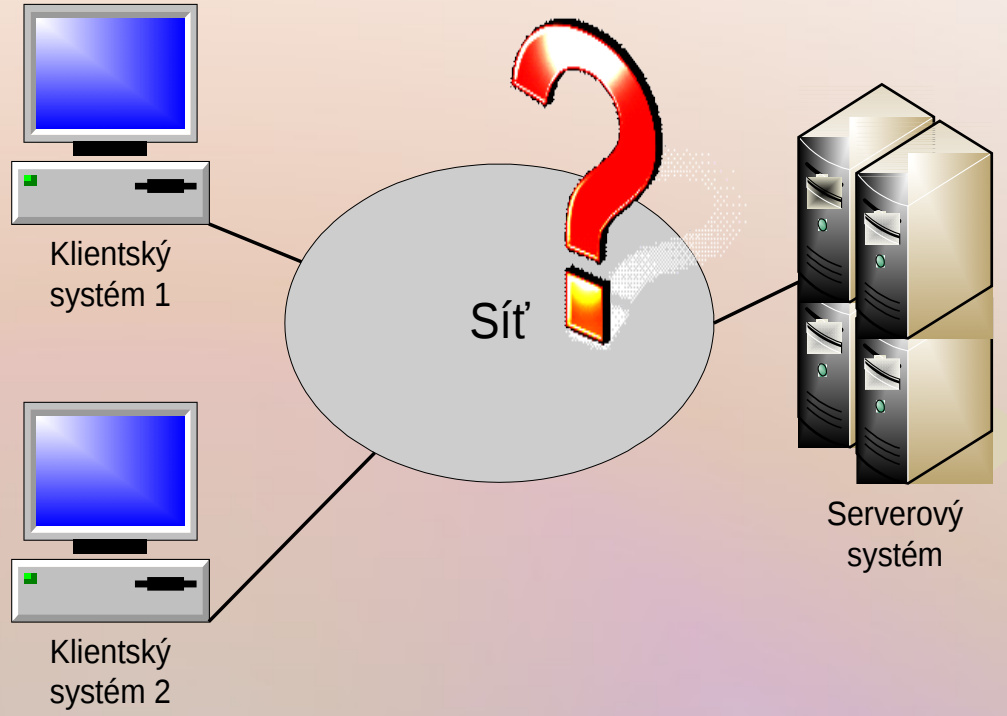
Operační systémy a sítě

Petr Štěpán, K13133
KN-E-229
stepan@labe.felk.cvut.cz

Téma 9. Základy počítačových sítí

Pohled na pojem „počítačová síť“

- Nejběžnější pohled na počítačovou síť
  - klient – server
- Předmětem našeho zájmu bude zejména:
  - Jak vypadají přenosy dat po síti (či sítích)
  - Jak se pozná, kdo co komu posílá
  - Jak se to zabezpečí
  - Jak vypadají procesy, které komunikují po síti
  - Co pro takové procesy poskytuje operační systém



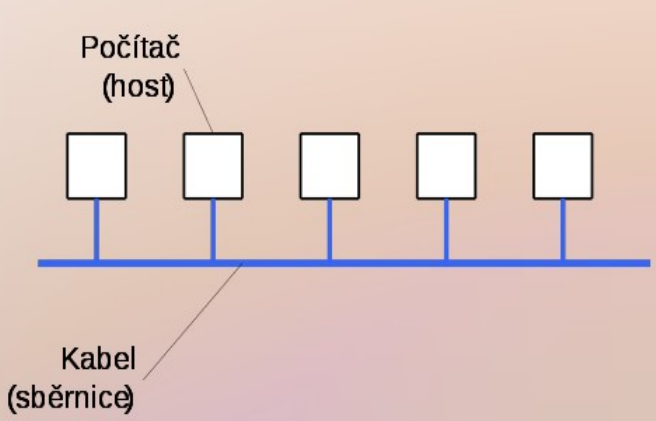
ISO-OSI síťový model

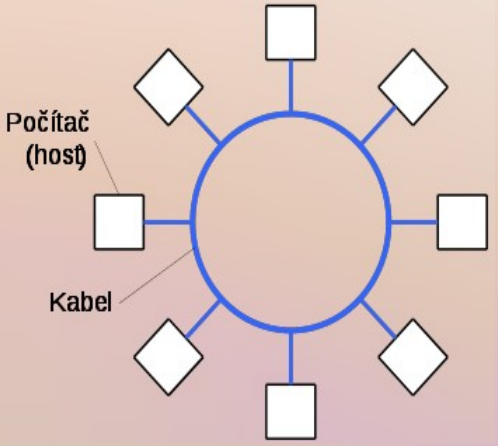
- Vzhledem ke komplexnosti přenosu dat po síti vždy vícevrstvá struktura
- OSI = *Open System Interconnect*
- Model o 7 vrstvách:

Datový element	Vrstva	Účel
Datový tok	Aplikační (*application*)	Koncové aplikace a s nimi spojené komunikační a formátové protokoly včetně síťového API (např. FTP, HTTP, DNS, TELNET, ...)
	Prezentační (*presentation*)	Transformace dat do tvaru, který používají aplikace
	Relační (*session*)	Organizace a synchronizace dialogu mezi spolupracujícími systémy a řízení výměny dat
Segment	Transportní (*transport*)	Pravidla pro přenos dat mezi dvěma počítači (koncovými body) včetně zabezpečení kvality přenosu (nesmí se nic ztratit, jindy jsou ale ztráty přípustné)
Paket	Síťová (*network*)	Tvorba a přenos logických jednotek (paketů), logické adresování, určování přenosových cest
Rámec	Spojová (*link*)	Tvorba fyzických přenosových jednotek (rámců), fyzická (hardwarová) adresace, LAN
Bit	Fyzická (*physical*)	Popis fyzického média (kabelů apod.), signálových úrovní, konektorů a dalších technických parametrů, reprezentace logických signálů

Základní struktury LAN

- Lokální sítě „s vysíláním“ (*broadcast networks*)
  - Sběrnicová a prstencová struktura
  - Každý počítač (uzel) je schopen oslovit všechny ostatní uzly v LAN

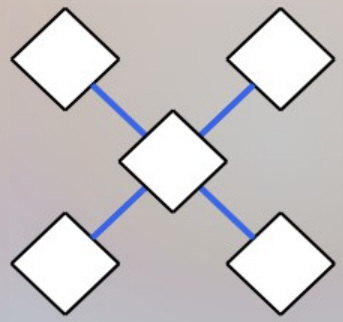




- Další možnosti užívané zejména pro propojování LAN
  - dvoubodové spoje (point-to-point)

Např. ADSL připojení doma

  - hvězdicová struktura (point-to-multi-point)



Např. připojení strojů k „přístupovému bodu“ WiFi. WiFi však simuluje sběrnicovou strukturu (umí „broadcast“)

Základní technologie LAN

- Technologie Token Ring (dnes již téměř historická technologie IBM)
  - TokenRing 4 Mbit/s TokenRing 16 Mbit/s
  - Prstencová topologie, předávání „tokenu“, 8-mi bitové adresy
  - Formát rámce TokenRing 4 Mbit

Začátek zprávy	Adresa odesilatele	Adresa příjemce	Typ rámce	Datový rámec	Konec zprávy	Parita	Odmítnutí
10 bitů	8 bitů	8 bitů	24 bitů	0-16352 bitů	9 bitů	1 bit	1 bit

- Technologie ethernet (nečastější LAN)
  - Časový multiplex CSMA/CD (= *Carrier Sense Multiple Access with Collision Detection*)
    - Každý uzel začne vysílat, kdykoliv potřebuje a poslouchá, zda slyší to, co říká. Pokud ne, došlo ke kolizi. Oba pak „zmlknou“ a za náhodnou dobu to zkusí znovu.
  - Adresování v ethernetu: jedinečné 48-bitové hardware adresy
  - Formát ethernetového rámce

Preambule	Adresa příjemce	Adresa odesilatele	Typ rámce	Datový rámec	Kontrolní součet
64 bitů	48 bitů	48 bitů	16 bitů	368-12000 bitů	32 bitů

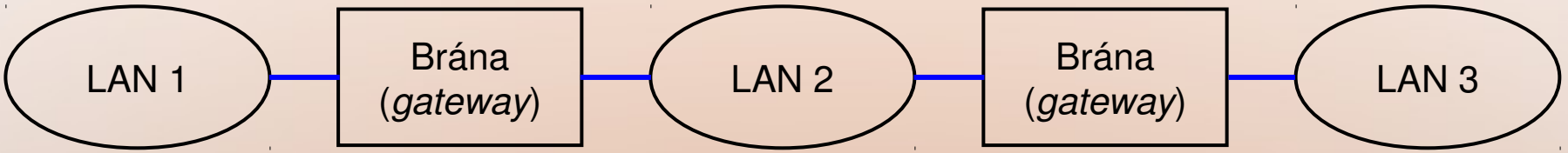
- Obě technologie podporují tzv. "broadcast"
  - tj. oslovení všech zařízení v lokální síti – NÁKLADNÉ

ARP protokol

- Pro doručení Ethernetového datagramu je nutné znát MAC adresu prvního počítače na cestě
- Je nutný převod IP na MAC adresu
- K tomu je protokol ARP, který slouží k přenosu informace o MAC adrese
- Standardní průběh komunikace:
  - Počítač vysílá broadcast s ARP dotazem Who has IP? Dotaz obsahuje MAC a IP adresu vysílajícího počítače.
  - Přijímající počítač si uloží získané informace MAC a IP od vysílajícího počítače pro případné další použití do ARP tabulky MAC ↔ IP.
  - Počítač s hledanou IP adresou odpoví ARP datagramem obsahující jeho MAC adresu.
  - Vysílající počítač si získanou MAC adresu uloží do ARP tabulky MAC ↔ IP.
  - Počítač může vyslat data na cestu internetem.

Architektura Internetu

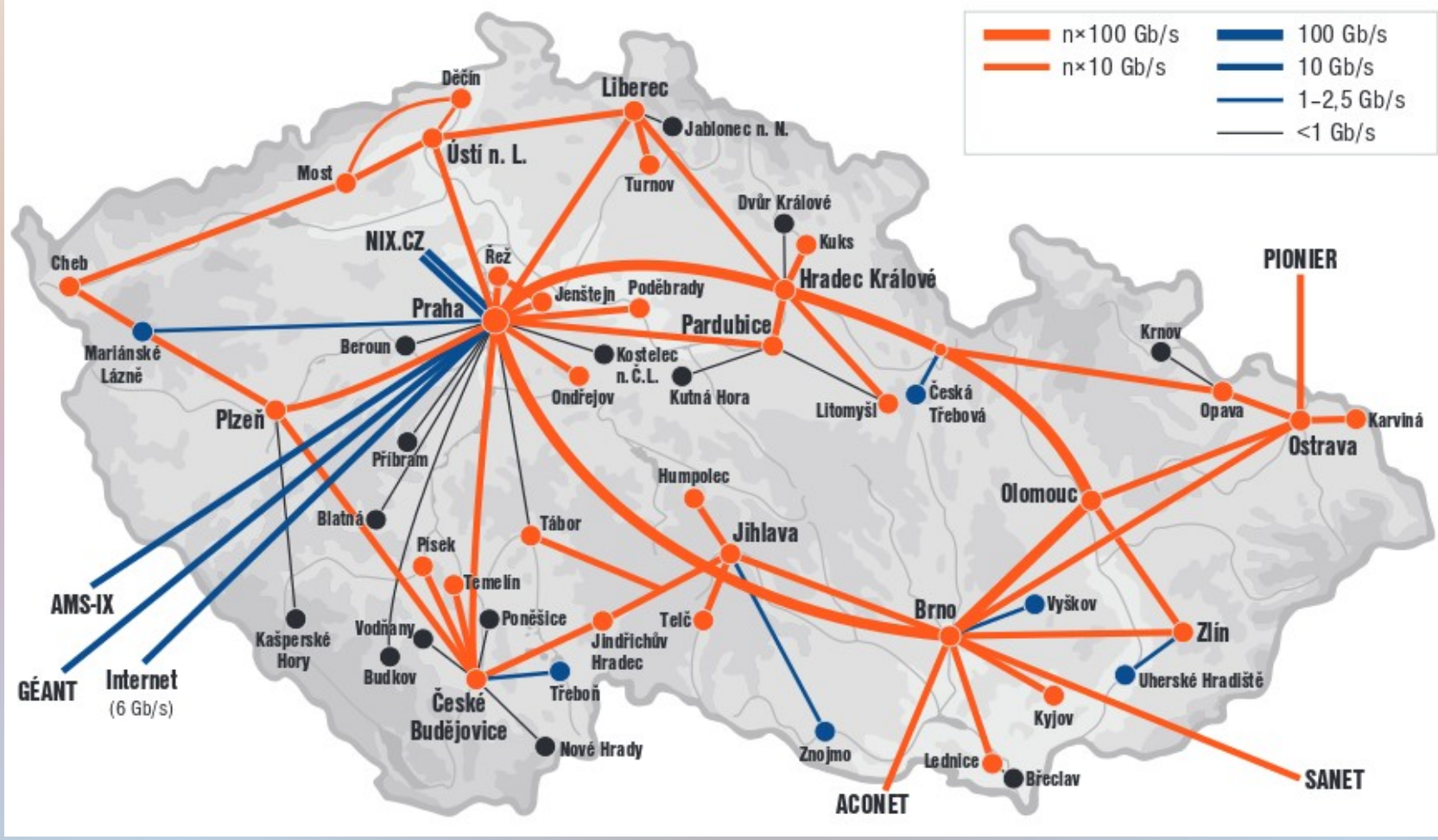
- Základní architektura Internetu (i internetů)
  - internet(working) s malým „i“ = obecné propojení několika LAN



  - Mosty (*bridges*), brány (*gateways*) a směrovače (*routers*) propojují fyzické lokální sítě
  - Mosty propojují segmenty LAN stejných fyzických technologií
    - Mnohdy oddělují provoz na segmentech adaptivním přeposíláním rámců na základě "naučených" fyzických adres
  - Brány propojují LAN s různými technologiemi
    - Velmi často jsou fyzicky integrovány se směrovači
  - Směrovače pracují s datovými jednotkami "vyšší úrovně"
    - Mosty i brány pracují na úrovni spojové (linkové) vrstvy ISO-OSI
    - Směrovače znají informace o sítích a posílají pakety (datagramy ➔) na základě "vyšších" (logických) adres.
    - Např. v IP staví na znalosti o cílové síti (nikoliv o cílovém stroji)
- IP protokoly považují všechny sítě za rovnocenné bez ohledu na jejich fyzickou technologii

CESNET

- CESNET znamená Czech Education and Scientific NETwork
- Cílem je výzkum a vývoj informačních a komunikačních technologií, budování a rozvoj e-infrastruktury CESNET určené pro výzkum a vzdělávání

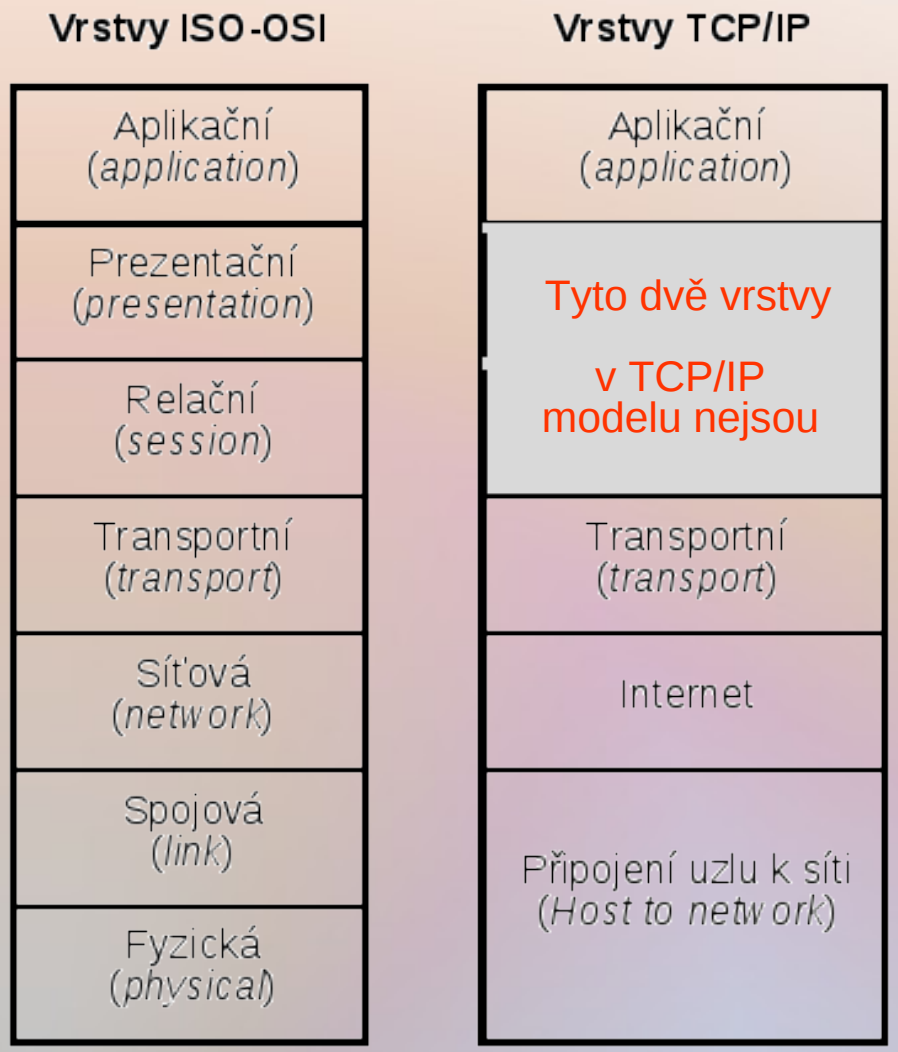


Internet a jeho charakteristiky

- TCP/IP Internet
  - Protokoly = formáty a pravidla pro zasílání zpráv po síti.
  - Protokoly zakrývají detaily komunikace.
- Služby Internetu
  - Aplikační služby (tzv. aplikační protokoly)
    - Elektronická pošta (SMTP), přenos souborů (TFTP, FTP), vzdálené terminály (telnet, ssh), informační služby (např. HTTP) a mnoho dalších.
- Služby transportní vrstvy
  - Služeb je celá řada, avšak z uživatelského pohledu jsou podstatné zejména transportní protokoly:
    - Služba „bezespojového“ zasílání paketů.
      Protokol **UDP** (= *User Datagram Protocol*)
    - Služba spolehlivého spojení.
      Protokol **TCP** (= *Transmission Control Protocol*)
- Charakteristiky Internetového TCP/IP
    - Nezávislost na technologii lokálních sítí a způsobu jejich propojování, potvrzování mezi koncovými účastníky spojení (na úrovni transportní vrstvy pro TCP nebo aplikační při UDP)
    - Standardizované aplikační protokoly nezávislé na hardwarových a softwarových platformách

Modely ISO-OSI a TCP/IP

- OSI model
  - používá
    - služby
    - rozhraní
    - protokoly
  - problémy s
    - časováním
    - technologiemi různých sítí
    - implementací a strategiemi
- TCP/IP model je jednodušší, ale hrubší
  - zejména
    - nerozlišuje služby, rozhraní a protokoly
    - neodděluje spojovou a fyzickou vrstvu
    - hlavní a pomocné protokoly se jsou chápány jako stejně důležité
- Přesto se podržíme TCP/IP
    - aplikačně nejdůležitější



Internet a jeho řízení

- Historické poznámky
  - ARPA/DARPA – projekt z počátku sedmdesátých let 20. stol.
  - BSD UNIX a systém symbolického adresování strojů prostřednictvím tzv. domén (*Domain Name System* = DNS) – 1984
  - Profesionalizace Internetu (devadesátá léta 20. stol.)
- Řízení Internetu
  - **IAB** = *Internet Architecture Board* – celková architektura
  - **IETF** = *Internet Engineering Task Force* – technologie, protokoly
  - **IANA** = *Internet Assigned Numbers Authority* – přidělování adres, čísel portů, atd.
  - ISOC = *Internet Society* – sdružení profesionálních firem
  - **IESG** = *The Internet Engineering Steering Group* – technická standardizace
- Dokumentace
  - RFC (*Request for Comment*) šířené volně po síti
    - např. http://www.ietf.org/rfc.html

Internetové adresy

- Základní adresování v Internetu
  - Každý stroj má svoji jednoznačnou identifikaci: tzv. IP adresu
- Současný Internet – v. 4 používá adresy 32 bitů
  - Konvence: 4 dekadická čísla à 8 bitů – 147.32.85.1
- Internet v. 6 má adresy 128 bitů
  - Proč rovnou v 6 – verze 5 byla v roce 1979 použita k definici Internet Stream Protocol, který se moc neprosadil
  - Dataily později
- IP adresa
  - Identifikuje každý jednotlivý síťový adapter
    - Stroj může mít i více adapterů („*multihomed*“ *host*)
    - Identifikace nemusí být jednoznačná: jeden adapter může mít více IP adres
  - Skládá se ze dvou částí
    - Identifikace (adresa) sítě – *netid* (bity vlevo)
    - Identifikace (adresa) stroje v síti – *hostid* (bity vpravo)

Internetové adresy

- Primární třídy IP adres

Třída	Bitový prefix	Počet bitů *čísla sítě*	Počet bitů *čísla stroje*	Počet sítí	Počet adres v síti	Rozsah adres
A	0	8	24	128 = 2^7	16.777.216 = 2^{24}	0.0.0.0 – 127.255.255.255
B	10	16	16	16 384 = 2^{14}	65 536 = 2^{16}	128.0.0.0 – 191.255.255.255
C	110	24	8	2 097 152 = 2^{21}	256 = 2^8	192.0.0.0 – 223.255.255.255
D	1110	nedefinováno		„Multicast" adresy (➔)		224.0.0.0 – 239.255.255.255
E	1111	nedefinováno		Experimentální rozsah		240.0.0.0 – 255.255.255.255

- Jiný pohled

Internetové adresy

- Konvence:
  - Adresa sítě je plná IP adresa s *hostid* = 0
  - Adresa tvořená číslem sítě a částí *hostid* tvořenou samými "1" je adresa oslovující všechny stroje v síti (*broadcast address*)
- Maska sítě:
  - „Adresová aritmetika“
    - Nutno znát binární reprezentace dekadických čísel a operace s binárními čísly!

$$IP\_Adresa \wedge Maska\_Sítě = netid$$

$$IP\_Adresa \wedge \neg(Maska\_Sítě) = hostid$$

- Adresování CIDR (= *Classless Inter-Domain Routing*)
  - Adresová aritmetika umožňuje efektivnější členění párů *netid* | *hostid* – hranice částí IP adresy může být kdekoliv
  - Maska sítě dána *n* (*n*=0 až 32) jedničkovými bity zleva
  - CIDR notace:
    - IP_Adresa/*n*; příklad: 147.32.85.128 – 147.32.85.191 = 147.32.85.128/26
      ale též = 147.32.85.183/26
    - LAN 192.168.200.64/30 obsahuje 4 adresy: 192.168.200.64 = netid, 192.168.200.65=stroj$_1$, 192.168.200.66=stroj$_2$, 192.168.200.67 = LAN broadcast

Internetové adresy

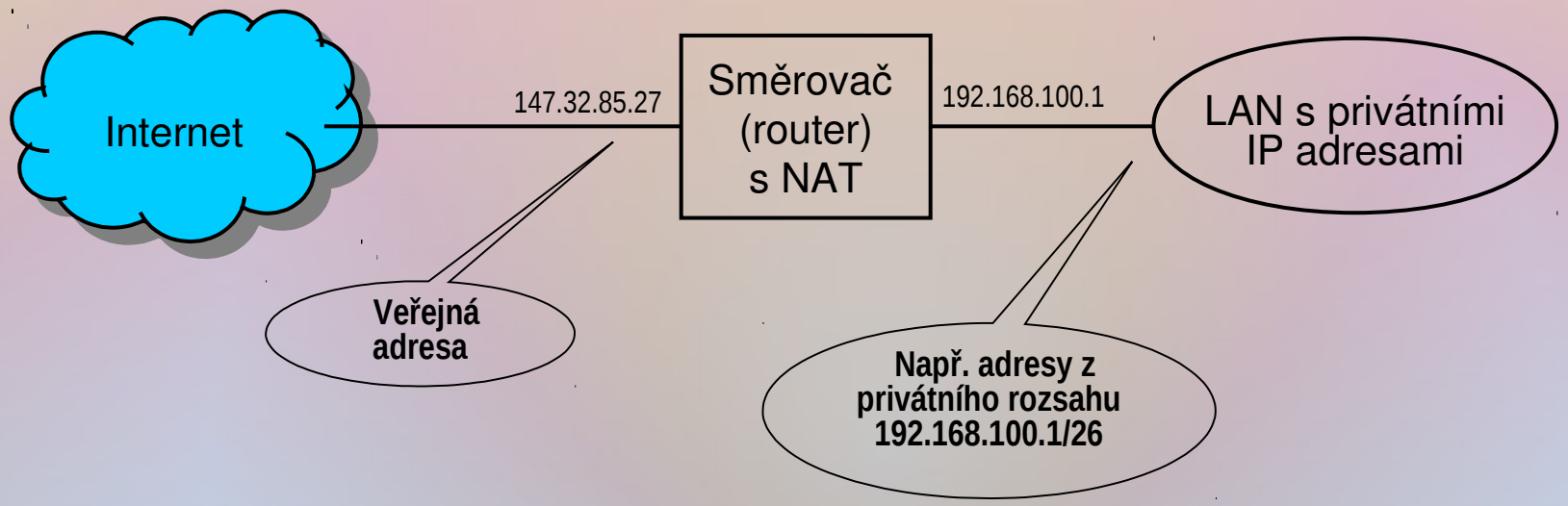
- Rezervované rozsahy IPv4 adres

CIDR notace	Rozsah adres	Počet adres	Účel
0.0.0.0/8	0.0.0.0 – 0.255.255.255	16.777.216	„Broadcast" v rámci dané (this) sítě (RFC 1700)
10.0.0.0/8	10.0.0.0 – 10.255.255.255	16.777.216	Privátní rozsah adres (RFC 1918)
127.0.0.0/8	127.0.0.0 – 127.255.255.255	16.777.216	„Loopback" adresy, stroj oslovuje „sám sebe" (obvykle se používá jen **127.0.0.1**)
169.254.0.0/16	169.254.0.0 – 169.254.255.255	65.536	Autokonfigurační rozsah, kdy stroj potřebuje zjistit svoji adresu, obvykle pomocí **DHCP** (➔) protokolu
172.16.0.0/12	172.16.0.0–172.31.255.255	1.048.576	Privátní rozsah adres (RFC 1918)
192.88.99.0/24	192.88.99.0–192.88.99.255	256	Pro mechanismus přechodné migrace mezi IPv4 a IPv6 (RFC 3068)
192.168.0.0/16	192.168.0.0–192.168.255.255	65.536	Privátní rozsah adres (RFC 1918)
198.18.0.0/15	198.18.0.0–198.19.255.255	131.072	Pro testování „inter-network" komunikací (RFC 2544)
224.0.0.0/4	224.0.0.0–239.255.255.255	268.435.456	Viz třída D – „Multicast", tj. komunikace 1:N (RFC 3171)
240.0.0.0/4	240.0.0.0–255.255.255.255	268.435.456	Viz třída E – Rezervováno pro experimentální vývoj protokolů dle zvláštního povolení IANA

- Speciální adresy
  - Privátní adresy
    - nesmí se šířit po Internetu – směrovače nesmí propustit datagramy s těmito adresami
  - "Multicast" adresy
    - jeden stroj rozesílá informace více zaregistrovaným strojům (např. internetová televize)

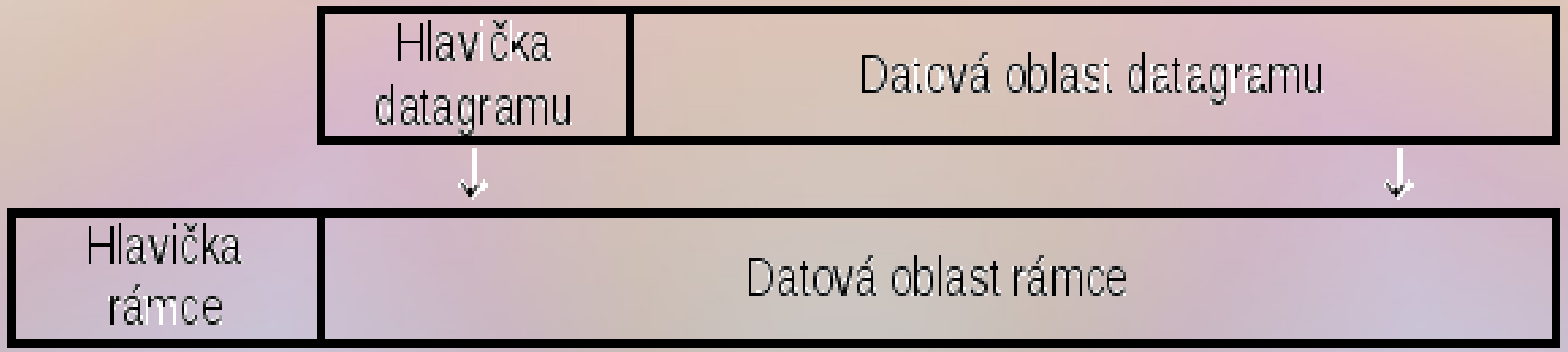
Internetové adresy (5)

- Šetření IP adresami
  - Užívání privátních adres a jejich překlad na adresy veřejné (NAT = *Network Address Translation*)
    - Množina privátních adres je překládáno na jedinou veřejnou adresu
    - Na privátním rozsahu (za NAT směrovačem) je problém se servery
    - Způsob práce NAT souvisí s IP protokoly, zejména pak s tzv. porty
    - Vrátíme se k tomuto problému a jeho řešení později



Internetové datagramy

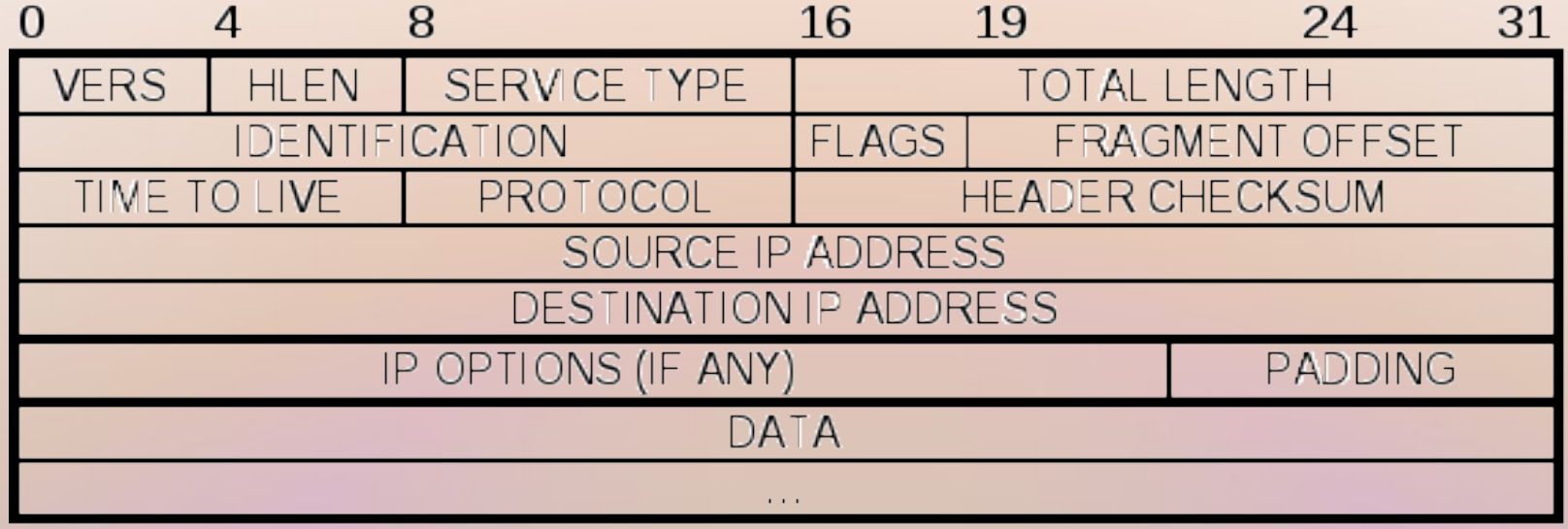
- Internet vytváří virtuální síť a přenáší tzv. IP datagramy
  - Síť představuje systém „s nejlepší snahou o doručování“ (*best effort delivery*)
  - Datagramy putují po různých fyzických sítích majících různou strukturu a velikost rámců
- Formát IP datagramu



Zapouzdření IP datagramu do rámce fyzické sítě

Hlavička IP datagramu

- Každý IP datagram má hlavičku nesoucí informace důležité pro přenos datagramu od odesilatele k adresátovi



Formát IP datagramu

Význam položek

- VERS: Verze IP protokolu – pro IP v. 4 VERS = 4
- HLEN: Délka hlavičky ve 32-bitových slovech (standardně 5).
- TOTAL LENGTH: Celková délka datagramu v bytech (oktetech) včetně hlavičky – max. 65535 bytů.
- SOURCE IP ADDRESS: IP adresa odesilatele
- DESTINATION IP ADDRESS: IP adresa adresáta

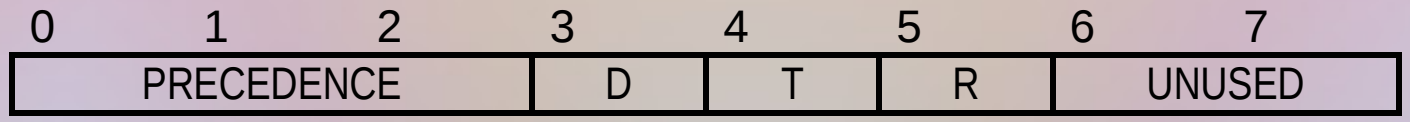
Hlavička IP datagramu (pokračování)

- IDENTIFICATION: obvykle sekvenční nebo náhodné číslo vygenerované odesilatelem datagramu.
- PROTOCOL: Identifikace protokolu IP datagramu (ICMP=1, UDP=17, TCP=6, ...). Definováno v RFC 1060

FLAGS, FRAGMENT OFFSET: Informace o fragmentaci datagramu

TIME TO LIVE (TTL): Určuje jak dlouho smí datagram putovat po Internetu. Každá brána dekrementuje tuto hodnotu; je-li TTL=0 odstraní datagram a pošle ICMP zprávu odesilateli

SERVICE TYPE: Osmibitové pole obsahující pokyny pro směrování paketu



Precedence datagramu:
0=normální,
7=řízení sítě

Malé zpoždění

Vysoká propustnost

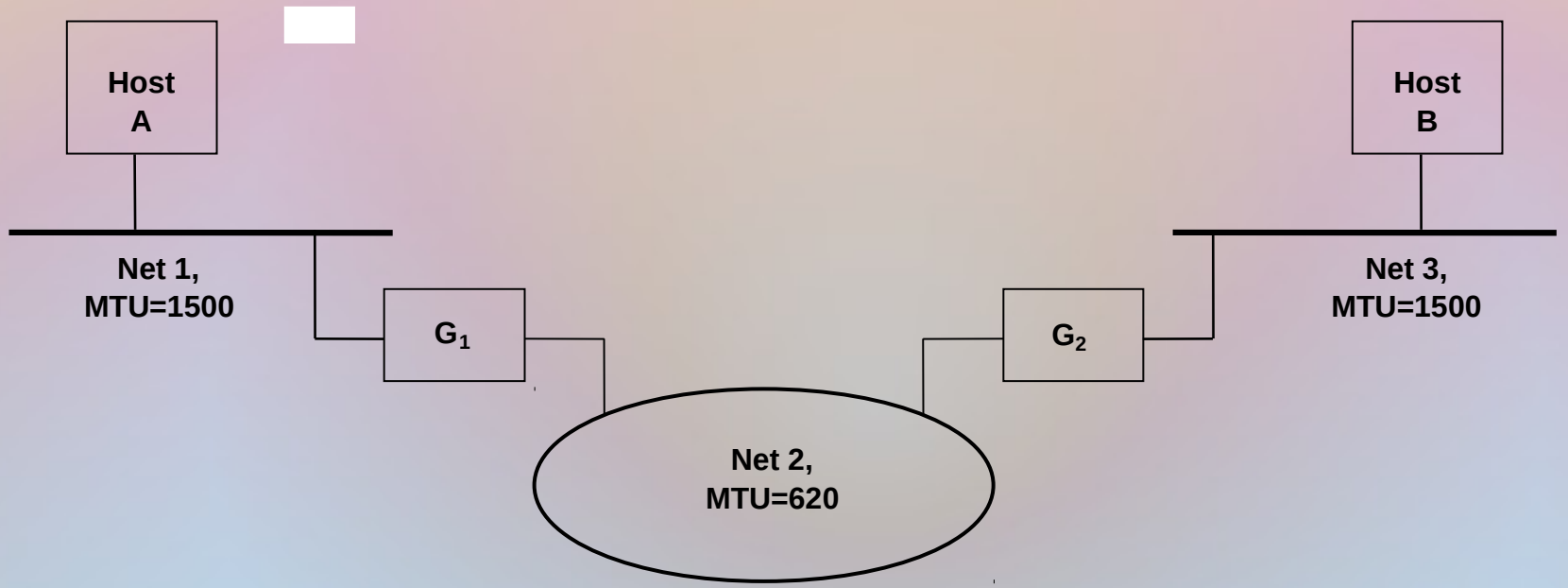
Vysoká spolehlivost

Fragmentace datagramů

- MTU: (*Maximum Transmission Unit*) určuje maximální velikost datagramu, kterou lze přenést po LAN s určitou technologií

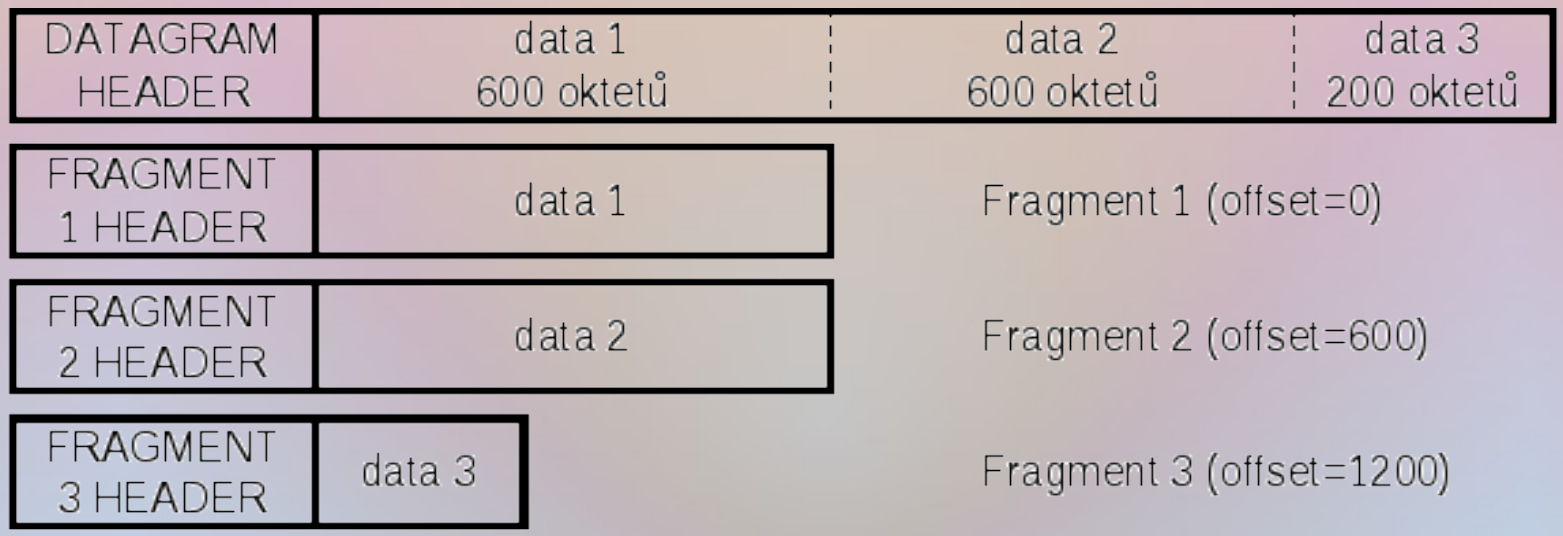
Síť	Implicitní MTU	Síť	Implicitní MTU
PPP	296	X.25	576
Ethernet	1 500	WiFi (IEEE 802.3)	1 492
TokenRing 4Mb	4 464	TokenRing 16Mb	17 914

- Internet – soustava LAN s různými MTU
  - Pokud je datagram větší než MTU, musí se fragmentovat



Fragmentace datagramů

- Fragmentace nastává kdekoliv po cestě datagramu
  - Je-li datagram fragmentován, neskládá se cestou, ale rekonstrukce datagramu je úkolem cílového stroje
  - Každý fragment putuje jako samostatný datagram:
    - Z hlavičky původního datagramu se okopírují pole: VERS, HLEN, SERVICE TYPE, IDENTIFICATION, PROTOCOL, SOURCE IP ADDRESS, DESTINATION IP ADDRESS
    - TOTAL LENGTH se změní na délku fragmentu a položka FRAGMENT OFFSET určuje polohu (offset) fragmentu v původním datagramu
    - Pole FLAGS obsahuje bit: "*more fragments*". Je-li tento bit 0, pak cílový stroj ví, že obdržel poslední fragment, a pomocí polí FRAGMENT OFFSET a TOTAL LENGTH může sestavit originální datagram

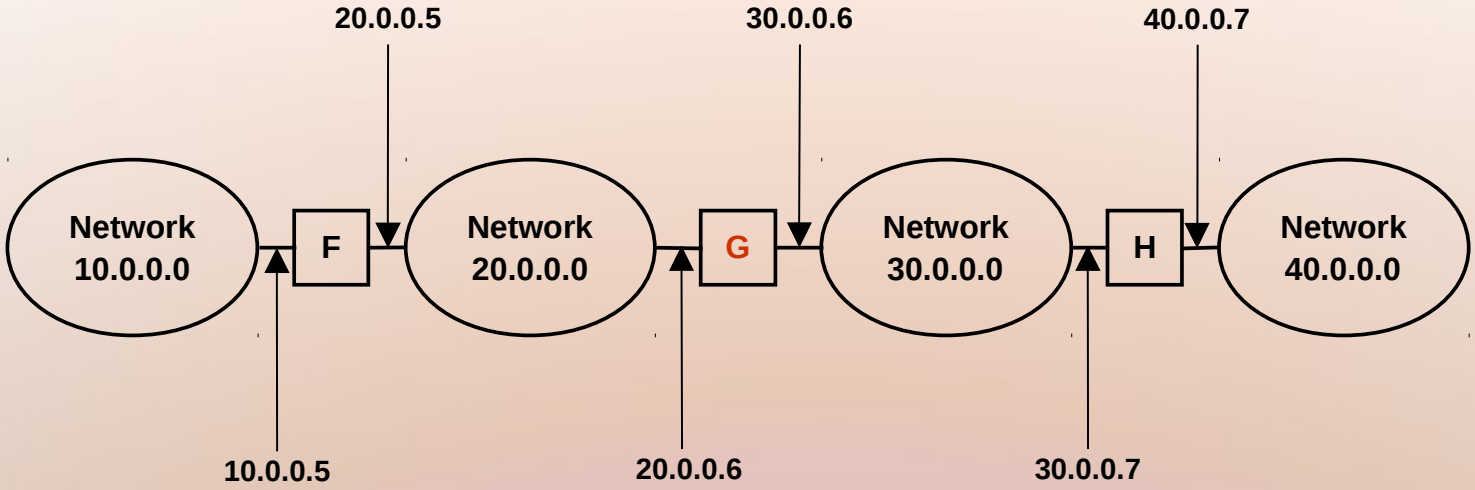


Fragmentace datagramu délky 1400 oktetů při průchodu sítí s MTU = 620

Směrování datagramů

- Směrování (*routing*) je proces rozhodování o cestě, kudy poslat datagram (nebo jeho fragment) k cíli
  - Za směrovač se považuje libovolný stroj schopný přijímat takové rozhodnutí
  - Směrování může být přímé nebo nepřímé
    - Přímé směrování nastává, když je cílový stroj součástí lokální sítě bezprostředně spojené se směrovačem
    - Jinak jde o směrování nepřímé
  - Směrovače v Internetu tvoří kooperativní propojenou strukturu. Datagramy putují od jednoho směrovače k druhému dokud nedosáhnou směrovače, který umí zaslat datagram přímo cílovému stroji
  - Tabulkou řízené směrování
    - Každý směrovač obsahuje tzv. směrovací tabulku tvořenou dvojicemi (*N*, *G*), kde *N* je *netid* cílové sítě a *G* je IP adresa "příštího" směrovače podél cesty k cílové síti *N*. „Příští směrovač“ musí být dosažitelný přímo.

Směrování datagramů



Při zasílání stroji na síti	Směruj na adresu
20.0.0.0	Adresuj přímo cílový stroj
30.0.0.0	Adresuj přímo cílový stroj
10.0.0.0	20.0.0.5
40.0.0.0	30.0.0.7

Tabulka směrovače G

- Implicitní směry (*default routes*)
  - Velmi často jsou LAN propojeny se "zbytkem Internetu" prostřednictvím jediného směrovače. Pak tento směrovač představuje pro tzv. default gateway, tj. adresu, kam všechny stroje v LAN posílají datagramy adresované vně LAN

Směrování datagramů

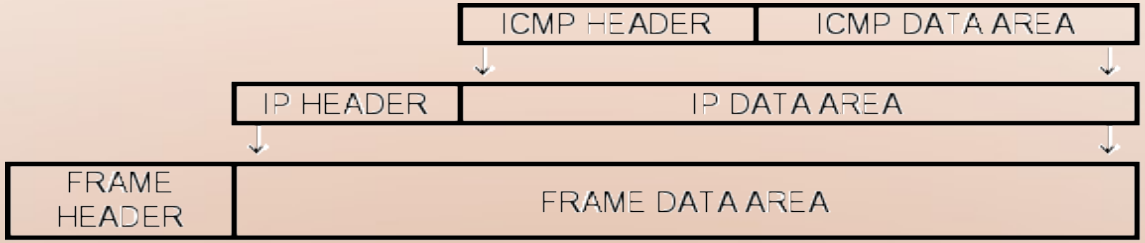
- Specializované směry ke strojům (*Host-Specific Routes*)
  - Někdy je výhodné přiřadit jednomu nebo několika strojům speciální směrovací informaci. Důvody mohou být bezpečnostní, administrativní i technické. Technickým důvodem je např. připojení samostatného stroje po point-to-point spoji (Internetový PPP protokol)
- Směrovací algoritmus:
  1. Vyjmi z datagramu cílovou IP adresu *ID* a s použitím síťové masky urči *netid* cílové sítě
  2. Pokud *ID* odpovídá některému spec. směru (*host-specific route*), pak pošli datagram přímo tomuto stroji
  3. Pokud *netid* se shoduje s některou přímo připojenou sítí, směruj přímo
  4. Pokud *netid* se nachází ve směrovací tabulce, pošli datagram odpovídajícímu směrovači
  5. Pokud bylo specifikováno implicitní směrování (*default route*), pošli datagram na "*default gateway*"
  6. Jinak oznam chybu směrování zasláním ICMP zprávy odesilateli (*Destination unreachable*)

Lokální doručení datagramu

- Přímé směrování musí doručit datagram lokálně
  - Totéž se děje při předání datagramu přímo dostupnému směrovači připojenému přes LAN (nikoliv při *point-to-point* spoji)
  - Datagram obsahuje IP adresu, avšak doručit je nutno na fyzickou adresu uvnitř LAN
- Mapování IP adres na fyzické adresy
  - ARP (= *Address Resolution Protocol*) – dynamické mapování
  - Řešení v "broadcast" LAN – zaslání datagramu strojem A s IP adresou I_A stroji B, který má IP adresu I_B
    - Odesilatel zná svoji IP adresou I_A a i fyzickou adresou F_A, a potřebuje zjistit fyzickou adresu F_B k jemu známé IP adrese I_B
    - Vyšle „ARP broadcast“ rámec, v jehož datové části bude vedle I_A i I_B. Tento rámec přijmou všechny stroje v LAN.
    - Stroj, který rozpozná svoji adresu I_B, na tuto „všeobecnou výzvu“ odpoví a sdělí tak odesilateli svoji fyzickou adresu F_B.
    - "Broadcast" však zatěžuje LAN, proto si tazatel získanou F_B jistou dobu (standardně 5 minut) pamatuje.
    - Vzhledem k tomu, že se dá očekávat brzká odpověď $B \rightarrow A$, stroj B získá a zapamatuje si z ARP rámce i adresy I_A a F_A.

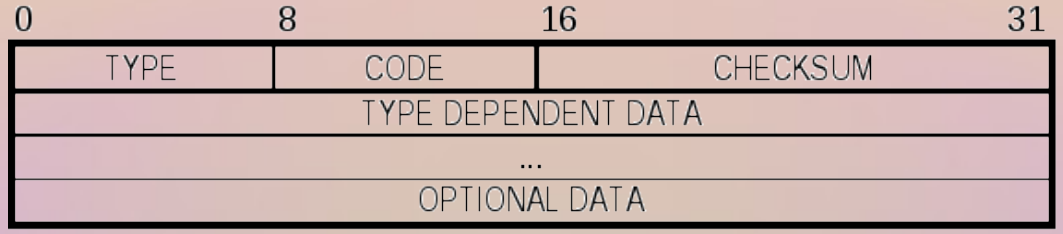
Protokol ICMP

- ICMP (= *Internet Control Message Protocol*)
  - Nejjednodušší protokol pro řízení sítě a předávání chybových hlášení



Zapouzdření ICMP v IP datagramu na fyzické síti

  - Hlavička ICMP datagramu nemá (kromě prvních 4 bytů) pevnou strukturu



  - Pole TYPE udává účel ICMP zprávy a určuje i formát a význam dalších polí – některé typy ICMP datagramů:
    - Standardizovaných typů je mnohem více (cca 40)

TYPE	Účel	TYPE	Účel
0	Echo reply	9	Router advertisement
3	Destination unreachable	10	Router discovery
5	Redirect (route change)	11	Datagram TTL exceeded
8	Echo request	12	Datagram parameter problem

Protokol ICMP – základní užití

- Operátorské použití
  - "Utilita" ping k testování dostupnosti cílového stroje je postavena na ICMP
    - "Náš" systém vyšle ICMP "Echo request" s cílovou adresou testovaného stroje. Navíc ping umí nastavit velikost zasílaného paketu a další příznaky v záhlaví datagramu (např. "don't fragment").
    - Dorazí-li ICMP datagram k cílovému stroji, ten odpoví pomocí ICMP "Echo reply", a když tento paket dorazí "k nám", víme, že cesta je OK.
  - "Utilita" traceroute (ve Windows tracert) dovolí trasovat cestu od "našeho" stroje k cíli
    - Využívá fakt, že každý směrovač po cestě datagramu dekrementuje pole TTL, a klesne-li hodnota tohoto pole na nulu, informuje zdrojový systém ICMP zprávou "Datagram TTL exceeded" (typ 11).
    - Posíláme tedy sérii datagramů ICMP "Echo request", kde první datagram má pole TTL=1, druhý TTL=2, atd. Tím se nám vrací datagramy ICMP type 11 od všech směrovačů po cestě "od nás" k cíli. Dosažení cíle je indikováno návratem ICMP "Echo reply".
    - Existují varianty traceroute užívající i jiných protokolů, ale princip s proměnným TTL je týž.
- ICMP se užívá i pro zjištění lokálního směrovače
    - Stroj na lokální síti vyšle ICMP 10 (Router discovery) s cílovou adresou 0.0.0.0 (*broadcast*) a směrovač odpoví ICMP 9 (Router advertisement)

To je dnes vše.

Otázky?