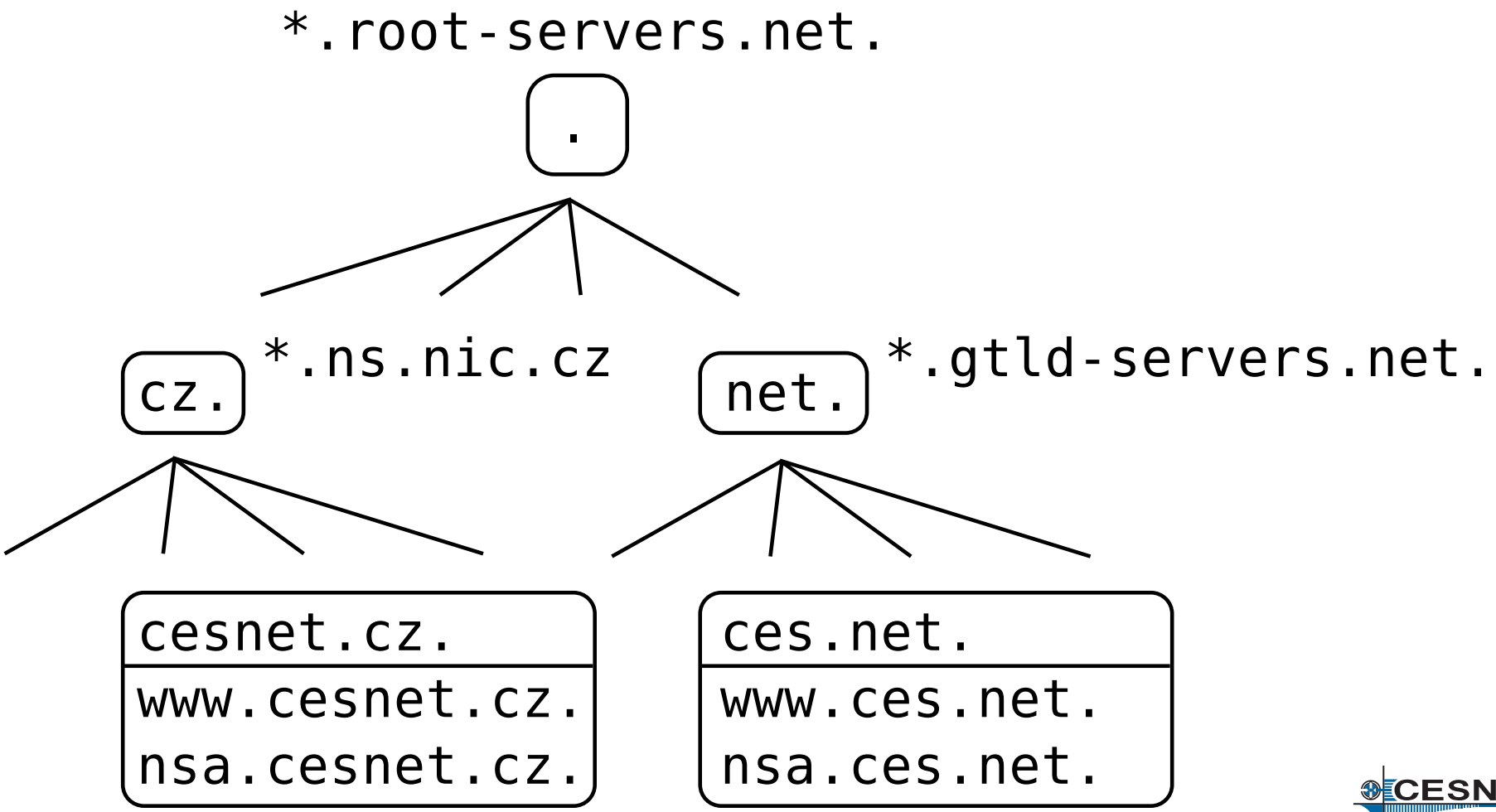
Ondřej Caletka

3. listopadu 2012

Domain Name System

- Navrženo v roce 1982 jako náhrada `HOSTS.TXT`.
- Hierarchická distribuovaná databáze.
- Důraz na dostupnost, namísto rychlých změn a plné konzistence.
- Binární protokol používající UDP a TCP spojení na známém portu 53.

Hierarchická struktura DNS zón



Pojmy z DNS

zóna Část globální databáze, samostatně spravovaná.

Např.: zóna `cz.` spravovaná CZ.NIC

autoritativní server Server poskytující odpovědi ze zón, které drží.

Např.: `a.ns.nic.cz.`

rekurzivní server/resolver Server, který dokáže postupnými dotazy zjistit odpověď na libovolný DNS dotaz.

Např.: Google Public DNS `8.8.8.8`

stub resolver Knihovní funkce, tvoří rozhraní mezi aplikací a rekurzivním serverem.

Např.: glibc

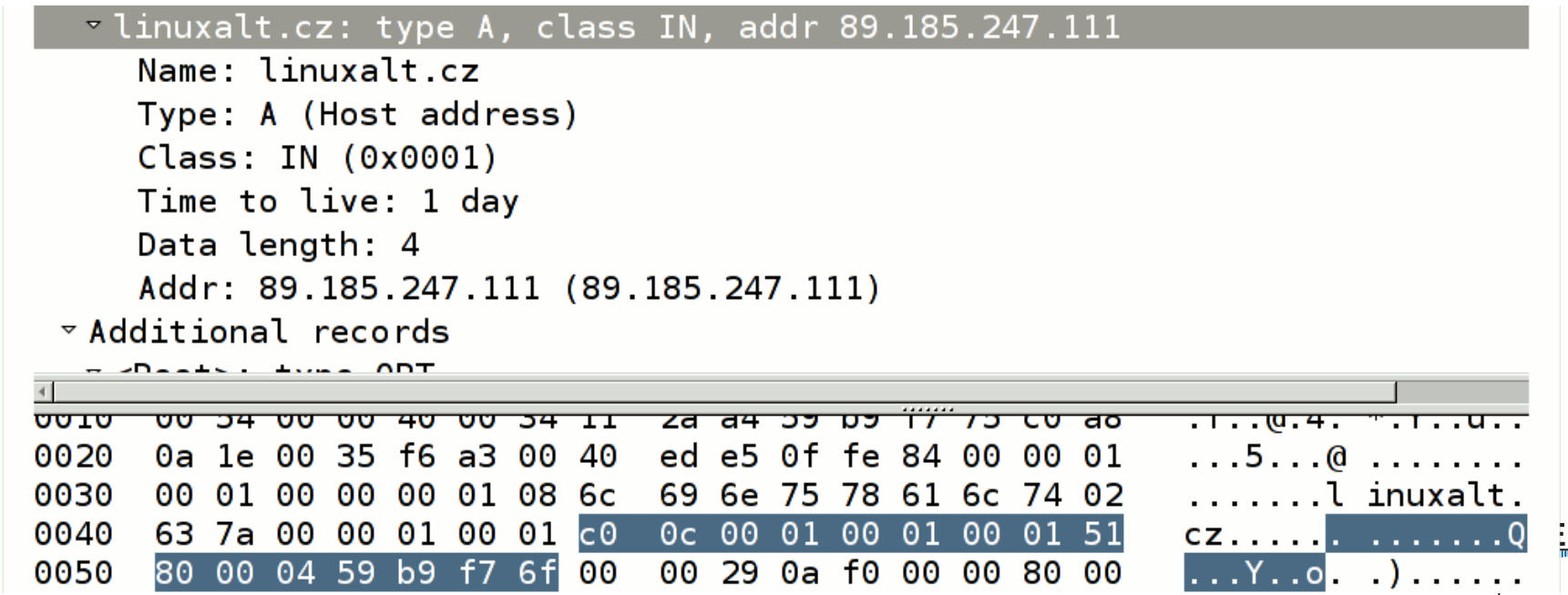
DNS zpráva

- Binární formát
- Společné záhlaví
  - ID transakce
  - Stavový kód
  - Příznaky
    - `AA` Authoritative Answer
    - `RD` Recursion Desired
    - `RA` Recursion Available
    - `TC` Truncated Message
- Čtyři sekce s *resource records*
  - `QUERY` dotaz
  - `ANSWER` konečná odpověď
  - `AUTHORITY` odkaz (referral)
  - `ADDITIONAL` doplňující informace

DNS Resource Record

- Udržován v cache po dobu TTL.
- Názvy domén jako spojový seznam *labels*.
- Komprese opakujících se názvů.

```
linuxalt.cz     86400    IN     A   89.185.247.111
```



Zónový soubor

- Textová podoba jedné DNS zóny.
- Začíná záznamem typu SOA (Start of Authority)
  - jméno primárního serveru
  - e-mail hostmastera
  - sériové číslo
  - časovací parametry
- Apex zóny (`@`)
  - Obsahuje SOA, NS, apod. pro doménu bez prefixu.
- Rídicí direktivy

  `$ORIGIN` doména připojená za relativní názvy

  `$INCLUDE` vložení dalšího souboru

  `$TTL` výchozí hodnota TTL
- Pozor na tečku na konci!

Zónový soubor – příklad

```
$TTL    3600
@         IN  SOA   nsa.cesnet.cz. ( ;primary nameserver
                    hostmaster.cesnet.cz. ;admin e-mail
                    2012072500 ; serial
                    28800      ; refresh  ( 8 hod)
                    7200       ; retry    ( 2 hod)
                    1814400    ; expire   (21 dni)
                    900 )      ; neg. TTL (15 min)

          IN  NS    nsa.cesnet.cz.
          IN  NS    nsa.ces.net.
          IN  NS    decsys.vsb.cz.
;
localhost IN  A     127.0.0.1
```

Delegace a subdelegace

- Způsob, jak je sestaven DNS strom.
- Nadřazená zóna obsahuje NS záznam s adresou serveru s zónou nižší úrovně.

  Např.: `cz. IN NS a.ns.nic.cz.`
- Pokud server pro zónu leží uvnitř stejné zóny, je třeba navíc GLUE záznam.

  Např.: `a.ns.nic.cz. IN A 194.0.12.1`
- Tyto informace se použijí pouze pro prvotní nasměrování (*priming*). Po spojení s delegovaným serverem jsou v cache přepsány informacemi z cílové zóny.

Zónové přenosy

- Synchronizace autoritativních serverů.
- Slave servery periodicky dotazují `SOA` master serveru.
- Došlo-li ke zvýšení sériového čísla, požádají pomocí TCP o záznam typu `AXFR`, nebo `IXFR`.
- Master server odpoví kompletním obsahem zóny (`AXFR`), nebo změnou proti předchozímu sériovému číslu (`IXFR`).
- Není-li master dlouho dostupný, zóna expiruje.
- Master může upozornit slave servery zprávou `NOTIFY`.

Typy záznamů

`A` IPv4 adresa

`AAAA` IPv6 adresa

`PTR` Reverzní záznam

Adresa se převrátí a připojí pod strom `in-addr.arpa.`, nebo `ip6.arpa.`

`MX` Mail eXchange - SMTP server

`CNAME` Canonical Name - alias

Nelze kombinovat s jiným typem RR pro stejné jméno. Neměl by se řetězit.

`SRV` Hledání služeb (SIP, XMPP, atd.)

`SSHFP` SSH finger print

`TLSA` TLS certifikát (DANE)

EDNS0, DNSSEC

- Historický limit UDP DNS paketu 512 B.
- Později přidána rozšiřující hlavička jako záznam typu `EDNS0` v poli `ADDITIONAL`.
  - Inzeruje podporovanou délku UDP paketu (např. 4096 B)
  - Obsahuje další příznak `DO` – DNSSEC OK
- DNSSEC – kryptografické ověření integrity dat
  - Nové typy RR: `DNSKEY`, `RRSIG`, `NSEC`, `DS`
  - Řádové navýšení objemu dat.
  - Resolver provádí validaci.

DNS v praxi

Teorie znamená, že o tom víte všechno, ale nic nefunguje. Praxe znamená, že všechno funguje, ale nikdo neví proč. U nás se teorie snoubí s praxí, protože nic nefunguje a nikdo neví proč.

Autoritativní servery

- Pestrá škála software:
  - BIND od ISC
  - NSD od NLnet Labs
  - Knot DNS od CZ.NIC Labs
  - YADIFA od EURid
  - PowerDNS od PowerDNS
- Obvykle pracují s předkompilovanými zónami.
- Dynamické změny za běhu – BIND, PowerDNS.
- Pozor na výchozí nastavení AXFR.
- Jeden ze serverů by měl mít nastavenu maximální velikost UDP zprávy méně než MTU.

  Kvůli sítím s rozbitým Path MTU Discovery.

Příklad konfigurace

BIND jako master server

```
zone "linuxalt.cz" {
    type master;
    file "/path/to/zones/linuxalt.cz";
}
```

BIND jako slave server

```
zone "linuxalt.cz" {
    type slave;
    file "/path/to/slaves/linuxalt.cz";
    masters {192.0.2.1; 2001:db8::1;};
}
```

Reverzní delegace

Adresa se převrátí (IPv4 po oktetech, IPv6 po nibblech) a připojí pod strom `in-addr.arpa.`, nebo `ip6.arpa.`

IPv4

```
server.example.com.      IN A   192.0.2.1
1.2.0.192.in-addr.arpa.  IN PTR server.example.com.
```

IPv6

```
server.example.com.      IN AAAA 2001:db8:123:456::1
1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.\
6.5.4.0.3.2.1.0.8.b.d.0.1.0.0.2.ip6.arpa.
                         IN PTR server.example.com.
```

Reverzní classless delegace

Problém: Rozsahy IPv4 adres jsou menší, než celá třída.

Zóna `2.0.192.in-addr.arpa.`

```
128/25   IN NS server.example.com.
         IN NS secondary.example.com.
128      IN CNAME 128.128/25
129      IN CNAME 129.128/25
...
255      IN CNAME 255.128/25
```

Zóna `128/25.2.0.192.in-addr.arpa`

```
129      IN PTR server.example.com
```

Cloudy a DNS

```
www.fffilm.name.   1800   IN CNAME ghs.google.com.
ghs.google.com.    590796 IN CNAME ghs.l.google.com.
ghs.l.google.com.  300    IN A     173.194.67.121
```

- CDN služby často intenzivně využívají `CNAME`
- *Problém:* Cloud hosting a doména bez `www`.
  - Nelze použít `CNAME` v apexu.
  - Tvrdé zadání `A` záznamu rozbíjí cloud.
  - Redirect server je asi nejlepší, ale špatné řešení.
  - Řešením by byly `SRV` záznamy pro službu `HTTP`, ale nikdo je nechce podporovat.

CESNET

Žolíkové DNS

- Funkce se obvykle nadužívá.
- Nebezpečné, je-li taková doména v prohledávací cestě.

  `www.google.com.example.com`
- Žolík pokrývá jen neexistující záznamy.

  Existující záznam, i jiného typu, vyhraje.

Odstrašující případ

```
@     IN     SOA    ...
      IN     NS     ...
      IN     A      ...
      IN     AAAA   ...
*     IN     CNAME  @
neco  IN     TXT    "neco"
```

Rekurzivní servery

- Nepříliš pestrá nabídka:
  - BIND od ISC
  - Unbound od NLnet Labs
  - PowerDNS Recursor od PowerDNS
- Neměly by se kombinovat s autoritativními.
- Neměly by sloužit pro celý Internet.
- Neměly by do Internetu přeposílat dotazy na privátní IP adresy.
- Zapnout DNSSEC validaci je jednoduché a bezpečné. Dělají to i velcí ISPs.

Stub resolver v glibc

- Konfigurace v `/etc/resolv.conf`
- Maximálně tři nameservery.
- Pořadí určuje prioritu.

  Rozkládání zátěže pomocí `option rotate`

Příklad kofigurace

```
search example.com
nameserver 8.8.8.8
nameserver 2001:4860:4860::8888
nameserver 8.8.4.4
option edns0 rotate
```

Temná strana DNS

- Každý může snadno odeslat UDP datagram s libovolnou zdrojovou adresou.
- Sítě často nefiltrují adresy podle BCP 38.
- DNS servery odpoví na falešnou adresu…
  - …a má obvykle dobrou konektivitu.
  - …a odpovědi jsou delší než dotazy.
  - …a obvykle nemají limity na počet dotazů od stejné adresy za určitou dobu.
- Zeptejte se svého dodavatele DNS software na *Request Rate Limiting* a zapněte ho.

Shrnutí

- Nepoužívejte žolíky, pokud je nepotřebujete.
- Pište `www.` na začátku a tečku na konci.
- Zkontrolujte platnost svých delegací a GLUE záznamů! Pokaždé, když děláte změny v síti.
- Nezapomeňte zvyšovat sériové číslo zóny.
- Implementujte ve své síti BCP 38.
- Uzavřete své rekurzivní servery pouze pro sebe.
- Dohlížejte servery a limitujte četnost dotazů.
- Nepropadejte panice!

CESNET

Závěr

Děkuji za pozornost.

CESNET