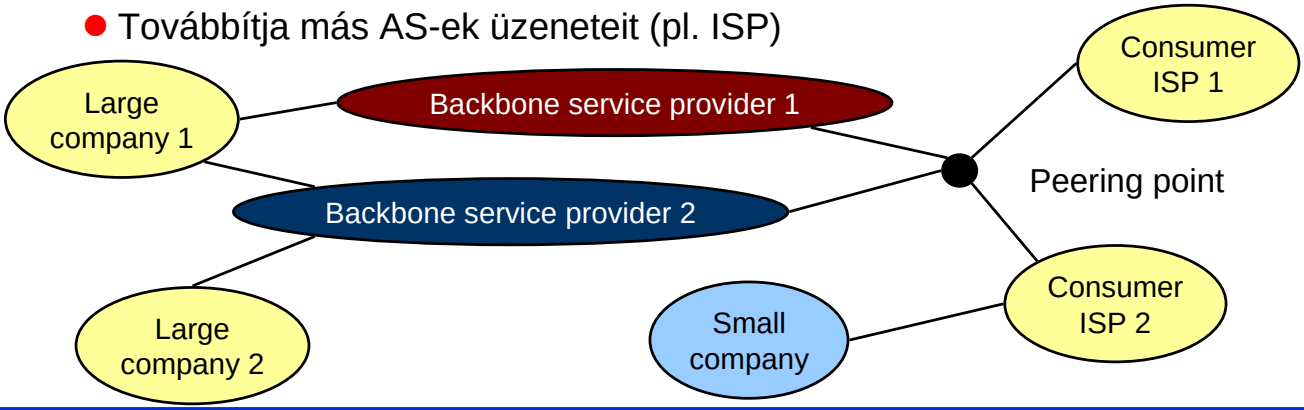
Számítógépes Hálózatok
2012

9. Hálózati réteg –
Inter-AS Routing, BGP, IP címzés, IPv6, DNS

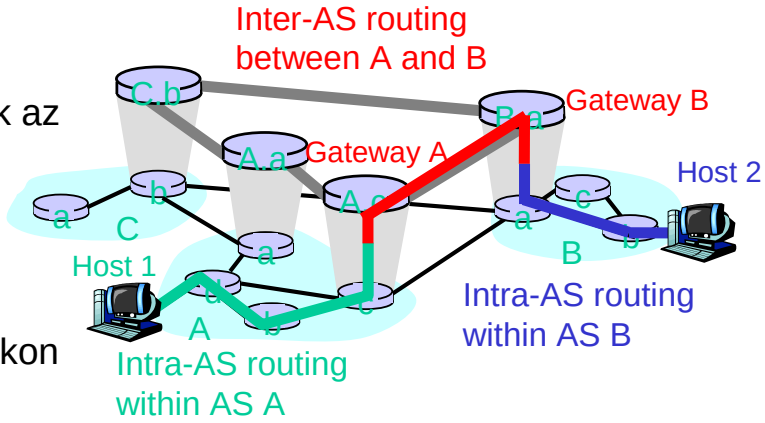
Autonóm rendszerek (AS) tipusai

- Stub-AS
  - Csak egy más AS-hez kapcsolódik
- Multihomed AS
  - Több AS-hez kapcsolódik
  - Nem továbbítja más AS-ek forgalmát
- Transit AS
  - Több kapcsolat
  - Továbbítja más AS-ek üzeneteit (pl. ISP)



Inter-AS-Routing

- Inter-AS-Routing nehéz...
  - Szervezetek megtagadhatják az üzenetek továbbítását (pl. csak fizető ügyfelek csomagjait továbbítja)
  - Politikai követelmények
    - Továbbítás más országokon keresztül?



  - Különböző AS-ek routing-metrikái sokszor nem összehasonlíthatók
    - Útvonal optimalizálás lehetetlen!
    - Inter-AS-Routing megpróbálja legalább a csomópontok elérhetőségét lehetővé tenni
  - Méret: inter-domain routereknek ma kb. 140.000 hálózatról kell tudni

Inter-AS routing: BGP (Border Gateway Protocol)

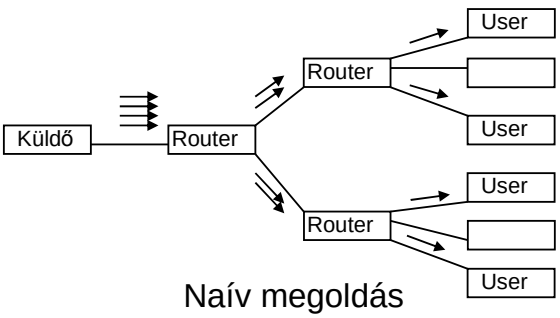
- Az inter-AS routing standard BGPv4
- Path Vector protokoll
  - Hasonló a Distance Vector protokollhoz
  - Minden Border Gateway meghirdeti minden szomszédjának (peers) az egész utat (AS-ek sorozata) a célig (advertisement)
  - TCP-t használ
- Amikor Gateway X az utat Z-hez Peer-Gateway W-nek küldi
  - akkor W választhatja ezt az utat, vagy éppen nem
  - Optimalizálási kritériumok:
    - költségek, politika, etc…
  - Ha W az X által meghirdetett utat választja, akkor meghirdeti
    - Path(W,Z) = (W, Path (X,Z))
- Megjegyzés
  - X tudja szabályozni a hozzá érkező forgalmat a meghirdetések által.
  - Komplikált protokoll

Broadcast és Multicast

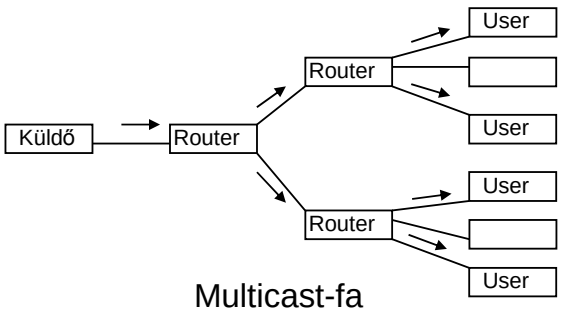
- Broadcast routing
  - Egy csomagot (másolatot) minden más csomópontnak el kell küldeni
  - Megoldások:
    - A hálózat elárasztása (flooding)
    - Jobb: Konstruáljunk egy minimális feszítőfát
- Multicast routing
  - Az adatokat egy küldőtől egyidejűleg több fogadóhoz kell eljuttatni
    - Real time Streaming, Video-On-Demand,
    - Telefon-, Videokonferencia (all-to-all multicast),…
  - IP D címosztály:
    - Egy multicast-csoport (group) minden tagja ugyanazt a címet használja
  - Megoldások:
    - Optimális: Minimális Steiner Fa Probléma
      - NP-teljes (2-approximáció O(n log n) idő alatt kiszámítható!)
    - Más (nem-optimális) fát konstruálni

Multicasting

- Naív megoldás: Multicast-via-Unicast:
  - A küldő egy külön másolatot küld az adatokról minden foadónak.
  - Nagyon inefficiens: A küldött csomagok száma sokkal nagyobb, mint ami szükséges lenne (különösen rossz all-to-all multicast esetén).
- Egy multicast-fa felepítése segítségével:
  - Minden linken csak egyszer továbbítódik egy csomag.
  - A routerek döntik el, hogy egy csomagot több linken is továbbítanak-e.



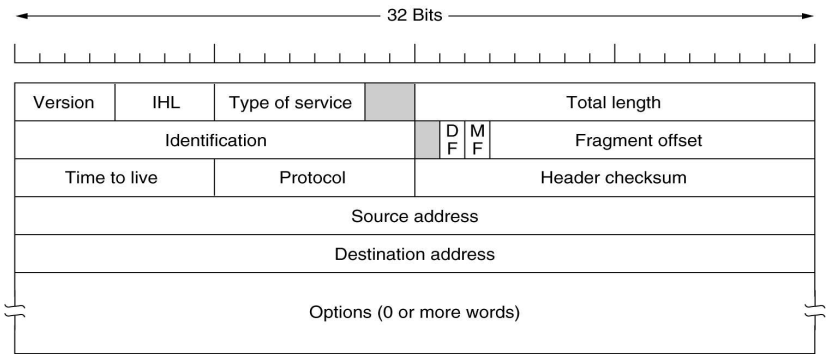
Naív megoldás



Multicast-fa

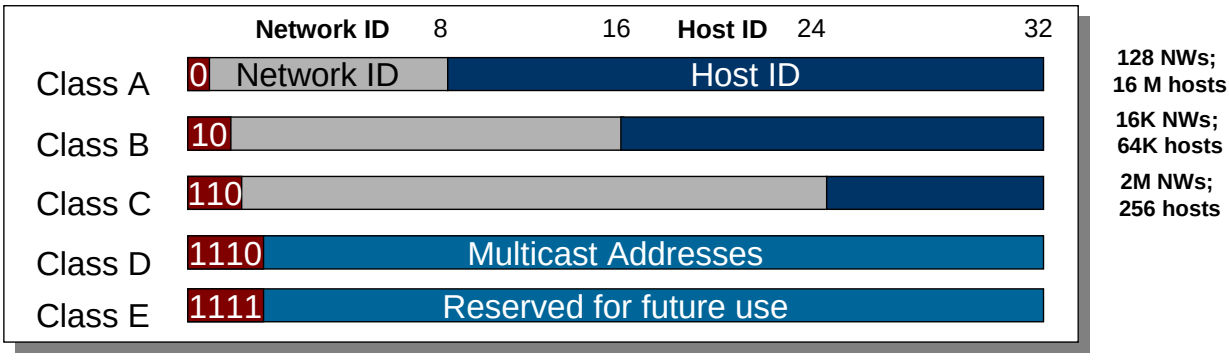
IPv4-Header (RFC 791)

- Version: 4 = IPv4
- IHL: fejléc hossz
  - 32 bites szavakban (>5)
- Type of Service
  - Optimalizálás „delay”, „throughput”, „reliability”, „monetary cost” szerint
- Checksum (csak a fejlécnek)
- Source és destination IP cím
- Protocol: azonosítja a megfelelő protokollt
  - pl. TCP, UDP, ICMP, IGMP
- Time to Live (TTL):
  - maximális hop szám



IPv4 címek

- Osztály alapú címzés (1993-ig)
  - 5 fix osztály, melyek mindegyikét egy prefix azonosítja
  - A, B, C osztály: fix hosszúságú hálózat prefix és host-ID
  - D a multicast osztály; E: lefoglalt

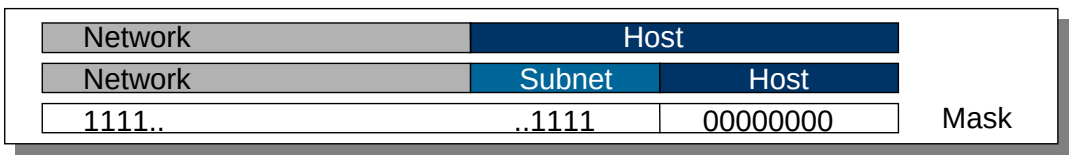


IPv4 címek

- Problémák:
  - A és B osztályú hálózatok sok állomást tartalmazhatnak, ami
    - a routerek számára nehezen kezelhető → **subnetting**
  - Elfogynak a címek: a címosztályok sok címet pazarolnak el
    - pl. egy szervezetnek 1000 állomással már B címosztályra lenne szüksége, ami elpazarol 64K-2K=62K címet → **classless addressing**

IPv4 címek -- Subnetting

- Képzeljünk el egy szervezetet, amely B címosztállyal rendelkezik és a szervezeten belül több LAN van (pl. Egyetem különböző Karai,…)
  - A központi routernek nem kell tudni minden állomásról, csak az egyes LAN-ok egy routeréről → új hierarchiaszint
- Subnetting
  - Bevezetünk alhálózatokat (subnet), melyeket az IP host-részéről leválasztott bitek azonosítanak (a hálózaton kívülről nem látható!)
  - A lokális routereknek tudni kell, hol van ez a leválasztás: ez a **subnet mask** által adható meg



  - Pl.: a subnet mask 11111111.11111111.11111111.00000000 azt jelenti, hogy
    - az IP cím 10000100.11100110.10010110.11110011
    - a 10000100.11100110 hálózatban a 10010110 alhálózatban
    - a 11110011 host-ot azonosítja

IPv4 címek -- Classless Inter Domain Routing (CIDR) (RFC 1519)

- A címoszályok nagyon sok IP címet pazarolnak
- 1993 óta: Classless Inter Domain Routing (CIDR)
  - A hálózat cím és a host-ID hossza szabadon adható meg hálózati maszk által.
    - Pl.: hálózati maszk 11111111.11111111.11111111.00000000
      - az IP cím 10000100.11100110.10010110.11110011
      - a 10000100.11100110.10010110 hálózatban
      - a 11110011 host-ot azonosítja
  - Ezek valódi hálózatokat jelentenek amit a többi router is lát – a subnetting-gel ellentétben
    - az IP csomagokba nem kell kiegészítés
    - többi routernek kezelni kell ezeket a változó hosszúságú hálózati címeket (forwarding és a routing algoritmus)

IPv4 címek -- CIDR

- Route aggregation
  - A BGP, RIP v2 és OSPF routing protokollok különböző hálózatokat egy prefix által kezelhetnek
    - Pl. minden hálózat, melynek a prefixe 10010101010* az X szomszédos routeren keresztül érhető el

IP cím lefordítása MAC címre: ARP (RFC 826)

- Address Resolution Protocol (ARP)
- IP cím MAC címre fordítása
  - Broadcast a LAN-ban, lekérdezni azt, hogy melyik állomáshoz tartozik az adott IP cím
  - Csak az a csomópont válaszol, amelyhez az IP tartozik, a MAC címmel
  - A router akkor a csomagot oda ki tudja szállítani

IPv6

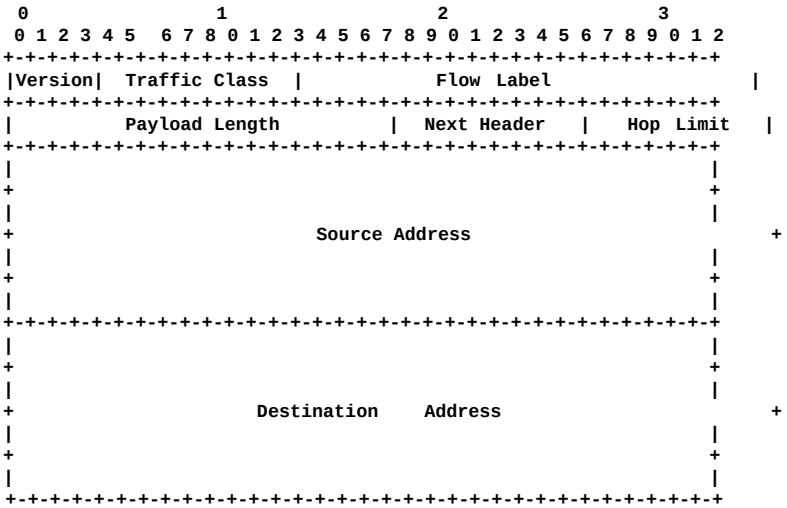
- 32 bites IP címek egyre szűkösebben állnak rendelkezésre
  - 4 milliárd ilyen IPv4 cím van (32 Bit) de
  - ezek statikusan hálózati és host-részre vannak osztva
    - címek mobil telefonoknak, hűtőszekrényeknek, autóknak, stb…
- Autokonfiguráció
  - DHCP, Mobile IP, átszámozás
- Új szolgáltatások
  - Biztonság (IPSec)
  - Quality of Service (QoS)
  - Multicast
- Egyszerűsítés a routernek
  - Nincs IP checksum
  - Nem partícionálja az IP csomagokat

A címek: DHCP

- DHCP (Dynamic Host Configuration Protocol)
  - Kézi hozzárendelés (hozzákötni a MAC címhez, pl. szervereknél)
  - Automatikus hozzárendelés (fix hozzárendelés, de nem előre beállított)
  - Dinamikus hozzárendelés (újrakiosztás lehetséges)
- Új számítógép kapcsolódása konfiguráció nélkül
  - A számítógép kér egy IP címet a DHCP szervertől
  - Az dinamikusan hozzárendel egy IP címet a számítógéphez
  - Miután a számítógép elhagyja a hálózatot, az IP cím újra kiosztható
  - Dinamikus hozzárendelés esetén az IP címeket „frissíteni” kell
  - Ha egy számítógép egy régi IP címet akar felhasználni, ami lejárt, vagy már újra ki van osztva
    - akkor a kéréseket vissza kell utasítani
  - Probléma: IP címek lopása

IPv6-Header (RFC 2460)

- Version: 6 = IPv6
- Traffic Class
  - QoS-hez (prioritásokhoz)
- Flow Label
  - QoS-hez, valós idejű alkamazásokhoz
- Payload Length
  - Az IP csomag fennmaradó részének (a datagrammnak) a hossza
- Next Header (mint IPv4-ben):
  - pl. ICMP, IGMP, TCP, EGP, UDP, Multiplexing, ...
- Hop Limit (Time to Live)
  - Hop-ok max. száma
- Source Address
- Destination Address
  - 128 Bit IPv6-Adresse



IPsec – Security Architecture for the IP (RFC 2401)

- Biztonsági protokollok
  - Authentication Header (AH)
    - Biztosítja az adat küldőjének authentifikációját,
    - kapcsolat mentes adat integritást,
    - védelemet Replay-támadásokkal szemben
  - Encapsulating Security Payload (ESP)
    - IP fejléc titkosítás nélkül, adatok titkosítva, authentifikálással
- Kulcs management:
  - IKE (Internet Key Exchange) Protokoll
  - Egy Security Association létrehozása
    - Security szolgáltatásokkal védett simplex kapcsolat két állomás, vagy egy állomás és egy security gateway (router, amely támogatja IPsec-et), vagy két security gateway között
    - Idientifikáció, kulcsok, hálózatok, megújítási időközök az authentifikációhoz és IPsec kulcsok rögzítése

IPsec

- IPsec transport üzemmódban (direkt kapcsolatokhoz)
  - IPsec fejléc az IP fejléc és az adatok között van
  - Megvizsgálják az IP routerek (azokban jelen kell lenni IPsec-nek)
- IPsec tunnel üzemmódban (ha legalább egy IPsec nélküli router közötte)
  - Az egész IP csomagot titkosítja és a IPsec fejléccel együtt egy új IP csomagba teszi
  - Csak a kapcsolat két végén kell hogy jelen legyen IPsec
- IPsec része az IPv6-nak
- porting IPv4-re létezik

IP címek és a Domain Name System (DNS)

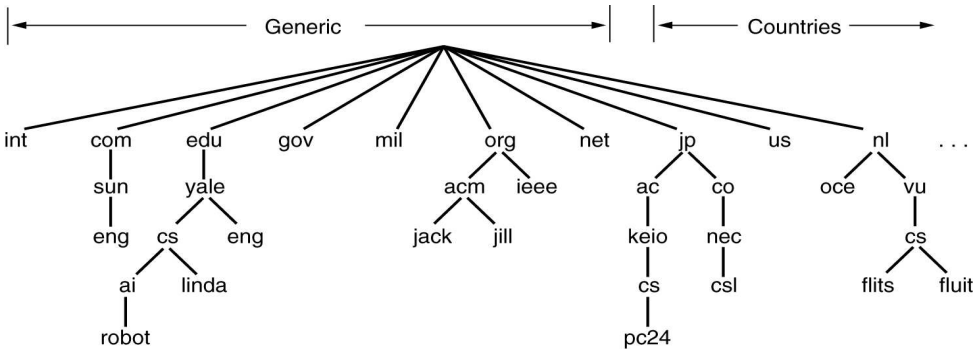
- IP címek
  - Minden hálózat interface egy hálózatban világszerte egyértelmű IP címmel rendelkezik
  - 32 bit, amely Net-ID és Host-ID-ra oszlik
  - Net-ID: az Internet Network Information Center adja ki
  - Host-ID: a helyi hálózat adminisztrátor adja ki
- Domain Name System (DNS)
  - Megfeleltet az IP-címnek egy nevet,
    mint pl. a 157.181.161.52 címnek a pandora.inf.elte.hu nevet
  - Elosztott robosztus adatbázis

Domain Name System (DNS)

- Az emberek számára 4 byte IPv4 cím nehezen kezelhetők:
  - 209.85.135.99 google.com-hoz
  - 157.181.151.154 az ELTE-hez
  - Mit jelent?
    - 207.46.19.30
    - 157.181.35.45
- Jobb: Természetes szavak az IP-címekhez
  - Pl. www.google.com
  - vagy www.elte.hu
- A Domain Name System (DNS)
  - lefordítja ezeket a címeket IP-címekre (és fordítva)
  - elosztott adatbázis

DNS – Felépítés

- DNS neveket képez le IP-címekre
  - Pontosabban: neveket erőforrás-bejegyzésekre
- A nevek hierarchikusan struktúráltak egy névtérben
  - Max. 63 jel komponensenként, összesen max. 255 jel
  - Minden domain-en belül, a domain tulajdonosa ügyeli fel a névteret a domain alatt



DNS Resource Record

- **Erőforrás bejegyzés** (resource record RR): a domain-ekről, egyes host-okról, stb... adnak információt
- RR formátum: (name, ttl, class, type, value)
  - name: pl. domain név vagy host név
  - ttl (time to live): érvényesség (másodpercben)
  - class: Internet esetén mindig “IN”
  - type: lásd a táblázatot
  - value: pl. IP-cím

Type	Meaning	Value
SOA	Start of Authority	Parameters for this zone
A	IP address of a host	32-Bit integer
MX	Mail exchange	Priority, domain willing to accept e-mail
NS	Name Server	Name of a server for this domain
CNAME	Canonical name	Domain name
PTR	Pointer	Alias for an IP address
HINFO	Host description	CPU and OS in ASCII
TXT	Text	Uninterpreted ASCII text

- RR Példa:
  pandora.inf.elte.hu. 43200 IN A 157.181.161.52

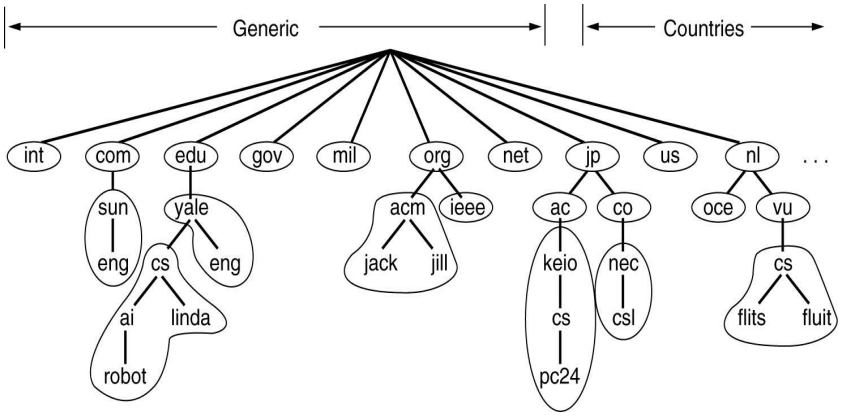
DNS Resource Records -- Példák

Példák RR tipusokra

- Type=A
  - name: egy végrendszer (host) neve
  - value: egy IP-cím
- Type=NS
  - name: egy domain (pl elte.hu)
  - value: a domain authoritative name server-jének az IP-címe
- Type=MX
  - value: a name-hez tartozó mail server neve
- Type=CNAME
  - name: egy alias név egy kanonikus névhez
  - value: a kanonikus név
- Type = SOA (start of authority)
  - name: a domain neve
  - value: szerverek neve, melyek a zónához tartozó mérvadó információkat rendelkezésre bocsátják, paraméterek a zónához
    - a zóna sorszáma,
    - frissítési intervallum a másodlagos szervernek,…

DNS Name Server

- A névtér **zónák**ra van osztva
- Minden zónához tartozik egy **Authoritativ Server** a mérvadó információval
  - Egy **Primary Name Server**
  - Továbbá egy vagy több **Secondary Name Server** a megbízhatóság miatt
- Minden Name Server ismeri
  - a saját zónáját
  - a gyermek-zónák Name-Server-jeit

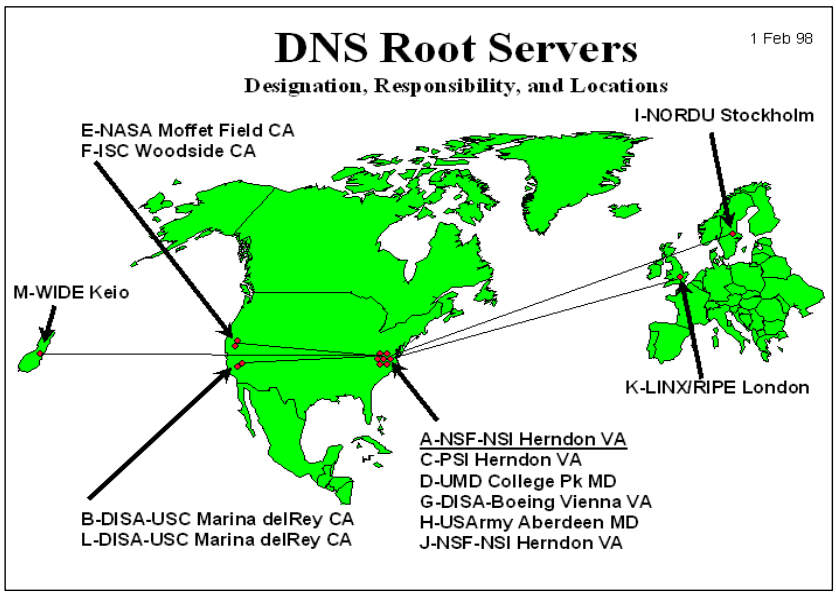


Servers/Resolvers

- Minden végrendszernek van egy „feloldója” (resolver)
  - Tipikusan egy könyvtár, amit felhasználásokhoz kapcsolhatunk
  - Lokális name-server-ek kézzel konfigurálva (pl. /etc/resolv.conf)
- Name servers
  - Tipikusan egy zónáért felelősek
  - Lokális szerverek
    - A lokális végrendszereknek végeznek lekérdezéseket távoli végrendszer nevekről
    - Megválaszolják a lekérdezéseket a lokális zónáról

DNS: Root Name Servers

- A “root” zonáért felelősek
- Jelenleg 13 root name server világszerte
  - A-M „számozva”
- Lokális szerverek kapcsolatba lépnek a root szerverrel, ha ők nem tudják megválaszolni a lekérdezést
  - Jól ismert root szerverekkel konfiguráltak



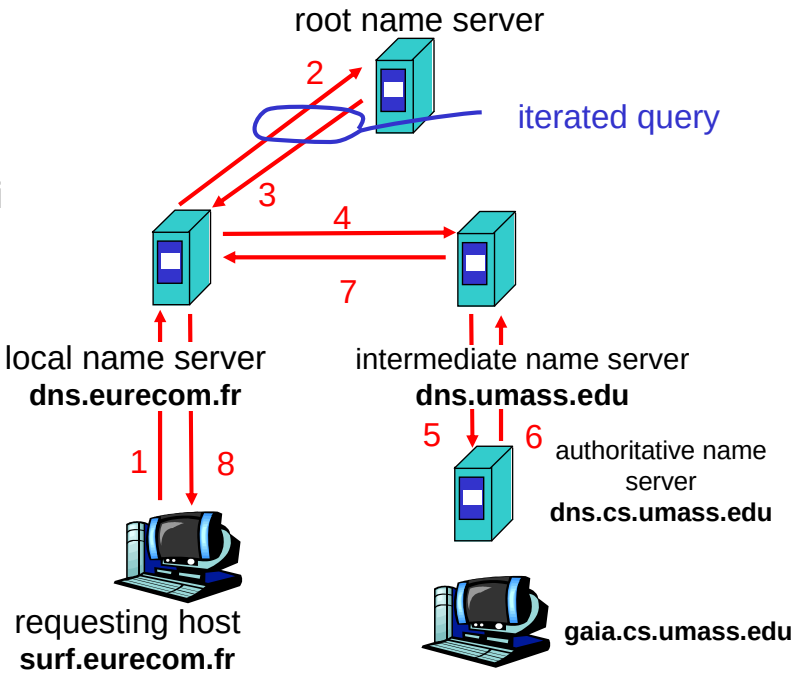
DNS lekérdezések

Iteratív lekérdezés:

- A megkérdezett szerver annyi információt ad a válaszban, amit ő maga tud
- Pl. annak a szervernek a nevét, akit meg kell kérdezni

Rekurzív lekérdezés:

- A megkérdezett szerver rekurzívan „kideríti” a hiányzó információt
- A lokális szerverek tipikusan rekurzív lekérdezési módban dolgoznak
- Root vagy távoli szerverek iteratívban



DNS üzenet formátum

12 bytes	Identification	Flags
	No. of Questions	No. of Answer RRs
	No. of Authority RRs	No. of Additional RRs
név, lekérdezés tipus mezeje	Questions (variable number of answers)	
Erőforrás bejegyzések a válaszban	Answers (variable number of resource records)	
Bejegyzések az „authoritative” szerverekhez	Authority (variable number of resource records)	
További „hasznos” információ	Additional Info (variable number of resource records	

Tipikus feloldási folyamat

- A www.inf.elte.hu név feloldásának lépései
  - A felhasználás hívja a gethostbyname() függvényt
  - A végrendszer lekérdezi a lokális name server-t (S_1)
  - S_1 lekérdezi a root server-t (S_2) a www.inf.elte.hu névvel
  - S_2 válaszol a elte.hu-hoz (S_3) tartozó NS bejegyzéssel
  - Honnan tudjuk meg az A bejegyzést S_3 -hoz
    - Erre való az „additional information section”
  - S_1 lekérdezi S_3 -t a www.inf.elte.hu névvel
  - S_3 válaszol a www.inf.elte.hu-hoz tartozó A bejegyzéssel
- Több A bejegyzés is érkezhet a válaszban → mit jelent ez?

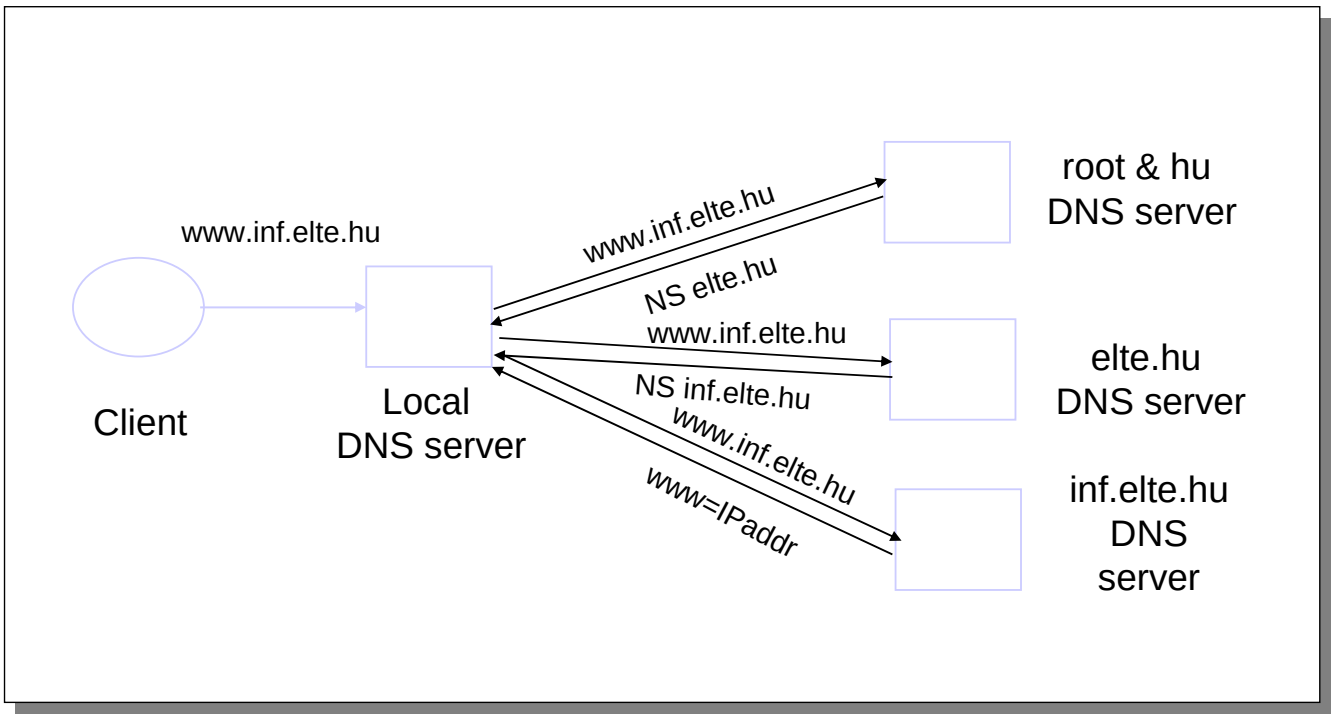
Caching

- DNS válaszok tárolódnak az érintett szervereken (caching)
  - Gyors válasz ismételt lekérdezés esetén
  - Más lekérdezések bizonyos részeket újra felhasználhatnak a válaszból
    - Pl. NS bejegyzéseket a domain-ekhez
- DNS negatív lekérdezések tárolódnak a cache-ben
  - Ne kelljen megismételni a kudarcot
  - Pl. elgépelés
- A cache-ben tárolt adatok érvényessége egy idő után lejár
  - Az érvényesség idejét (TTL) az adat tulajdonosa határozza meg
  - Minden bejegyzés tartalmaz TTL-t

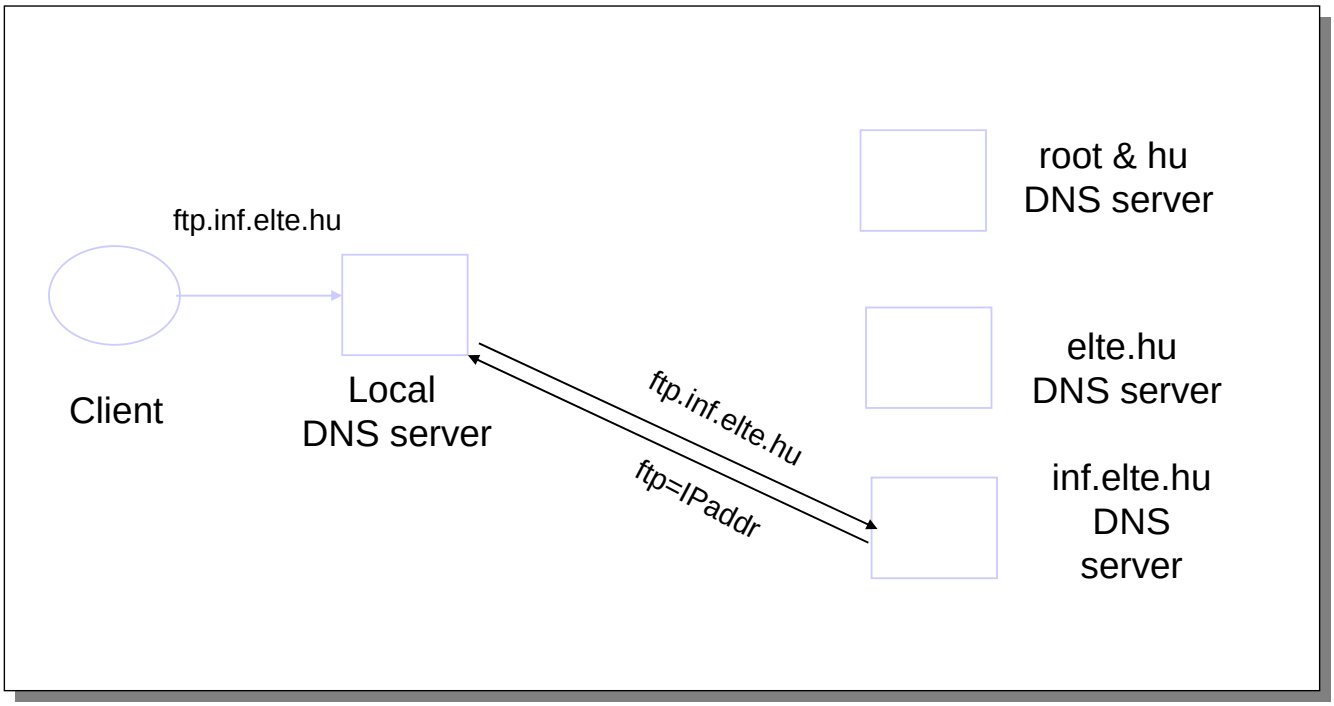
Prefetching

- Name server minden válaszhoz adhat további adatokat
- Tipikusan prefetcing-hez használják
  - CNAME/MX/NS tipikusan más végrendszer nevére mutat
  - Válaszok tartalmazzák a végrendszerek címeit, amelyekre mutatnak az “additional section” részben

DNS lekérdezés példa



Példa egy későbbi lekérdezésre

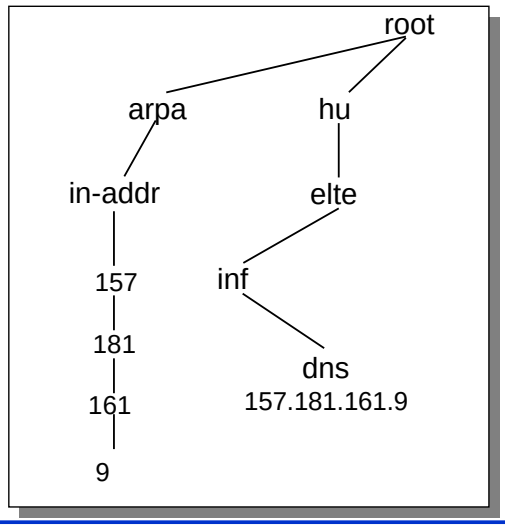


Megbízhatóság, rendelkezésre állás

- DNS szerverek replikáltak
  - A name service müködik, ha egy replika működik
  - A lekérdezések kiegyensúlyozhatók a replikák között (load balancing
- UDP-t használ a lekérdezéshez
  - Megbízhatónak kell lenni → Miért nem TCP?
  - Timeout esetén alternatív szervert próbál
  - „Exponential backoff”, ha visszatér ugyanahhoz a szerverhez
  - Ugyanaz az azonosító minden lekérdezéshez
    - Mindegy melyik szerver válaszol

Reverse Name Lookup

- Melyik számítógéphez tartozik az 157.181.161.9 IP-cím?
  - Lekérdezés: 9.161.181.157.in-addr.arpa
  - Miért van megfordítva a cím?
  - dns.inf.elte.hu



Dinamikus DNS

- Probléma
  - Időlegesen hozzárendelt IP-címek
  - Pl. DHCP által
- Dinamikus DNS
  - Amint egy csomópont egy új IP-címet kap, regisztrálja azt azon a DNS-szerveren, amely őérte felelős
  - Rövid TTL bejegyzések biztosítják azt, hogy a bejegyzések gyorsan aktualizálódjanak
    - egyébként a lekérdezések rosz számítógépre irányítódnának
- Felhasználás
  - Egy privát domain regisztrálása
  - lásd www.dyndns.com