3. Síťová vrstva

PB156: Počítačové sítě

Eva Hladká

Fakulta informatiky Masarykovy univerzity

jaro 2010

Struktura přednášky

L3. Síťová vrstva – Přehled

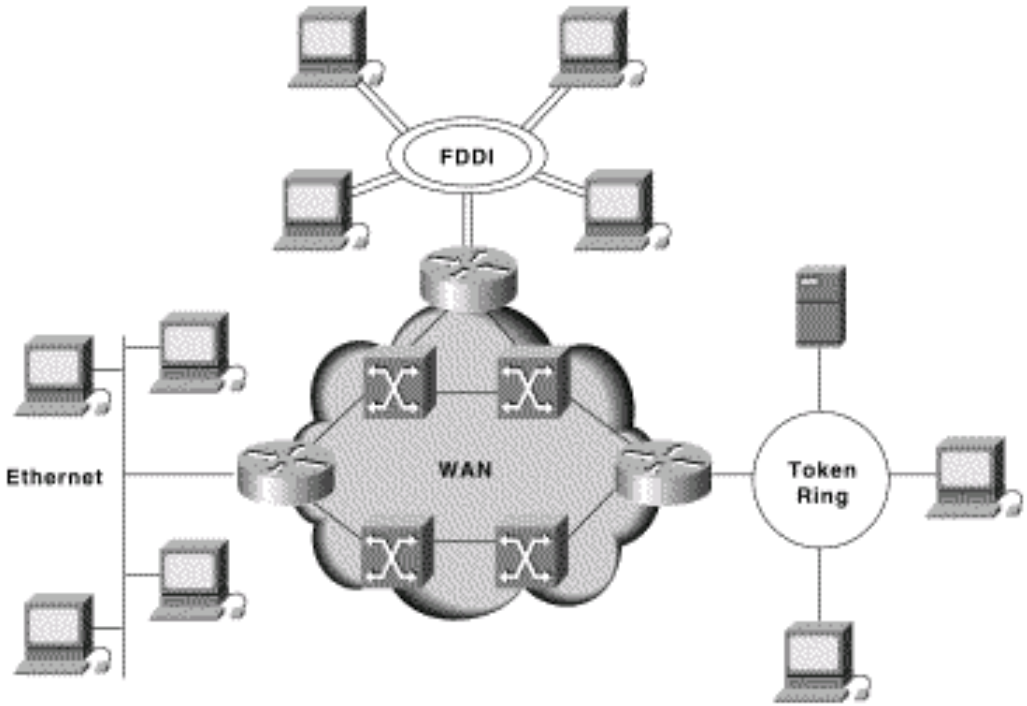
Proč nestačí L2?

- nemožnost vybudování geograficky libovolně rozlehlé sítě
- neuniformní prostředí

Co nás nyní čeká...

- představení L3, poskytované služby
- Internetworking, modely zajištění síťových služeb
- adresace na L3, přidělování adres
- protokoly IPv4, ARP, ICMP
- protokoly IPv6, ICMPv6
- směrování, směrovací techniky

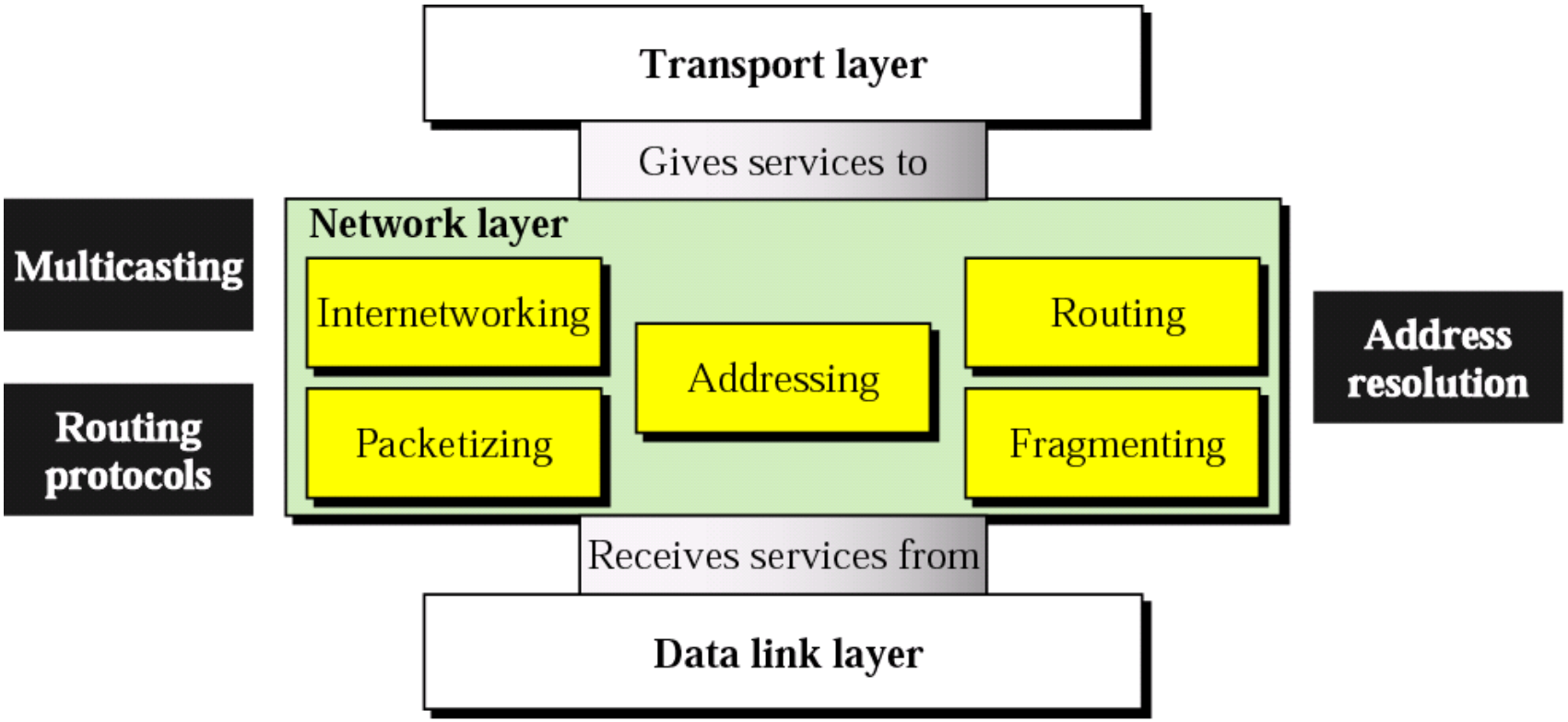
L3 z pohledu sítě – kde se pohybujeme?



- propojování lokálních sítí do větších, komplexních sítí (např. Internet)
- možnost ustavení komunikačního kanálu mezi „libovolnými“ stanicemi v Internetu
  - skrze více samostatných fyzických sítí (LANs)
  - tzv. *host-to-host delivery*

Úvod

- **síťová vrstva:**
  - poskytuje služby pro *transportní vrstvu*:
    - přijímá *segmenty* od transportní vrstvy, které transformuje do *paketů*
    - ve spolupráci s vrstvou datového spoje zajišťuje přenos paketů mezi komunikujícími uzly *(i mezi různými fyzickými LAN sítěmi)*
  - logicky spojuje samostatné heterogenní LAN sítě
    - vyšším vrstvám poskytuje iluzi uniformního prostředí jediné velké sítě (*WAN – Wide Area Network*)
  - poskytuje možnost jednoznačné identifikace (adresace) každého PC/zařízení na Internetu
  - zajišťuje *směrování* procházejících paketů
  - ve spolupráci s vrstvou datového spoje mapuje adresy síťové vrstvy na fyzické adresy (MAC adresy)
  - další služby: multicast



Obrázek: Ilustrace služeb síťové vrstvy.

Služby

- *Propojování fyzických sítí (Internetworking)*
  - iluze uniformního prostředí jediné velké sítě
- *Tvorba paketů (Packetizing)*
  - přijaté segmenty transformovány na pakety (IP protokol)
- *Fragmentace paketů (Fragmenting)*
  - rozdělování segmentů na pakety s délkou závislou na vlastnostech/schopnostech sítě
- *Adresace (Addressing)*
  - adresy entit síťové vrstvy – tzv. *IP adresy*, jedinečné skrze celou síť
  - pakety obsahují zdrojovou a cílovou IP adresu komunikujících entit
- *Mapování IP adres na/z fyzické adresy (Address Resolution)*
  - ARP, RARP protokoly
- *Směrování (Routing)*
  - nalezení nejvhodnější cesty mezi komunikujícími entitami, reakce na chyby
- *Metody základního monitoringu stavu sítě (Control Messaging)*
  - základní informace o nedoručitelnosti paketů, stavu sítě, uzlů, atp. – ICMP protokol

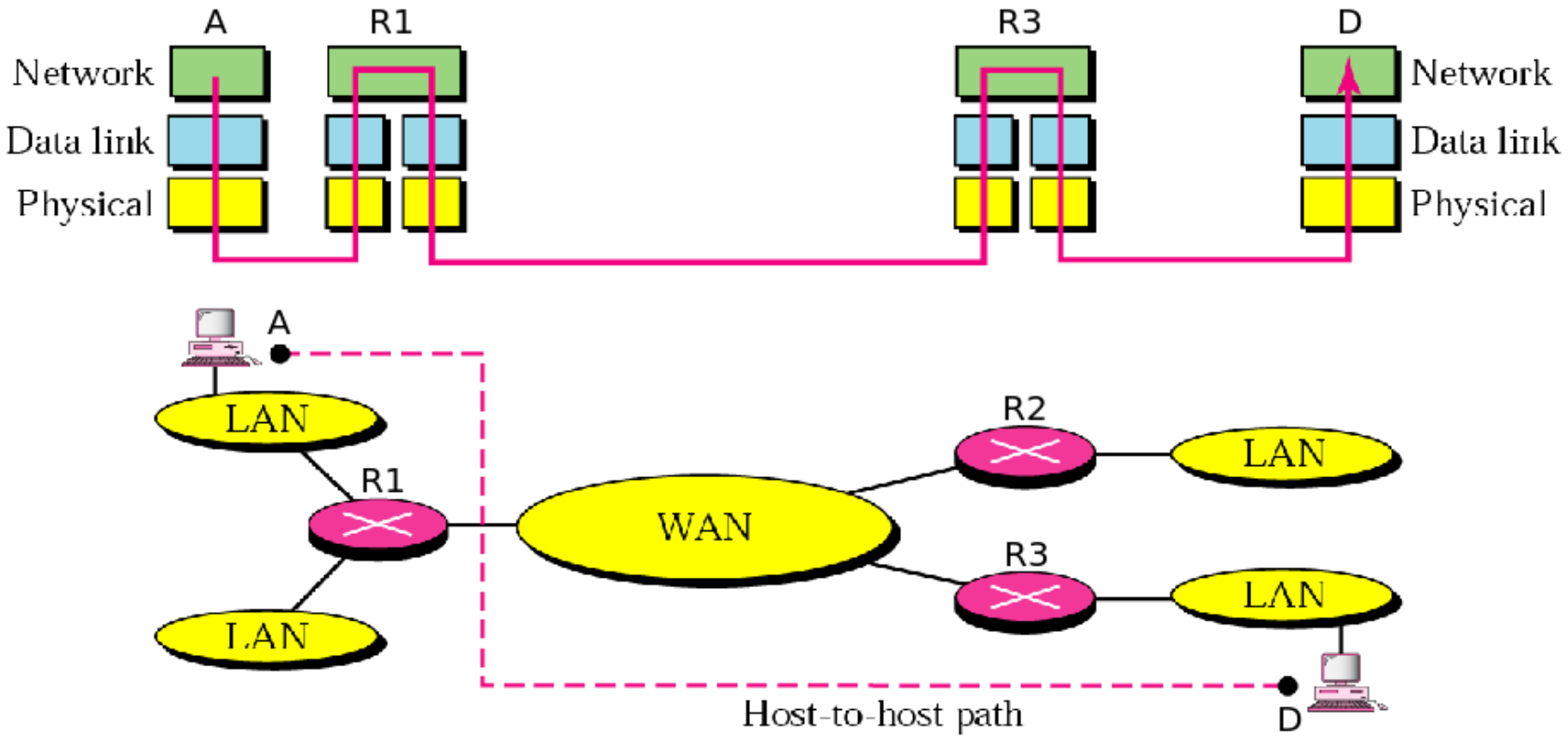
Propojování sítí (Internetworking)

- vzájemné propojování celých sítí i jednotlivých kabelových segmentů (hierarchie)
- propojením vzniká tzv. *internetwork*, zkráceně *internet*
  - **internet** = jakékoliv propojení dvou či více sítí
  - **Internet** = jméno jedné konkrétní sítě (celosvětového Internetu)
- důvody pro internetworking:
  - překonání technických omezení/překážek – např. omezený dosah kabelových segmentů
  - optimalizace fungování sítě – snaha regulovat tok dat, zamezení zbytečného šíření provozu
  - zpřístupnění vzdálených zdrojů – přístup ke vzdáleným serverům
  - zvětšení dosahu poskytovaných služeb – elektronická pošta, internetové telefonování, ...

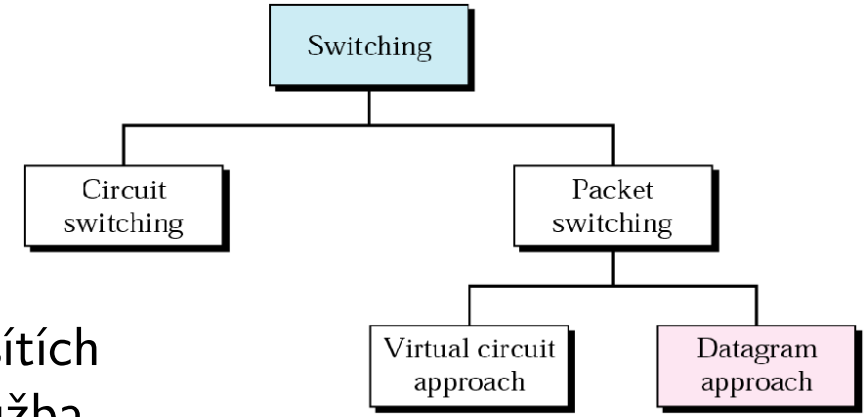
Internetworking – obecná podstata

- rozdíl dle vrstvy, na které propojovací zařízení operuje:
  - fyzická vrstva: *opakovač (repeater)* – viz minule
  - vrstva datového spoje: *můstek (bridge)*, *přepínač (switch)* – viz minule
  - síťová vrstva: *směrovač (router)* – dnes
  - aplikační vrstva: *brána (gateway)* – v budoucnu

Internetworking na L3



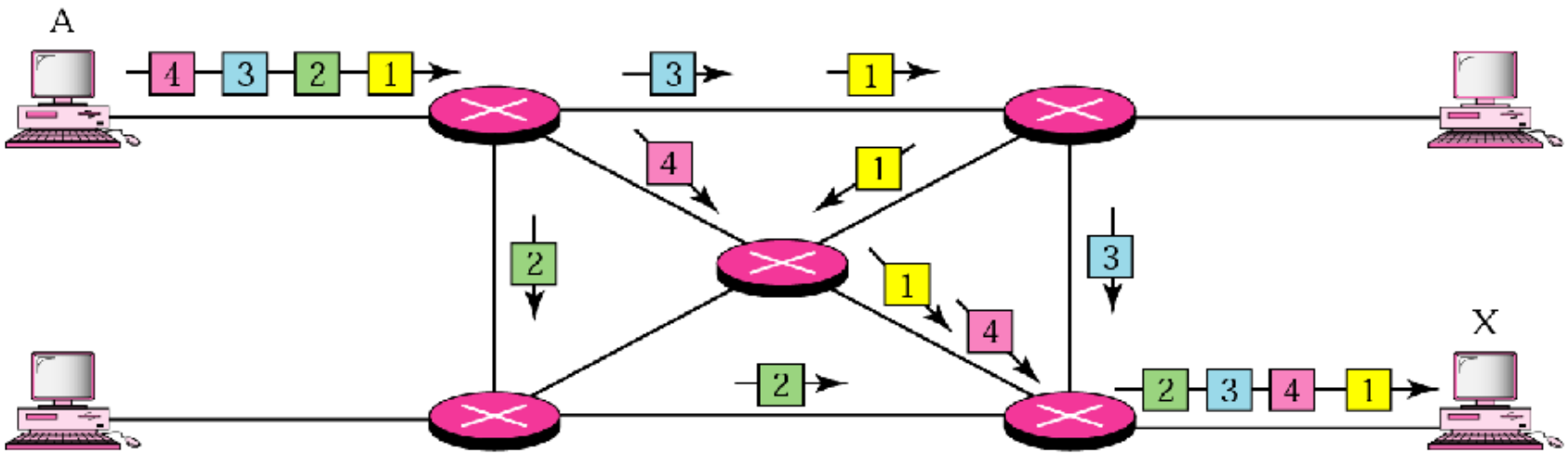
Internetworking – modely zajištění síťových služeb



- přepínání okruhů (*Circuit Switching*):
  - ustavení přímého fyzického spojení mezi odesílatelem a příjemcem
  - bez potřeby paketizace
  - vrstva L1, využito ve spojovaných sítích
  - spojovaná (*connection-oriented*) služba
- přepínání paketů (*Packet Switching*):
  - zasílání nezávislých datových jednotek (paketů)
  - *virtuální kanály (Virtual Circuits Approach):*
    - na začátku přenosu ustavena cesta (implementováno na L2/L3)
    - všechny pakety jedné relace putují po stejné trase
    - využito ve WANs, Frame Relay, ATM (viz *PV169: Základy přenosu dat*)
    - spojovaná (*connection-oriented*) služba
  - *datagramový přístup (Datagram Approach):*
    - každý paket obsluhován zcela nezávisle na ostatních
    - nespojovaná (*connectionless*) služba
    - pakety jsou zde nazývány *datagramy*
    - Internet

Internetworking – Datagram Approach

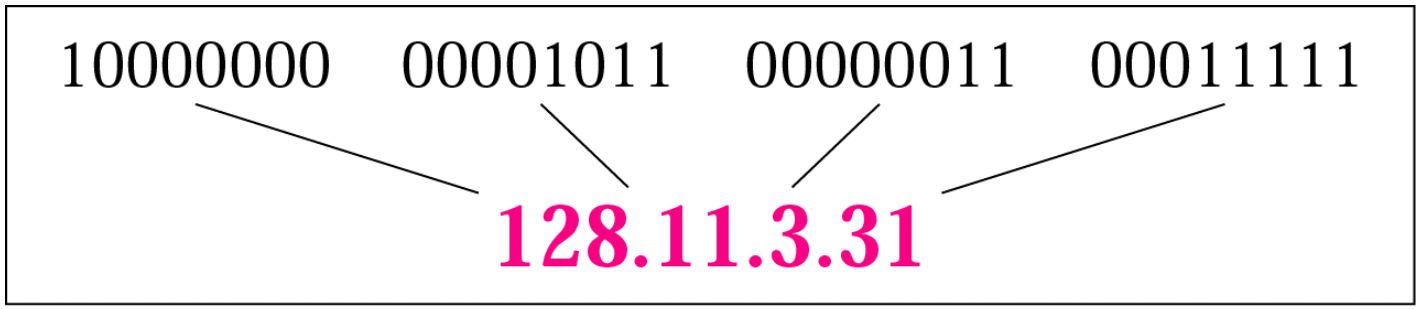
Internet **na síťové vrstvě** využívá *datagramový přístup* k přepínání paketů, komunikace je *nespojovaná*.



Obrázek: Ilustrace datagramového přístupu k přepínání paketů.

Adresace na L3

- požadavek *jednoznačné identifikace* každého zařízení připojeného k Internetu
- nutnost *systematického přidělování adres*
  - za účelem snadnějšího směrování
- každému zařízení/rozhraní přiřazena *Internetová adresa* (*IP adresa*)
  - *IPv4 adresa* (32 bitů) vs. *IPv6 adresa* (128 bitů)

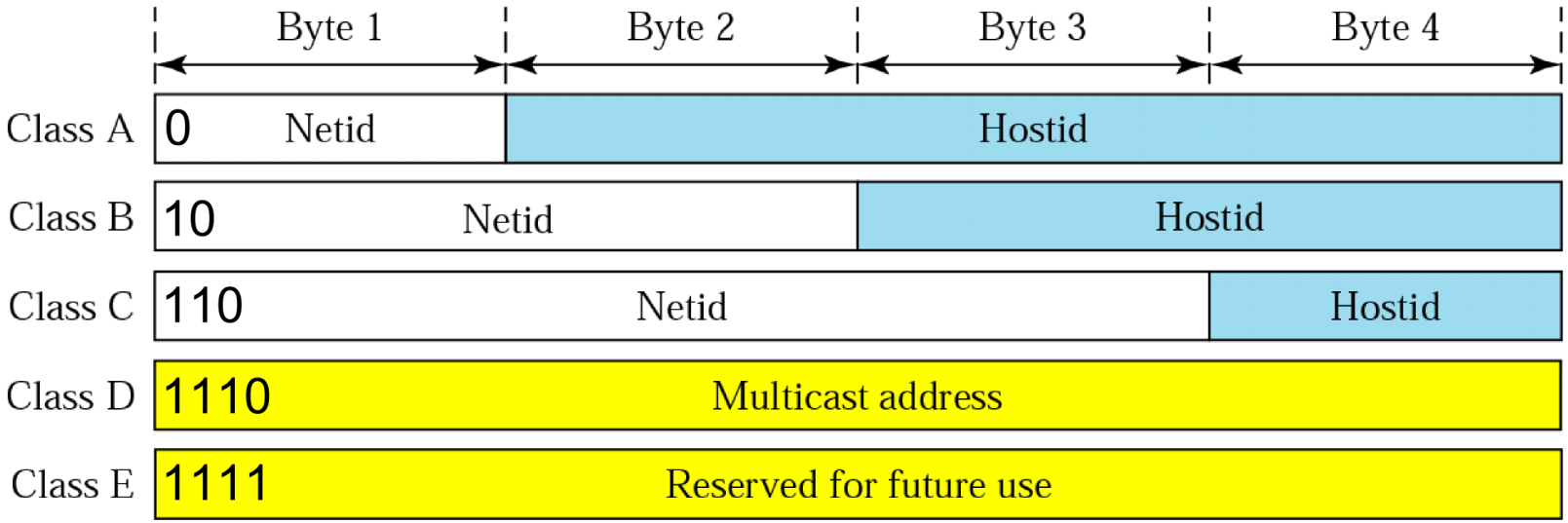


IPv4 – typy adres

- *Individuální (unicast) adresy* – identifikace jednoho síťového rozhraní
  - identifikace jediného odesílatele/příjemce
- *Broadcast adresy* – slouží pro zasílání dat všem možným příjemcům na dané LAN („all-hosts broadcast“)
  - zdrojová adresa datagramu (identifikace odesílatele) je unicastová
- *Skupinové (multicast) adresy* – slouží pro adresování skupiny příjemců (síťových rozhraní), kteří o data **projevili zájem**
  - data směrovači rozesílána všem členům skupiny
  - zdrojová adresa datagramu (identifikace odesílatele) je unicastová

Přidělování adres – Classful Addressing

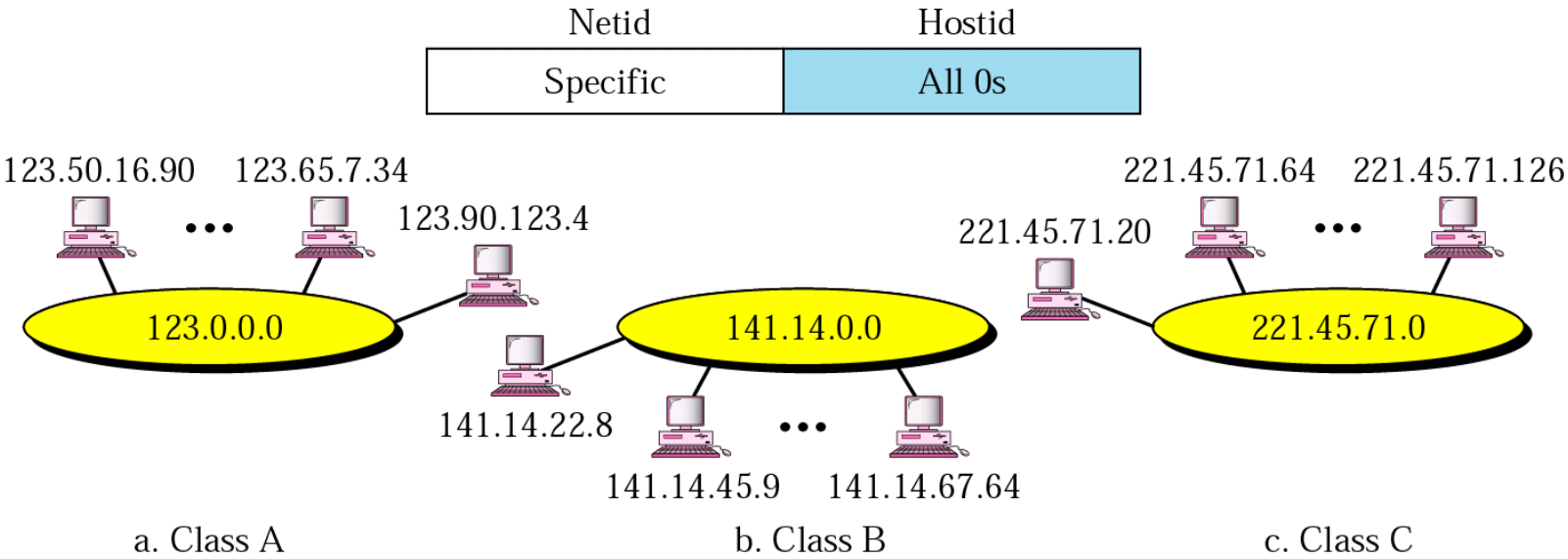
- *Classful Addressing*:
  - zcela první metoda přidělování adres
  - adresní prostor rozdělen do 5 tříd:
    - **třída A:** 2^7 sítí, každá z nich 2^{24} uzlů
    - **třída B:** 2^{14} sítí, každá z nich 2^{16} uzlů
    - **třída C:** 2^{21} sítí, každá z nich 2^8 uzlů
    - **třída D:** multicastové adresy
    - **třída E:** rezervovaný prostor



Přidělování adres – Classful Addressing

NetID vs. HostID

- *Adresa sítě (NetID)*:
  - identifikuje danou síť (nemůže být přidělena uzlu/rozhraní)
  - tuto identifikaci lze využít pro směrování (viz později)
- *Adresa uzlu/rozhraní (HostID)*:
  - identifikuje jedinečný uzel v síti NetID



Příklad: HostID = 147.251.48.1 $\Rightarrow$ třída B $\Rightarrow$ NetID = 147.251.0.0

Problémy Classful adresování

- nedostatečná granularita – každá třída rozdělena na pevný počet sítí s pevnou maximální velikostí
  - = plýtvání adresním rozsahem
  - organizace chce využít 10 IP adres? Dostane C třídu (256 adres)
  - organizace chce využít 270 IP adres? Dostane B třídu (65536 adres)
  - organizace chce využít 70000 IP adres? Dostane A třídu (2097152 adres)
  - *možné řešení:* přidělování více síťových adres menší třídy
- popsané řešení generuje nárůst *směrovacích tabulek*
  - roste objem směrovacích informací, které musí být zpracovávány při rozhodování o volbě dalšího směru procházejícího paketu
  - nutnost prohledávání tabulek (lineární složitost)

Ilustrace problému: organizace s 1500 uzly

1. přidělena adresa třídy B $\Rightarrow$ zabráno 65536 adres $\Rightarrow$ 1 záznam ve sm.tabulce
2. přiděleno 8 adres třídy C $\Rightarrow$ zabráno 2048 adres $\Rightarrow$ 8 záznamů ve sm.tabulce

Problémy Classful adresování – řešení

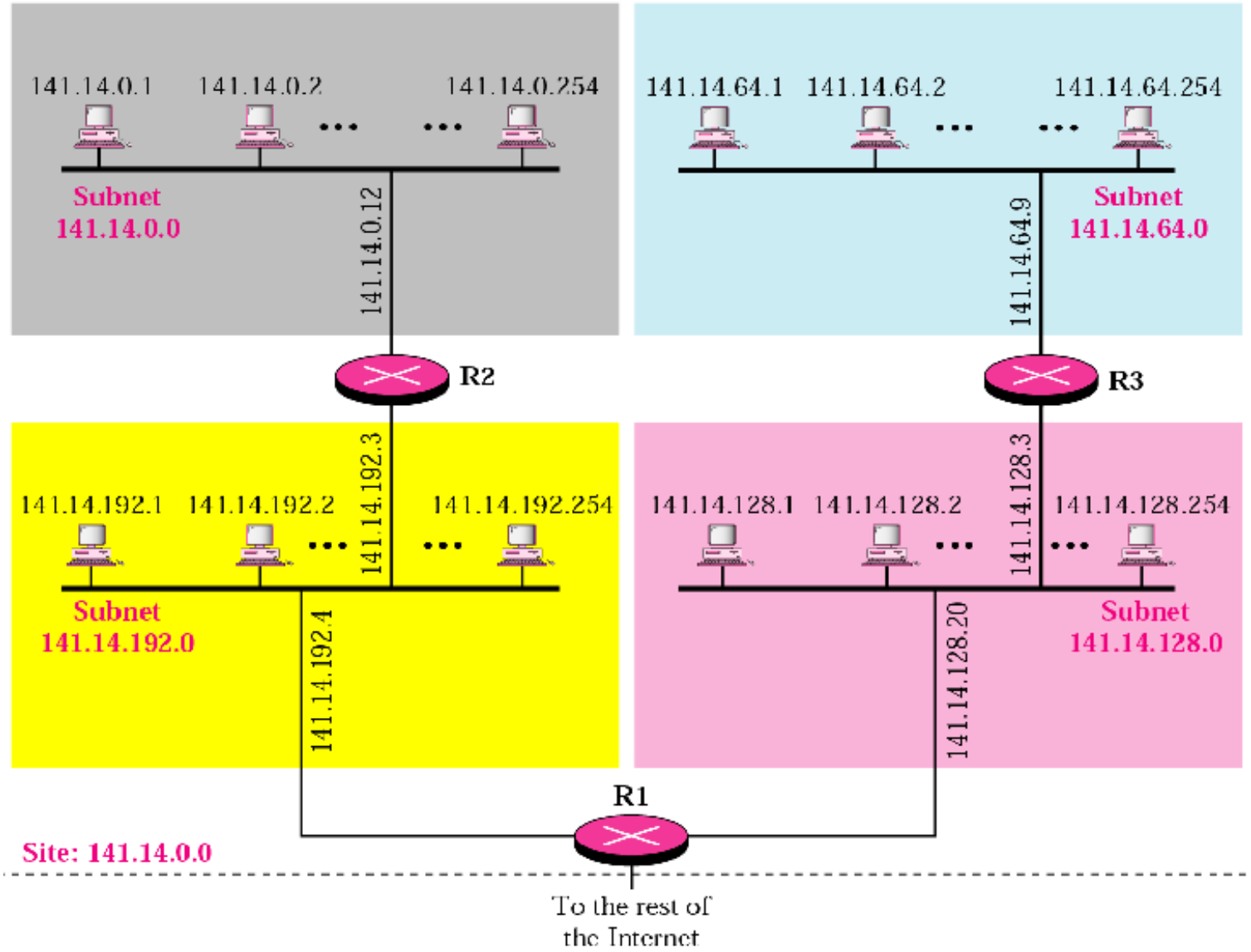
Subnetting, Supernetting

- Lze přidělenou adresu sítě dále dělit do menších podsítí?
  - např. rozdělení sítě dle organizačních složek v rámci jedné organizace
  - *Subnetting*
- Lze využít skutečnosti, že organizace má přidělen souvislý blok adres určité třídy?
  - a snižovat tak velikost směrovacích tabulek
  - *Supernetting*

Classful adresování – Subnetting

- standardní IP adresa poskytuje dvouúrovňovou hierarchii
  - adresa sítě a adresa uzlu
- *Subnetting* zavádí možnost tříúrovňové hierarchie
  - adresa sítě, adresa podsítě a adresa uzlu
  - využitelné v nějaké geograficky omezené oblasti (velké organizace, univerzity, ISPs)
  - síť rozdělena na menší podsítě (*subnetworks (subnets)*)
  - důležitý princip *uzavřenosti*:
    - „zvenčí“ (z pohledu Internetu) se jeví jako 1 síť (1 záznam ve sm. tabulkách), podsítě se rozlišují až na hraničním směrovači
    - tj. má pouze lokální platnost, nikoli platnost globální

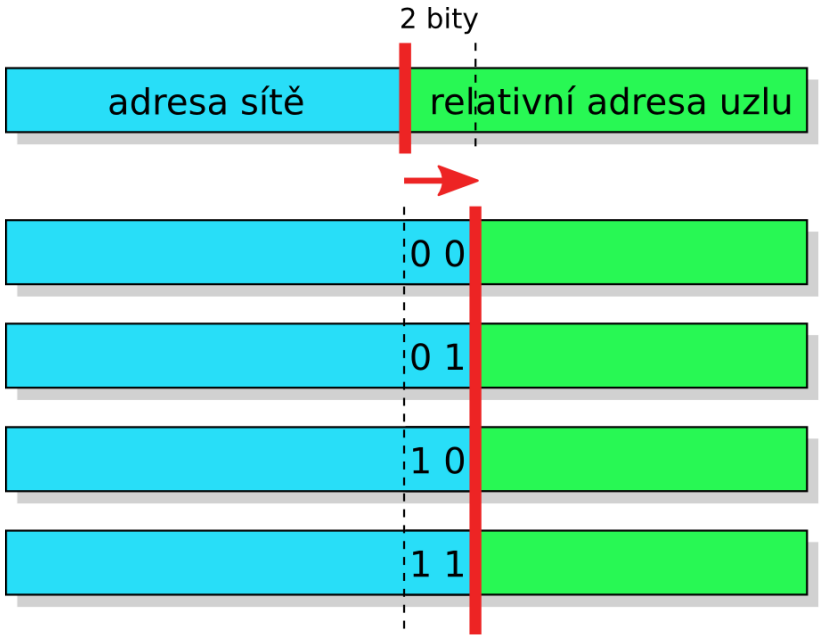
Classful adresování – Subnetting



Classful adresování – Supernetting

- *Supernetting*:
  - pravý opak subnettingu, posouvá pomyslnou dělící čáru mezi oběma složkami IP adresy směrem k vyšším bitům
  - spojuje (agreguje) několik původně samostatných síťových IP adres v jednu výslednou
  - musí však jít o „sousední“ síťové adresy
    - síťové IP adresy se musí shodovat v určitém počtu vyšších bitů své síťové části
    - a musí vyčerpávat všechny bitové kombinace v příslušném počtu nižších bitů (své síťové části)

Classful adresování – Subnetting vs. Supernetting



(a) Subnetting

(b) Supernetting

Classful adresování – Maska sítě/podsítě

- oba způsoby vyžadují mechanismus pro identifikaci bitů, které identifikují síť
  - v rámci subnettingu nezbytné jen na hraničních směrovačích
  - v rámci supernettingu nezbytné na všech směrovačích
- využitý mechanismus – *maska sítě*
  - 32-bitový řetězec (v rámci IPv4)
  - obsahuje 1 v těch bitech, které odpovídají síťové části adresy, 0 tam, kde jde o relativní adresu uzlu v rámci sítě
  - IP adresa uzlu && maska sítě = adresa sítě

Class	Binary form	Decimal form	Using slash
A	11111111 00000000 00000000 00000000	255.0.0.0	/8
B	11111111 11111111 00000000 00000000	255.255.0.0	/16
C	11111111 11111111 11111111 00000000	255.255.255.0	/24
---	11111111 11111000 00000000 00000000	255.248.0.0	/13
---	11111111 11111111 11111111 10000000	255.255.255.128	/25

Přidělování adres – Classless Addressing

- do poloviny 90. let adresy přidělovány pouze v rámci tříd
  - nejmenší počet přidělených adres – 256 (třída C)
- *Classless Addressing*:
  - zobecnění a rozšíření subnettingu/supernettingu
  - zavádí zcela variabilní délku bloku adresy sítě
    - identifikace sítě = adresa sítě a maska sítě
  - adresy se přidělují *hierarchicky*
    - umožnění agregace směrování (viz později) $\Rightarrow$ snaha o minimalizaci velikosti směrovacích tabulek
  - opodstatnění subnettingu zůstává

Classless Addressing – *Classless Inter-Domain Routing (CIDR)*

- konvence popisující „pravidla hry“ – použití IP adres, významu masek, supernetting a subnetting
- nahrazuje původní „třídní“ charakter IP adres (třídy A, B a C)
  - IP adresy přidělovány po tzv. *CIDR blocích*
- velikost CIDR bloku dána příslušnou maskou
  - možno velmi pružně přizpůsobovat
- $\Rightarrow$ snížení tempa vyčerpávání adresového prostoru

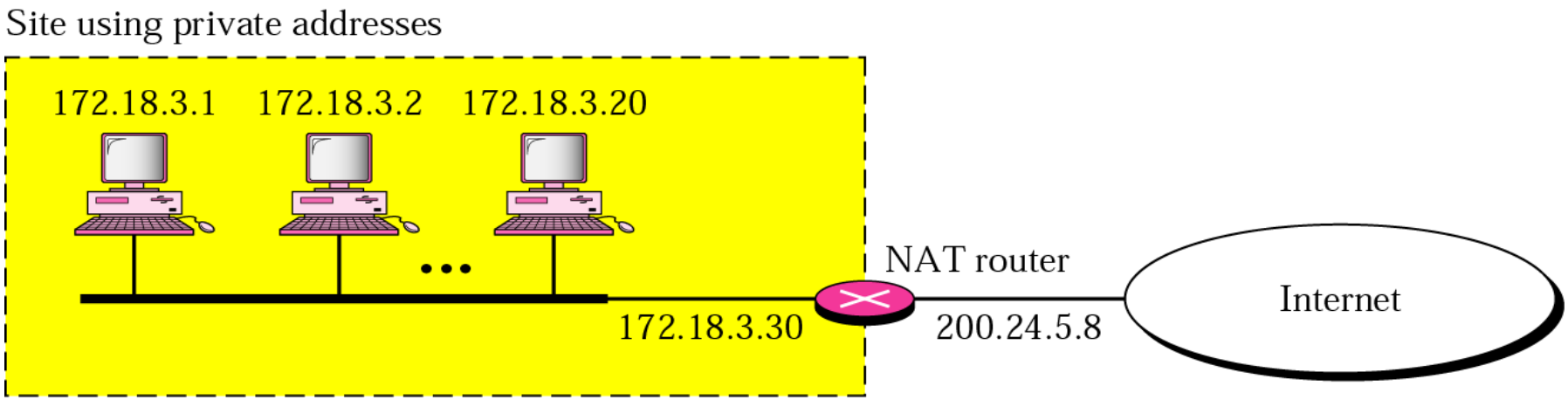
- *Důsledek CIDRu:* adresy závislé na poskytovateli
  - původně IP adresy nezávislé na způsobu jejich připojení
  - zavedení závislosti
    - poskytovatel získává CIDR blok, který si rozděluje dle svého uvážení
    - vnější směrovače směrují jen na základě CIDR bloku
    - při změně poskytovatele je potřeba síť přeadresovat (přečíslovat)

Network Address Translation (NAT)

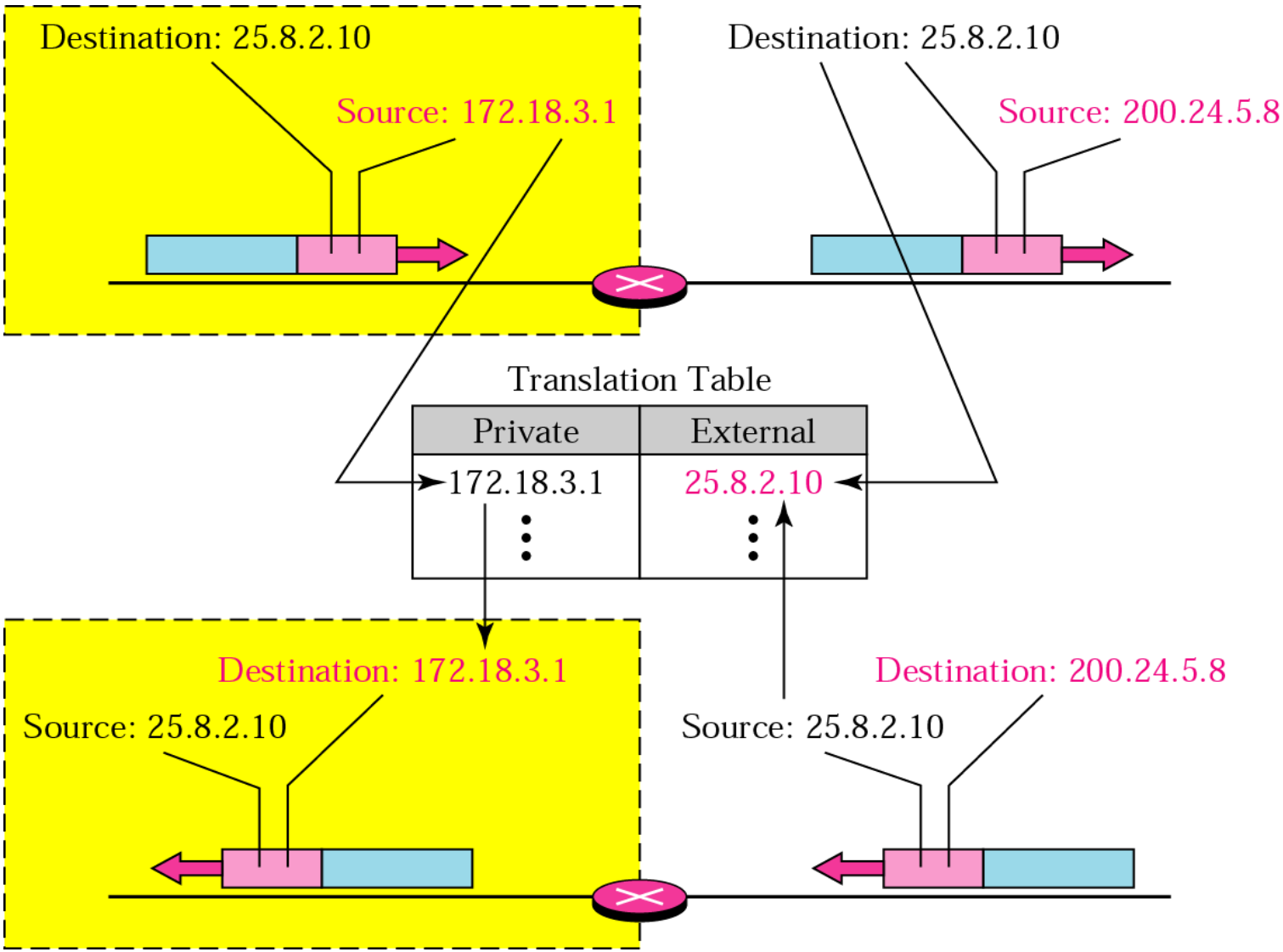
- další mechanismus pro snížení tempa vyčerpávání adresového prostoru
- určeno zejména pro domácí uživatele
  - původně připojováni modemy → možnost dynamického přidělování adres
  - nyní ADSL, kabelová připojení – (většinou) trvalá alokace adres
  - časté požadavky na přidělení více IP adres
- **řešení:** *Network Address Translation (NAT)*
  - „skrývání“ vnitřní sítě za jednu/několik externích adres
    - v rámci vnitřní sítě možnost využít mnoho interních adres
    - rezervované privátní adresy (viz obrázek), unikátní v rámci organizace
    - *vedlejší efekt:* ochrana vnitřní sítě
  - *překlad adres* procházejících síťovým prvkem (např. NAT směrovačem)

Range			*Total*
10.0.0.0	to	10.255.255.255	2^{24}
172.16.0.0	to	172.31.255.255	2^{20}
192.168.0.0	to	192.168.255.255	2^{16}

Network Address Translation (NAT) – ilustrace



Network Address Translation (NAT) – překlad adres



Network Address Translation (NAT) – překlad adres II.

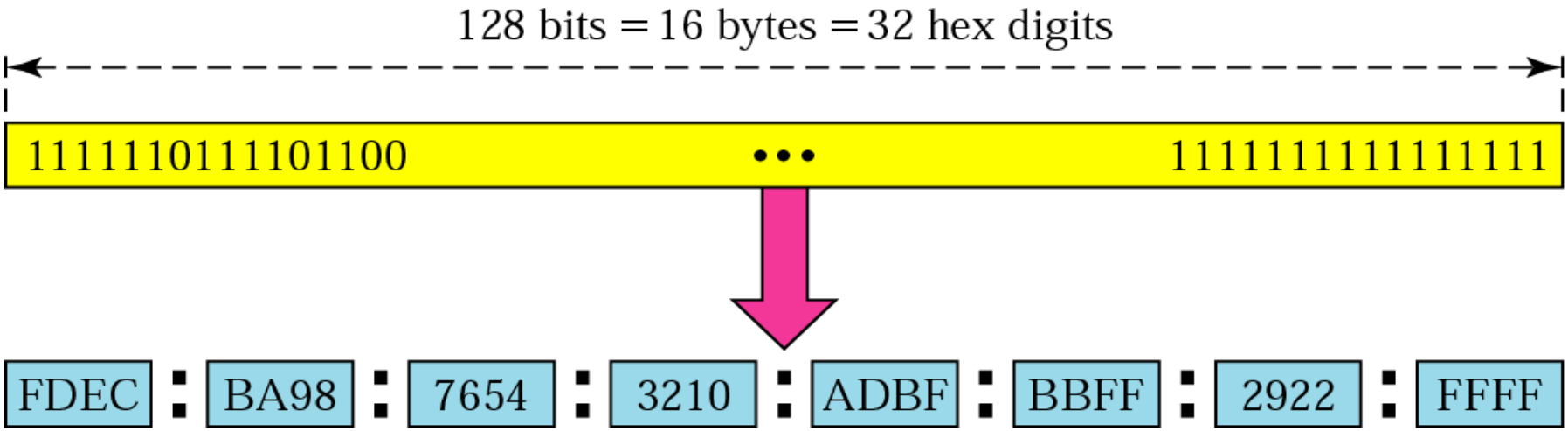
- překlad adres odchozích paketů je triviální
- překlad adres příchozích paketů vyžaduje dodatečné informace:
  - kterému stroji z vnitřní sítě mají být data přeposlána?
  - *překladové tabulky (translation tables)*

Private Address	*Private Port*	*External Address*	*External Port*	*Transport Protocol*
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
...	...	...	...	...

Obrázek: Ukázka překladové tabulky.

IPv6 adresy

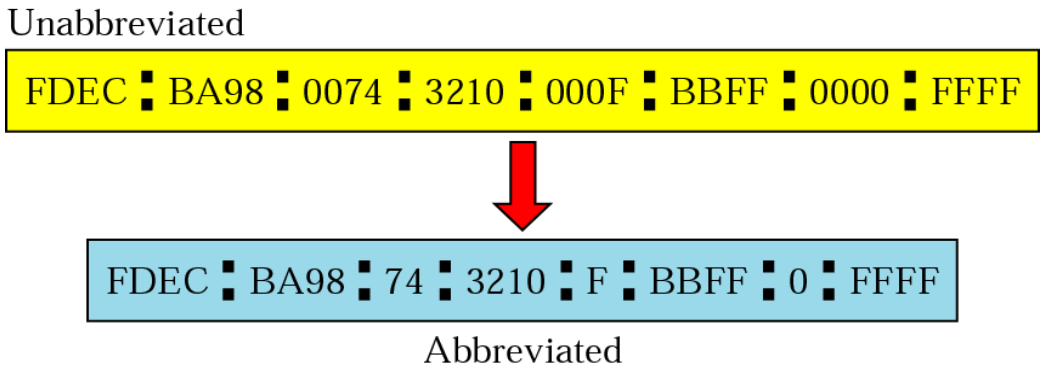
- adresy využívané protokolem IPv6 (viz dále)
- (prozatím) finální řešení nedostatku IP adres
- IPv6 adresa má 128 bitů (= 16 bajtů):
  - 2^{128} možných adres ($\approx 3 \times 10^{38}$ adres $\Rightarrow \approx 5 \times 10^{28}$ adres na každého obyvatele Země)
  - hexadecimální zápis místo dekadického (po dvojicích bajtů oddělených znakem „:“)



IPv6 adresy – zkracování zápisu

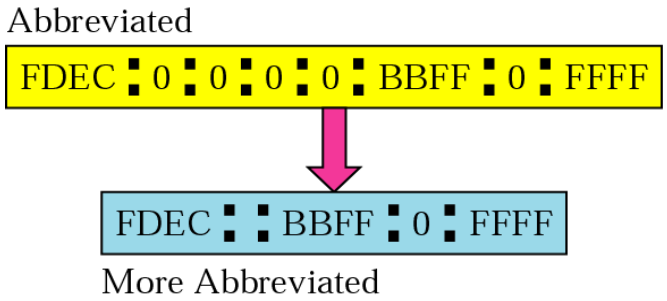
Úvodní nuly lze ze zápisu každé skupiny vynechat:

- 0074 lze psát jako 74, 000*F* jako *F*, ...
- 3210 **nelze** zkracovat!



Po sobě jdoucí nulové skupiny lze vynechat:

- vždy však **pouze jednu** takovou nulovou skupinu!



IPv6 adresy – hierarchie

- cílem opět usnadnění směrování
- strukturu individuálních IPv6 adres definuje RFC 3587
- základní struktura:

n bitů	64-n bitů	64 bitů
globální směrovací prefix	adresa podsítě	adresa rozhraní

- globální směrovací prefix $\approx$ adresa sítě
- adresa podsítě obvykle 16 bitů $\Rightarrow$ globální prefix 48 bitů
  - prvních 16 bitů obsahuje hodnotu 2001_{16}
  - dalších 16 bitů přiděluje regionální registrátor (RIR)
  - dalších 16 bitů přiděluje lokální registrátor (LIR)

16 bitů	16 bitů	16 bitů	16 bitů	64 bitů
2001	přiděluje RIR	přiděluje LIR	adresa podsítě	adresa rozhraní

IPv6 adresy && CIDR

- IPv6 adresace je pouze *classless*, třídy neexistují
- sítě v IPv6 popisovány s využitím notace CIDR (stejně jako v IPv4)
- např. *FDEC:0:0:0:0:BBFF:0:FFFF/60*

IPv6 adresy – typy adres

- *Individuální (unicast) adresy* – totéž co v IPv4, identifikace jednoho síťového rozhraní
- *Skupinové (multicast) adresy* – totéž co v IPv4, slouží pro adresování skupin počítačů či jiných síťových zařízení
  - data jsou vždy doručena všem členům skupiny
  - prefix `ff00::/8`
- *Výběrové (anycast) adresy* – novinka v IPv6
  - také označují skupinu příjemců
  - data se však doručí jen jedinému jejímu členovi (tomu, který je nejblíže)

- broadcast adresy IPv4 protokolu se v IPv6 nevyužívají
  - nahrazeny speciálními multicastovými skupinami (např. všechny uzly na dané lince)

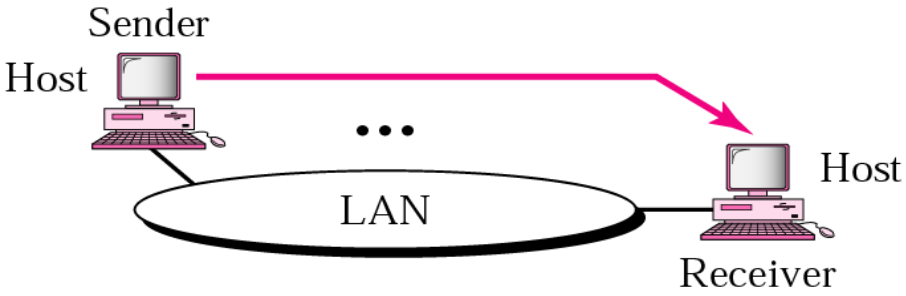
Interakce L3 se spojovou vrstvou (L2) – mapování adres

- mechanismus doručení dat v IP sítích – *hop-by-hop*
- vlastní předání/doručení zprávy na základě *fyzických (MAC) adres*
- 2 alternativy:
  - příjemce na stejné LAN jako odesílatel
    - IP datagram obsahuje IP adresu příjemce, rámec L2 vrstvy MAC adresu příjemce
  - příjemce na jiné LAN než odesílatel
    - IP datagram obsahuje IP adresu příjemce, rámec L2 vrstvy MAC adresu směrovače
    - směrovač po přijetí (a zpracování) datagramu jej vloží do nového rámce s MAC adresou dalšího směrovače ve snaze přiblížit se cíli (odtud hop-by-hop)
    - po dosažení cílové LAN platí alternativa 1 (lokální „odesílatel“ = poslední směrovač)

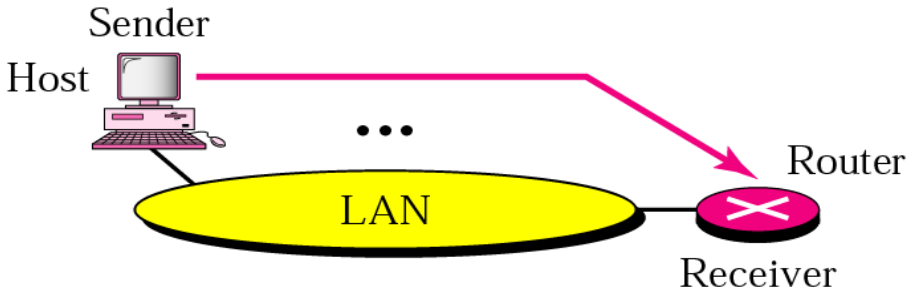
Interakce L3 se spojovou vrstvou (L2) – mapování adres II.

- $\Rightarrow$ nutnost mapování IP adres na fyzické (MAC) adresy
  - *statické mapování*
    - vytvoření statické tabulky párů *(IP adresa, MAC adresa)*
    - obtížně spravovatelné
  - *dynamické mapování*
    - **Address Resolution Protocol (ARP)**

Interakce L3 se spojovou vrstvou (L2) – mapování adres III.

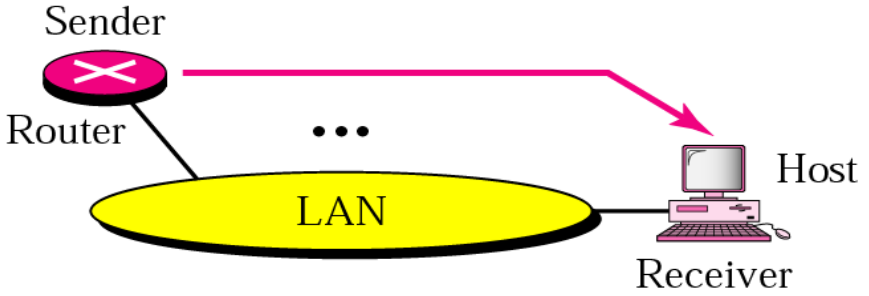


Case 1. A host has a packet to send to another host on the same network.



Case 2. A host wants to send a packet to another host on another network.
It must first be delivered to the appropriate router.

Case 3. A router receives a packet to be sent to a host on another network.
It must first be delivered to the appropriate router.

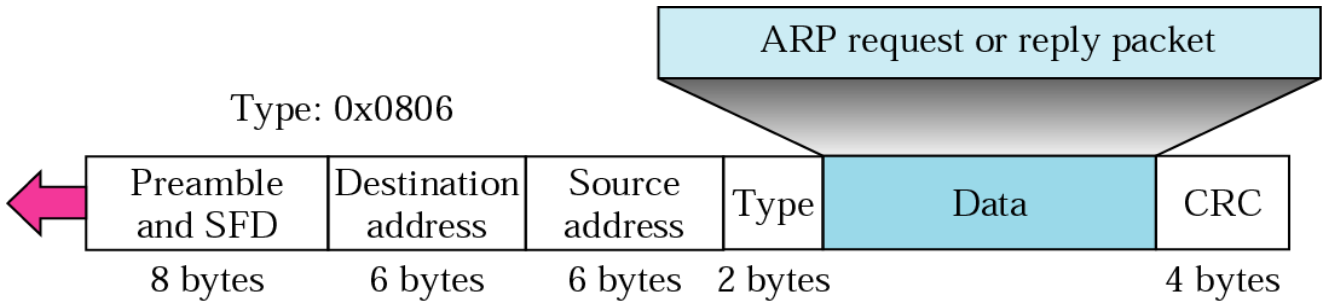


Case 4. A router receives a packet to be sent to a host on the same network.

Obrázek: Případové ilustrace využití ARP protokolu (*hop-by-hop* doručení).

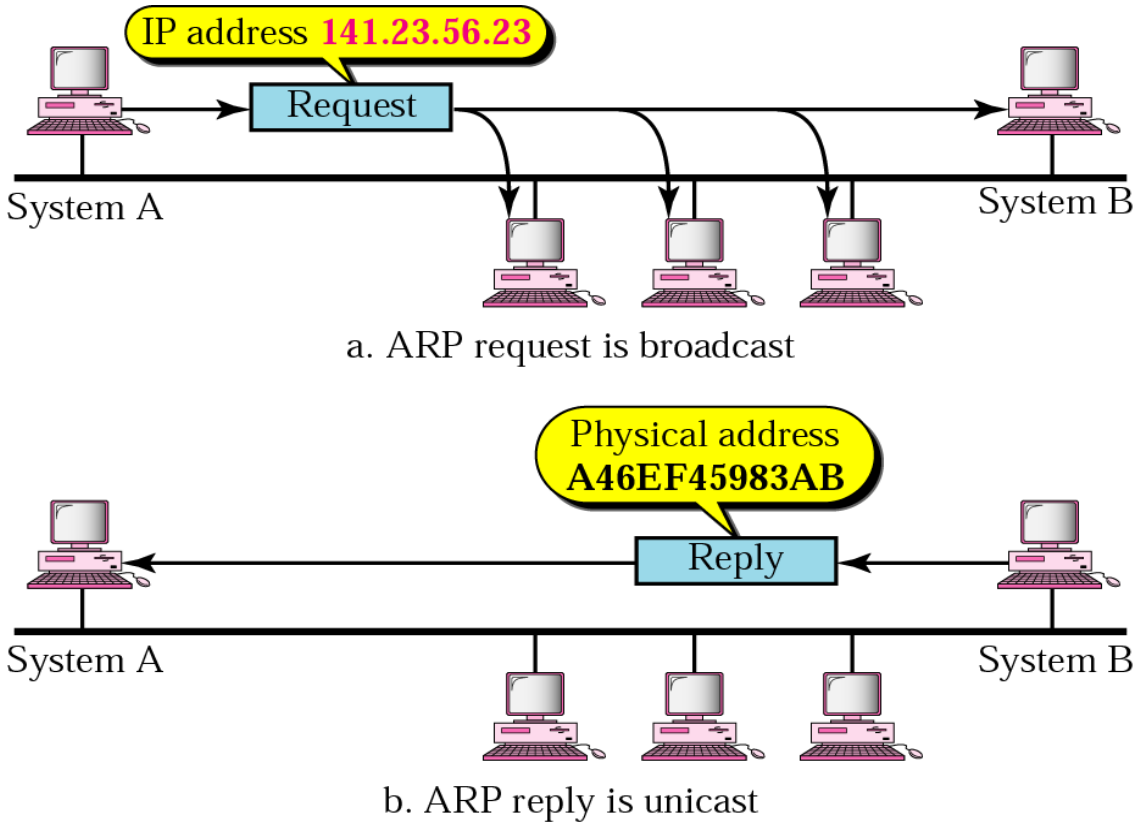
Interakce L3 se spojovou vrstvou (L2) – ARP protokol

- protokol pro zjištění MAC adresy uzlu/směrovače na základě IP adresy
- mechanismus:
  1. zaslání tzv. *ARP request* paketu **všem** uzlům na dané LAN (broadcast)
     - paket obsahuje IP & MAC adresu odesílatele a IP adresu hledaného uzlu
  2. paket zpracován všemi uzly; odpoví jen ten, jehož IP adresa se shoduje s hledanou
     - ostatní paket zahodí
  3. hledaný uzel žadateli odpovídá tzv. *ARP reply* paketem
- ARP pakety baleny přímo do rámců L2 vrstvy



- protokol **RARP (Reverse Address Resolution Protocol)**
  - zpětný překlad MAC adres na IP adresy; již se nevyužívá

Interakce L3 se spojovou vrstvou (L2) – ARP protokol II.

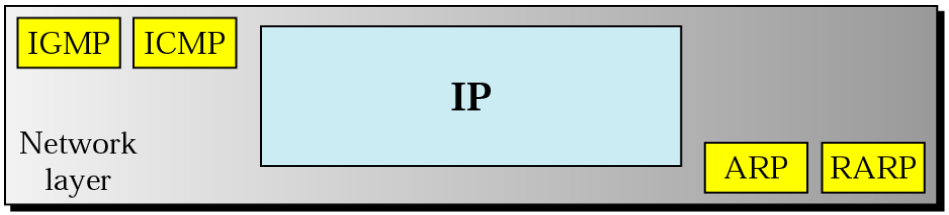


Obrázek: Ilustrace mechanismu operace ARP protokolu.

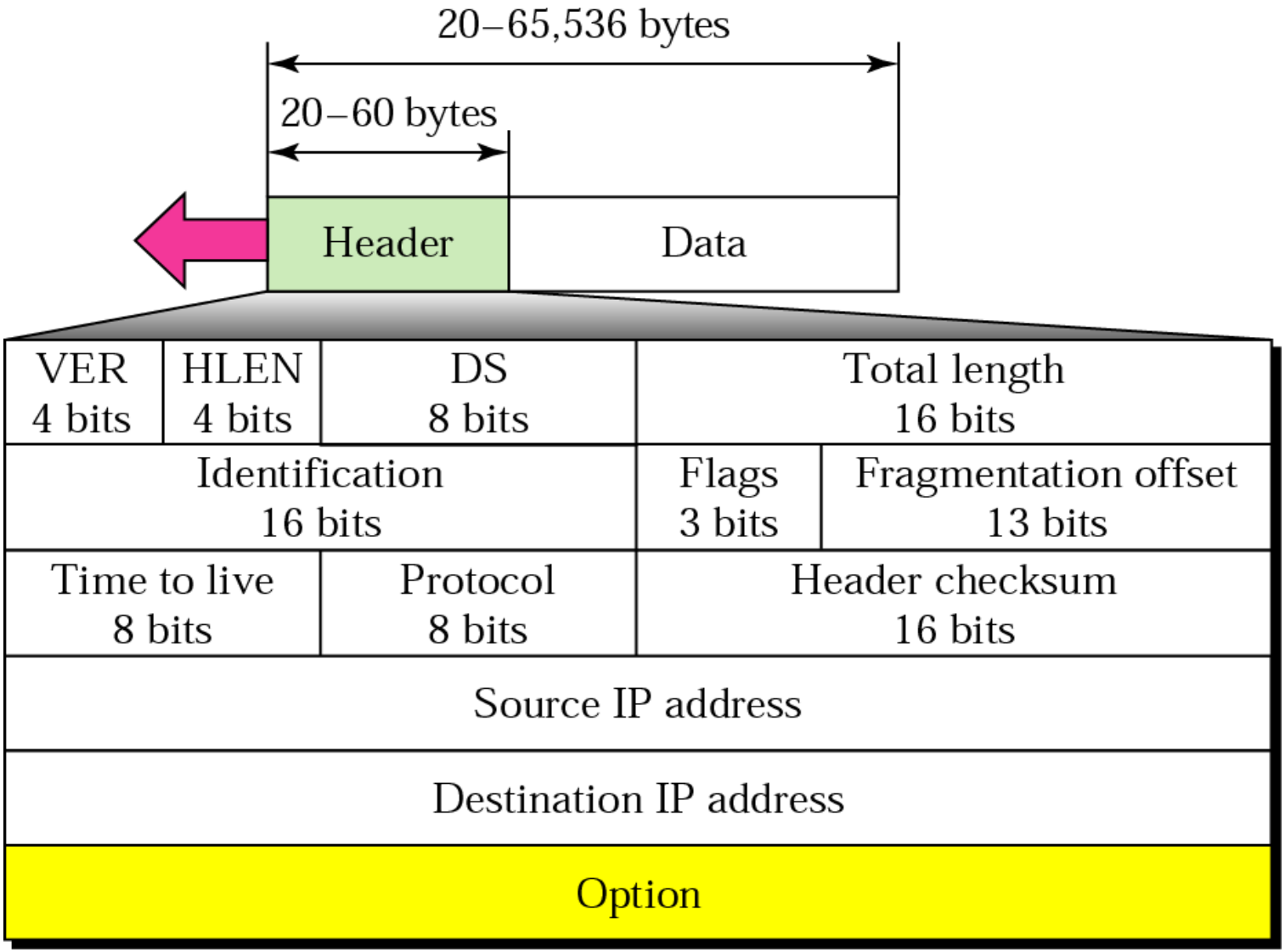
- více viz animace:
  `http://frakira.fi.muni.cz/~jeronimo/vyuka/OsiSchool_ARP.swf`

Internet Protocol (IP protokol)

- nejrozšířenější protokol síťové vrstvy
  - doprava dat (datagramů) na místo jejich určení, a to i přes mezilehlé uzly (směrovače) – *host-to-host delivery*
    - uzly/rozhraní v rámci IP protokolu jednoznačně identifikovány IP adresami
    - využívá *datagramový přístup* k přepínání paketů, komunikace je *nespojovaná*
    - $\Rightarrow$ směrování (příští přednáška)
  - poskytuje nespolehlivou (tzv. *best-effort*) službu
  - doplněn dalšími podpůrnými protokoly (ICMP, ARP, RARP, IGMP)
    - ošetření nestandardních situací, šíření informací potřebných ke korektnímu směrování, identifikace rozhraní na LAN atd.
- navržen a standardizován ve dvou verzích:
  - *Internet Protocol verze 4 (IPv4)* – 1981, RFC 791
  - *Internet Protocol verze 6 (IPv6)* – 1998, RFC 2460



IPv4 datagram

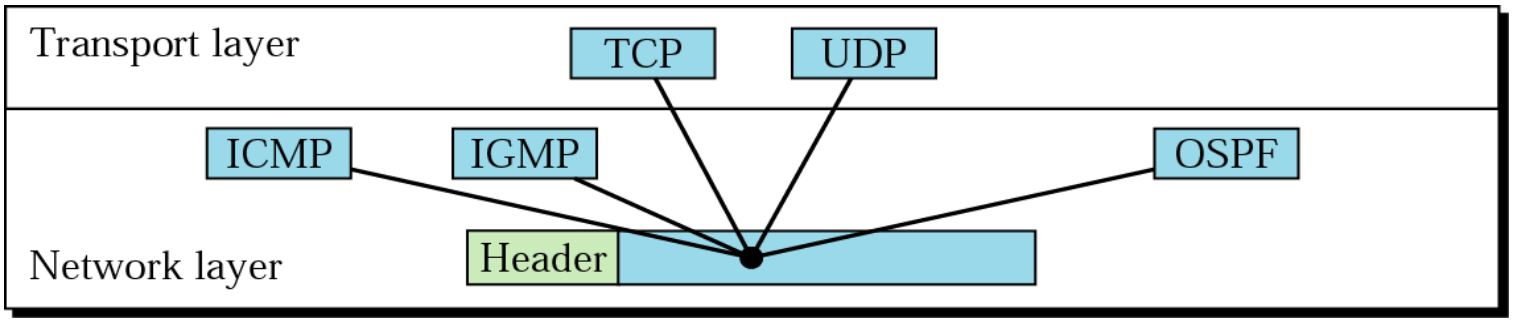


IPv4 datagram II.

- **Version (VER)** – verze IP protokolu
- **Header length (HLEN)** – délka hlavičky IP datagramu (ve 4B slovech)
  - nezbytné kvůli poli *Option* (proměnná délka datagramu)
- **Differentiated services (DS)**, také **Type of service (TOS)** – třída datagramu v rámci kvality služby (QoS)
  - nezbytné pro odlišení „důležitých“ (řídící datagramy, provoz v reálném čase) a „méně důležitých“ datagramů
  - později (konec semestru)
- **Total length** – délka celého IP datagramu (v B)
  - max. $2^{16} - 1 = 65535$ bajtů
- **Identification**, **Flags**, **Offset** – viz Fragmentace v IPv4, slide 52
- **Time to live (TTL)** – řízení maximálního počtu skoků (= směrovačů) navštívených datagramem
  - odesílací uzel vloží číslo ($\approx 2\times$ největší počet skoků mezi libovolnými dvěma uzly)
  - po průchodu směrovačem TTL dekrementováno o 1
  - pokud po dekrementování platí $TTL = 0$, datagram je zahozen

IPv4 datagram III.

- **Protocol** – identifikace protokolu vyšší vrstvy využívajícího služeb IP vrstvy
  - nezbytné pro specifikaci cílového protokolu, kterému má být datagram doručen
    - forma multiplexingu/demultiplexingu
  - identifikátory určeny v online databázi asociace IANA
    - např. 1 = ICMP, 2 = IGMP, 6 = TCP, 17 = UDP, atd.
    - viz `http://www.iana.org/assignments/protocol-numbers`

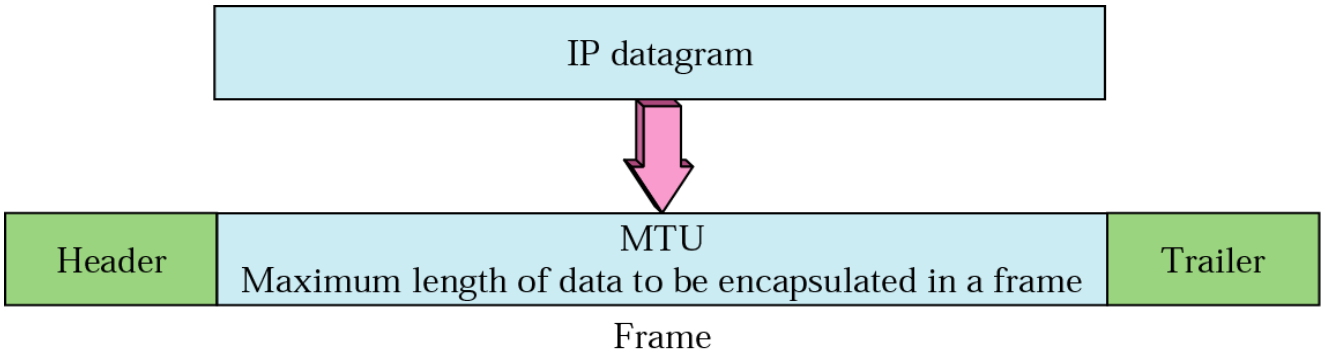


IPv4 datagram IV.

- **Header checksum** – kontrolní součet *hlavičky* IP datagramu
  - bez dat
    - data (resp. transportní protokoly) mají vlastní kontrolní součty
  - hlavní důvod pro zdvojení:
    - nutnost přepočítávání kontrolního součtu na směrovačích díky proměnlivým polím IP datagramu (např. TTL)
    - $\Rightarrow$ počítání kontrolního součtu jen hlavičky = úspora času (data se stejně nemění)
- **Source IP address**, **Destination IP address** – 32-bitová IPv4 adresa identifikující odesílací/přijímající uzel
- **Options** – volitelná součást IP datagramů, určeno zejména pro budoucí rozšíření IPv4
- **Data** – vlastní přenášená data

IPv4 – fragmentace datagramů

- datagram při cestě k cíli prochází různými sítěmi
- ne všechny sítě (resp. využité L2 protokoly) mohou přenášet data stejné velikosti
- *Maximum Transfer Unit (MTU)* – maximální velikost dat, které lze přenést využitým L2 protokolem
  - určuje maximální velikost přenositelného IP datagramu (*Total size*)



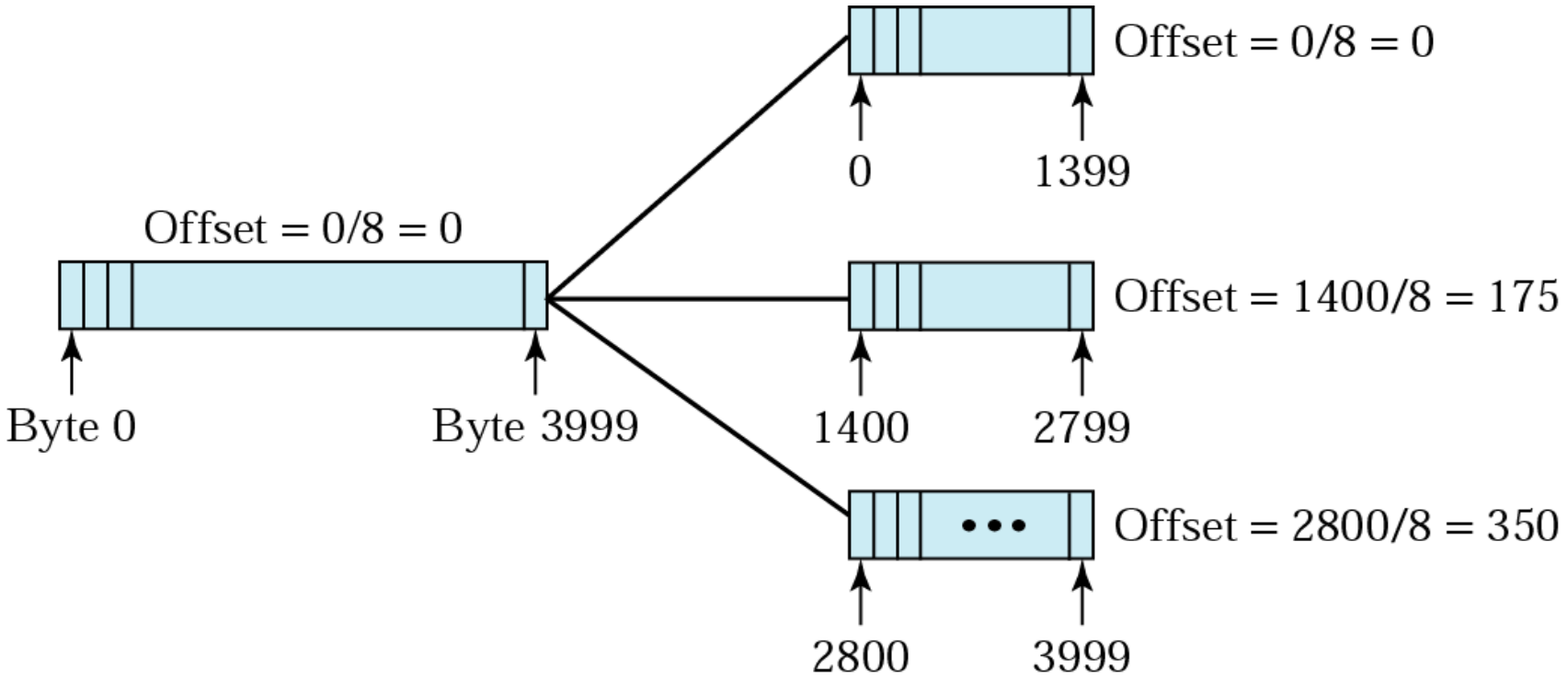
IPv4 – fragmentace datagramů II.

- *situace:*
  - zdrojový uzel chce odeslat datagram, který je větší než MTU výstupní linky
  - směrovač přijme datagram, který je větší než MTU výstupní linky
- *řešení:* provedení tzv. *fragmentace IP datagramu*
  - původní datagram je rozdělen na několik menších datagramů (tzv. *fragmenty*)
  - každý fragment získá svou vlastní IP hlavičku (= stane se z něj nový, plnohodnotný datagram)
  - fragmenty na cílovém uzlu složeny do původního datagramu (před předáním transportnímu protokolu)
- složení fragmentů do původního datagramu vyžaduje:
  - identifikaci datagramu, kterému fragmenty náleží
  - znalost počtu fragmentů
  - znalost pozice každého fragmentu v původním datagramu
- využití polí IP hlavičky: *Identification*, *Flags* a *Offset*

IPv4 – fragmentace datagramů III.

- **Identification** – pole identifikuje původní datagram, kterému fragmenty náleží
  - tj. všechny fragmenty jednoho datagramu mají stejné identifikační číslo
- **Flags** – 3-bitová hodnota:
  - 1 bit rezervovaný
  - *do-not-fragment bit* – hodnota 1 = datagram nesmí být fragmentován (v případě nutnosti generována ICMP zpráva – viz dále)
  - *more-fragment bit* – hodnota 1 = fragment není posledním fragmentem (0 určuje poslední fragment daného datagramu)
- **Offset** – relativní pozice fragmentu v původním datagramu
  - 13 bitů $\Rightarrow$ offset max. 8191 $\Rightarrow$ nelze pokrýt větší datagramy
  - $\Rightarrow$ jednotka offsetu stanovena na 8 B

IPv4 – fragmentace datagramů IV.



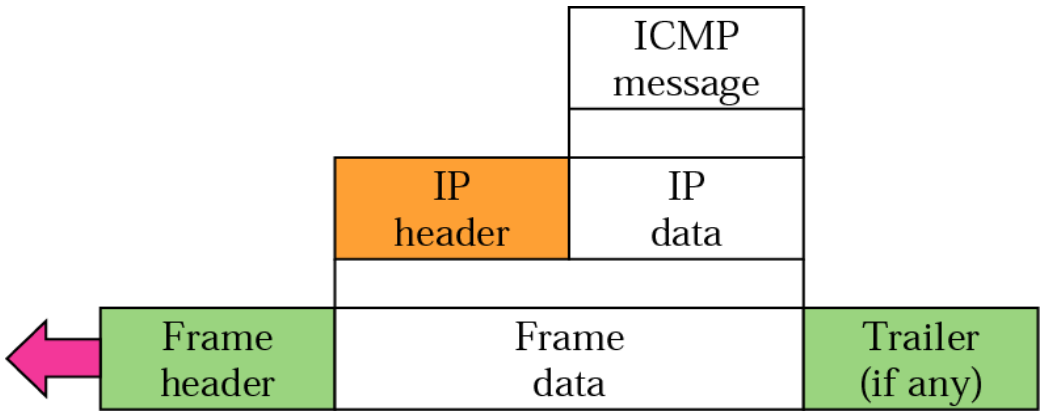
Obrázek: Ukázka fragmentace 4000B datagramu do 3 fragmentů.

IPv4 – fragmentace datagramů V.

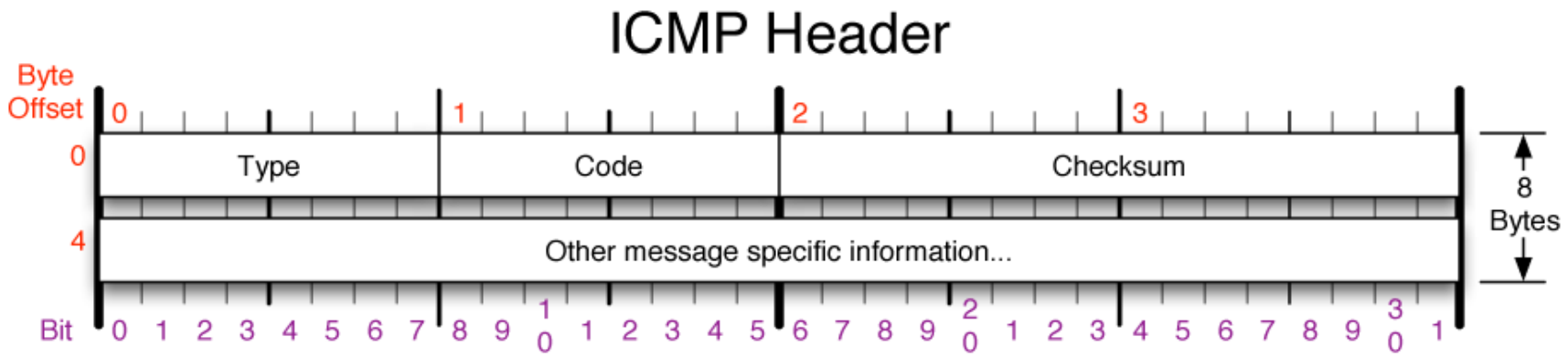
- Kde se fragmentace provádí?
  - na zdrojovém uzlu
  - na směrovači/směrovačích
- Kde se provádí skládání fragmentů?
  - **pouze** na cílovém uzlu
    - ztráta fragmentu = ztráta datagramu
  - na směrovačích nelze skládat ze dvou důvodů:
    - zbytečná zátěž směrovače
    - fragmenty putují sítí nezávisle na sobě (tj. i jinými cestami)
- Možno provádět vícenásobnou fragmentaci
  - „fragmentaci fragmentu“
  - *otázka:* jak bude vypadat hlavička fragmentů fragmentu?

Internet Control Message Protocol (ICMP)

- IP protokol poskytuje nespolehlivou (best-effort) službu
  - bez mechanismů pro informování odesílatele o vzniklých chybách
  - bez podpůrných mechanismů pro zjišťování stavu sítě
- *Internet Control Message Protocol (ICMP)*
  - RFC 792
  - doprovodný protokol IP protokolu
  - poskytuje informace o chybách při přenosu IP datagramů
  - poskytuje základní informace o stavu sítě
- přestože je ICMP protokolem síťové vrstvy, zprávy nejsou předávány linkové vrstvě, ale baleny do IP protokolu
  - hodnota pole *Protocol* v hlavičce IP datagramu nastavena na 1



Internet Control Message Protocol (ICMP) – hlavička



ICMP Message Types

Type	Code/Name
0	Echo Reply
3	Destination Unreachable
	0 Net Unreachable
	1 Host Unreachable
	2 Protocol Unreachable
	3 Port Unreachable
	4 Fragmentation required, and DF set
	5 Source Route Failed
	6 Destination Network Unknown
	7 Destination Host Unknown
	8 Source Host Isolated
	9 Network Administratively Prohibited
	10 Host Administratively Prohibited
	11 Network Unreachable for TOS
	12 Host Unreachable for TOS
	13 Communication Administratively Prohibited
4	Source Quench
5	Redirect
	0 Redirect Datagram for the Network
	1 Redirect Datagram for the Host
	2 Redirect Datagram for the TOS & Network
	3 Redirect Datagram for the TOS & Host
8	Echo
9	Router Advertisement
10	Router Selection
11	Time Exceeded
	0 TTL Exceeded in Transit
	1 Fragment Reassembly Time Exceeded
12	Parameter Problem
	0 Pointer indicates the error
	1 Missing a Required Option
	2 Bad Length
13	Timestamp
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply
30	Traceroute

Checksum

Checksum of entire UDP segment and pseudo header (parts of IP header) (for UDP)

Checksum of ICMP header (for ICMP)

RFC 768 and 792

Please refer to RFC 768 for the complete User Datagram Protocol (UDP) Specification, and to RFC 792 for the Internet Control Message protocol (ICMP) specification.

Copyright 2004 - Matt Baxter - mjb@fatpipe.org

Aktuální přehled definovaných typů ICMP zpráv dostupný na adrese `http://www.iana.org/assignments/icmp-parameters`

Internet Control Message Protocol (ICMP) – příklady zpráv

- oznamy o chybách:
  - *Destination unreachable* – „Destination“ může být protokol, port, uzel nebo celá síť
  - *Time exceeded* – informace o vypršení TTL či informace o vypršení času pro znovusložení fragmentů IP datagramu
- dotazy na stav sítě/uzlu:
  - *Echo request/reply* – požadavek na odpověď

- zprávy obsahují část paketu, který
  - způsobil chybu
  - na který se váže odpověď

- přímé využití ICMP v aplikacích:
  - program `ping` – využití ICMP Echo request/reply
  - program `traceroute` – využití ICMP Time exceeded (TTL expired)

Internet Control Message Protocol (ICMP) – omezení

- ochrana proti rekurzivnímu generování:
  - Chybový ICMP paket *není generován* jako reakce na:
    - ICMP chybu
    - broadcast nebo multicast zprávu
    - poškozenou IP hlavičku (špatná cílová adresa)
    - chybu fragmentu (kromě prvního)
- generování ICMP zpráv často výkonnostně omezeno

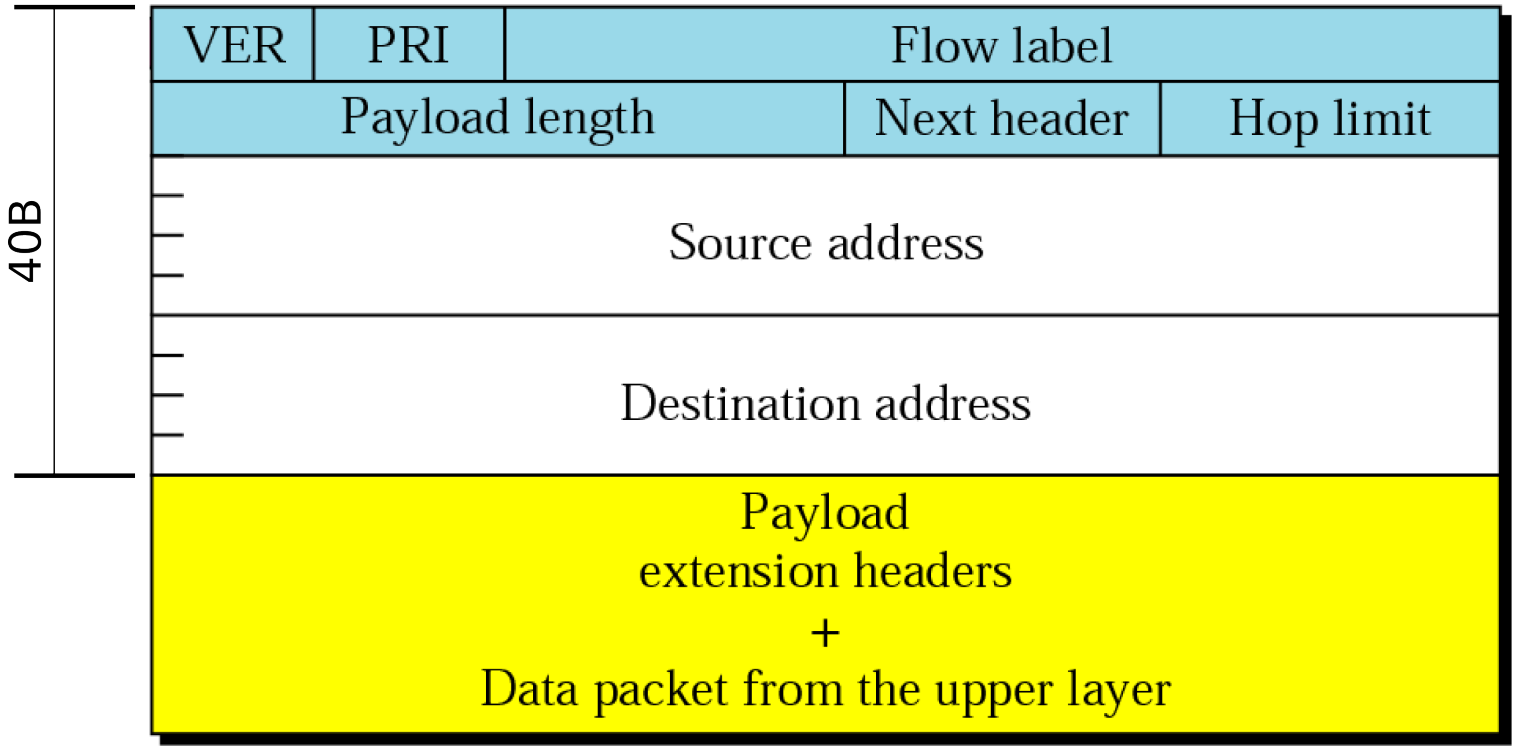
IP protokol verze 6 (IPv6) – Proč nový protokol?

- *hlavní impulz pro návrh nového IP protokolu:* relativně rychlé vyčerpávání adresního prostoru IPv4 protokolu
- další důvody: problémy IPv4, které vyvstaly s rozvojem Internetu, zejména
  - slabá podpora přenosů aplikací reálného času
  - žádná podpora zabezpečené komunikace na úrovni IP
  - žádná podpora autokonfigurace zařízení
  - žádná podpora mobility
  - atp.
- (mnoho vlastností do IPv4 zpětně doimplementováno)

IP protokol verze 6 (IPv6) – vlastnosti

- *rozšířený adresní prostor* – 128-bitová IPv6 adresa, 2^{128} jedinečných adres
- *jednodušší formát hlavičky* – základní 40B hlavička obsahující pouze nejnutnější informace
- *možnosti dalšího rozšíření* – skrze tzv. *rozšiřující hlavičky*
- *podpora přenosů reálného času* – značkování toků, prioritizace provozu
- *podpora zabezpečení přenosu* – podpora autentizace, šifrování a verifikace integrity přenášených dat
- *podpora mobility* – skrze tzv. *domácí agenty*
- *podpora autokonfigurace zařízení* – stavová a bezstavová konfigurace

IPv6 datagram – základní hlavička

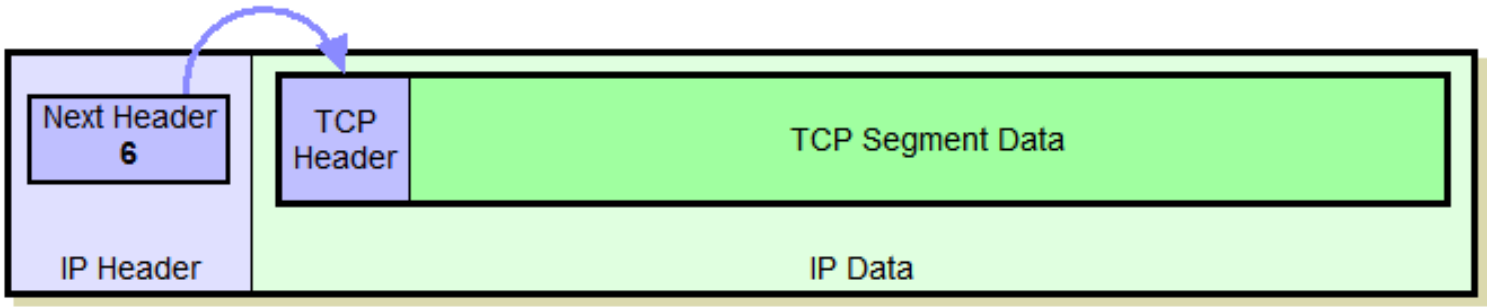


- pevná velikost základní hlavičky (40 B)
- kontrolní součet, volby (*options*) a fragmentační informace nejsou součástí základní hlavičky
  - volby (options) a fragmentační informace možno zajistit skrze rozšiřující hlavičky
  - kontrolní součet na L3 zmizel bez náhrady (zajištěn na L2 a L4)

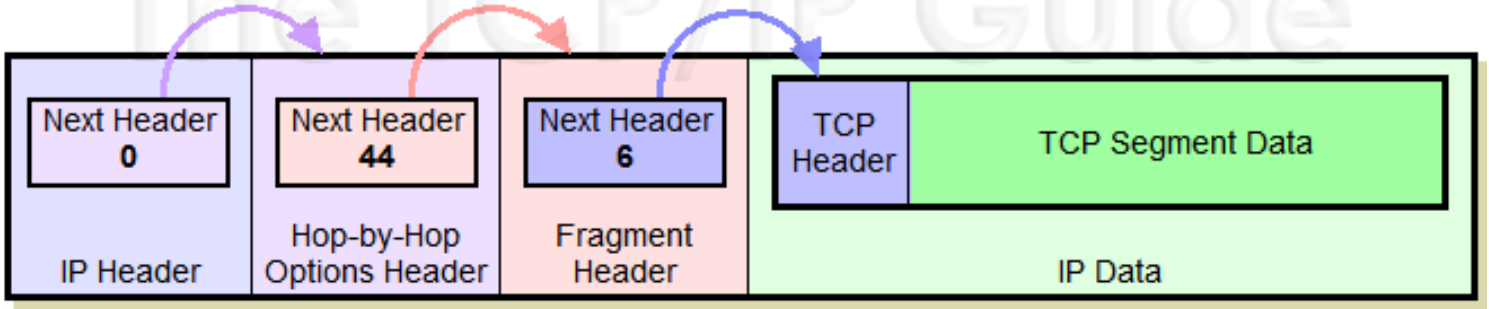
IPv6 datagram – základní hlavička II.

- **Version (VER)** – verze IP protokolu (nyní 6)
- **Priority (PRI)**, také *Traffic Class* – priorita datagramu (zařazení do určité přepravní třídy)
- **Flow label** – identifikace proudu datagramů od jednoho odesílatele ke stejnému cíli se stejnými vlastnostmi
  - původně pro podporu aplikací v reálném čase, aktuálně nevyužito
- **Payload length** – celková délka IPv6 datagramu (bez základní hlavičky)
- **Next header** – hlavička transportního protokolu nebo rozšiřující hlavička
- **Hop limit** – $\approx$ TTL v IPv4
- **Source/Destination address** – IPv6 adresa zdrojového/cílového uzlu

IPv6 datagram – rozšiřující hlavičky



IPv6 Datagram With No Extension Headers Carrying TCP Segment



IPv6 Datagram With Two Extension Headers Carrying TCP Segment

Definováno několik rozšiřujících hlaviček

- např. Hop-By-Hop Options (volby pro všechny), Routing (směrování), Fragment (fragmentace), Encapsulating Security Payload (šifrování obsahu), Authentication Header (autentizace), atd.

IPv6 – podpora zabezpečených přenosů

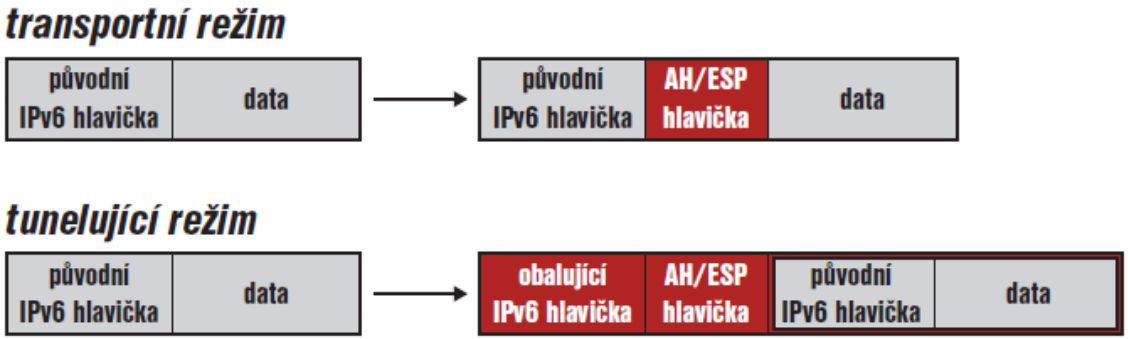
- implementace zabezpečené komunikace na síťové vrstvě
  - označováno jako IPSec
  - původní IPv4 zcela ignoruje (doimplementováno dodatečně)
  - v IPv6 povinná
- poskytované služby:
  - *autentizace dat* – cílem je ověřit, že data odeslal skutečně ten, kdo to o sobě tvrdí. Navíc zaručuje, že obsah datagramu je původní a nebyl během průchodu sítí změněn.
  - *šifrování dat* – umožňuje utajit obsah korespondence (data nesená v zašifrovaných datagramech dokáže rozluštit jen jejich příjemce)
- dvě rozšiřující hlavičky:
  - *AH (Authentication Header)* – autentizace datagramu (ověření pravosti jeho adres a obsahu)
  - *ESP (Encapsulating Security Payload)* – autentizace datagramu + možnost šifrování obsahu

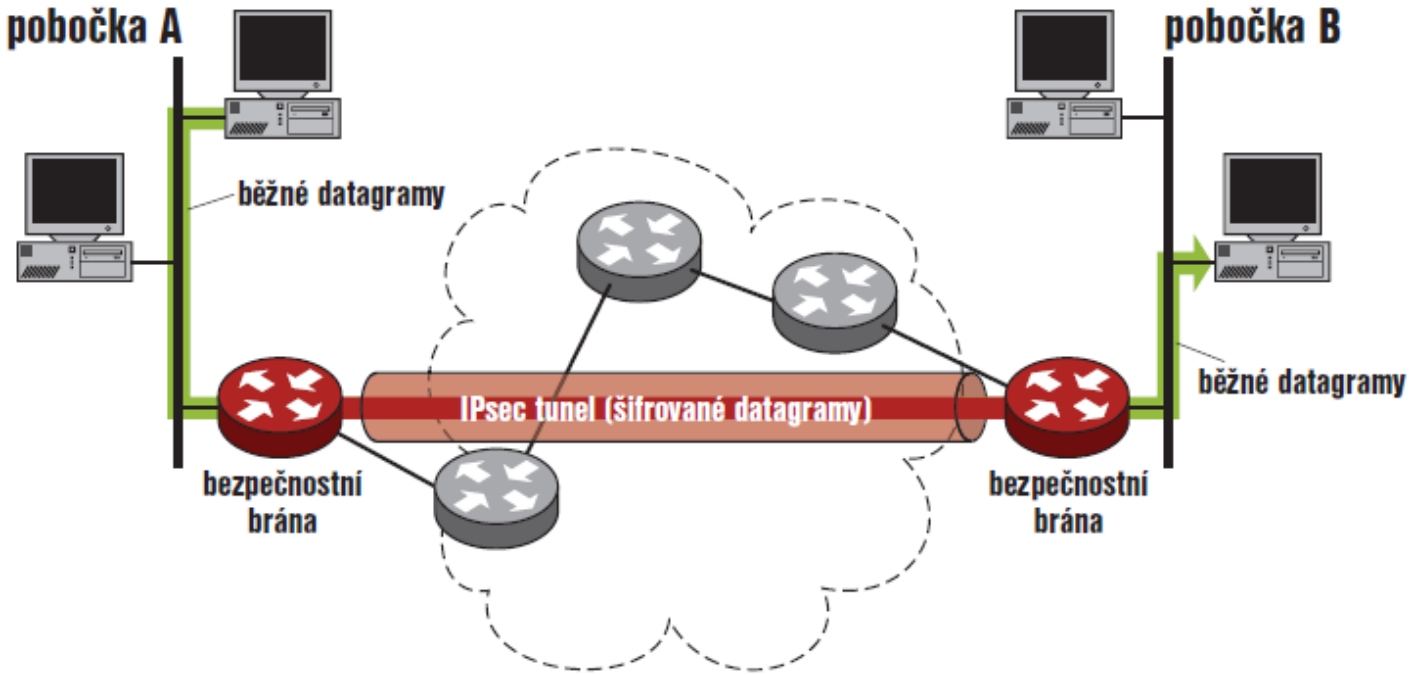
IPv6 – podpora zabezpečených přenosů II.

- *AH (Authentication Header)*
  - ověření totožnosti odesílatele
  - možnost ochrany před opakovaným vysíláním téhož
    - aby vetřelec nemohl jednoduše odeslat ještě jednou sekvenci paketů, které se mu podařilo zachytit
- *ESP (Encapsulating Security Payload)*
  - širší služby než AH
  - umožňuje buď řešit šifrování paketu nebo ověřování totožnosti odesílatele, avšak ne současně

- 2 režimy ochrany:
  - *transportní režim* – bezpečnostní hlavičky se vkládají přímo jako součást datagramu mezi jeho rozšiřující hlavičky
  - *tunelující režim* – celý stávající datagram se zabalí jako data do nového datagramu, který je opatřen novými hlavičkami, včetně bezpečnostních

IPv6 – podpora zabezpečených přenosů III.





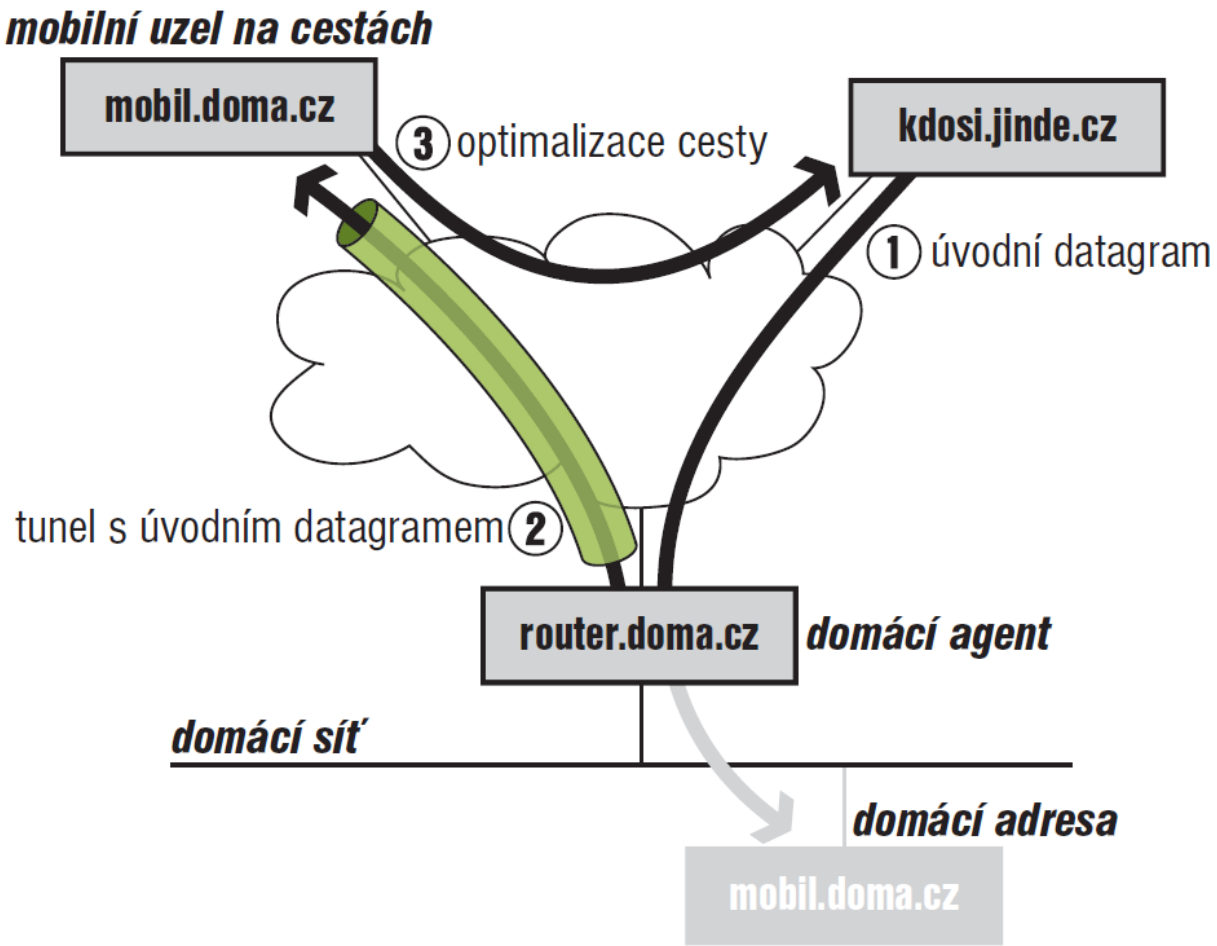
IPv6 – podpora zabezpečených přenosů IV.

- *bezpečnostní asociace (Security Association, SA)*
  - virtuální spojení dvou počítačů, které zajišťuje zabezpečený přenos dat
  - součástí jsou všechny potřebné informace
    - použitý bezpečnostní protokol (AH nebo ESP, nikoli oba) a jeho režim, šifrovací algoritmus a klíče platné pro toto spojení, čítače, doba životnosti, ochranné prvky proti opakování, atp.
  - jsou jednosměrné
- správa bezpečnostních asociací:
  - dříve:
    - *Internet Security Association and Key Management Protocol* (ISAKMP, RFC 2408) pro obecný rámec vzájemné dohody o parametrech bezpečnostních asociací, a
    - *Internet Key Exchange (IKE) verze 1* (RFC 2409) pro výměnu klíčů
  - nyní: *Internet Key Exchange (IKEv2) Protocol* (RFC 4306)
    - kompletní funkce potřebné pro správu bezpečnostních asociací a nastavení jejich parametrů i používaných klíčů

IPv6 – podpora mobility

- *nosná myšlenka:* i pohyblivé zařízení je někde „doma“
  - existuje pro něj tzv. *domovská síť*
- adresy:
  - *domácí adresa (Home Address)* – neměnná adresa, pod níž je stroj trvale dostupný (i když není v domovské síti)
  - *dočasná adresa (Care-of Address)* – měnící se adresa (závislá na síti, kde se aktuálně zařízení nachází)
- *domácí agent (home agent)* – jeden ze směrovačů v domácí síti, jehož prostřednictvím je mobilní zařízení trvale dosažitelné
  - stahuje na sebe datagramy směřující k mobilnímu uzlu a předává mu je tunelem
- optimalizace cesty – seznámení vzdálené strany s aktuální dočasnou adresou mobilního uzlu
  - cílem je zefektivnění komunikace
  - není nezbytná (komunikace může po celou dobu probíhat prostřednictvím domácího agenta)

IPv6 – podpora mobility II.



Obrázek: Ilustrace funkce domácího agenta v IPv6. (Satrapa P., IPv6)

IPv6 – podpora autokonfigurace

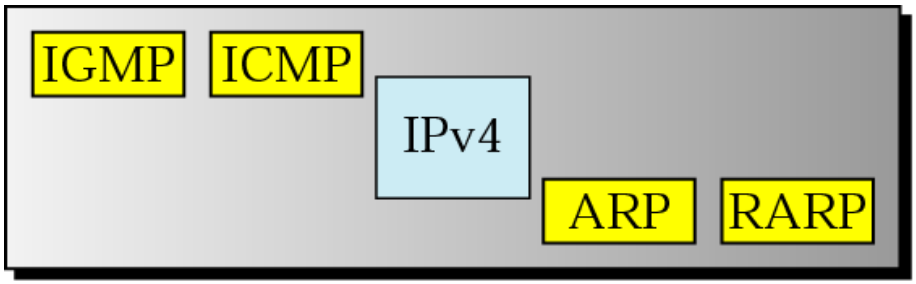
- *stavová konfigurace* – základem server spravující konfigurační parametry, které pak na požádání sděluje klientům
  - mechanismus ala RARP → BOOTP → DHCP
  - navrženo DHCPv6
- *bezstavová konfigurace* – zcela nový způsob konfigurace IPv6 klientů
  - předpokládá se, že v síti sídlí „ctnostní mudrcové“ (směrovače), kteří vědí vše potřebné
    - čas od času sdělí, jaká je situace v sítí – tzv. *ohlášení směrovače (Router Advertisements)*
    - ohlášení informují o všem potřebném (informace o síti – prefix, implicitní směrovač, atp.)
    - nově příchozí klient čeká na ohlášení nebo si ohlášení aktivně vyžádá
    - na základě ohlášení si vypočte vlastní IPv6 adresu (prefix + L2 adresa)
    - nezbytné doplnit mechanismem pro oznamy lokálních DNS serverů (např. skrze DHCPv6)

IPv6 – fragmentace datagramů

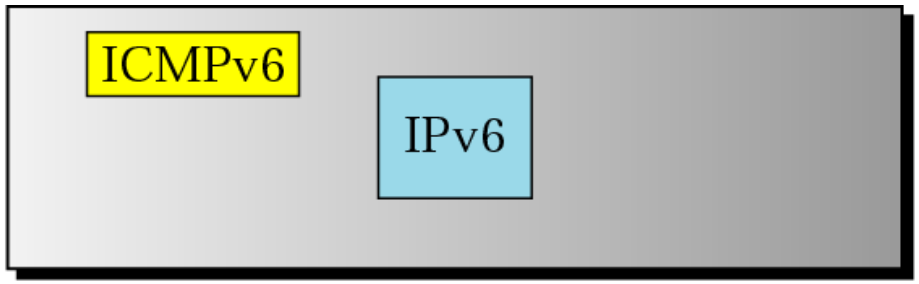
- stejný mechanismus jako v IPv4
- **rozdíl:** vnitřní uzly (směrovače) *nesmí* fragmentovat
  - fragmentovat smí pouze zdrojový uzel
  - cílem je snížení zátěže vnitřních uzlů
- $\Rightarrow$ nutnost zjištění maximální velikosti paketů
  - skrze celou cestu k cíli
  - mechanismus *Path MTU Discovery*
    - = zjištění minimálního MTU využitelného pro přenos dat mezi dvěma uzly
    - provedeno před vlastní komunikací
    - využití *Packet too big* chybových zpráv protokolu ICMP (ICMPv6) – obsahují informaci o vyžadovaném MTU
  - problém s dynamickými cestami
    - při déletrvajících přenosech nutnost pravidelného opakování *Path MTU Discovery*

IPv6 – podpůrné protokoly

- ICMP protokol verze 6 (ICMPv6)
  - založen na stejných principech/mechanismech jako ICMPv4
  - navíc zahrnuje funkcionalitu protokolů ARP a IGMP
    - s využitím *Neighbour Discovery* protokolu operujícím nad ICMPv6



Network layer in version 4



Network layer in version 6

ICMPv6

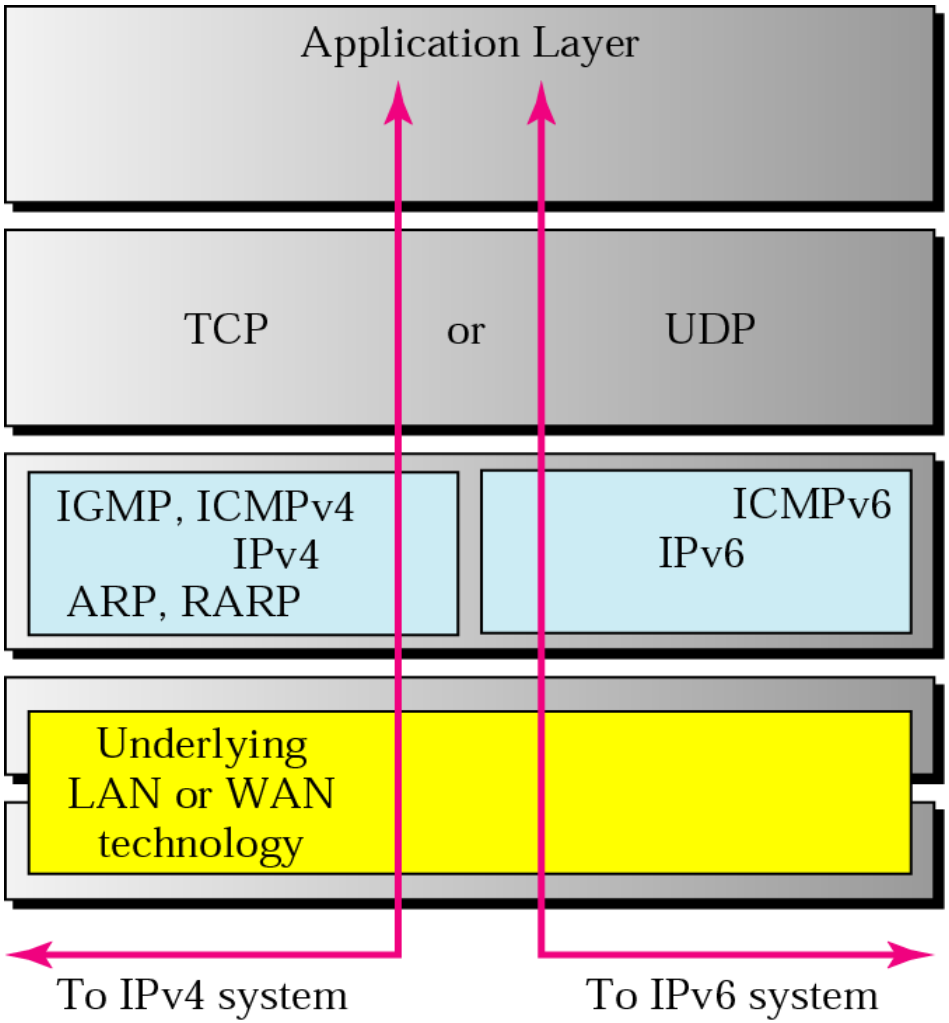
- v IPv6 hlavičce identifikován hodnotou 58 v položce *Next header*
- formát ICMPv6 hlavičky shodný s ICMPv4
  - zprávy identifikovány dvojicí *(typ, kód)*
- zprávy rozděleny do dvou tříd:
  - *chybové* – typ leží v intervalu $(0, 127)$
  - *informační* – typ leží v intervalu $(128, 255)$
- aktuální přehled definovaných typů ICMPv6 zpráv dostupný na adrese `http://www.iana.org/assignments/icmpv6-parameters`

Mechanismy pro podporu přechodu IPv4 → IPv6

- při návrhu IPv6 se počítalo s pozvolným přechodem z IPv4
  - ⇒ nezbytný mechanismus koexistence IPv4 a IPv6
- 3 základní skupiny:
  - *Dvojí zásobník* – příslušné zařízení podporuje jak IPv4, tak IPv6
  - *Tunelování* – IPv6 datagramy zabaleny jako data do IPv4 datagramu, který daná síť dokáže přepravit
  - *Translátory* – zařízení pro překlad IPv6 datagramů do IPv4 datagramů (směr klient → server) a pro překlad odpovědi serveru (naopak, z IPv4 do IPv6)

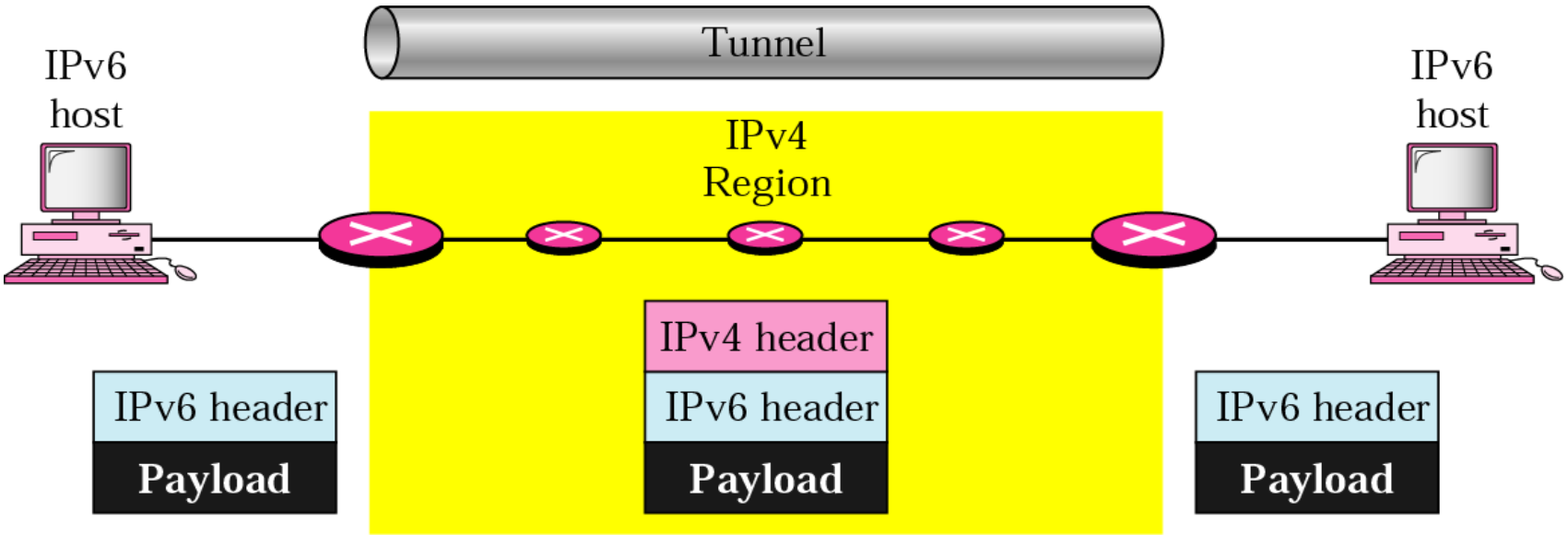
Mechanismy pro podporu přechodu IPv4 → IPv6

Dvojí zásobník



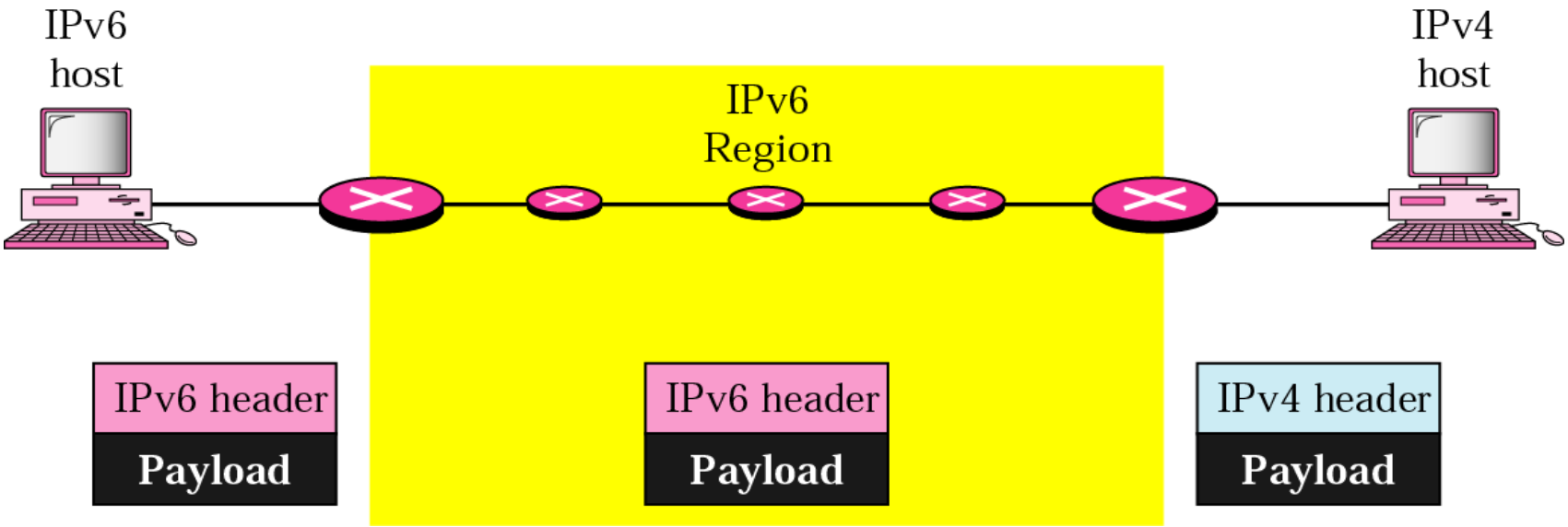
Mechanismy pro podporu přechodu IPv4 → IPv6

Tunelování



Mechanismy pro podporu přechodu IPv4 → IPv6

Translátory



IPv6: Literatura

- příslušná RFC
- Satrapa P.: *IPv6.* Sdružení CZ.NIC, 2008.
  Dostupné online: `http://knihy.nic.cz/files/nic/edice/pavel_satrapa_ipv6_2008.pdf`
- Blanchet M.: *Migrating to IPv6.* John Wiley & Sons, Ltd., 2005.
- `http://www.ipv6.cz`