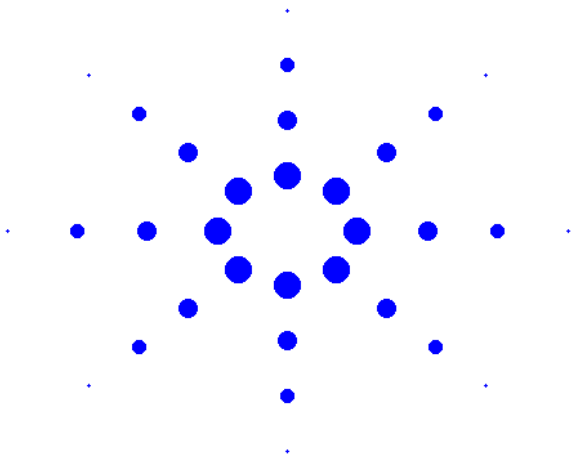
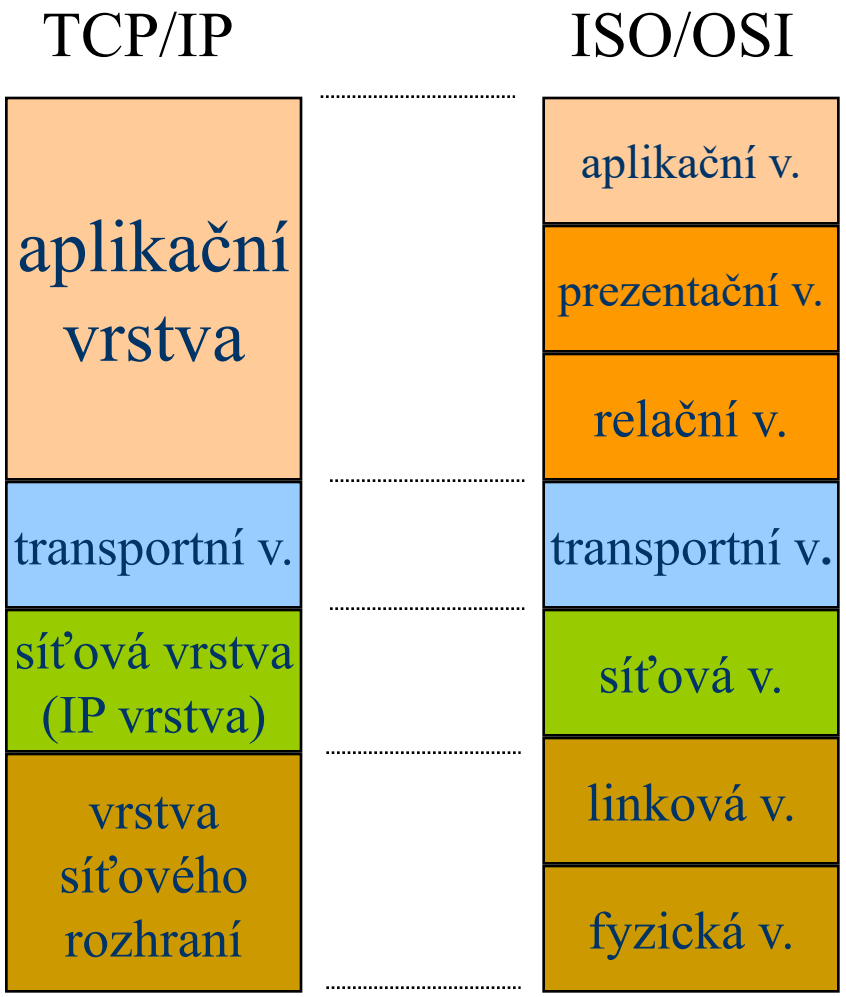
Počítačové sítě - B4B38PSIA

- Protokolový zásobník
- IP adresy
- IPv4, IPv6
- ICMP, ARP a NDP
- UDP, TCP
- DNS



Historie TCP/IP, ISO/OSI Model

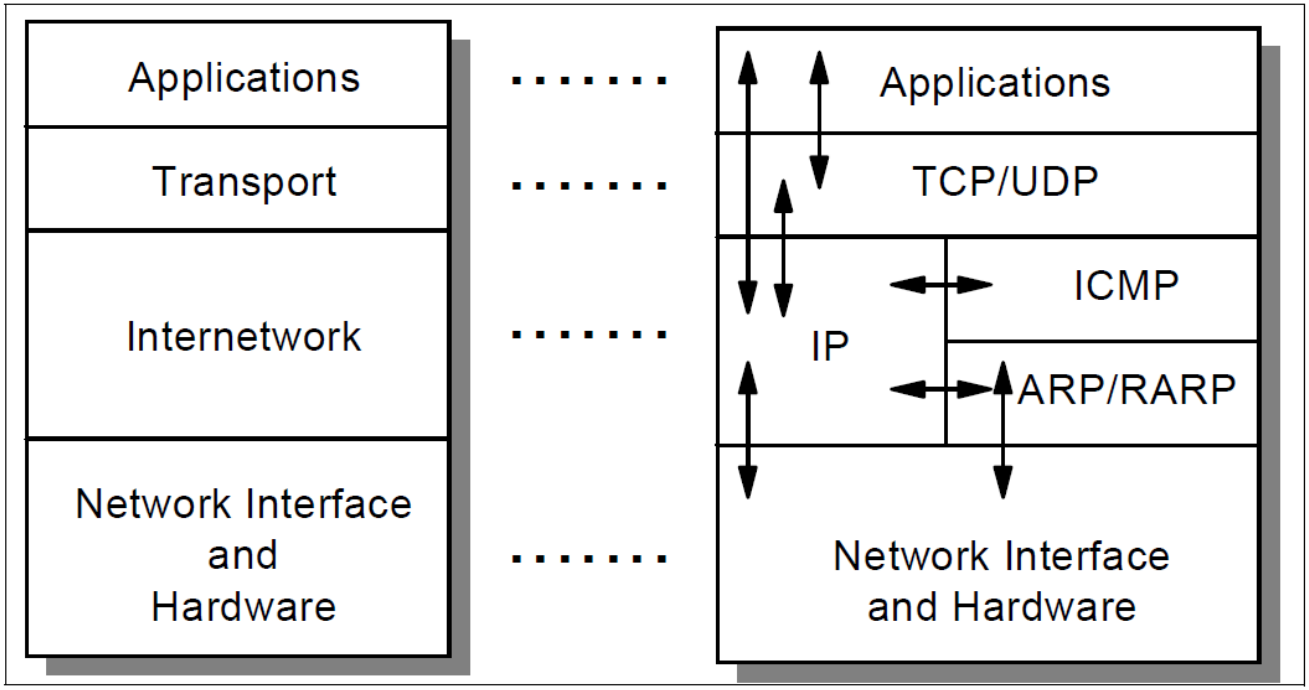
- TCP/IP je starší
  - neodpovídá plně ISO/OSI
- Vývoj v 70. letech minulého století
- 1971 - ARPANET working group
  - financováno DARPA
  - spolupráce ITU, ISO
- 1978 – dokončeno
- 1980 – ARPANET přechází na TCP/IP
- 1983 – otevřená TCP/IP implementace v BSD Unixu



Další vývoj TCP/IP

- NSFNET (National Science Foundation)
  - univerzitní a výzkumná síť
  - propojení s ARPANETem na Carnegie Mellon University
  - uzly sítě propojeny prostřednictvím pronajatých linek
    - 1. generace – 56 kbit/s
    - 2. generace – 448 kbit/s
    - 3. generace (1989) – 1.544 Mbit/s (T1)
    - 1992 – migrace na T3 – 45 Mbit/s
  - 1995 – skončila podpora vlády USA
- Projekt Internet2
  - http://www.internet2.edu
- Komerční páteřní sítě
- …

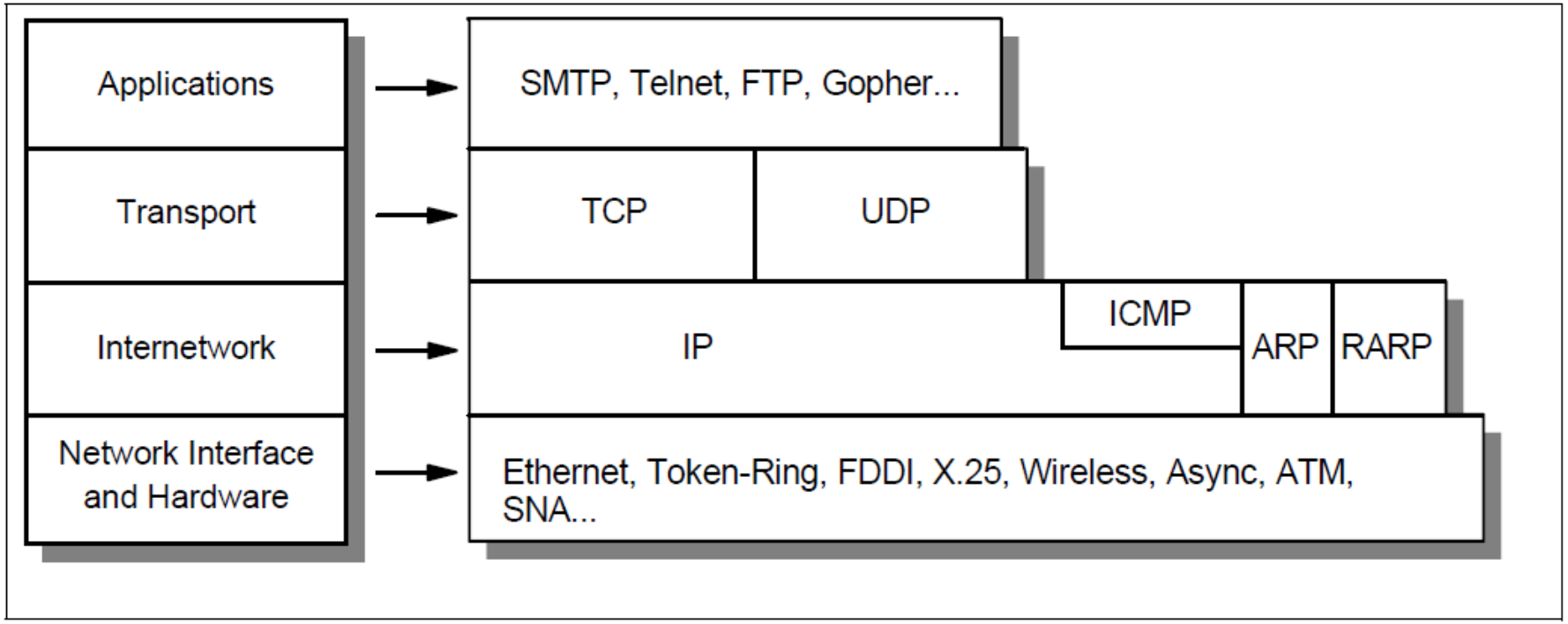
Protokolový zásobník TCP/IP



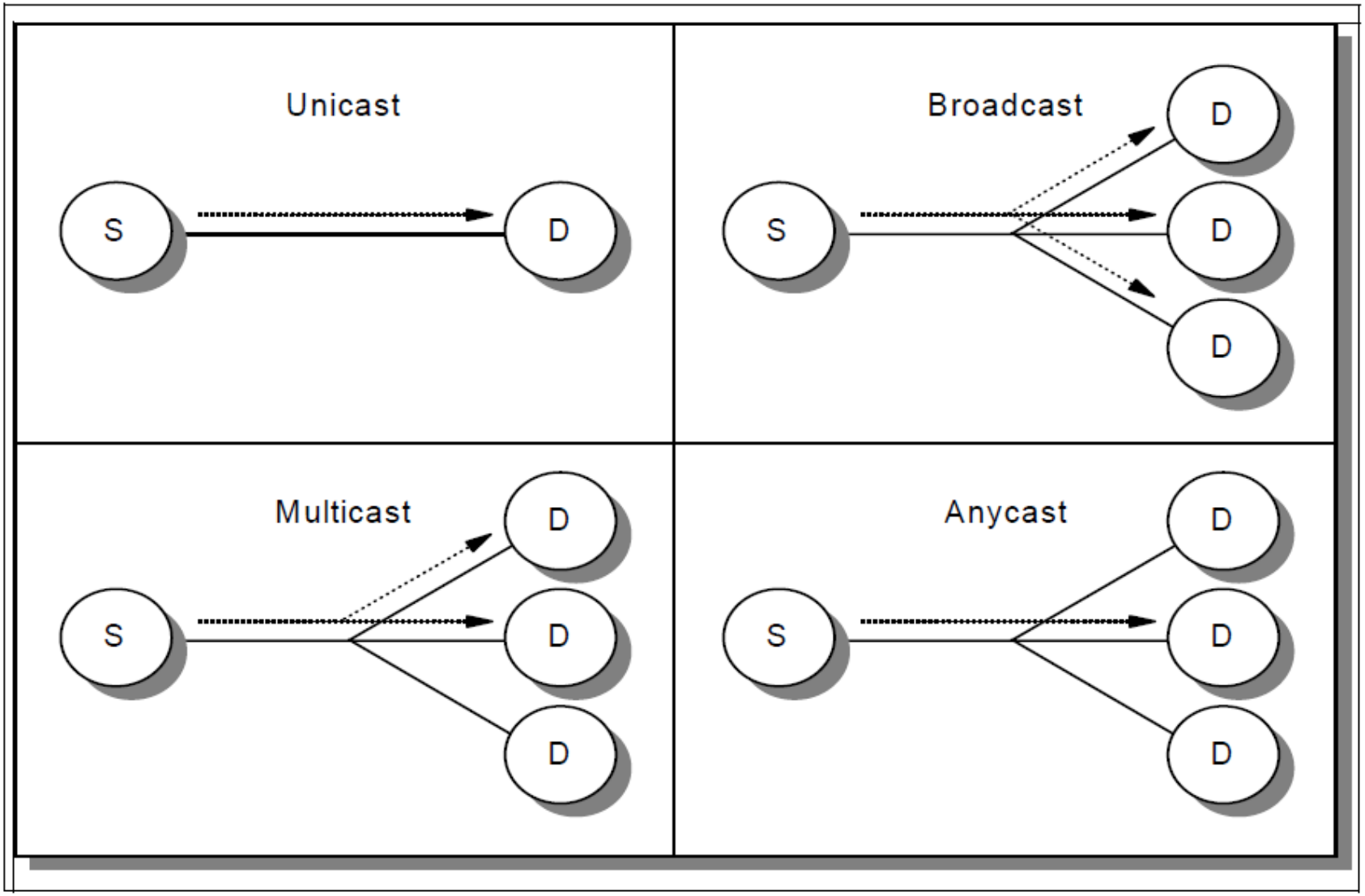
- Nebyl navržen dle OSI modelu, ale
  - Network interface and hardware ≈ fyzická a spojová vrstva
  - Internetwork ≈ síťová vrstva
  - Transport ≈ transportní vrstva
  - Applications ≈ relační až aplikační vrstva

Protokolový zásobník TCP/IP

- Příklady protokolů jednotlivých vrstev
  - nejedná se o kompletní sady protokolů TCP/IP



IP Protokol – Doručení dat na IP adresu



IP Protokol – Doručení dat na IP Adresu

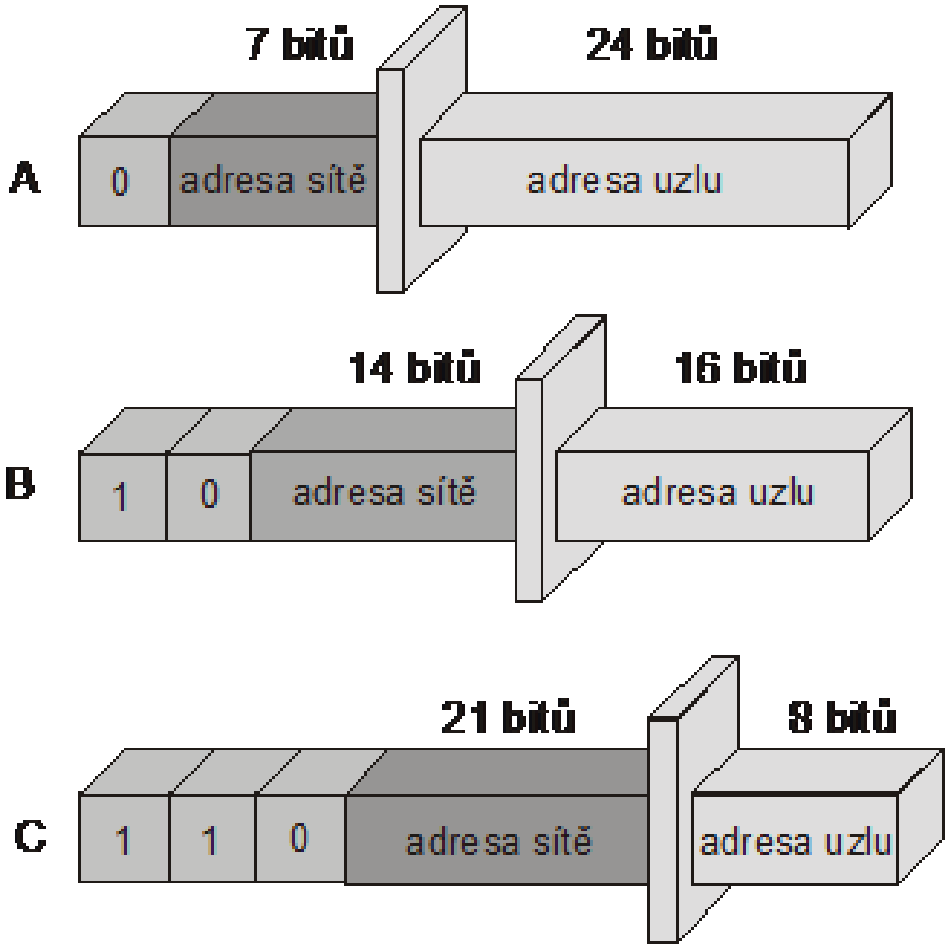
- Unicast
  - právě jeden příjemce
  - jediná možnost pro spojově-orientované technologie
- Broadcast
  - limited broadcast: 255.255.255.255, ff0x::1
    - dosah je omezen směrovači
  - directed broadcast: např. 169.47.255.255
    - směrován do cílové sítě
- Multicast
  - doručení vybrané skupině hostitelů
  - adresy třídy D nebo multicastový prefix ff00::/8
- Anycast
  - nejbližší příjemce ze skupiny možných poskytovatelů služby (IPv6)

IP Protokol – IP Adresy

- IPv4 … IPv6 adresy
  - 4 resp. 16 bajtů
  - Správa IP adres je regionální
    - American Registry for Internet Numbers (ARIN, Ameriky + Afrika)
    - Reseaux IP Europeans (RIPE, Evropa, střední východ, Afrika)
    - Asia Pacific Network Information Centre (APNIC, Asie, Austrálie …)
- IPv4 adresa 10.0.0.15 (4 dekadická čísla v rozsahu 0 ÷ 255)
  - IP address = <network number><host number>
  - původně se používaly tzv. třídy adres
    - A – 2^7-2 (126) sítí, 2^{24}-2 (16777214) hostitelů
    - B – 2^{14}-2 (16382) sítí, 2^{16}-2 (65534) hostitelů
    - C – 2^{21}-2 (2097150) sítí, 2^8-2 (254) hostitelů

IP Protokol – třídy adres

- Původně 3 třídy adres
  - nízká granularita
  - plýtvání adresovým prostorem
- Třída adres D
  - 1110…
  - pro multicastovou komunikaci
- Třída adres E
  - 11110…
  - pro budoucí definice a experimenty



IP Protokol – IP Adresy

- Vylepšená správa IPv4 adres
  - hranice položky <host number> definována 32bitovou maskou podsítě
  - např. 255.255.252.0 – 2^{10} – 2 uzlů v síti
- Classless InterDomain Routing - CIDR
  - IP adresa = <network number><subnetwork number><host number>
  - maska podsítě v síti třídy B je např. 255.255.255.240
    - prvních 12 bitů definuje číslo podsítě, zbývající 4 adresu hostitele
    - 2^{12}-2 (4094) možných podsítí, pouze 2^{4}-2 (14) hostitelů v podsíti
  - statická definice (všechny podsítě mají shodnou velikost)
  - definice s proměnnou délkou (různé velikosti podsítí)
  - neviditelné pro uzly vně místní sítě, využívané jen uzly v rámci sítě
  - lokální management

IP Protokol – IP Adresy

- Vyhrazené IPv4 adresy
  - netid, hostid=0: adresa sítě (147.229.0.0)
  - netid=0, hostid: adresa uzlu (0.0.0.5)
  - 255.255.255.255 limited broadcast (omezená všeobecná adresa), neprochází přes router
  - netid, hostid=11..1 directed broadcast (147.229.255.255), řízená všeobecná adresa
  - 127.x.x.x loopback, softwarová zpětnovazební adresa (komunikace mezi procesy počítače bez vysílání na síť)
- Privátní (neveřejné) IPv4 adresy
  - 10.0.0.0: síť třídy A
  - 172.16.0.0 až 172.31.0.0 - 16 sítí třídy B
  - 192.168.0.0 až 192.168.255.0: 256 sítí třídy C
  - nejsou směrovány ve veřejném Internetu

IP Protokol – IP Adresy

- IPv6 adresy (16 bajtů)
  - individuální (unicast)
  - skupinová (multicast)
  - výběrová (anycast)
- Jedno rozhraní uzlu může (musí) mít několik adres
- Zápis IPv6 adresy
  - 8 skupin po 4 šestnáctkových číslicích
    - např. fedc:ba98:7654:3210:fedc:ba98:7654:3210
  - lze vynechat počáteční nuly
    - místo 0123:0000:0000:0000:fedc:ba98:7654:3210 lze psát 123:0:0:0:fedc:ba98:7654:3210
  - lze nahradit skupinu nul
    - místo 123:0:0:0:fedc:ba98:7654:3210 lze psát 123::fedc:ba98:7654:3210
    - dvojici „::“ lze použít jen jednou

IP Protokol – IP Adresy

- Zápis IPv6 adresy v tzv. kanonickém tvaru
  - šestnáctková čísla malými písmeny
  - vynechání počátečních nul je povinné
  - konstrukce „::“ musí mít maximální efekt
    - dvojici „::“ se musí použít na nejdelší (případně první) sekvenci nul
- Při použití v URL se IPv6 adresa uzavírá do hranatých závorek
  - http://[2002:d91f:cd32::1]/
- Určení příslušnosti k síti nebo podsíti vyjadřuje tzv. prefix
  - IP adresa/délka prefixu
    - 12ab:0:0:cd30:0:0:0:0/60
  - prefix lze zapsat v kanonickém tvaru
    - 12ab:0:0:cd30::/60

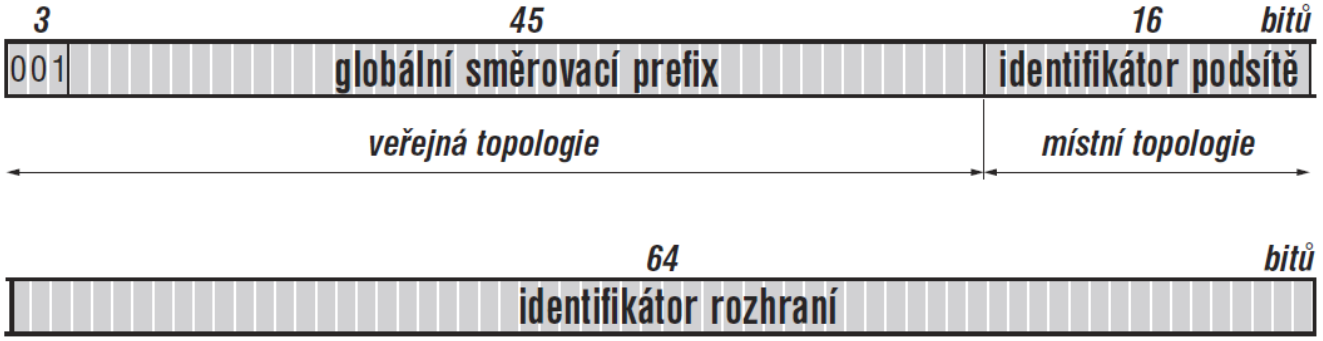
IP Protokol – IP Adresy

- Typy IPv6 adres
  - ::/128 - nedefinovaná adresa
  - ::1/128 - loopback
  - fc00::/7 - unikátní individuální lokální adresa
  - fe80::/10 - individuální lokální linková adresa
  - ff00::/8 - skupinová adresa
  - ostatní - individuální globální adresy
    - využívá se prefix 2000::/3
  - aktuální info lze nalézt na
    - http://www.iana.org/assignments/ipv6-address-space

- Známé prefixy
  - 64:ff9b::/96 - adresy s vloženou IPv4 (IPv6 klient, IPv4 server – NAT64)
  - 2002::/16 - přenos IPv6 přes IPv4 sítě

IP Protokol – IP Adresy

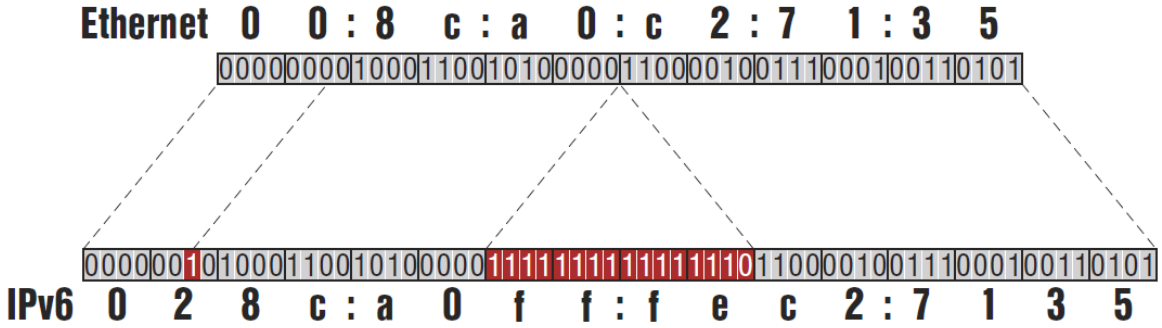
- Globální individuální IPv6 adresa
  - 2000::/3



- Globální směrovací prefix
  - typicky 48 bitů
    - i delší pro malé sítě
  - identifikuje koncovou síť, přidělen zmocněnou autoritou zvenčí
- Identifikátor podsítě
  - typicky 16 bitů
- Identifikátor rozhraní
  - vždy 64 bitů (kromě adres s prefixem ::/3)

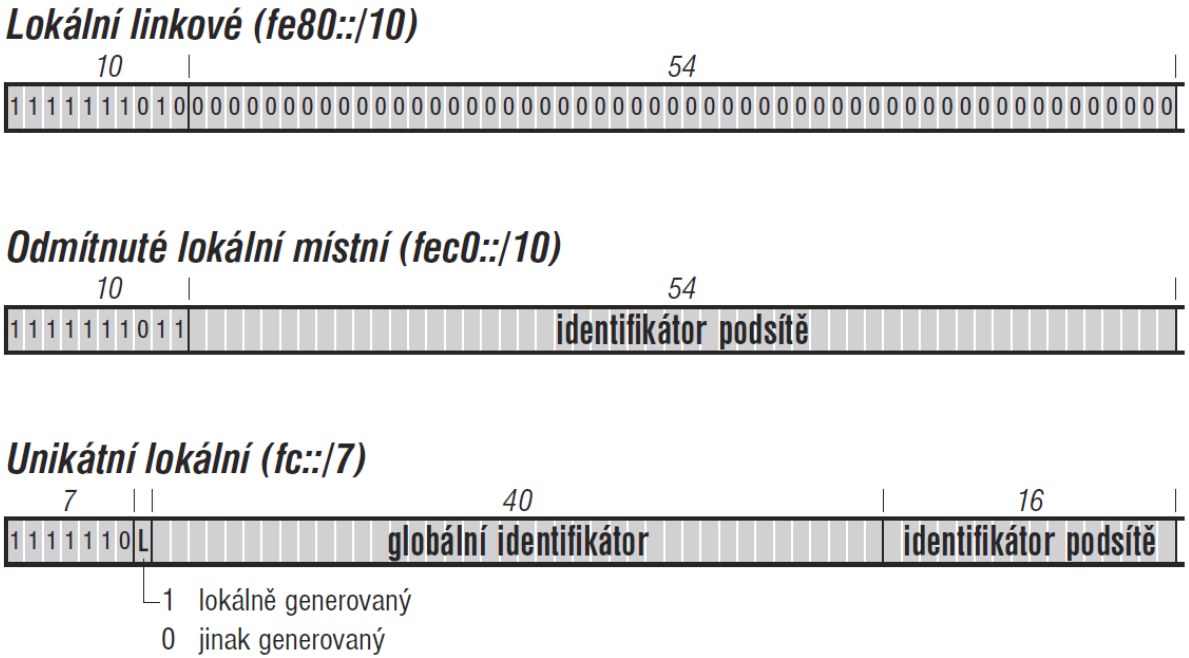
IP Protokol – IP Adresy

- Identifikátor rozhraní
  - vždy 64 bitů (kromě adres s prefixem ::/3)
  - odvozen dle IEEE EUI-64
- Pokud rozhraní již EUI-64 identifikátor má
  - převezme se s invertovaným druhým nejméně významným bitem nejvyššího bajtu
    - tzv. modifikované EUI-64
    - kvůli snadnému vytváření jednoduchých lokálních identifikátorů
  - příznak globality
    - EUI-64 - 0 celosvětově globální, 1 lokální identifikace
    - modifikované EUI-64 – opačně
- Vytvoření modifikovaného EUI-64 z Ethernet MAC adresy



IP Protokol – IP Adresy

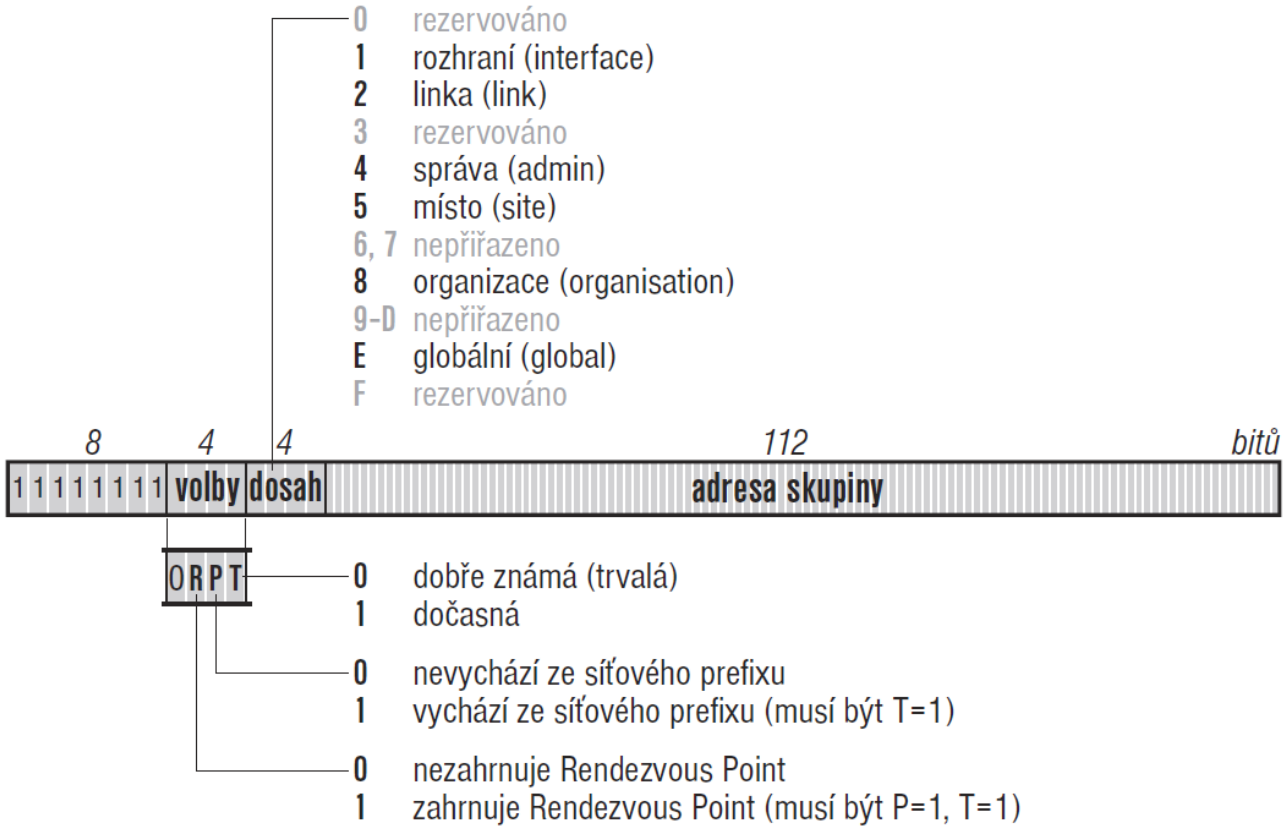
- Individuální lokální IPv6 adresy
  - podobné neveřejným adresám v IPv4



  - následuje modifikovaný EUI-64 identifikátor, hostitel si je generuje sám
- Unikátní lokální adresy (náhrada site-local)
  - příznak L = 0 pro budoucí centrální autoritu, zatím vše lokálně
  - „unikátní“ globální identifikátor (hash adresy a času generující stanice)

IP Protokol – IP Adresy

- Skupinové IPv6 adresy

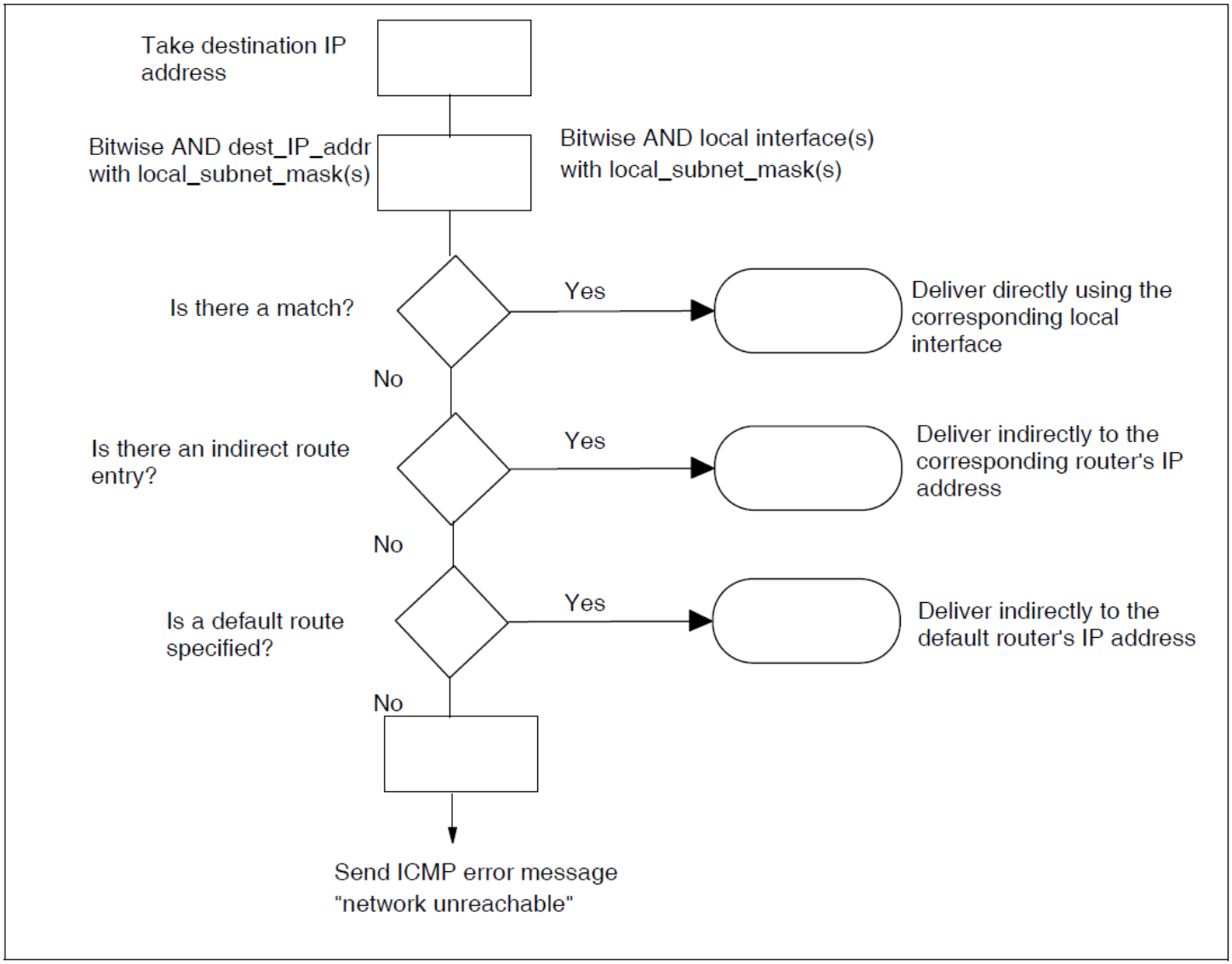


  - nesmí být použita jako adresa odesilatele, adresy ff0x::0 jsou rezervovány

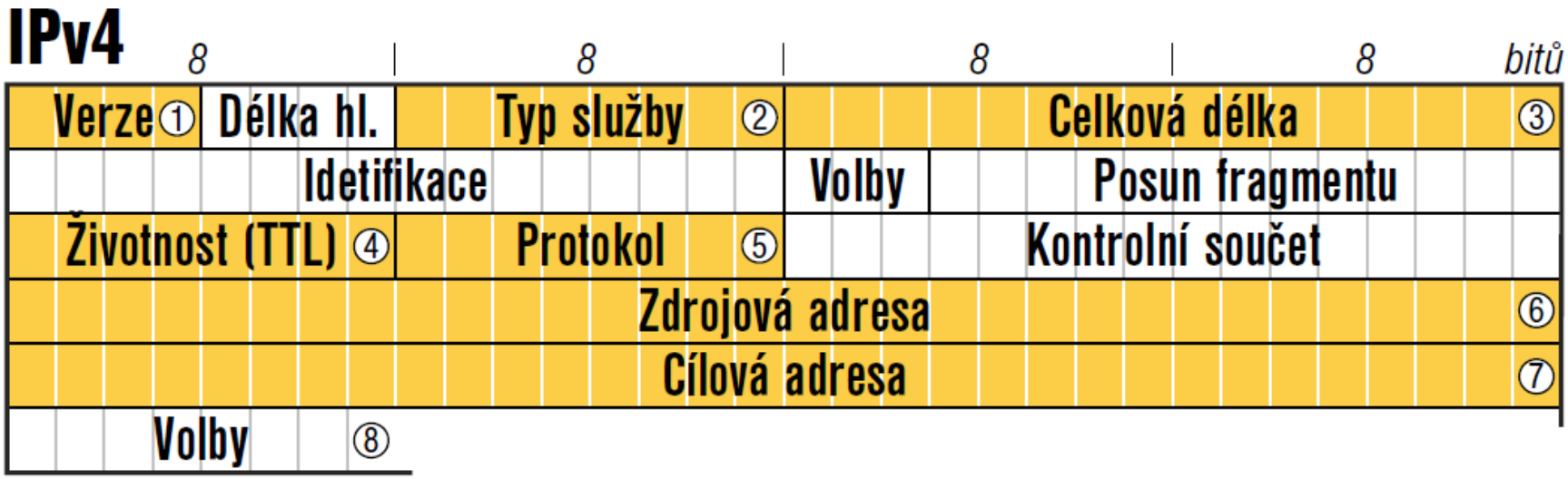
IP Protokol – IP Adresy

- Skupinové IPv6 adresy
  - dobře známé adresy jsou přidělovány centrálně (IANA)
    - dočasné si volí aplikace
  - dosah (ff0*x*::101 je skupinová adresa pro NTP servery)
    - ff02::101 NTP servery na stejné lince (např. Ethernetu)
    - ff05::101 NTP servery v daném místě (lokalitě)
    - ff0e::101 NTP servery v celém Internetu
  - přidělování skupinových adres
    - 0–3fff:ffff skupiny přidělené IANA (definice celé adresy, RFC 2375)
    - 4000:0000–7fff:ffff identifikátory přidělené IANA
    - 8000:0000–ffff:ffff dynamické, volně k použití
  - existuje několik způsobů tvorby skupinových adres
    - odvozeno z individuálního prefixu sítě
    - část z nich pro tzv. Source Specific Multicast (IP rádio či televize, prefix ff3x::/96)
    - odvozeno od identifikátoru rozhraní
    - odvozeno od tzv. rendezvous pointu (prefix ff70::/12)
  - předdefinované adresy
    - ff01::1 resp. ff02::1 – broadcast v daném dosahu
    - ff01::2 až ff05::2 – broadcast na všechny směrovače v daném dosahu

IP Protokol – IPv4 „Transmission Routing“



IPv4 paket

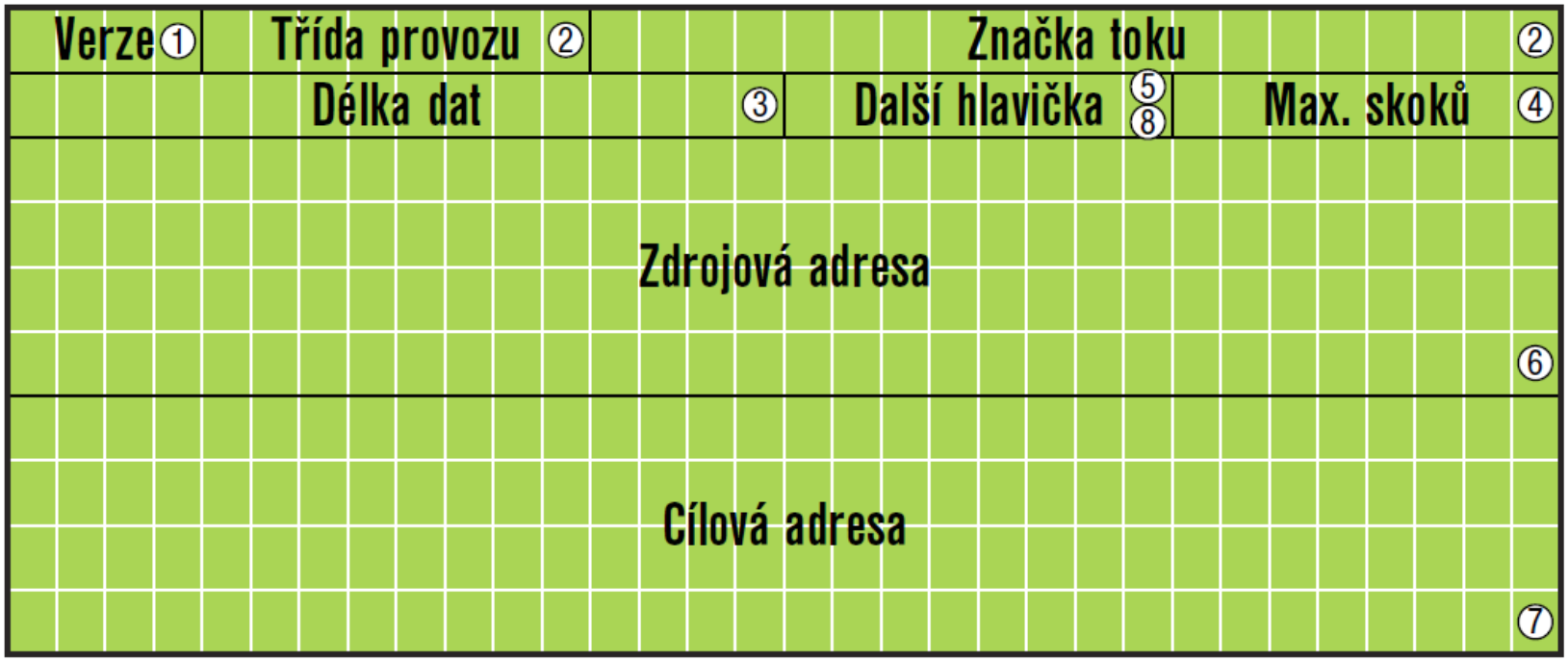


- Verze – 4 pro IP v4
- Délka hlavičky – délka hlavičky v 32 bitových slovech, obvykle 5
- Typ služby – často ignorováno, obsahuje požadavky na prioritu atp.
- Celková délka – celková délka IP paketu včetně hlavičky (až 64 kB)
- Identifikce – použit pro identifikaci fragmentů téhož IP paketu

IPv4 paket

- Volby – tříbitové pole, první bit vždy 0, druhý označuje možnost fragmentace (0 – ano, 1 – ne), třetí označuje, zda se jedná o poslední fragment paketu (0 – ano, 1 – ne)
- Posun fragmentu – offset dat od počátků původního paketu v násobcích 8 bajtů
- Životnost (Time To Live) – omezuje počet průchodů paketu směrovači a zamezuje jeho nekonečnému kroužení v síti. Každý směrovač snižuje hodnotu o 1, je-li nulová, je paket zahozen.
- Protokol – udává protokol nesený v IP paketu (hodnota 1 pro ICMP, 6 pro TCP …)
- Kontrolní součet – kontrolní součet hlavičky
- Zdrojová adresa – IP adresa odesilatele
- Cílová adresa – IP adresa příjemce
- Volby – rozšíření o další funkce, běžně se nepoužívá
- Padding – zarovnání pole Volby na 32 bitů

IPv6 paket



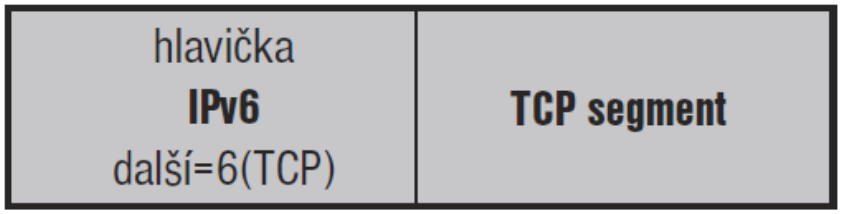
- Verze – 6 pro IPv6
- Třída provozu – možnost různého zacházení s IP pakety sítí (*diffserv*)
- Značka toku – možnost označit sekvenci IP paketů
- Délka dat – celková délka IP paketu bez standardní hlavičky (až 64 kB)
- Další hlavička – typ následující rozšiřující hlavičky
- Max. skoků – maximální počet průchodů paketu směrovači (TTL)

IPv6 Paket – zřetězené hlavičky

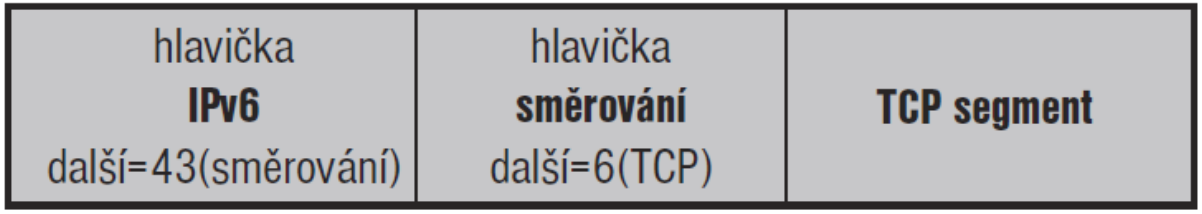
- Rozšiřující hlavičky – položka Další hlavička (Next header)
  - 0 – volby pro všechny (hop by hop options)
  - 43 – směrování
  - 44 – fragmentace
  - 50 – šifrování
  - 51 – autentizace
  - 59 – poslední hlavička
  - 60 – volby pro cíl (destination options)
- Nahrazuje položku Protokol (IPv4)
  - 6 – TCP
  - 8 – EGP
  - 9 – IGP
  - 17 – UDP
  - 46 – RSVP
  - 47 – GRE
  - 58 – ICMP

IPv6 Paket – zřetězené hlavičky

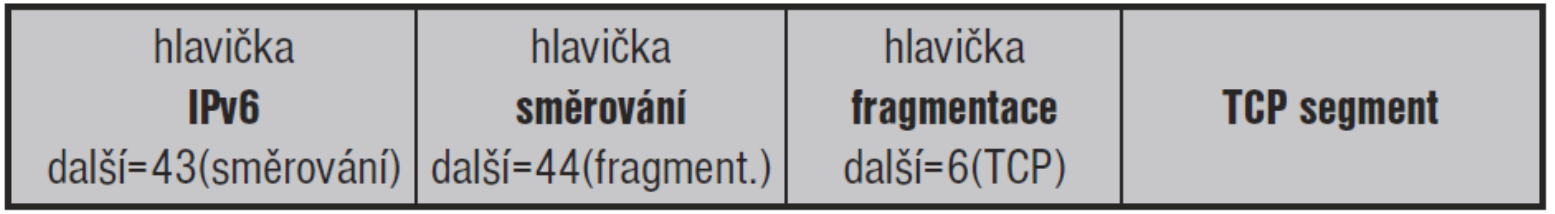
- Rozšiřující hlavičky – položka Další hlavička (Next header)
  - pořadí je předepsáno (kvůli náročnosti a rychlosti zpracování)



a) bez rozšiřujících hlaviček



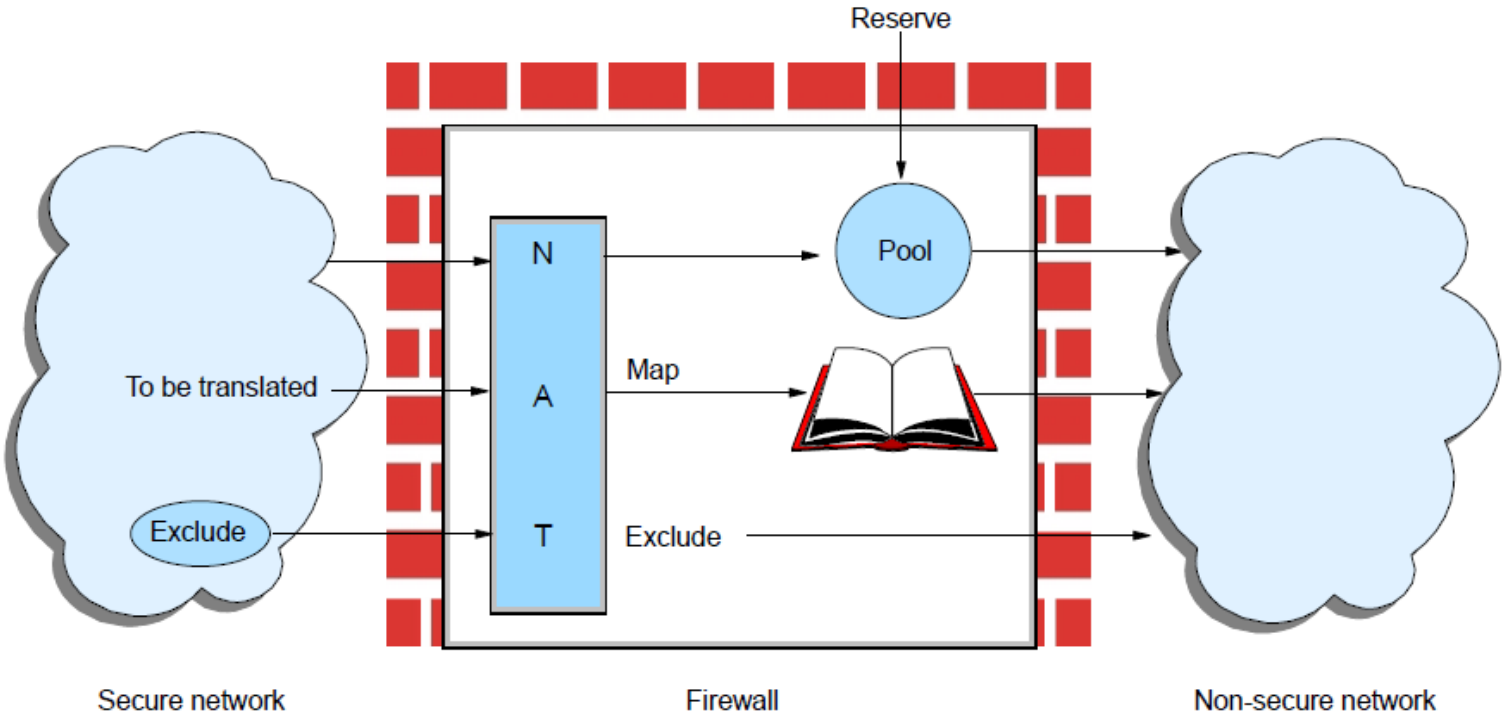
b) s hlavičkou *Směrování*



c) s hlavičkami *Směrování* a *Fragmentace*

NAT – Network Address Translation

- Předpokladem je, že pouze nízký počet uživatelů z privátní sítě potřebu současně přístup do veřejného Internetu
- Jejich privátní adresa je při přechodu do veřejné části sítě nahrazena adresou veřejnou (přeložena) a naopak



Problémy s překladem adres

- Network Address Translation (NAT)
  - NAT by měl být transparentní pro komunikující uzly
    - vyžaduje nový výpočet CRC v hlavičkách vyšších vrstev
    - problém např. s FTP - IP adresa je i v datovém poli aplikačního paketu, musí být také správně přeložena
    - I další aplikace mohou mít obdobné problémy
  - otázka – kdy vracet veřejnou adresu do bloku adres volných pro překlad?
    - IP protokol je bezestavový – nejsme schopni zjistit, zda bude komunikace pokračovat
    - TCP poskytuje informaci o stavu spojení, UDP ne
    - vrací se, pokud není nějakou dobu využita – timeout (obvykle desítky sekund až jednotky minut)
  - NAT musí být přednastaven pro komunikaci iniciovanou z veřejné sítě
    - např. interní mail server

Varianta překladu NAPT

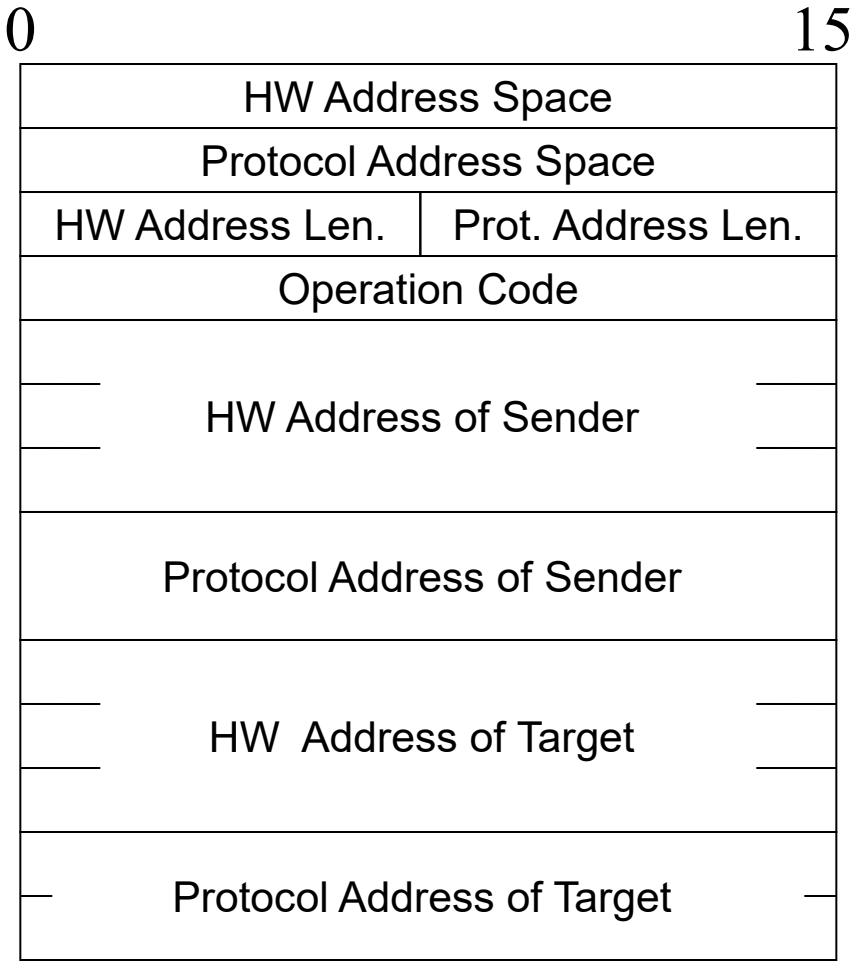
- Network Address Port Translation (NAPT)
  - kromě IP adresy se překládá i číslo portu (buď TCP nebo UDP) nebo číslo dotazu (query ID u ICMP)
- Celá privátní síť je tak z pohledu veřejné sítě vnímána jako jeden uzel s mnoha současně komunikujícími procesy
- NAPT není schopen správně překládat fragmentované pakety v případě komunikace více uzlů z privátní sítě s jedním uzlem sítě veřejné
  - číslo portu portu je nahrazeno „fragmentation ID“ ve fragmentech
  - uzel ve veřejné síti pak není schopen rozlišit pakety s tímtéž „fragmentation ID“ od dvou různých privátních uzlů
    - nepravděpodobné, leč možné

ARP Protokol

- Address Resolution Protocol
  - poskytuje konverzi IP adresy na MAC adresu
    - spojová vrstva neumí s IP adresou pracovat
    - spojová vrstva posílá rámce na MAC (linkovou) adresu
  - uzel sítě si udržuje tabulku s relacemi IP→MAC
    - ARP cache
  - pokud MAC pro požadovanou IP není nalezena, ARP ji vyžádá
    - pošle dotaz broadcastem na spojové vrstvě (MAC broadcast)
    - příjemci vyhodnotí, zda je dotazovaná IP adresa jejich
  - ARP odpověď obsahuje požadovanou MAC adresu
    - nový záznam (aktualizace) ARP cache
    - ARP odpověď je adresovaná odesilateli ARP dotazu
  - ARP cache lze aktualizovat při příjmu ARP žádostí
    - možná jen aktualizace, nikoliv vytvoření nového záznamu

ARP Protocol

- Address Resolution Protocol



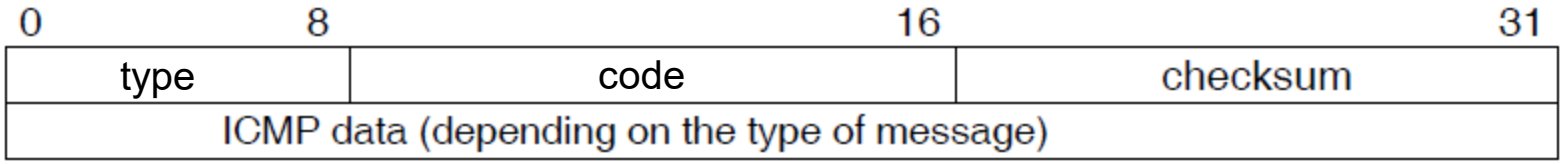
- HW Addr. Space – 1 pro Ethernet
- Prot. Addr. Space – 0x0800 pro IP
- HW Addr. Len. – 6 pro Eth. MAC
- Prot. Addr. Len. – 4 pro IPv4
- Operational Code
  - 1 – žádost, 2 – odpověď
- HW Address of Sender
  - MAC adresa odesilatele
- Protocol Address of Sender
  - IP adresa odesilatele
- HW Address of Target
  - cílová MAC adresa
- Protocol Address of Target
  - cílová IP adresa

ICMP Protocol

- Slouží k hlášení problémů při zpracování IP paketů směrovači
  - ICMP zpráva je v datovém poli IP paketu (protocol = 1, další hlavička = 58)
  - povinná část implementace IP
  - posílá se jen pro první fragment (v případě fragmentace)
  - není určen k zajištění spolehlivosti komunikace !!!
  - nepoužívá se pro broadcast nebo multicast
  - nepoužívá se, pokud zdrojová IP adresa není unikátní adresa hostitele (např. samé 0, samé 1, loopback …)
  - generování ICMP zpráv je volitelné
    - téměř vždy směrovači
    - u hostitelů záleží na implementaci
  - některé ICMP pakety mohou být zahozeny směrovači
    - např. kvůli bezpečnosti

ICMP Protocol

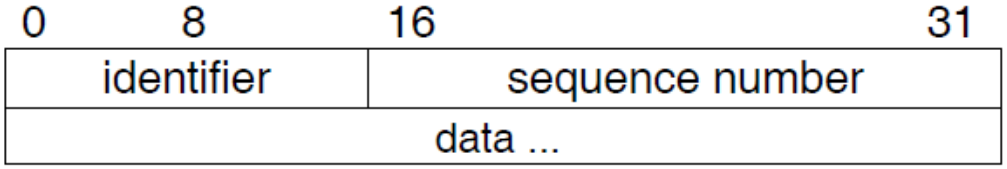
- ICMPv4 zpráva



  - type definuje typ ICMP zprávy
    - 0 – echo reply
    - 3 – destination unreachable
    - 4 – source quench
    - 5 – redirect
    - 8 – echo
    - 9 – router advertisement, 10 – router solicitation
    - 17 – address mask request, 18 – address mask reply
    - …
  - code definuje specifický kód chyby

ICMP Protokol

- Echo a Echo reply
  - struktura datového pole ICMP zprávy



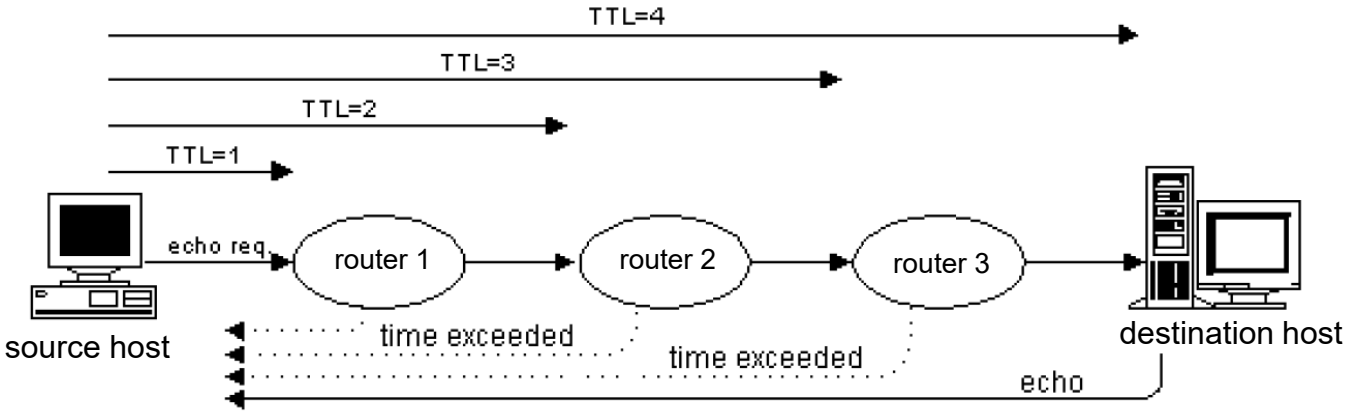
  - používá např. aplikace ***ping***
    - odesilatel vyšle zprávu echo (8) k cíli
    - cíl odpoví zprávou echo reply (0) odesilateli
    - lze např. testovat, zda jsou hostitel či router aktivní
  - zpracování zprávy echo může být administrátorem z důvodu bezpečnosti zakázáno
    - na hostiteli
    - na routeru
  - využito také aplikací ***tracert*** jako testovací paket
    - viz následující slajd

ICMP Protocol

- Time Exceeded
  - struktura níže je v datové části ICMP zprávy

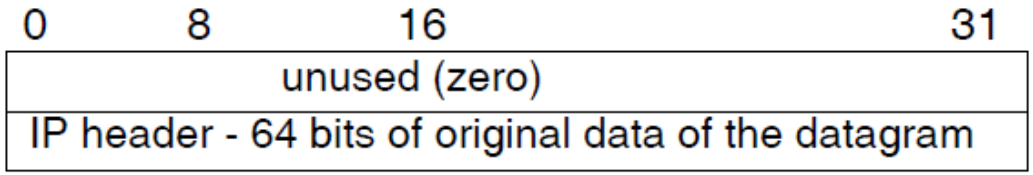
0	8	16	31
unused (zero)			
IP header - 64 bits of original data of the datagram			

  - pole *code* poskytuje podrobnou informaci
    - 0 – transit TTL exceeded
    - 1 – reassembly TTL exceeded
  - využito v programu tracert (Win) nebo traceroute (Unix)

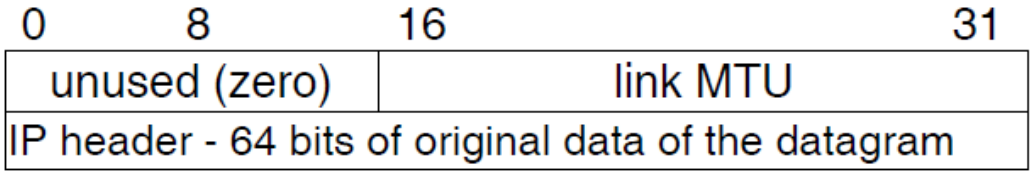


ICMP Protocol

- Destination Unreachable
  - struktura níže je v datové části ICMP zprávy



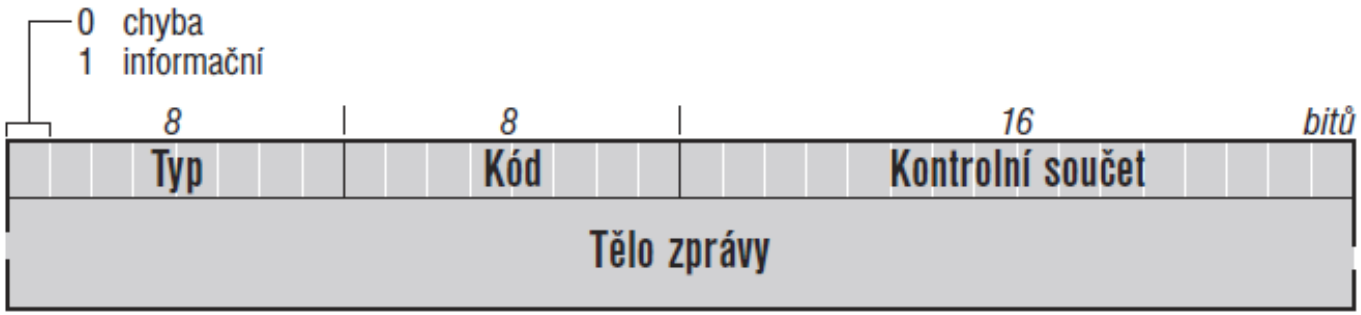
  - pole *code* poskytuje podrobnou informaci
    - 0 – network unreachable
    - 1 – host unreachable
    - 2 – protocol unreachable
    - 3 – port unreachable
    - 4 – fragmentation needed but disabled



    - ….

ICMP Protocol

- ICMP v6 zpráva



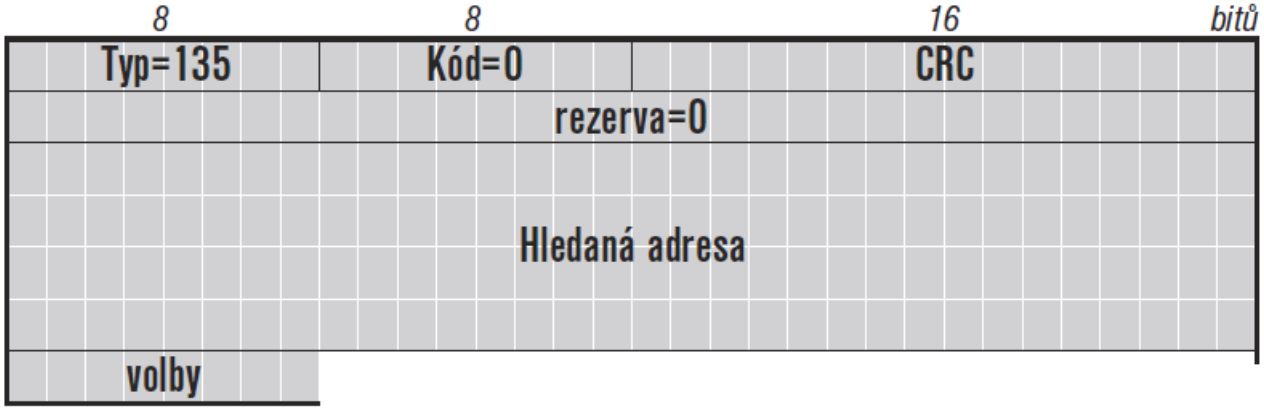
  - 1 – cíl nedosažitelný
  - 2 – příliš velký paket
  - 3 – vypršela životnost paketu
  - 4 – problém s parametry (chybné kódy v hlavičce)
  - 128 – echo request
  - 129 – echo reply
  - …
- kód definuje specifický kód chyby

NDP Protokol

- Neighbor Discovery Protocol
  - umožňuje
    - zjistit linkovou adresu z IP adresy
    - detekovat změny v linkových adresách
    - hledat směrovače
    - přesměrování
    - zjišťování parametrů sítě (např. prefixů) pro automatickou konfiguraci
    - ověření dosažitelnosti sousedů
    - detekci duplicitních adres
  - využívá ICMP v6
    - router solicitation (výzva směrovači)
    - router advertisement (ohlášení směrovače)
    - neighbor solicitation (výzva sousedovi)
    - neighbor advertisement (ohlášení souseda)
    - redirect (přesměrování)

NDP Protokol

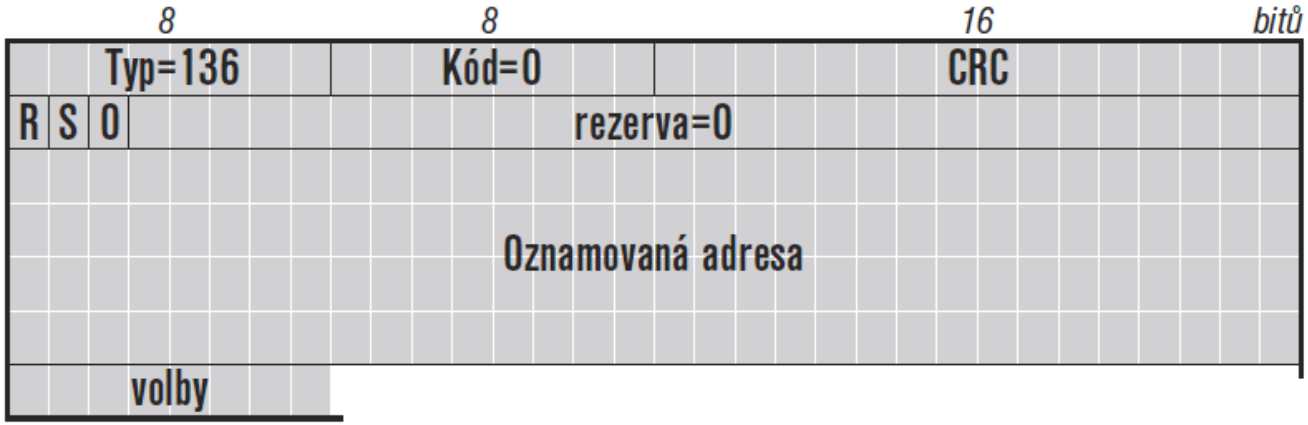
- Neighbor Discovery Protocol
  - hledání souseda
    - skupinové adresy s prefixem ff02:0:0:0:0:1:ff00::/104
    - posledních 24 bitů hledané adresy se připojí za tento prefix
  - příklad (hledáme linkovou adresu pro 2001:db8:1:1:022a:fff:fe32:5ed1)
  - skupinová adresa pro zaslání dotazu je ff02::1:ff32:5ed1



  - ve volitelných položkách může poslat vlastní linkovou adresu
    - kvůli odeslání odpovědi

NDP Protokol

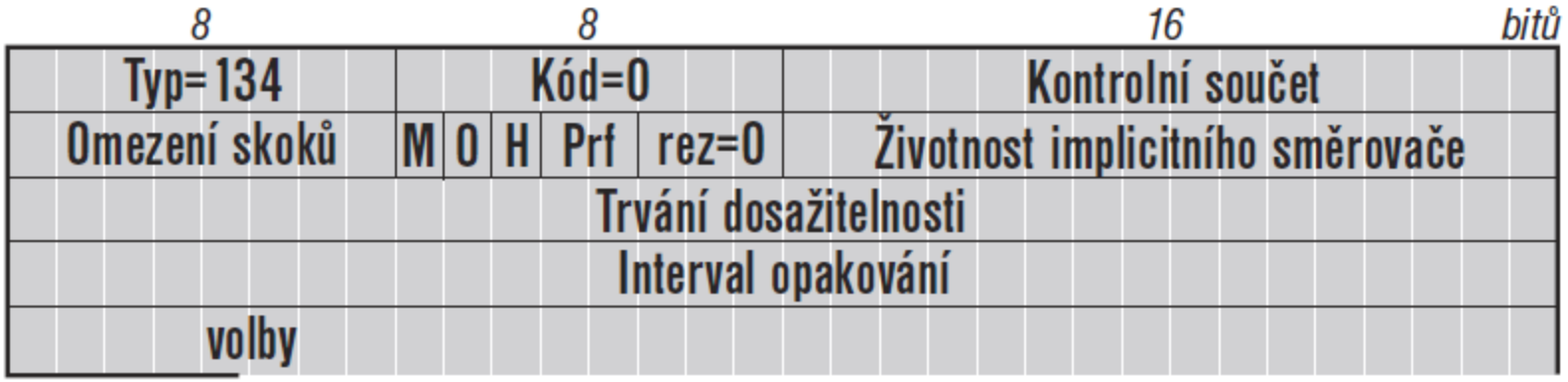
- Neighbor Discovery Protocol
  - ohlášení souseda
    - odpověď na výzvu



    - R – router, S – solicited, O - override
    - oznamovaná adresa je IP adresa
    - linková adresa je ve volitelných položkách
  - může být i nevyžádané (změna linkové adresy)
    - multicast na ff02::1

NDP Protokol

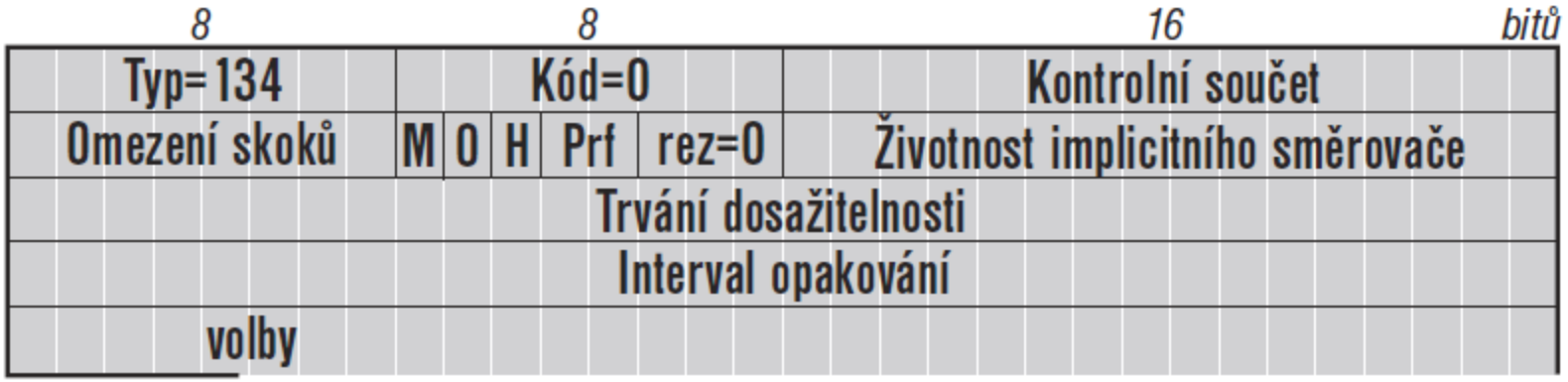
- Neighbor Discovery Protocol
  - ohlášení směrovače – bezestavová konfigurace



  - Omezení skoků – doporučení pro IP hlavičku
  - M | O – Managed | Other statefull
    - 1 | X – vše prostřednictvím DHCPv6
    - 0 | 1 – adresa, prefix, směrování, ostatní přes DHCPv6
    - 0 | 0 – DHCPv6 není k dispozici
  - H – podpora mobility
  - Prf – priorita implicitního směrovače (pokud je životnost nenulová)

NDP Protokol

- Neighbor Discovery Protocol
  - ohlášení směrovače – bezestavová konfigurace



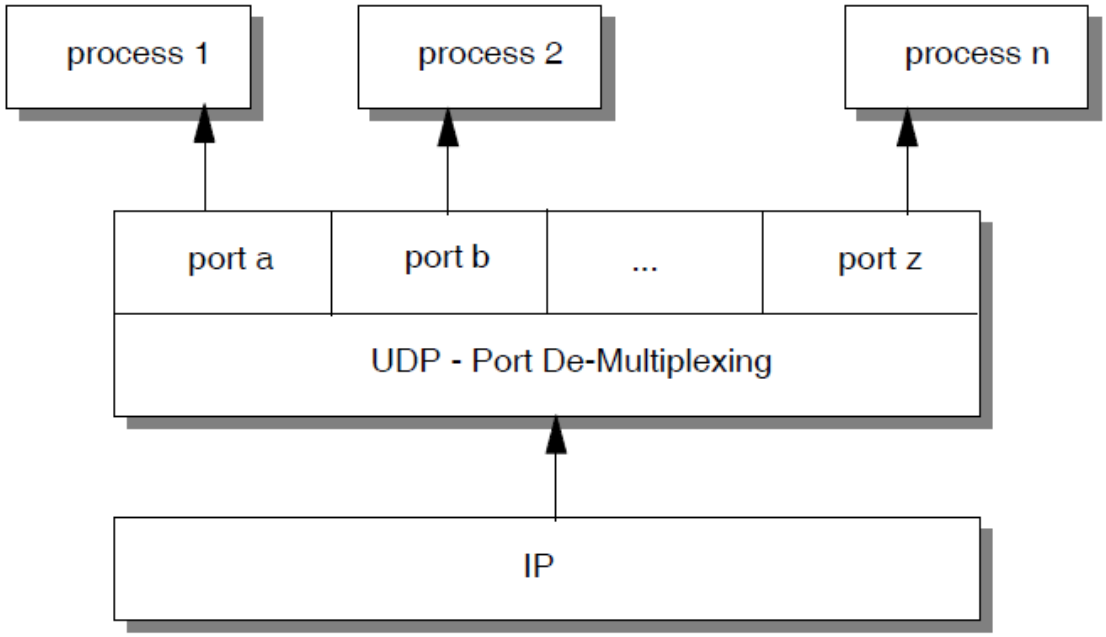
  - životnost implicitního směrovače
    - čas v sekundách, jak dlouho lze využívat
  - trvání dosažitelnosti
    - jak dlouho považovat uzel za dosažitelný po ověření dosažitelnosti
  - interval opakování
    - interval mezi dvěma výzvami sousedovi při ověřování dosažitelnosti
  - ve volbách může být linková adresa směrovače, prefix sítě, jeho časová platnost, MTU …

Porty a Sokety

- Port se využívá jako lokální multiplexor pro současně běžící aplikace (procesy)
  - data (pakety) jsou doručeny aplikaci dle trojice:
    - protokol (tedy UDP/IP, TCP/IP …)
    - síťová adresa (tedy IP adresa)
    - multiplexor na transportní vrstvě (tedy číslo portu)
  - port v IP síti je šestnáctibitové číslo
    - různé instance pro TCP a UDP protokoly
  - čísla portů < 1024
    - dobře známé porty
    - definované aplikace (např. 23 pro Telnet, 20 a 21 pro FTP …)
    - původně definovány organizací IANA, dnes přidělovány ICANN
  - soket (socket)
    - struktura definující komunikační kanál pro aplikaci (výše popsaná trojice)
      - první implementace v BSD soketech (BSD Unix)

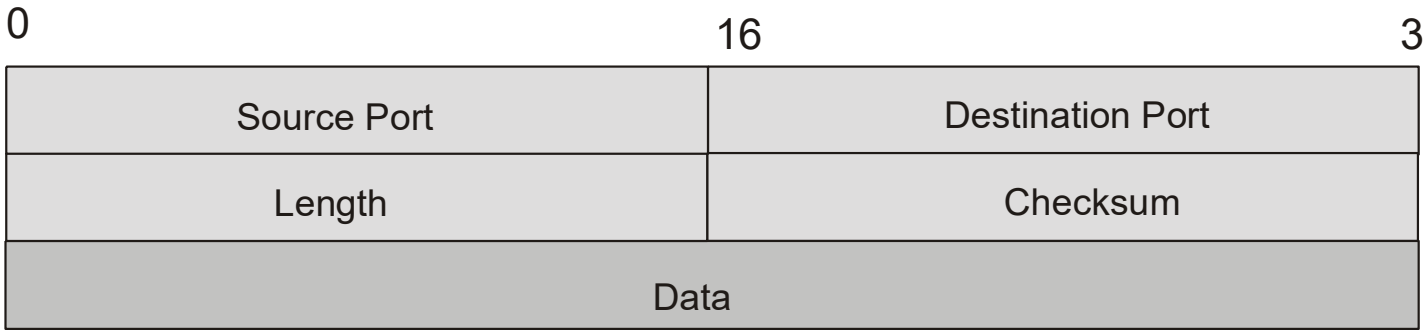
UDP Protokol

- User Datagram Protocol
  - povinná implementace
  - poskytuje aplikační rozhraní k IP
    - Jediná přidaná hodnota je multiplexování podle čísla portu
    - **neposkytuje** spolehlivost, emulaci spojení, řízení datového toku ….



UDP Packet

- Hlavička paketu



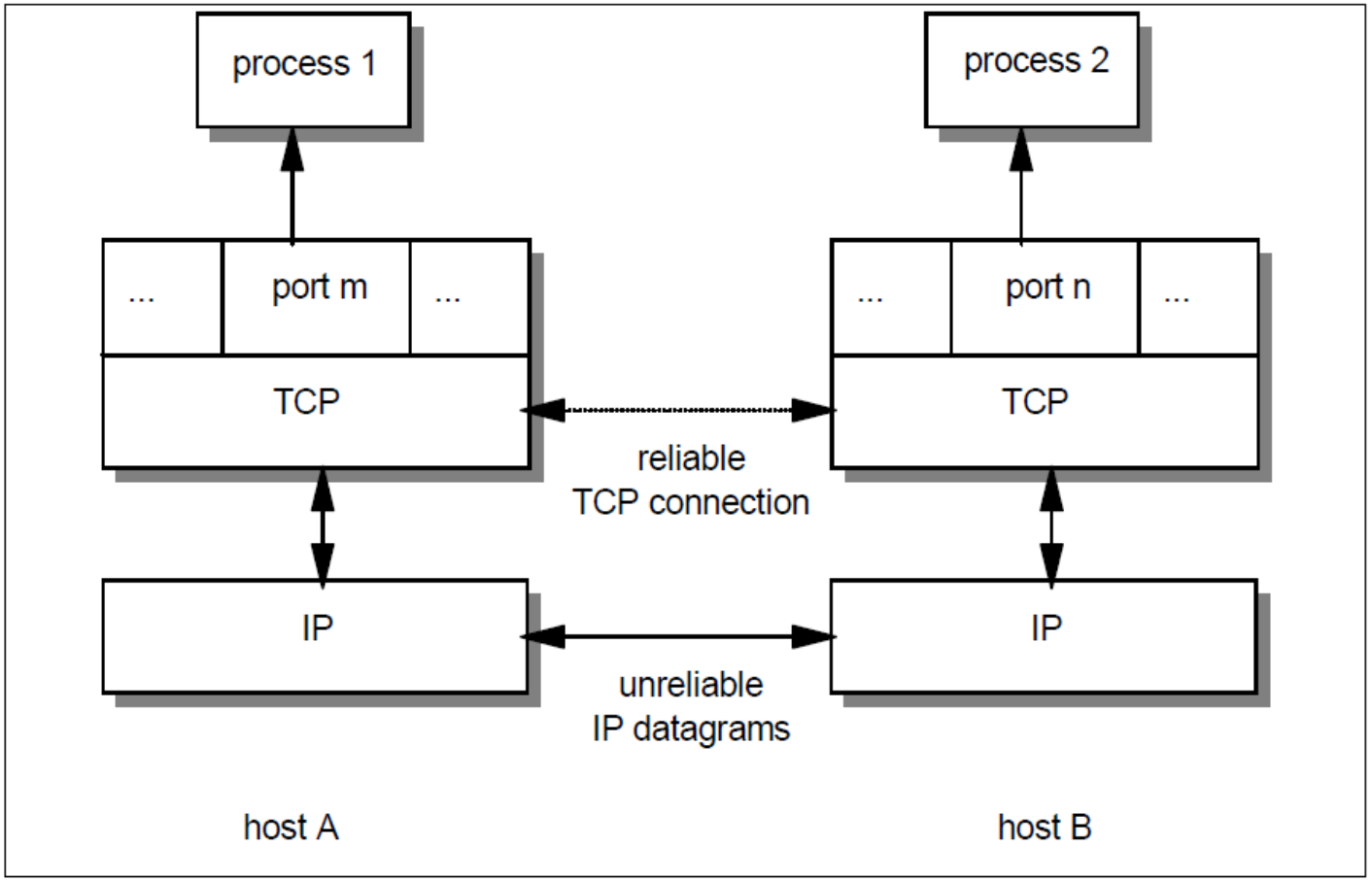
  - velikost 8 bajtů
  - Source Port – číslo portu odesilatele
  - Destination Port – číslo portu příjemce
  - Length – délka UDP paketu
    - včetně UDP hlavičky a datového pole
  - Checksum – kontrolní součet (volitelný)
    - včetně tzv. pseudo-IP hlavičky, UDP hlavičky a dat

TCP Protokol

- Transmission Control Protocol
  - povinná implementace
  - multiplexování dle čísla portu
  - spolehlivá komunikace
    - sekvenční čísla
    - generování potvrzení (ACK), opakování při chybě
  - emulace spojení
    - aplikační rozhraní jako proud bajtů
    - garantuje pořadí přenášených dat
    - emulace plně duplexní komunikace
  - řízení datového toku
    - mechanismus okének
  - segmentace dat do paketů

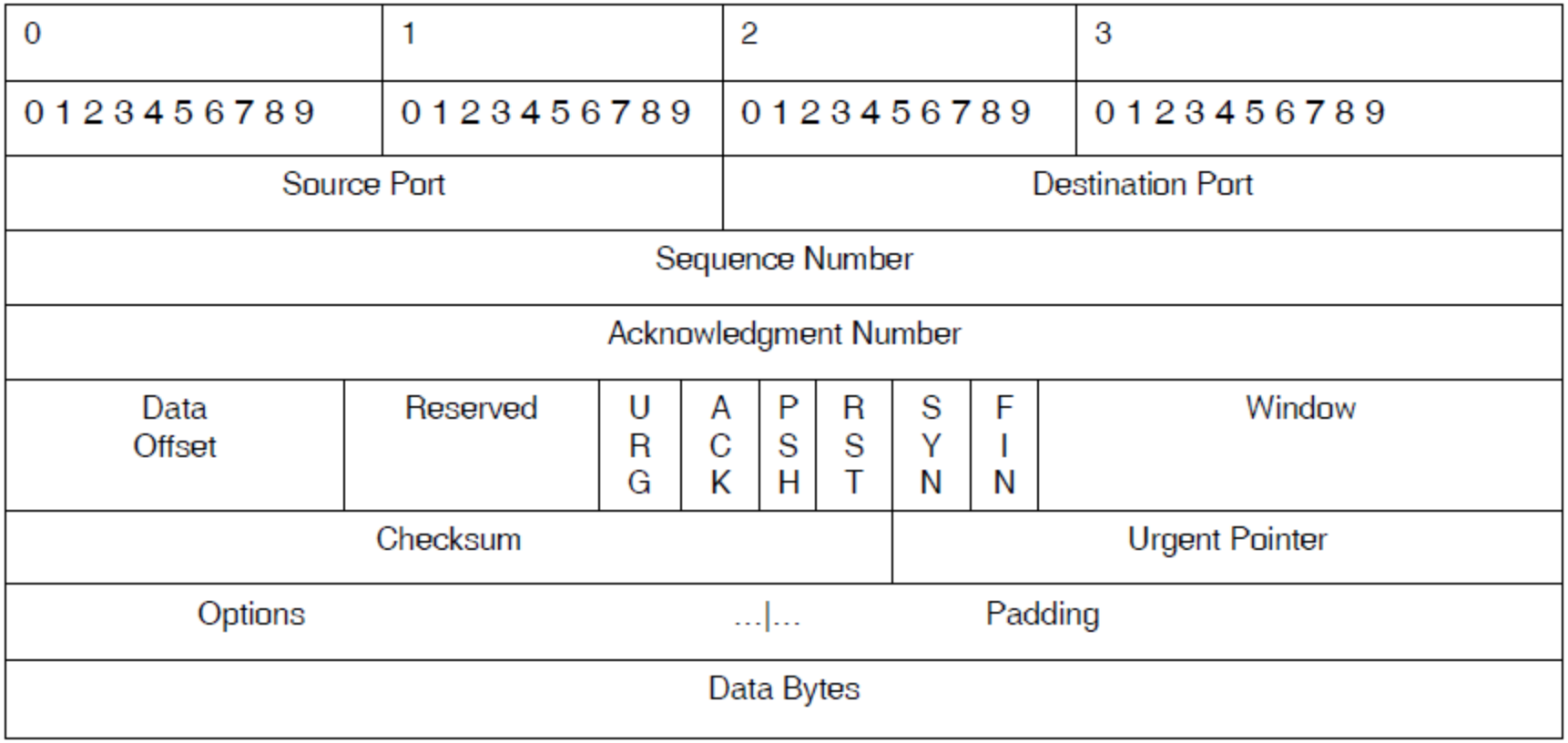
TCP Protokol

- Spolehlivý spojovaný plně duplexní komunikační kanál



TCP Paket

- Hlavička TCP paketu



– source port
  - odesilatel

– destination port
  - příjemce

TCP Paket

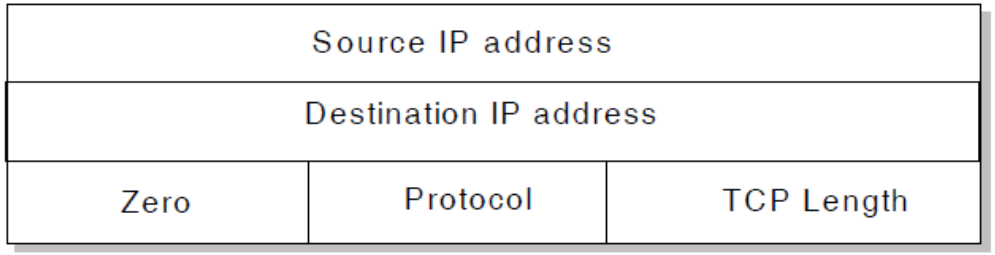
- Hlavička TCP paketu
  - sequence number – pořadové číslo
    - pořadové číslo prvého datového bajtu v paketu
    - pokud je nastaven příznak SYN, SN(n) znamená, že následující první datový bajt bude mít SN = n+1
  - acknowledgement number – číslo potvrzení
    - platné pokud je nastaven ACK příznak
    - obsahuje další očekávané SN, AN(n+1) potvrzuje přijetí až po bajt (n)
  - data offset
    - délka TCP hlavičky v DWORDech (32 bit)
    - ukazuje, kde začíná datové pole paketu
  - reserved
    - 6 nulových bitů
  - URG flag
    - je použito pole urgent pointer – ukazatel na důležitá data

TCP Paket

- Hlavička TCP paketu
  - ACK flag
    - potvrzovací číslo (acknowledgement number) je platné
  - PSH flag
    - push function – vynucené odeslání dat
  - RST flag
    - reset spojení
  - SYN flag
    - synchronizace pořadových čísel při otevírání spojení
  - FIN flag
    - ukončení spojení
  - window
    - platné pouze při nastaveném ACK
    - definuje množství dat, které je příjemce schopen přijmout (počínaje bajtem s potvrzovacím číslem v témž paketu)

TCP Paket

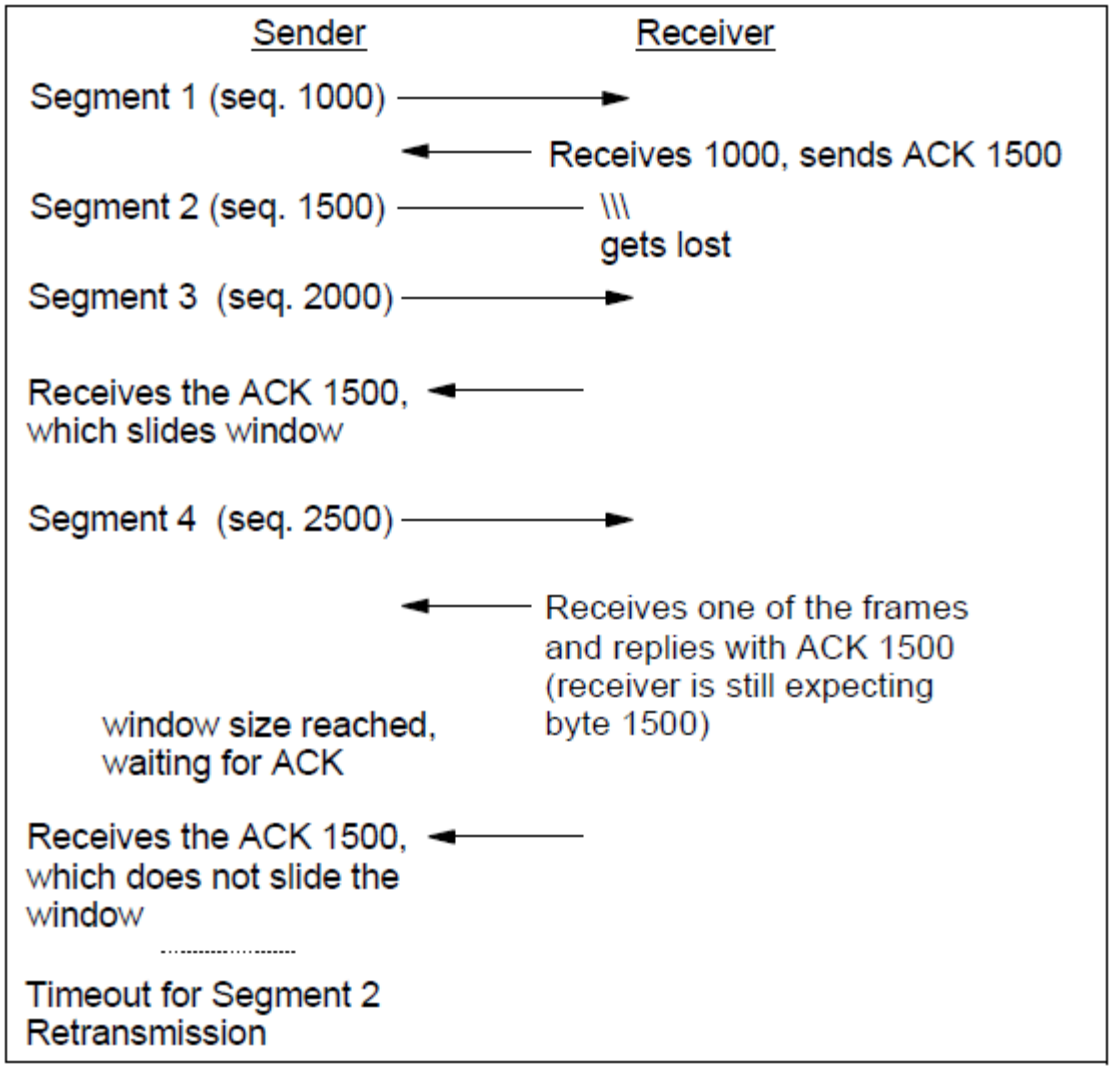
- Hlavička TCP paketu
  - checksum – kontrolní součet
    - včetně pseudo-IP hlavičky, TCP hlavičky a datového pole



  - urgent pointer – ukazatel na důležitá data
    - definuje oblast důležitých dat (příznak URG musí být nastaven)
  - options
    - volitelné položky hlavičky
    - buď pevná (1 bajt) nebo proměnná délka
    - např. časová značka, max. velikost segmentu, koeficient položky window …
  - padding
    - doplňuje velikost hlavičky na celý DWORD

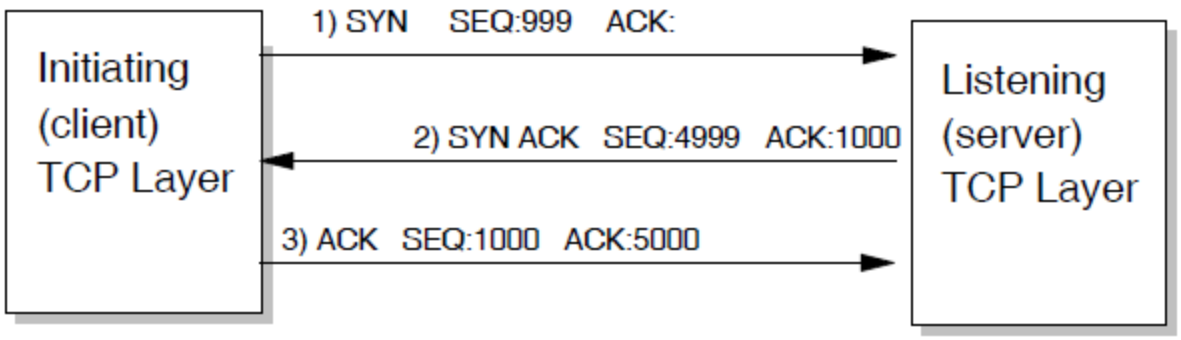
TCP – Potvrzovací mechanismus

- Sekvenční a potvrzovací číslo
- Velikost segmentu např. 500 bajtů
- Hodnota timeoutu odvozena od RTT (round trip time)



Vytvoření a zrušení TCP spojení

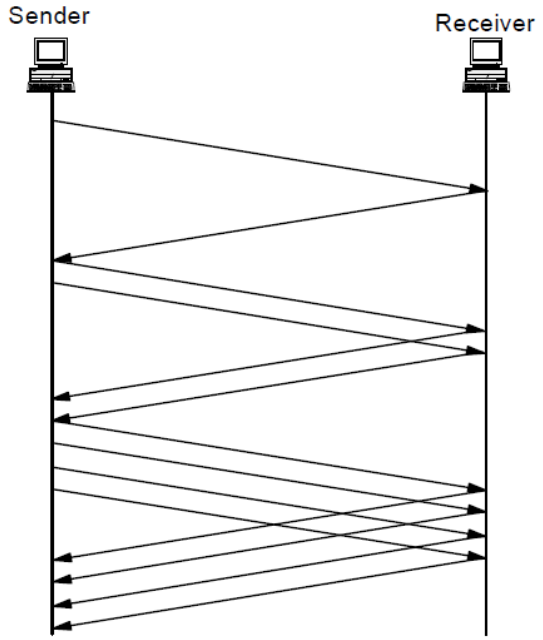
- Vytvoření: tzv. 3-way handshake



- Zrušení (z pohledu iniciátora zrušení):
  - FIN odeslán
  - ACK přijat zpět, čekání na FIN z druhé strany
  - FIN přijat, ACK odeslán
  - ACK přijat – spojení oboustranně zrušeno

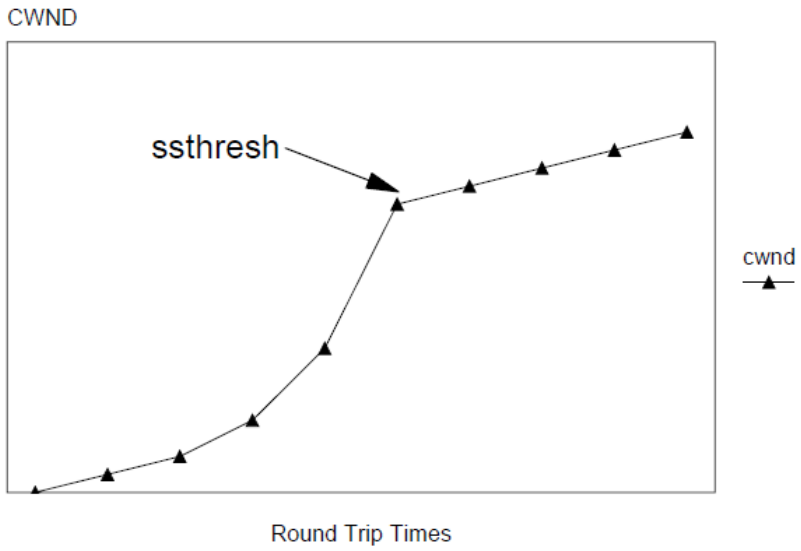
TCP Congestion Control

- Několik současně spolupracujících mechanismů
- Slow start
  - rychlost odesílání na novém spojení je pomalu navyšována s cílem nalézt limit (max. packet rate)
  - velikost vysílacího okna (congestion window size - *cwnd*) začíná na velikosti jednoho segmentu (*segsize*)
    - zvyšuje se po příjmu ACK dokud nedojde ke ztrátám paketů
  - exponenciální růst
    - 1 (1 ACK – zvýšení na 2)
    - 2 (2 ACK – zvýšení na 4)
    - 4 (4 ACK – zvýšení na 8)
    - …



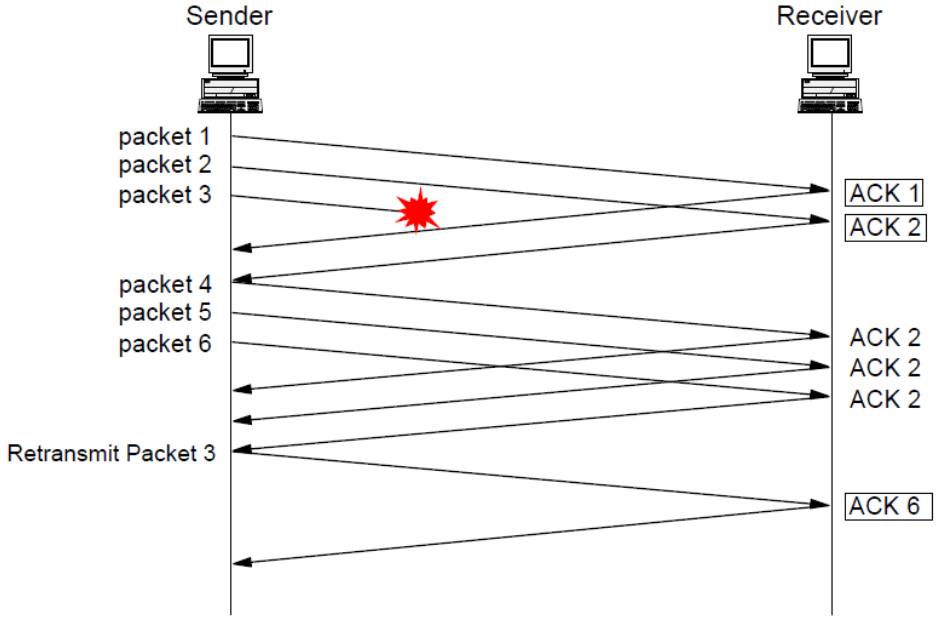
TCP Congestion Control

- Congestion avoidance
  - kombinováno s pomalým startem
    - proměnná *slow start threshold* (*ssthres*), inicializována na 0xffff
  - v případě ztráty dat (indikováno timeoutem nebo duplikovaným ACK)
    - hodnota *ssthres je nastavena na polovinu cwnd*
    - pokud nastal timeout, *cwnd* = *segsize*
  - pokud *cwnd* < *ssthres* (fáze pomalého startu)
    - exponenciální růst *cwnd*
  - jinak fáze předcházení zahlcení
    - *cwnd* += *segsize*segsize/cwnd*
    - lineární růst



TCP Congestion Control

- Fast retransmit
  - není třeba vyčkávat na timeout, pokud dojde ke ztrátě paketu
  - využití duplikovaných potvrzení (ACK)
    - nízký počet duplikovaných ACK → předpokládá se změna v pořadí paketů, nikoliv ztráta
    - vyšší počet (3 a více) → předpokládá se ztráta paketu a vysílání se opakuje aniž by se čekalo na timeout



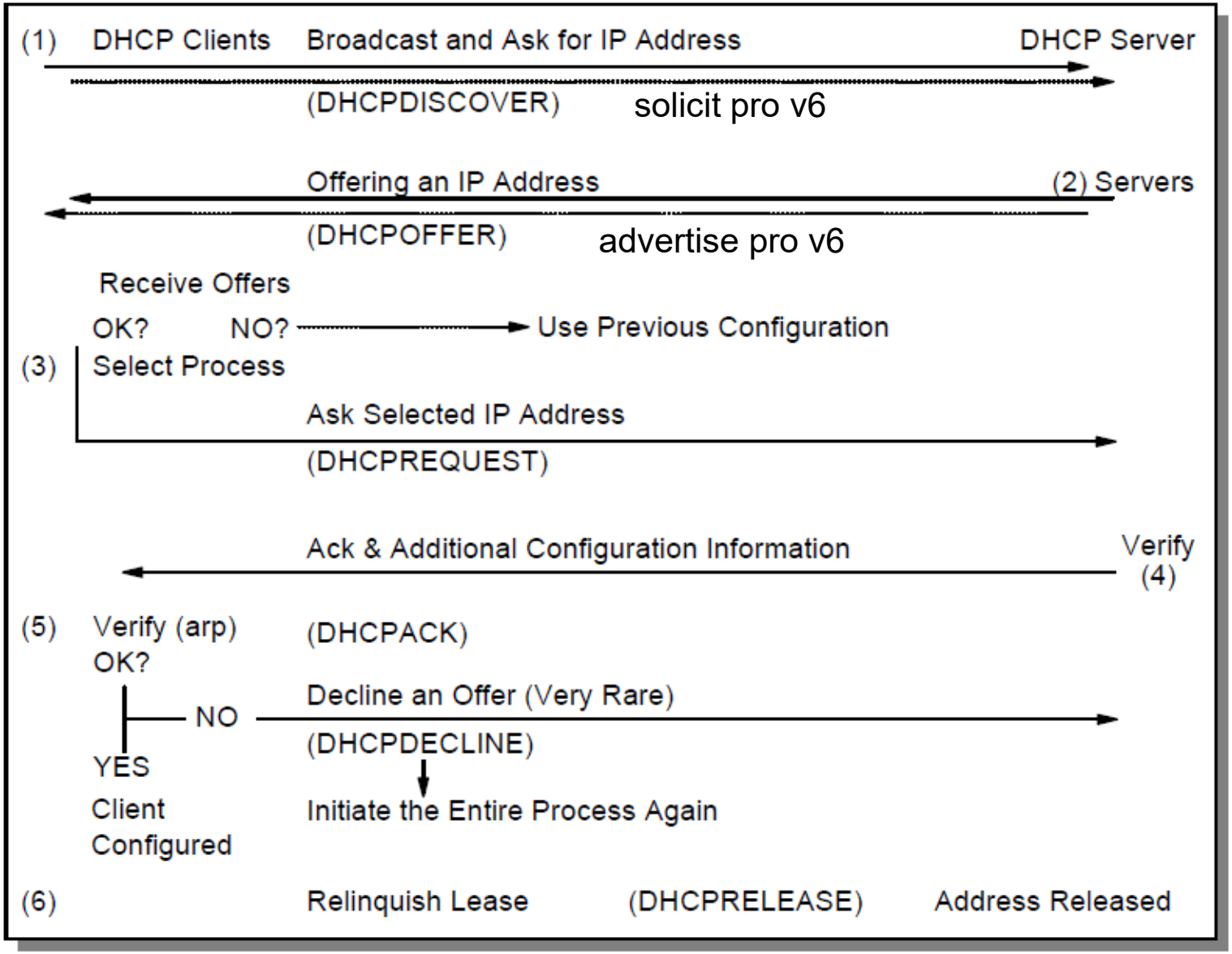
TCP Congestion Control

- Fast retransmit with fast recovery
  - pokud jsou přijaty 3 duplikovaná potvrzení
    - *ssthres* = *cwnd*/2 (ale minimálně 2**segsize*)
    - Ztracený segment je odeslán znovu
    - *cwnd* = *ssthres* + *3*segsize* (3 duplikáty ACK indikují že 3 segmenty po tom ztraceném dorazily OK)
    - pokud je přijat další duplikát ACK, *cwnd* += *segsize*
    - dovolí-li to hodnota *cwnd*, je vyslán další segment
  - pokud přijde ACK pro novější segment
    - všechny segmenty poslané po tom ztraceném jsou potvrzeny
    - *cwnd* = *ssthres*
      - předchází zahlcení – *cwnd* je nyní poloviční než když došlo ke ztrátě paketu

DHCP Protocol

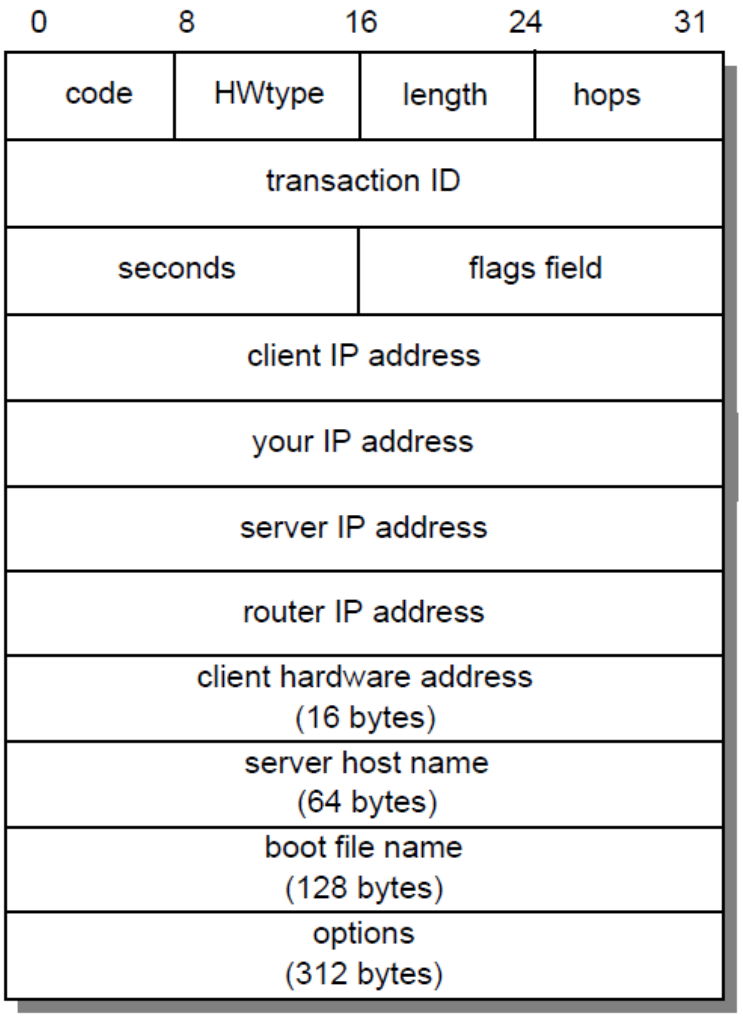
- Dynamic Host Configuration Protocol
  - distribuce konfigurace TCP/IP
  - umožňuje bootování bezdiskových hostitelů
  - vychází ze staršího protokolu BOOTP
    - BOOTP klienti mohou být obslouženi DHCP serverem
  - DHCP zprávy se přenášejí v UDP paketech
    - nezabezpečené
    - možné útoky prostřednictvím neautorizovaného DHCP serveru
  - klasické DHCP využívá broadcast
  - DHCPv6 skupinové adresy
    - všichni DHCP agenti a servery ff02::1:2
    - všechny DHCP servery ff05::1:3

DHCP Protokol



DHCP Protocol

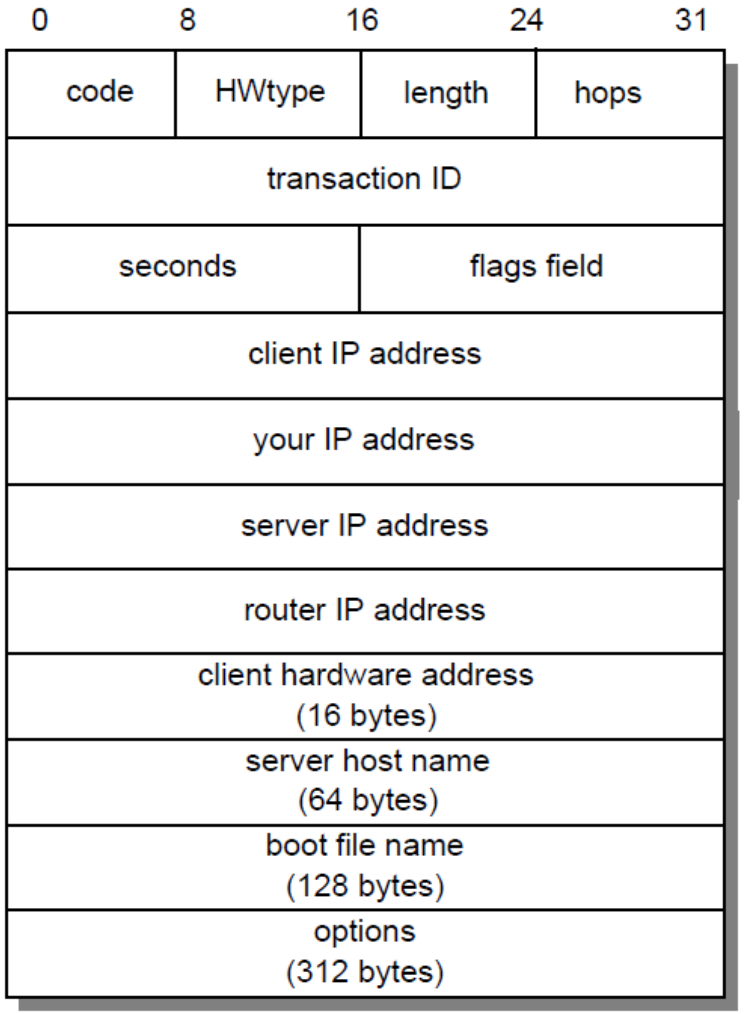
- DHCP Packet



- code: 1 – žádost, 2 – odpověď
- HWtype: 1 – Eth., 6 – IEEE802
  …
- length: délka HW adresy
- hops: klient nastaví na 0, inkrementováno směrovačem při předání – detekce smyček
- transaction ID – náhodné číslo
- seconds: čas od počátku procesu bootu
- flags: MSB pro žádost o zaslání odpovědi broadcastem (když klient ještě nezná svoji adresu)

DHCP Protocol

- DHCP Packet



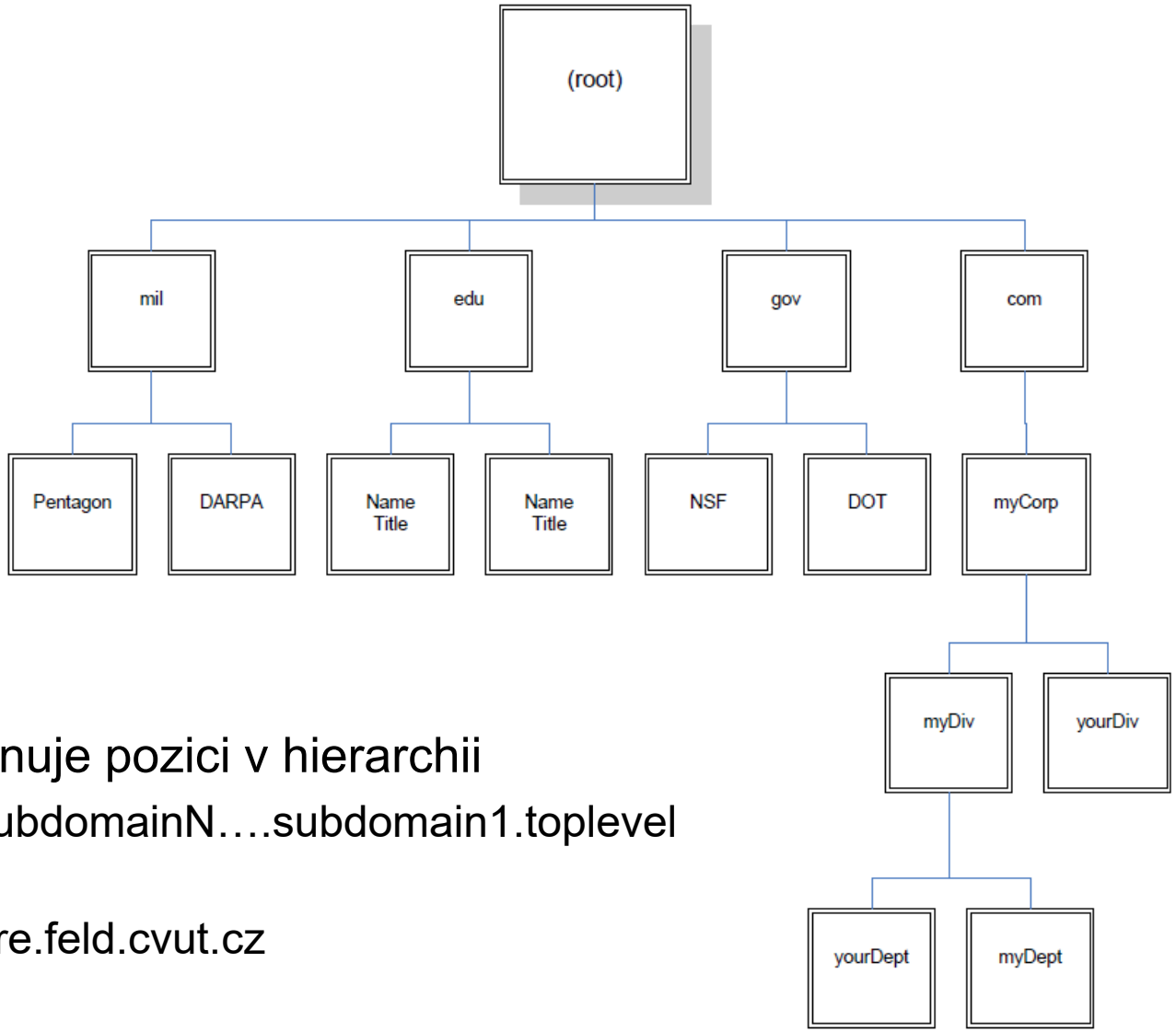
- client address: IP nebo 0.0.0.0
- your address: IP přiřazená serverem, pokud je v žádosti 0.0.0.0
- server address: nastavuje server
- router address: adresa relay routeru, nikoliv konfigurovaný směrovač
- client hw address: nastavuje klient
- server host name: volitelný řetězec zakončený nulou
- boot file name: volitelné, poskytuje server
- options: volitelné položky

DNS (Domain Name System)

- Proč jmenný systém?
  - je jednodušší používat jména počítačů než jejich IP adresy
    - pro lidi, nikoliv pro stroje ☺
  - pokud je uzel přesunut do jiné sítě
    - IP adresa se změní
    - jmenný název zůstává
  - původně plochý jmenný prostor
    - jednoduchá jména uzlů, např. “Jupiter”
    - lokální soubor “hosts.txt”, obsahuje databázi jmen a příslušných adres, později stahovaný z FTP
    - složitá (nemožná) metoda pro velké sítě s distribuovanou správou

DNS (Domain Name System)

- DNS
  - hierarchický jmenný systém
  - distribuovaný systém s distribuovanou správou
  - obvykle kopíruje organizační struktury
  - název uzlu definuje pozici v hierarchii
    - “hostname.subdomainN….subdomain1.toplevel domain”
    - např. measure.feld.cvut.cz



DNS (Domain Name System)

- Top-level domény
  - generické →
  - států – např. cz, sk …
  - www.icann.org

Domain name	Meaning
aero	The air transport industry
biz	Business use
cat	The Catalan culture
com	Commercial organizations
coop	Cooperatives
edu	Educational organizations
gov	U.S. governmental agencies
info	Informational sites
int	International organizations
jobs	Employment-related sites
mil	The U.S. military
mobi	Mobile devices sites
museum	Museums
name	Family and individual sites
net	Network infrastructures
org	Non-commercial organizations
pro	Professional sites
travel	The travel industry

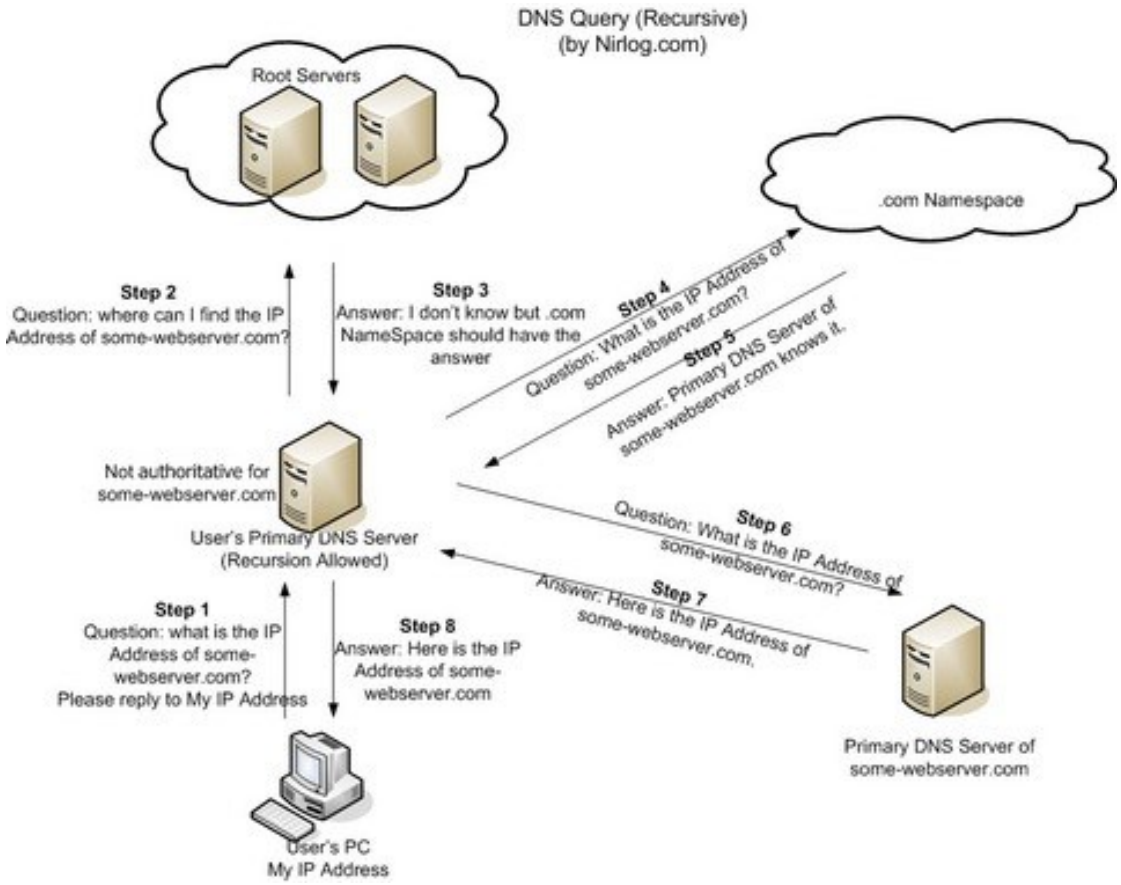
DNS (Domain Name System)

- Rozděleno na zóny
  - pravomoc (authority) k zóně má odpovídající jmenný server
  - jediný server může být zodpovědný za více zón
  - pravomoc k dílčímu podstromu v hierarchii může být delegována
    - a to rekurzivně
  - pro TLD zóny jsou k dispozici kořenové jmenné servery, koordinované ICANN
    - podpora anycastu v IPv6
    - podpora „load sharingu“

- DNS poskytuje více než jen překlad jmen → IP
  - překlad IP → jméno
  - získání dalších informací o hostiteli (např. typ OS …)
  - informaci o konfiguraci mailu pro konkrétní doménu
  - …

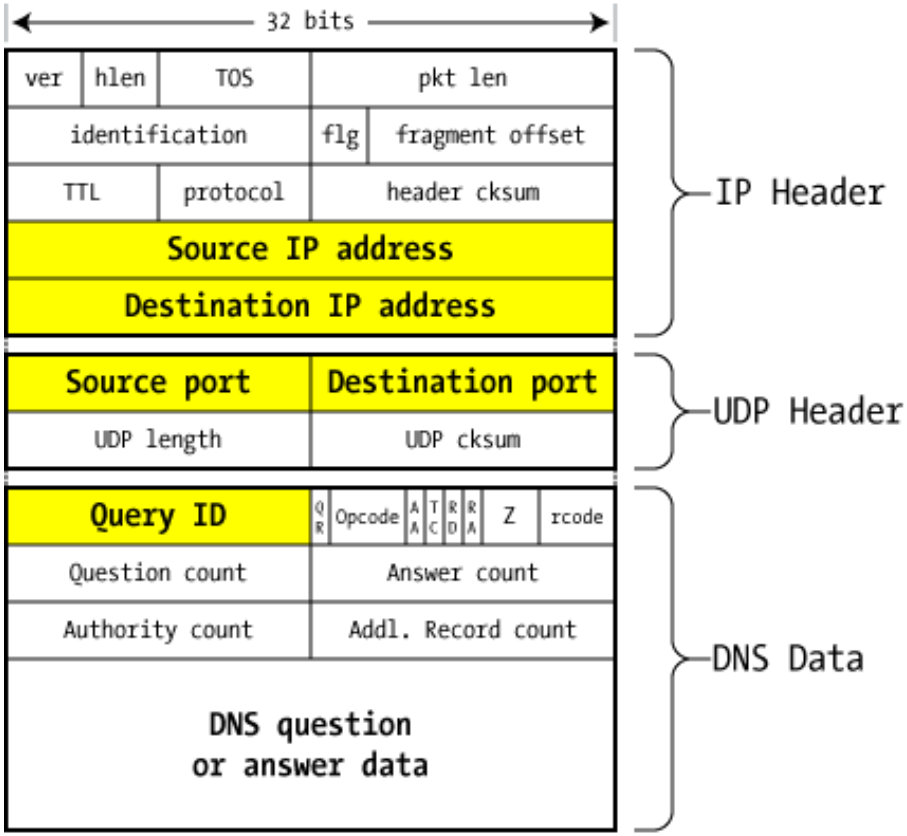
DNS (Domain Name System)

- Zpracování DNS dotazu
  - rekurzivní nebo **iterativní** (většinová varianta)



DNS (Domain Name System)

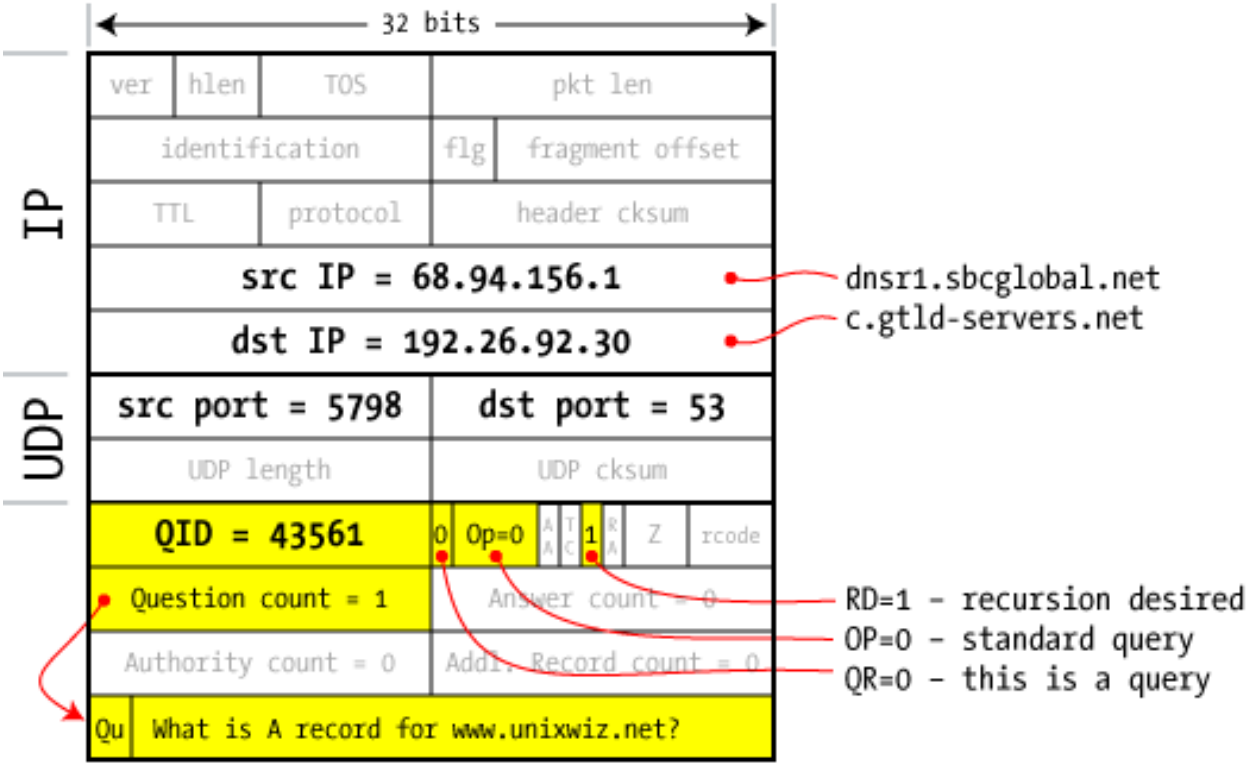
- Formát DNS paketu
  - query ID: náhodné číslo
  - QR: 0 pro dotaz, 1 pro odpověď
  - op.code:
    - 0 – standardní dotaz
    - 1 – inverzní dotaz
    - 2 – dotaz na status serveru
  - AA: autoritativní odpověď
  - TC: zkráceno
    - příliš dlouhá data pro UDP
  - RD: požadavek rekurze
  - RA: potvrzení rekurze
  - Z: vše 0
  - rcode: kód odezvy
  - question, answer authority a additional rec. count
    - počty jednotlivých typů záznamů



DNS packet on the wire

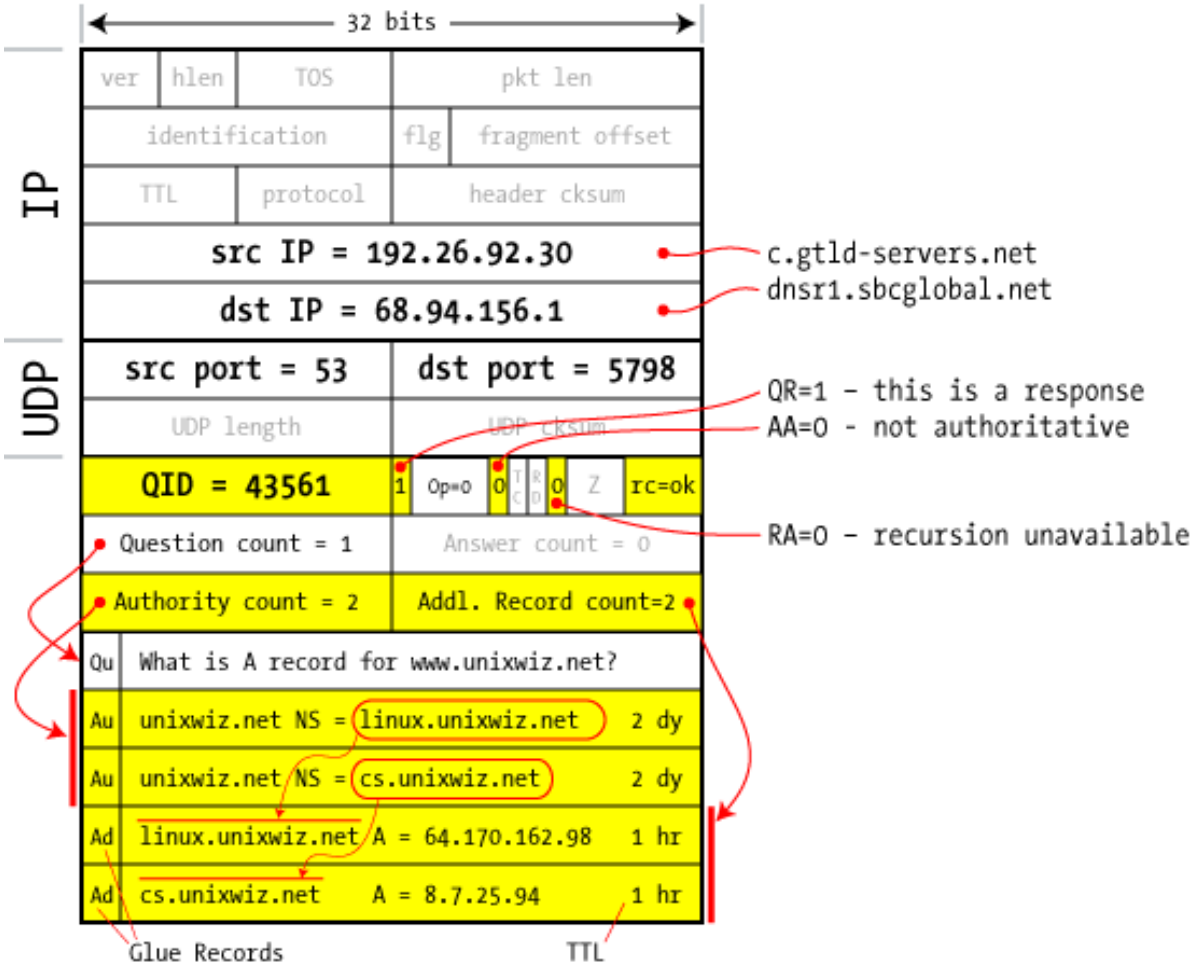
DNS Query / Response Sequence

- Krok 4 (o 2 slajdy výše) – dotaz na server



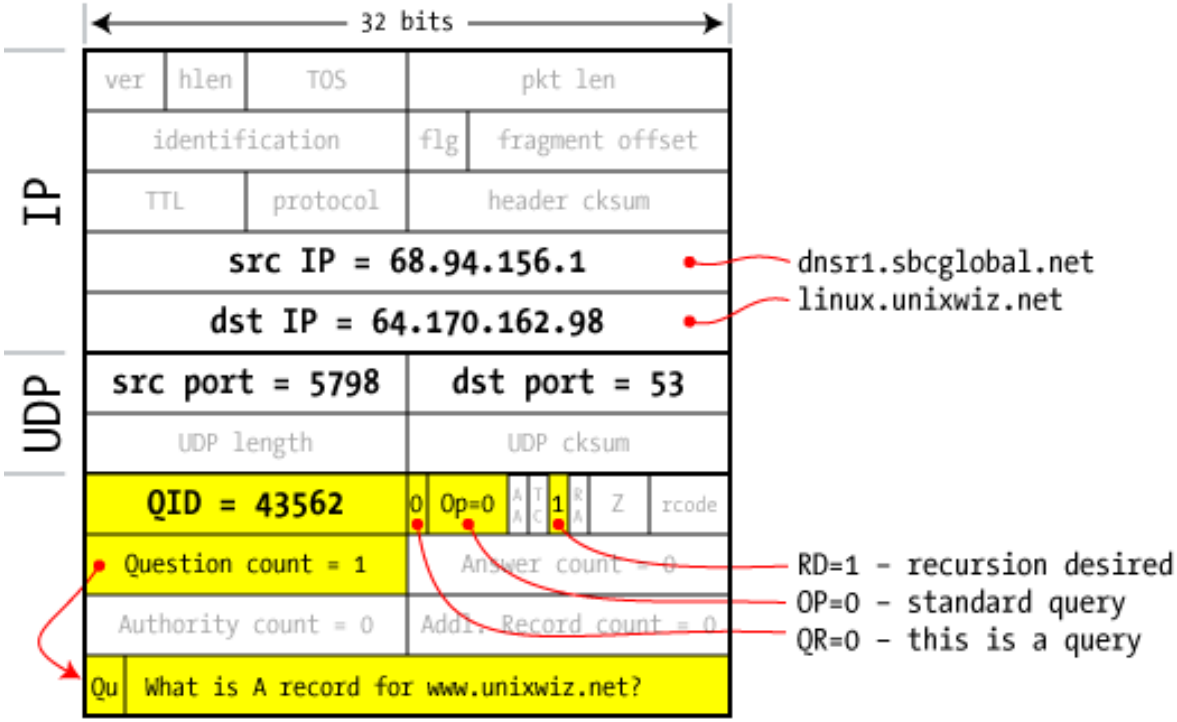
DNS Query / Response Sequence

- Krok 5 – odezva od serveru



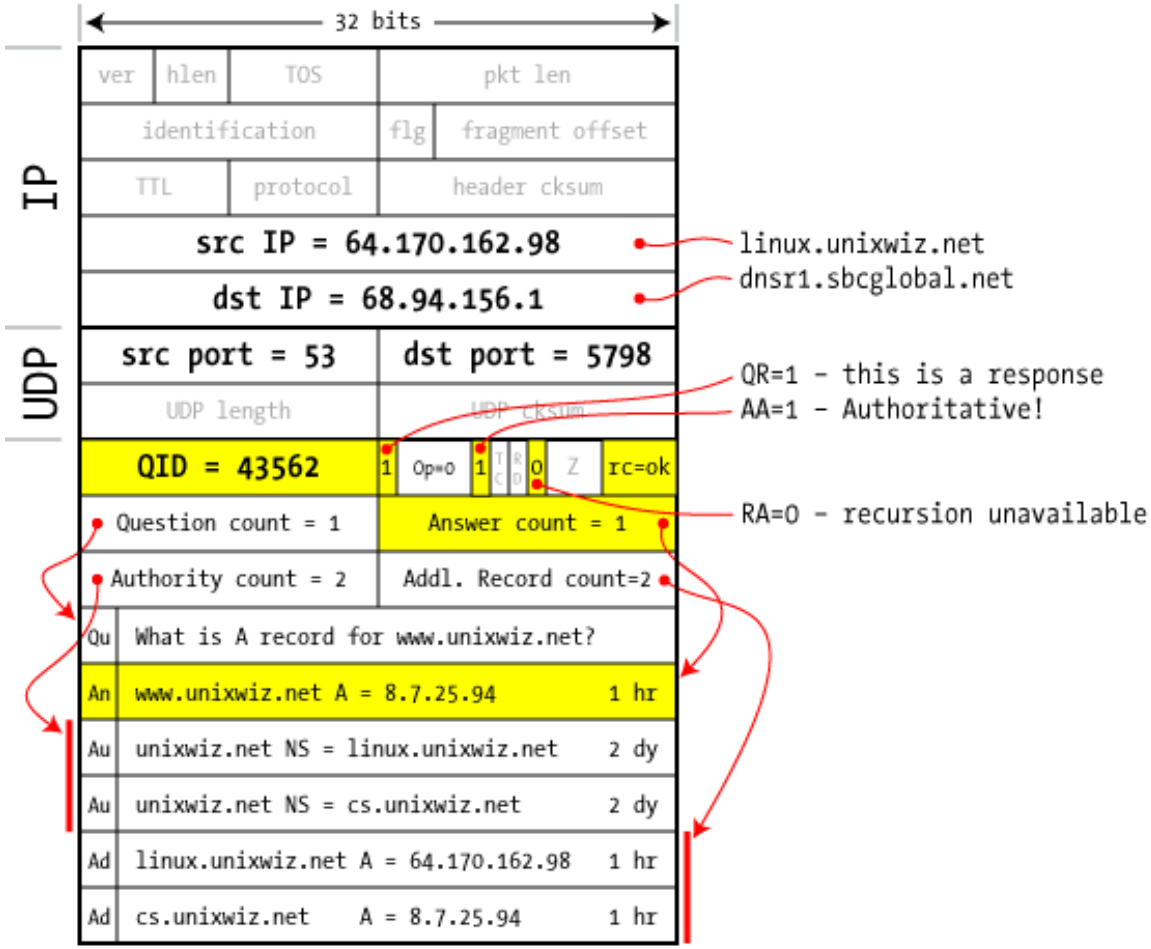
DNS Query / Response Sequence

- Krok 6 – dotaz na autoritativní server



DNS Query / Response Sequence

- Krok 7 – odpověď od autoritativního serveru

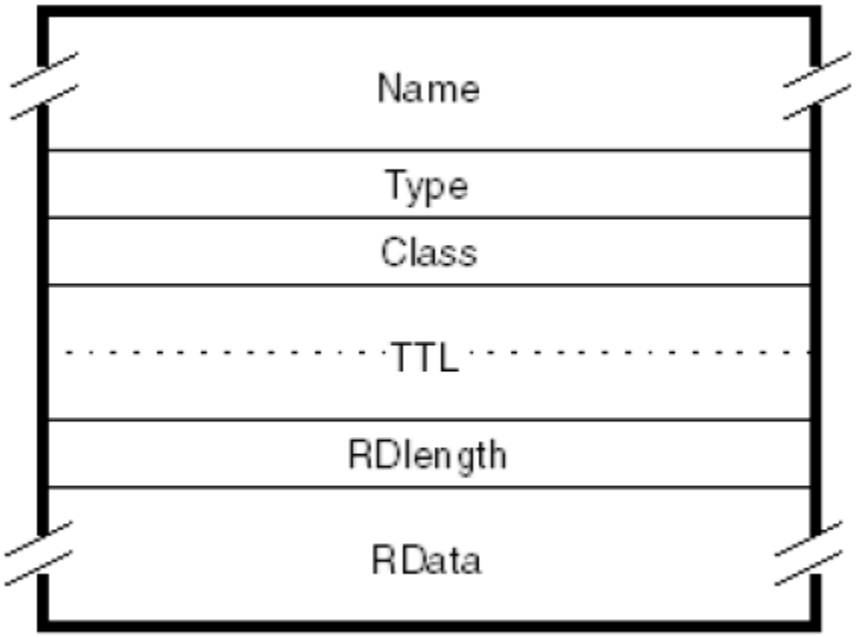


Typy DNS Serverů

- Primární
  - má autoritu nad zónou
  - probíhá na něm editace záznamů
- Sekundární
  - má autoritu nad zónou
  - přebírá data od primárního serveru
    - tzv. přenos zóny v pravidelných časových intervalech (typicky jednotky hodin)
- „Caching only“
  - poskytují neautorizované odpovědi
  - získávají informace z primárních nebo sekundárních serverů
  - slouží k omezení zátěže těchto serverů (a také k redukci síťového provozu)
- Tentýž fyzický server může pracovat jako DNS server různého typu pro různé zóny

Typy DNS záznamů

- Záznamy (Resource records)
  - obecný formát záznamu



  - name
    - sekvence jmenovek (každá až 63 znaků, 1. musí být písmeno), oddělené znakem tečka)
    - malá a velká písmena se nerozlišují

Typy DNS záznamů

- Resource records
  - type

Type	Value	Meaning	RFC def
A	1	A host address	1035
NS	2	An authoritative name server	1035
CNAME	5	The canonical name for an alias	1035
SOA	6	Marks the start of a zone of authority	1035
MB	7	A mailbox domain name (experimental)	1035
MG	8	A mail group member (experimental)	1035
MR	9	A mail rename domain name (experimental)	1035
NULL	10	A NULL resource record (experimental)	1035
WKS	11	A well-known service description	1035
PTR	12	A domain name pointer	1035
HINFO	13	Host information	1035
MINFO	14	Mailbox or mail list information	1035
MX	15	Mail exchange[a]	1035
TXT	16	Text strings	1035
RP	17	Responsible person record	1183
AFSDB	18	Andrew File System database	1183
X25	19	X.25 resource record	1183
ISDN	20	ISDN resource record	1183
RT	21	Route Through resource record	1183
NSAP	22	Network Service Access Protocol record	1348

Typy DNS záznamů

- Resource records
  - type

Type	Value	Meaning	RFC def
NSAP-PTR	23	NSAP Pointer record	1348
KEY	25	The public key associated with a DNS name	2535
AAAA	28	An IPv6 address record	3596
LOC	29	GPS resource record	1876
SRV	33	Defines the services available in a zone	2872
CERT	37	Certificate resource records	4398
A6	38	Forward mapping of an IPv6 address	2874
DNAME	39	Delegation of IPv6 reverse addresses	2672
DS	39	Delegated Signer record (DNS security)	4034
RRSIG	46	Resource record digital signature	4034
NSEC	47	Next Secure record (DNS security)	4034
DNSKEY	48	Public Key record (DNS security	4034

Typy DNS záznamů

- Záznamy (Resource records)
  - class
    - aktuálně pouze hodnota IN (Internet system), ostatní jsou zastaralé
  - TTL
    - specifikuje čas v sekundách, po který je záznam platný v DNS cache
    - typická hodnota pro záznam A je 86400 (1 den)
  - RDlength
    - délka pole RData
  - RData
    - řetězec popisující daný záznam
    - formát závisí na typu (a třídě) záznamu
- Přístup do DNS – program NSLOOKUP
  - zkuste si to ….
- Dynamické DNS záznamy
  - spolupráce DHCP ↔ DNS