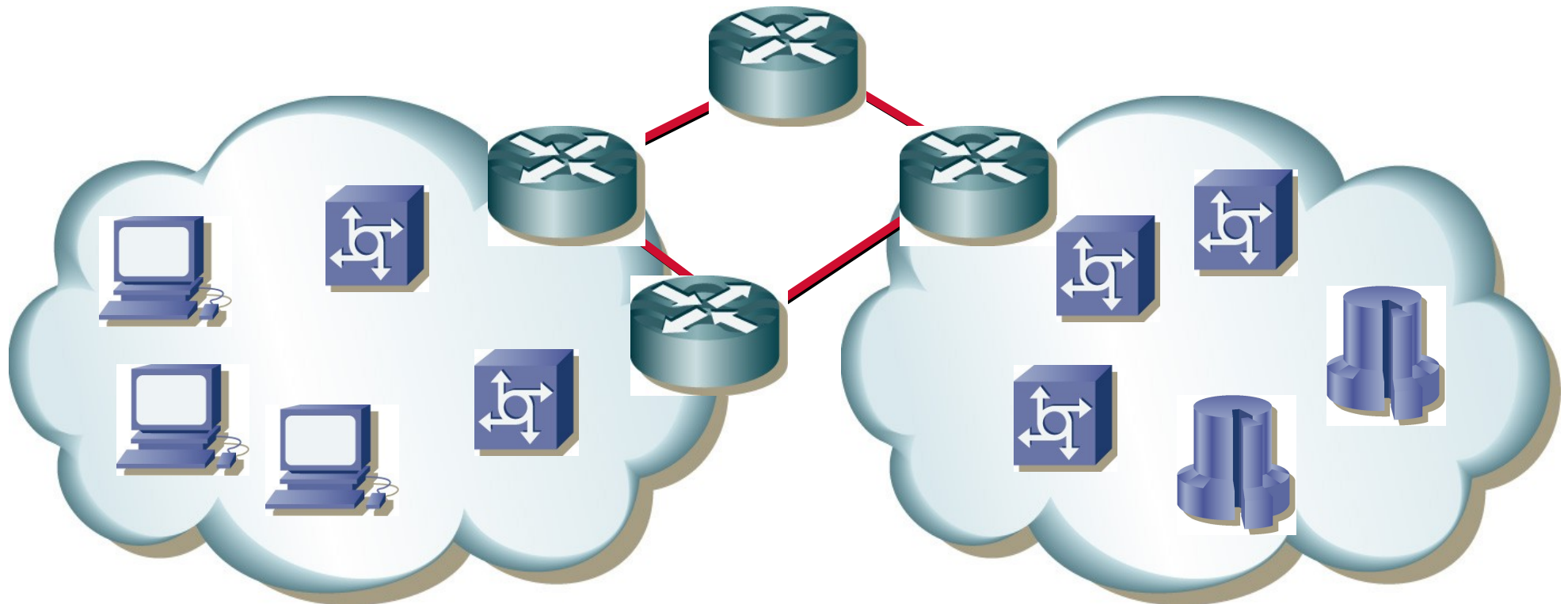


Počítačové sítě II

16. Domain Name System

Miroslav Spousta, 2006

<qiq@ucw.cz>, <http://www.ucw.cz/~qiq/vsfs/>



Domain Name System

- DNS = Doménový jmenný systém
- IP používá číselné adresy (32 nebo 128 bitů)
 - těžko zapamatovatelné
 - Tyto adresy stačí pro funkci IP, TCP, UDP, téměř celého Internetu (až na lidi)
- překlad snadno zapamatovatelných jmen na IP adresy
 - a obráceně
- DNS zahrnuje
 - hierarchickou databázi, kde jsou uloženy údaje
 - mechanismus pro dotazování se do této databáze
 - pravidla pro vytváření jmen

Historie

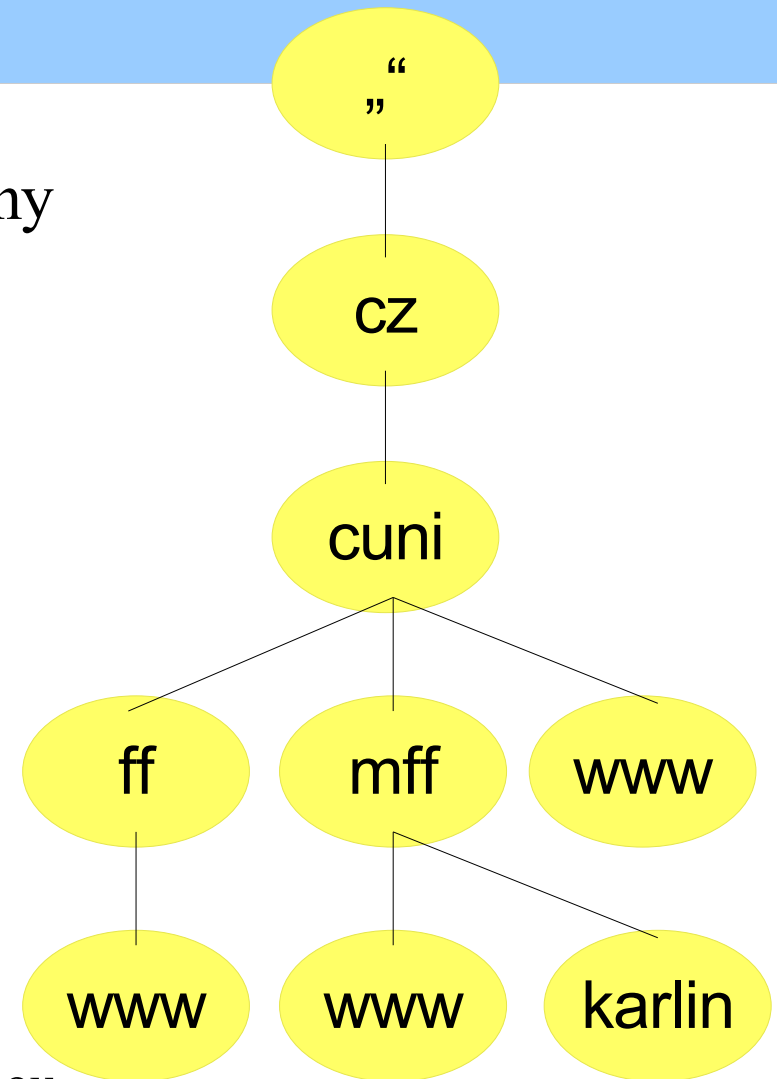
- původní řešení: centralizované
 - v zárodečném ARPANETu
 - existovala centrální autorita, která přidělovala jména (a jejich přiřazení k IP adresám)
 - aktualizace informací probíhala pomocí distribuce tzv. hosts souboru
- s nárůstem velikosti souboru přestala tato možnost vyhovovat
 - problém s velikostí a rychlostí aktualizací
 - bylo potřeba databázi distribuovat (rozložení zátěže i pravomocí)
- počátkem 80. let začíná vznikat DNS
- dodnes existuje soubor /etc/hosts
 - obsahující statické přiřazení jmen k adresám
 - ale spravuje se pouze lokálně (pro jeden uzel)
 - hodí se pro malé/ad-hoc sítě (nebo pro blokování reklamních serverů)

Vlastnosti DNS

- DNS je jedna z nejdůležitějších služeb Internetu
 - bez ní Internet fakticky funguje, pro lidi je však nepoužitelný
 - ale ani např. e-maily nebudou fungovat prakticky vůbec
- robustnost
 - odolnost proti výpadkům uzlů (DNS serverů) – replikace
 - distribuované místně – odolnost proti výpadkům sítě/připojení sítě
 - rozložení zátěže
- distribuce pravomocí
 - mnoho subjektů, které si chtějí spravovat jména samy (instituce, firmy, „„)
 - adresy nepřiděluje centrální autorita
- snadno zapamatovatelné, logické názvy
 - jednoduchá hierarchická struktura, jména jsou *jednoznačná*
 - rozdělení prostoru na několik částí – úrovní

Hierarchie

- jmenný prostor je rozdělený na části – domény
- vytváří stromovou strukturu
 - kořen je počátek prostoru
 - cesta od kořene k listu udává plné jméno uzlu
 - neboli *fqdn* – fully qualified domain name
- doména je jedna úroveň stromu
 - je jednoznačně určena cestou ke kořeni „“
 - jsou v ní definovány další domény nebo jména – nižší úrovně
 - úrovně se číslují od 1 (ale říká se jí nejvyšší)
 - domény se zapisují za sebou, jednotlivé úrovně jsou odděleny tečkou
 - zapisuje se zprava doleva (vpravo je nejvyšší doména – úroveň 1)



Doménová jména

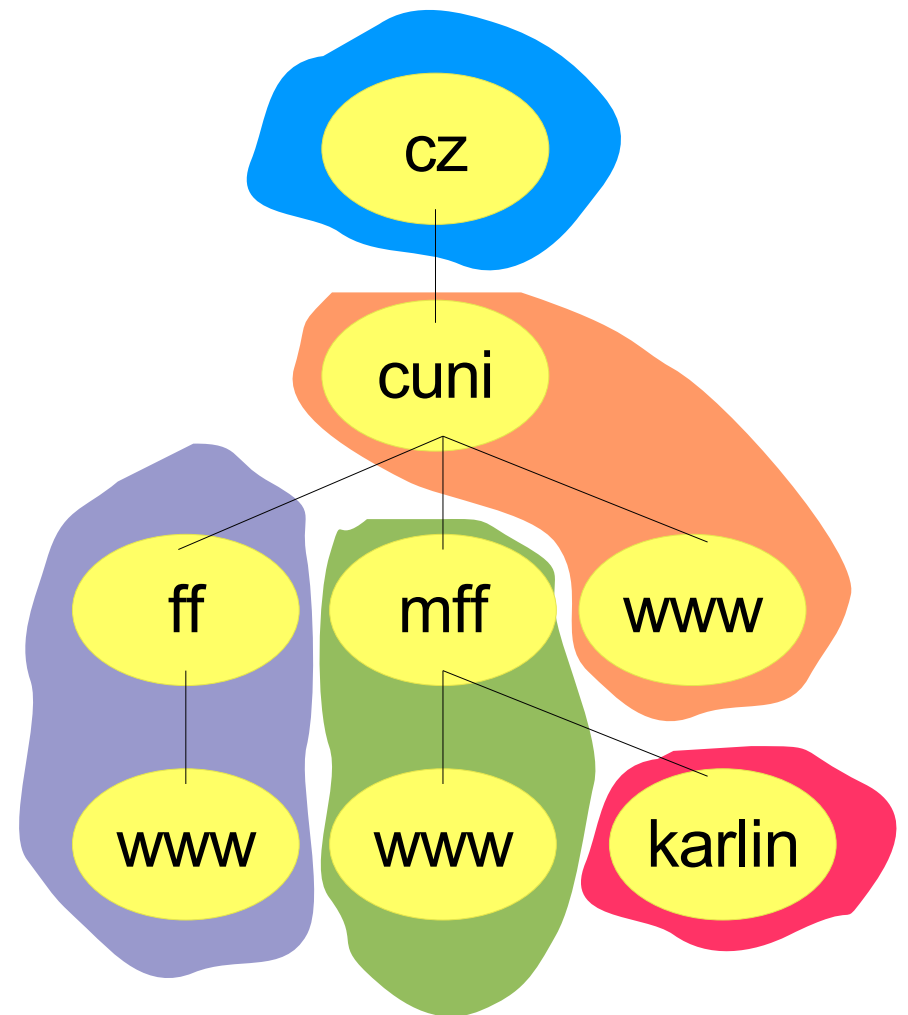
- mohou obsahovat velká a malá písmena anglické abecedy, číslice a pomlčku
 - pomlčka nesmí stát na začátku ani na konci jména
- maximální velikost na jedné úrovni je 63 znaků
- velká a malá písmena se nerozlišují (Sun.COM == sun.com)
- počet vnoření (poddómén) není explicitně omezen
 - ale fqdn musí mít maximálně 255 znaků
- čemu odpovídá jedna úroveň (doména)?
 - organizačnímu členění, prostorovému rozdělení, ...
 - není explicitně určeno, záleží na uživateli
- velikost domény souvisí s její udržitelností
 - v rámci domény musí být jména jednoznačná
 - doména je většinou spravovaná centrálně, z jednoho místa

Domény

- TLD (top level domain – domény nejvyšší (první) úrovně)
 - jsou centrálně přidělené a spravované
- např. cz, sk, uk, com, org, net, us
- jejich správa (přidělování domén nižší úrovně) je delegováno
- zájemci o nižší (druhou) úroveň se obracejí na národní registrátory
 - CZ: NIC.CZ, poskytovatelé připojení
- v rámci své domény si subjekty přidělují poddomény libovolně
 - malé organizace mají většinou všechna jména přímo ve svojí doméně
 - větší doménu mohou dělit (cuni.cz => ff.cuni.cz, natur.cuni.cz, mff.cuni.cz, ...)
- pro každé doménové jméno může v databáze existovat N záznamů různých typů

Zóny

- zóna: doména s (některými) poddoménami, která je jednotně spravovaná
 - tedy jedním subjektem (např. ISP, firmou, ...)
- z podstromu mohou být „vykousnuty“ některé podstromy
 - které spravuje někdo jiný, tedy pravomoce (autorita) jsou delegovány na jiný subjekt
- zóna je zpravidla uložena v jednom souboru (tzv. zónový soubor)
 - v něm jsou položky vztahující se k dané zóně – resource records (RR)



Resource records

- obsah DNS databáze: položky – RR (resource records)
 - přísluší vždy k nějakému doménovému jménu
- každá položka RR se skládá z několika částí:
 - jméno (identifikace položky)
 - TTL (jak dlouho může client držet záznam v cache)
 - class (třída) – pro Internet IN
 - typ (A, NS, MX, ...)
 - RDATA – data, interpretace záleží na typu
- položky jsou většinou uloženy v jednoduché textové databázi
 - zónovém souboru, jedna položka na jednom řádku
 - pochází z BIND (Berkeley Internet Name Daemon)

Typy RR

- SOA: Start of Authority – informace o zóně
 - na začátku každého zónového souboru
 - e-mail na osobu zodpovědnou za danou zónu
 - časy pro synchronizaci mezi primárním a sekundárním NS
- A: překlad na IP adresu
 - hostname: doménové jméno, které má A záznam
- CNAME: alias pro jiné jméno
- NS: Name Server pro danou doménu (FQDN)
- MX: Mail eXchange
 - mail server pro doménu
 - může jich být více, liší se prioritou, čím nižší číslo, tím vyšší priorita
- PTR: ukazatel jinam do jiné části DNS stromu
 - pro reverzní překlad

CNAME

- CNAME uvádí, že nějaké doménové jméno je alias za jiné
- mohou se řetězit, ale neměly by se (zdržuje)
- pro doménu, která je aliasem *nesmí* existovat záznamy jiných typů
- MX a NS záznamy *nesmí* ukazovat na alias (ani na ostatní by neměly)
- reverzní (PTR) záznam by neměl ukazovat na alias, ale na kanonické jméno
- hodí se pro reverzní záznamy při CIDR delegaci
- např.: `www CNAME host1.domena.tld`

Zpětný překlad

- chceme přeložit IP (zpět) adresu na doménové jméno
- využívá se záznam typu PTR
- a speciální doména: `in-addr.arpa`
 - pod ní se použijí bajty IP adresy v obráceném pořadí
 - např. `150.0.0.10.in-addr.arpa`
 - záznam pak vypadá: `150.0.0.10.in-addr.arpa PTR host1`

Reverzní záznamy

- chceme obrácenou službu: z IP adresy získat doménové jméno
 - např. web/mail server chce do logu zapsat, odkud přišel požadavek
- využívá se záznam typu PTR
- a speciální doména: in-addr.arpa
 - v ní je po jednotlivých bytech uveden celý IPv4 rozsah
 - např. pro adresu 195.113.31.125 je to 125.31.113.195.in-addr.arpa
 - bajty jsou uvedeny v obráceném pořadí (!)
 - pro toto jméno existuje překlad pomocí PTR na jméno: artax.karlin.mff.cuni.cz
 - záznam by měl ukazovat na doménové jméno, které má A záznam
 - tedy ne na alias (CNAME)

```
1.33.168.192.in-addr.arpa    1D  IN  PTR test1.aups.cz.
2.33.168.192.in-addr.arpa    1D  IN  PTR test2.apus.cz.
3.33.168.192.in-addr.arpa    1D  IN  PTR test3.apus.cz.
```

Zónový soubor

```
$TTL 3D
@          IN      SOA     ns.linux.xxx. hostmaster.linux.xxx. (
                        199802151      ; serial, todays date + serial #
                        8H              ; refresh, seconds
                        2H              ; retry, seconds
                        4W              ; expire, seconds
                        1D )            ; negative TTL
;
;
;          NS       ns              ; Inet Address of name server
;          MX       10 mail.linux.xxx. ; Primary Mail Exchanger
;          MX       20 mail.friend.xxx. ; Secondary MX
;
localhost A        127.0.0.1
ns         A        192.168.196.2
mail      A        192.168.196.4
```

DNS protokol

- běží nad UDP (TCP), port 53
 - na dotazy na jméno preferován UDP
 - na transfery zón se používá TCP
- stejný formát paketu pro dotaz i odpověď
- Header: hlavička požadavku
 - jestli se jedná o dotaz či odpověď, kolik položek je v které části, ...
- Question: dotaz (třída, typ, jméno)
- Answer: odpověď (RR odpovědi, může jich být více)
- Authority: NS, odkud záznam(y) pochází
- Additional: další informace, které souvisí s dotazem
 - např. překlad NS/MX na IP adresu

Header
Question
Answer
Authority
Additional

DNS odpověď

```
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 3018
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 3, ADDITIONAL: 3
;; QUESTION SECTION:
;idnes.cz.                IN      A
;; ANSWER SECTION:
idnes.cz.                2940    IN      A      194.228.207.200
;; AUTHORITY SECTION:
idnes.cz.                2940    IN      NS     ns2.mafra.cz.
idnes.cz.                2940    IN      NS     ns.mafra.cz.
idnes.cz.                2940    IN      NS     ns2.tel.cz.
;; ADDITIONAL SECTION:
ns.mafra.cz.             1579    IN      A      194.228.207.7
ns2.tel.cz.              1712    IN      A      194.228.2.1
ns2.mafra.cz.            1579    IN      A      62.209.231.7
```


nslookup, dig

- příkazy pro testování DNS
- **nslookup** může fungovat také v interaktivním módu
- hodí se pro jednoduché testování DNS
- umí automaticky převádět i obráceným směrem: IP -> jméno
- původem z BIND (Berkeley Internet Name Domain)
- najdete je všude (Linux, UNIX, Windows, ...)
- pokročilejší nástroj: **dig**

```
bug:~# nslookup
> idnes.cz

Server:                62.24.64.2
Address:               62.24.64.2#53

Non-authoritative answer:
Name:   idnes.cz
Address: 194.228.207.200

> quit

Server:                62.24.64.2
Address:               62.24.64.2#53

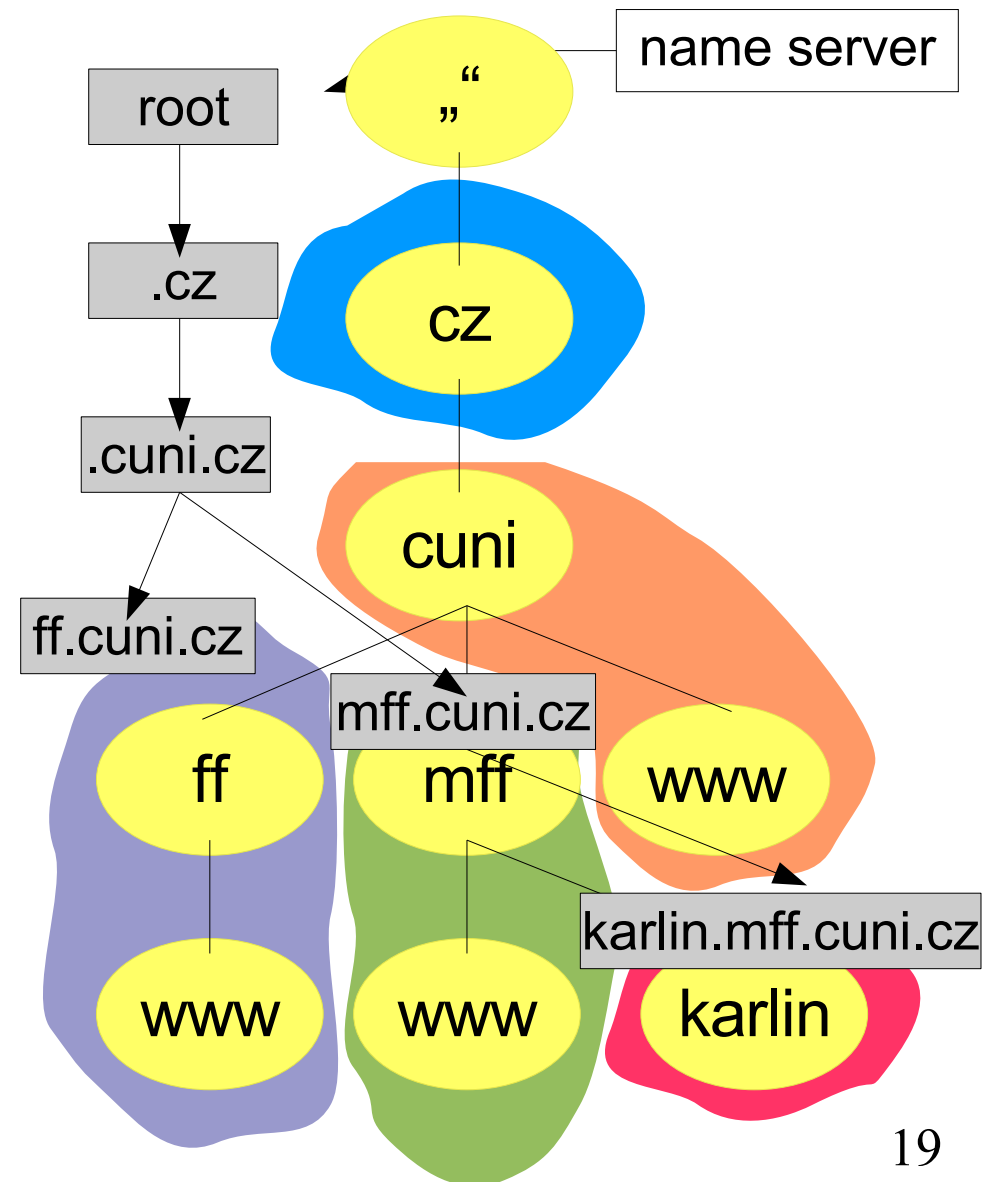
** server can't find quit: NXDOMAIN
>
```

Name servers

- pro každou doménu je potřeba uchovat různé informace
 - překlad na IP adresy (uzel se jménem procite má adresu 1.2.3.4)
 - který uzel přijímá poštu pro danou doménu
 - které name servers jsou dostupné pro doménu/poddoménu
 - ...
- informace jsou distribuovány po celém Internetu
 - jsou soustředěny na *name serverech*
 - které mají informace o určité doméně (zóně)
 - a odpovídají na dotazy klientů
- každá doména má svůj name server
 - ale může ho sdílet s jinou doménou (jeden name server pro více domén)

Name servery

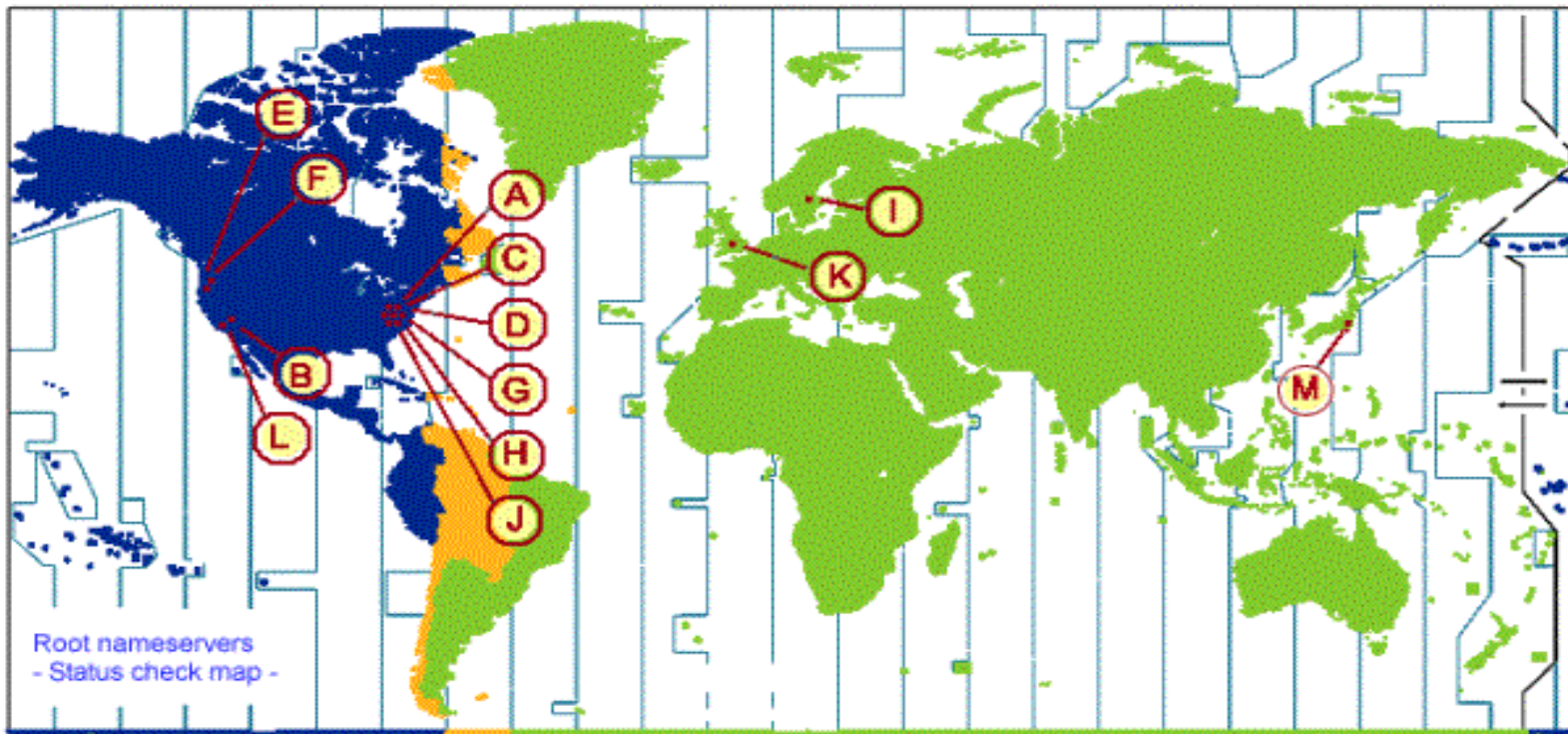
- name servery tvoří stromovou hierarchii
 - kořen stromu je tzv. kořenový (root) name server, který uchovává odkazy na všechny name servery pro TLD
 - ve skutečnosti je těchto name serverů více (a.root-servers.net – m.root-servers.net)
 - odkazy na nižší zóny se řeší pomocí tzv. *glue records*
- name server uchovává informace pro celou zónu
- pro jednu doménu existují aspoň dva name servery
 - ochrana proti výpadkům uzlu/sítě



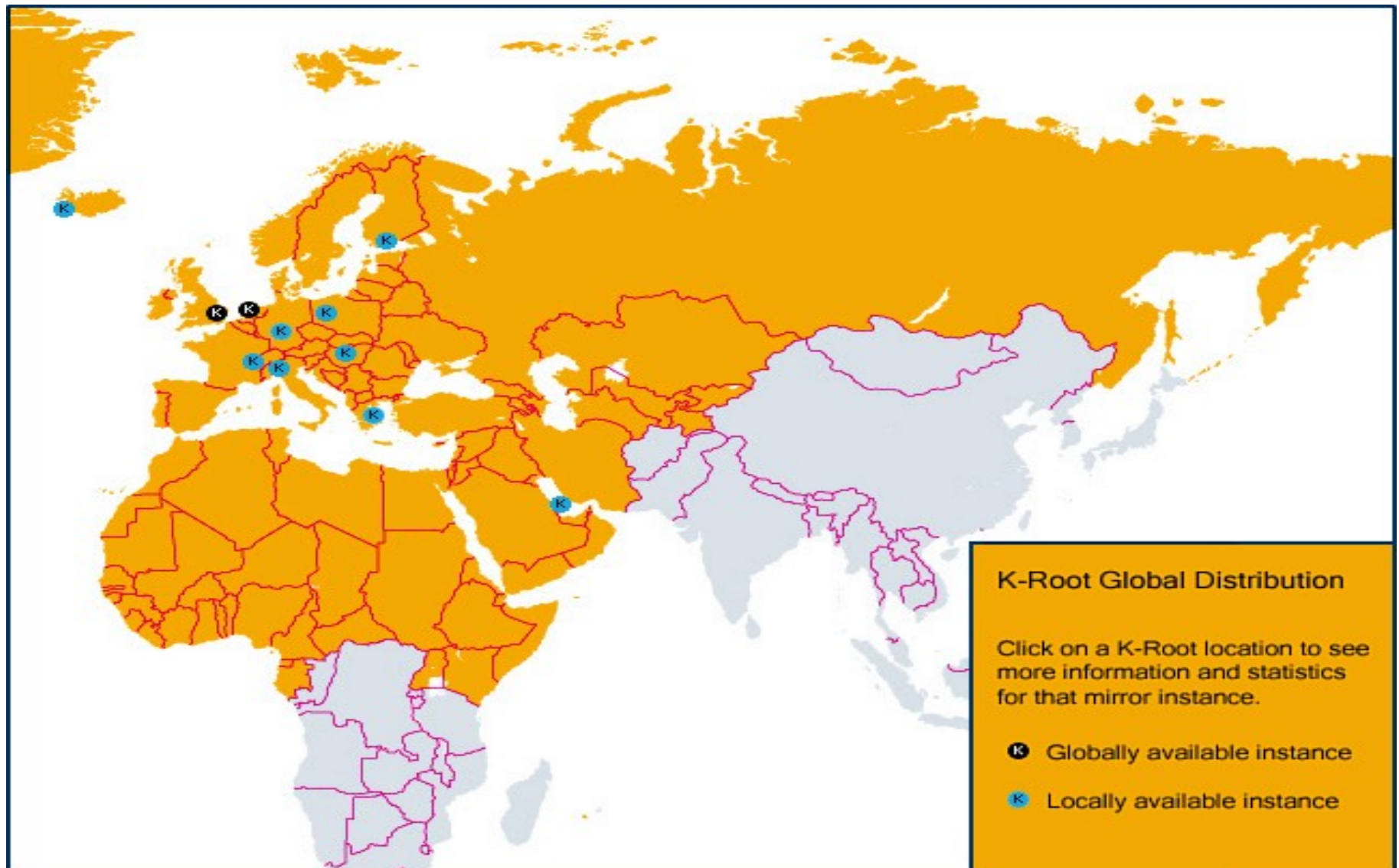
Root name servers

-
-
-

Map of the Root Servers



Root name servers



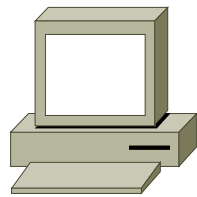
Primární a sekundární NS

- primární NS pro doménu
 - pro každou doménu existuje právě jeden
 - je na něm uložen zónový soubor
 - slouží jako primární zdroj dat
- sekundární NS
 - obsahuje obraz dat pro doménu primárního NS
 - pravidelně si aktualizuje tento obraz dat (pomocí zone transfer)
 - odpovídá na dotazy stejně jako primární NS
 - měl by být od primárního NS dostatečně vzdálen (v jiné síti)
 - pro každou doménu by měl existovat alespoň jeden
- primární i sekundární NS jsou *autoritativní* pro danou doménu
 - neboli vždy vrací platné informace (aspoň by měly)

DNS dotaz

Q: www.karlin.mff.cuni.cz
A: cz.: ns.tld.cz

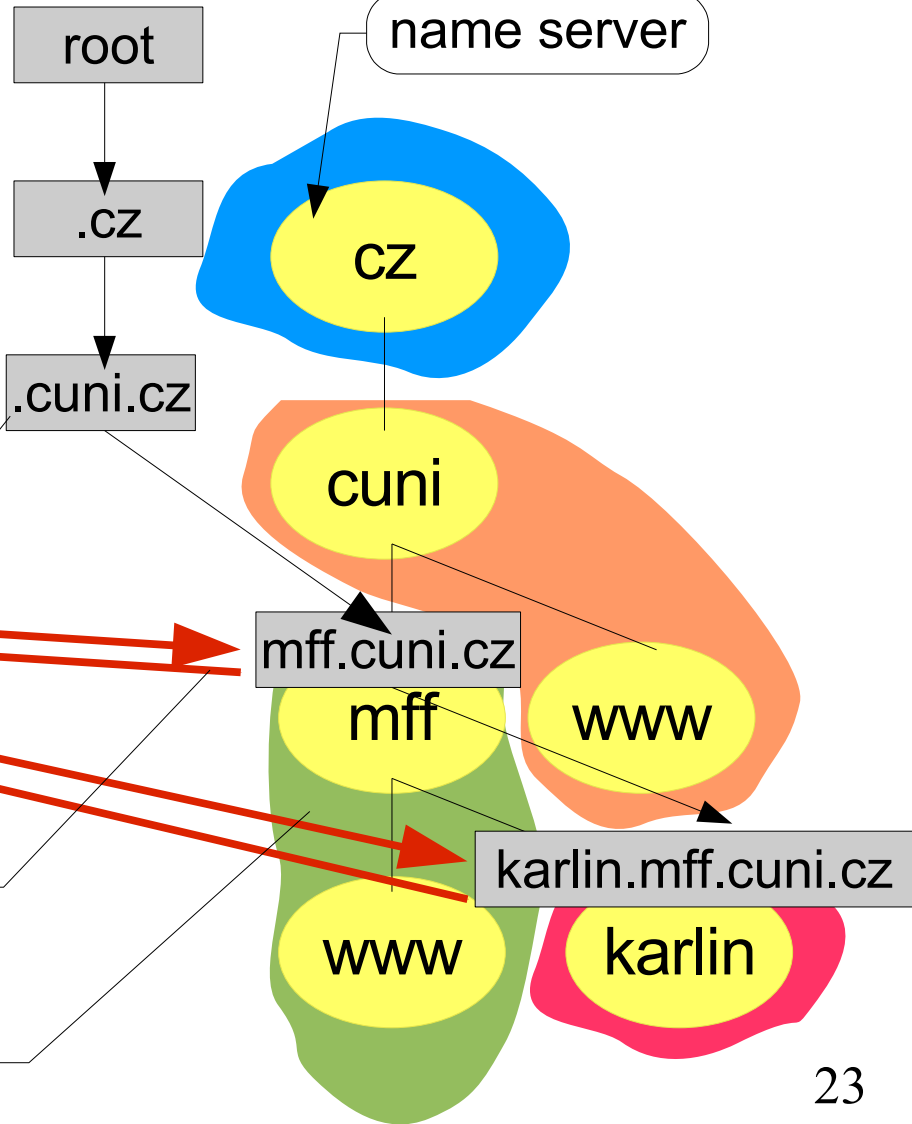
Q: www.karlin.mff.cuni.cz
A: cuni.cz.: golias.ruk.cuni.cz



Q: www.karlin.mff.cuni.cz
A: mff.cuni.cz.: ns.ms.mff.cuni.cz

Q: www.karlin.mff.cuni.cz
A: karlin.mff.cuni.cz.: csmat.karlin.mff.cuni.cz

Q: www.karlin.mff.cuni.cz
A: www.karlin.mff.cuni.cz.: 195.113.30.214



DNS dotaz: www.cuni.cz

```
; <<>> DiG 9.2.1 <<>> +trace www.cuni.cz
;; global options:  printcmd
.                272471  IN      NS      A.ROOT-SERVERS.NET.
.                272471  IN      NS      B.ROOT-SERVERS.NET.
.                272471  IN      NS      C.ROOT-SERVERS.NET.

.                272471  IN      NS      I.ROOT-SERVERS.NET.
.                272471  IN      NS      J.ROOT-SERVERS.NET.
.                272471  IN      NS      K.ROOT-SERVERS.NET.
.                272471  IN      NS      L.ROOT-SERVERS.NET.
.                272471  IN      NS      M.ROOT-SERVERS.NET.
;; Received 324 bytes from 195.113.0.2#53(195.113.0.2) in 36 ms
cz.              172800  IN      NS      NS2.NIC.FR.
cz.              172800  IN      NS      NS.RIPE.NET.
cz.              172800  IN      NS      SUNIC.SUNET.SE.
cz.              172800  IN      NS      NS-EXT.VIX.COM.
cz.              172800  IN      NS      NS.TLD.cz.
cz.              172800  IN      NS      NSS.TLD.cz.
;; Received 269 bytes from 198.41.0.4#53(A.ROOT-SERVERS.NET) in 122 ms
cuni.cz.         18000   IN      NS      goliias.ruk.cuni.cz.
cuni.cz.         18000   IN      NS      ns.ces.net.
;; Received 94 bytes from 192.36.125.2#53(SUNIC.SUNET.SE) in 37 ms
www.cuni.cz.     86400   IN      CNAME   tarantula.ruk.cuni.cz.
tarantula.ruk.cuni.cz. 86400   IN      A       195.113.0.123
ruk.cuni.cz.     86400   IN      NS      cucc.ruk.cuni.cz.
ruk.cuni.cz.     86400   IN      NS      goliias.ruk.cuni.cz.
ruk.cuni.cz.     86400   IN      NS      ruzenka.prf.cuni.cz.
;; Received 187 bytes from 195.113.0.2#53(goliias.ruk.cuni.cz) in 0 ms
```


DNS resolver

- knihovny, které zajišťují kontakt s DNS serverem
 - DNS: klient – server architektura
 - klient přijímá od aplikací dotazy a pokládá je NS
 - vyhodnocuje a vrací odpověď aplikaci
 - obvykle forma sdílené knihovny
- name server
 - odpovídá na dotazy ohledně vlastních domén (vlastní zóny)
 - a řeší dotazy od klienta, hledá a vrací odpověď
 - musí obsahovat také resolver
 - dotazy si ukládá v cache (non-authoritative answer)

Optimalizace DNS

- redundantní name servery
 - primární a sekundární
 - rozkládání zátěže
- cachování dotazů
 - nameserver, který vyřeší dotaz pro klienta si ho uloží do dočasné paměti
 - každá položka má dobu, po kterou může zůstat nacachovaná
 - při dalším dotazu vrátí záznam z dočasné paměti (největší optimalizace)
 - tzv. neautoritativní odpověď
- caching only name server
 - NS, který pouze řeší dotazy
 - není autoritativní pro žádnou zónu
- forwarding name server
 - ani neřeší dotazy, jen přeposílá na jiný NS

Glue records

- rekurzivní provádění dotazu do DNS
 - začneme u kořenových NS, tážeme se na celé jméno
 - dostáváme odkazy na NS pro domény nižších řádů
- při vyhodnocování dotazu můžeme narazit na záznam o NS, který je součástí hledané domény
 - např. **seznam.cz NS ns.seznam.cz**
- pak je potřeba, aby součástí zóny byl i překlad na IP adresu pro daný name server
 - v odpovědi se vrací v Additional sekci
 - **ns.seznam.cz A 212.80.76.20**

CIDR a reverzní záznamy

- klasický zpětný překlad se příliš nehodí pro CIDR
 - členění na podsítě nemusí být po byte, je problém s autoritou nad danou zónou
- řešení: vytvoří se umělá doména, místo PTR ukazatelů se uvede alias (CNAME) do této domény

```
1.33.168.192.in-addr.arpa. CNAME 1.0-15.33.168.192.in-addr.arpa.  
2.33.168.192.in-addr.arpa. CNAME 2.0-15.33.168.192.in-addr.arpa.  
3.33.168.192.in-addr.arpa. CNAME 3.16-31.33.168.192.in-addr.arpa.  
0-15.33.168.192.in-addr.arpa. NS nekde.nejaky.ns
```

```
@ 1D IN SOA 0-15.33.168.192.hostmaster.xxx. (0 3600 120 3600 3600)  
@ 1D IN NS ns  
ns 1D IN A 192.168.33.1  
1.0-15.33.168.192.in-addr.arpa 1D IN PTR apus.example.  
2.0-15.33.168.192.in-addr.arpa 1D IN PTR prunella.example.
```

Diakritika v DNS

- v původním návrhu DNS se s národními znaky nepočítalo
 - pouze anglická abeceda, čísla a „-“
- byla snaha povolit národní znaky
 - např.: košíčky.cz
- řešení: IDN (Internationalized Domain Names)
 - RFC 3490
 - v doménových jménech je možné používat podmnožinu UNICODE
 - nadstavba nad DNS, překládá se na straně klienta (neboli v DNS se nic nemění)
- jméno se nejprve přeloží do speciální formy (ne-ASCII znaky se zakódují podobně jako UTF-7)
- toto zakódované jméno se použije pro DNS dotaz
 - již je v čistém ASCII, jen pro člověka nečitelné

IDN Překlad

- nejprve se jméno (v UNICODE) převede na kanonický tvar
 - sjednocení variant zápisu (velká a malá písmena)
- poté se jméno převede na tzv. punycode
 - algoritmus převodu je reverzibilní (je možné převést punycode zpět na UNICODE)
 - přidá se prefix xn--, za ním znaky bez diakritiky, zakódované znaky nakonec
 - např. košíčky.cz => xn—koky-wpa6qow.cz
- takto převedené jméno (vyhovuje původním požadavkům DNS) se pošle jako dotaz
- problémy:
 - je potřeba registrovat domény speciálního tvaru
 - ne všude je možné zadat jméno v UNICODE => pak je potřeba používat přeloženou formu
 - phishing (jméno obsahuje podobné znaky)

Whois služba

- protokol pro dotazování do databáze, např. databáze domén
 - zlaté stránky pro domény
 - různé údaje o doménách: subjekt, administrátor, technická podpora, od kdy do kdy je doména platná, etc.
- registrátoři domén poskytují whois službu (někdy pouze přes web)
- ČR: whois.nic.cz

```
$ whois idnes.cz

domain:          idnes.cz
descr:           MAFRA a. s.
descr:           Prague 1
admin-c:         JC607-RIPE_XX
tech-c:          JC607-RIPE_XX
reg-c:           REG-INTERNET-CZ
nserver:         ns.mafra.cz ns2.mafra.cz
created:         19971130
changed:         20060113
expire:          20061020
...
```