

Počítačové sítě

Jan Outrata



KATEDRA INFORMATIKY
UNIVERZITA PALACKÉHO V OLMOUCI

přednášky

Tyto slajdy byly jako výukové a studijní materiály vytvořeny za podpory grantu FRVŠ 1358/2010/F1a.



Síťová vrstva



■ linkové protokoly

- vyměňují data mezi **sousedními uzly** v rámci lokální nebo rozlehlé sítě, pomocí sdíleného komunikačního média (např. Ethernet) nebo dvoubodovými linkami (např. PPP)
- standardizované normami IEEE 802.x

■ síťové protokoly

- vyměňují data mezi **libovolnými (nesousedními) uzly** v rozlehlé síti tvořené mnoha lokálními sítěmi
- data jsou **směrována (routing)** rozlehlou sítí pomocí směrovačů – nejdůležitější funkce síťové vrstvy (protokolu)
- dříve různá řešení (TCP/IP, ISO OSI, firemní), dnes de facto standard TCP/IP

■ směrovač (router):

- propojuje lokální sítě (LAN) na úrovni síťové vrstvy, umožňuje libovolné topologie sítě (v praxi propojené hvězdicové)
- řeší směrování z lokální sítě k následujícímu směrovači nebo koncovému uzlu (**next hop**), rozhoduje na základě svých **směrovacích tabulek**
- = běžný počítač nebo specializované zařízení (směrovač, router) s **více síťovými rozhraními**, předávající si data mezi rozhraními – **forwarding**
- „vybaluje“ data (síťový paket) z linkového rámce a „zabaluje“ do jiného linkového rámce – i když jsou linkové protokoly sítí stejné!
- nemění síťový paket (např. adresy)!, až na výjimky, např. položka TTL, fragmentace, volitelné položky aj.

■ koncové uzly – vysílají a přijímají síťové pakety „zabalené“ do linkových rámců



- fyzická a linková vrstva implementována na síťové kartě (HW) a jejím ovladačem (driver, SW)
- **rozhraní ovladače** – standardizovaný způsob přístupu ze síťové vrstvy k linkové, funkce:
 - výběr linkového protokolu (rámce)
 - identifikace a přepínání síťového protokolu (buffer, např. SSAP, DSAP)
 - „zabalování“ síťových paketů a „rozbalování“ linkových rámců
 - služby podvrstvy LLC (správa linkových spojů)
- standardizovaná rozhraní: PKDRV (Packet Driver, pro TCP/IP), **NDIS** (Network Driver Interface Specification, Microsoft/IBM, vyžaduje linkový protokolový ovladač, kterému NDIS ovladač předává rámce)



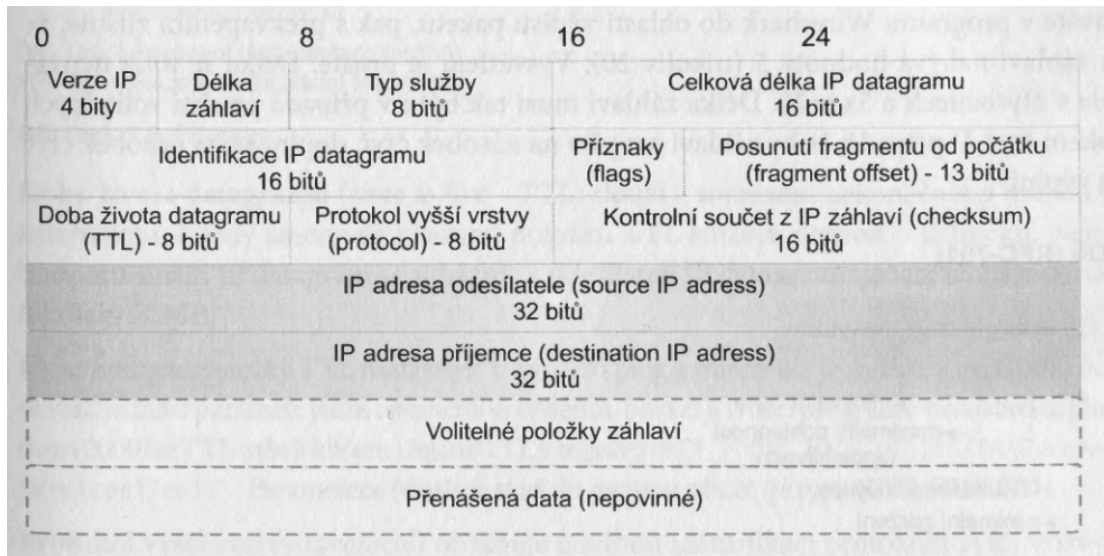
- 1980 RFC-760, 1981 RFC-791
- poskytuje „**nespolehlivou**“ **nespojovanou službu** – nevytváří spojení, nepotvrzuje příjem paketů
- spojuje lokální sítě do celosvětové sítě **Internet**
- tvořen několika dílčími protokoly: vlastní IP a služební **ICMP** (diagnostika a signalizace mimořádných stavů), **IGMP** (skupinové adresování), **ARP** a **RARP** (zjištění linkové adresy k IP adrese a opačně)
- síťové rozhraní uzlu má alespoň jednu **síťovou IP adresu**

- základní jednotka dat přenášených IP
- = záhlaví 20 B povinných položek + volitelné položky, data, max. délka 64kB

Obrázek: Obrázek průvodce 131

- délka záhlaví: v jednotkách 4 B, tzn. max. 60 B
- **typ služby (TOS)**: původně specifikace kvality přenosu (bity pro prioritu, min. zdržení a cena, max. výkon a dostupnost), dnes **DS (Differentiated Services)** – požadavky garance šířky pásma, protokol RSVP
- identifikace, příznaky a posunutí **fragmentu**: pro účely **fragmentace paketu**, bity příznaků pro zakázání fragmentace (DF) a indikaci dalších fragmentů (MF, tento není poslední)

IP paket (datagram)





- **doba života (TTL):** zamezení nekonečného „toulání“ paketu, každý směrovač snižuje alespoň o 1 (a musí tedy změnit kontrolní součet záhlaví), při 0 se paket zahazuje a odesílateli je to signalizováno protokolem ICMP, nastavena v OS
- **protokol vyšší vrstvy:** čísla přiděluje IANA, např. ICMP 1, IGMP 2, IP 4, TCP 6, UDP 17, tunelování protokolů, např. IP over IP (privátní sítě, IPv6 over IPv4), IPX over IP

CVIČENÍ: zachytávání a inspekce IP paketů



- každé síťové rozhraní počítače (síťová karta) může mít jednu nebo více **jednoznačných** IP adres
- přidělení adresy síťovému rozhraní staticky pomocí programu `ipconfig` (MS Windows) nebo `ifconfig/ip` (UNIX, GNU/Linux)

CVIČENÍ: zjištění IP adresy síťového rozhraní a jeho změna

- = číslo délky **4 B** (pro protokol IPv4), notace zápisu s hodnotami bytů v desítkové soustavě oddělenými tečkou, např. **158.194.80.13** =
10011110.11000010.01010000.00001101

= dvě části adresy: **adresa sítě** a **adresa uzlu (rozhraní)** v síti

Classfull adresace

- od počátku Internetu až do roku 1993: RFC 796
- jaká část pro síť určují počáteční bity prvního bytu, dělení sítí do 5 (základních) **tříd**:
 - třída **A**: adresa začíná (bitem) 0, 1 byte pro síť, 126 sítí (s hodnotami prvního bytu) 1 až 126 (0 a 127 mají zvláštní význam), $2^{24} - 2$ uzlů (0 a 255 mají zvláštní význam)
 - třída **B**: začíná 10, 2 byty pro síť, 2^{14} sítí 128 až 191, $2^{16} - 2$ uzlů
 - třída **C**: začíná 110, 3 byty pro síť, 2^{21} sítí 192 až 223, 254 uzlů
 - třída **D**: začíná 1110, nedělí se, 2^{28} skupinových adres 224.0.0.0 až 239.255.255.255 (**IP multicast**, RFC 1112)
 - třída **E** (a další): začíná 1111, 2^{28} adres 240.0.0.0 až 255.255.255.254 původně rezervovaných pro speciální a experimentální účely, dnes již také přidělené

Speciální adresy

- celá = 0: tento uzel (= loopback, bez přidělené adresy)
- uzel = 0: **adresa sítě**
- síť = 0: uzel na této síti (nepoužívá se)
- uzel samé 1: **všesměrová adresa sítě (network broadcast)**
- samé 1 (255.255.255.255): **všesměrová adresa lokální sítě (local broadcast)**, nesměruje se
- 127.cokoliv: programová (lokální, SW) **smyčka (loopback)**, typicky **127.0.0.1**, odeslaný paket „ihned přijde“

CVIČENÍ: zjištění všech uzlů na lokální síti pomocí programu ping

Classless adresace

- od roku 1993: RFC 1517–1520
- sítě se nerozlišují podle tříd, ale podle **síťové masky**:
 - = 4B číslo (notace IP adres), bity = 1 určují v IP adrese adresu sítě
 - určení adresy sítě: bitový součin IP adresy a síťové masky
 - počet uzlů v síti = $2^{(\text{počet } 0 \text{ v masce})} - 2$
 - masky odpovídající třídám adres = **standardní** síťové masky, pro třídu A **255.0.0.0**, pro třídu B **255.255.0.0**, pro třídu C **255.255.255.0**
 - **notace sítě spolu s maskou**: adresa sítě/maska, např. 158.194.0.0/255.255.0.0
 - v binárním vyjádření ji tvoří (de facto) zleva souvislá řada 1 → notace adresa sítě/počet 1 v masce, tzv. **CIDR formát** (Classless Inter-Domain Routing), např. 158.194.0.0/16

Subsítě

- = **část sítě určená maskou**: část adresy pro uzel rozdělena na část pro subsítě a pro uzel, síťová maska pokrývá část adresy pro síť i subsítě
- výjimka: síť s maskou /32 je adresou samostatného uzlu
- např. síť 158.194.0.0/16 může být rozdělena např. do 256 subsítí s adresami 158.194.0.0/24 až 158.194.255.0/24
- ! **nejednoznačnosti**: subsítě samé 0 (adresa uzlu samé 0) – adresa subsítě nebo celé sítě?, subsítě samé 1 (adresa uzlu samé 1) – všesměrová adresa subsítě nebo celé sítě (tj. všech subsítí)? → nepoužívají se
- síť může být na subsítě rozdělena pomocí **konstantní síťové masky** (všechny subsítě mají stejnou, viz příklad výše) nebo **variabilní síťové masky** (subsítě mají různou masku, např. 158.194.1.0/30, 158.194.80.0/20, 158.194.92.0/22) – POZOR na adresy subsítí při daných maskách (korektní, síť rozsahem adres disjunktní nebo jedna celá částí druhé)!
- subsítě je možné opět pomocí „prodloužení“ masky opakovaně rozdělit do (sub)subsítí

Supersítě a autonomní systémy

- **supersítě** – síťová maska nepokrývá celou adresu sítě, duální k subsíti
 - použití pro **agregaci adres sítí** – výhodné pro směrování, administrativu přidělování adres apod. např. na tzv. peeringových uzlech (síťových center vzájemně propojujících sítě poskytovatelů internetového připojení a poskytovatelů obsahu a služeb)
 - např. síť 158.194.92.0/24 je součástí supersítě 158.194.0.0/16
- z hlediska dopravy IP paketů (směrování) se Internet dělí na tzv. **autonomní systémy (AS)** = supersítě spravované největšími poskytovateli internetového připojení, bloky IP adres v rámci AS přidělují regionální a lokální **Internet Registry**
 - např. síť 158.194.0.0/16 (UPOL-TCZ) je součástí autonomního systému AS2852 (CESNET2), který je součástí bloku AS2830 – AS2879 patřícího **RIPE NCC** (regionální Internet Registry pro Evropu a přidružené země)
 - přidělené bloky adres pro (super)sítě a autonomní systémy a informace o nich lze zjišťovat programem `whois`, např. `whois 158.194.80.13`, `whois AS2852`

- **Intranet** = lokální síť (pro informační systém), obvykle uzavřená nebo s omezením provozu z vnější sítě dovnitř, příp. i ven
- v síti (Internetu) musí být IP adresy jednoznačné, v lokální síti:
 - libovolné adresy (jednoznačné v rámci lokální sítě) a **NAT (Network Address Translation)** = **překlad adres** lokální sítě na adresy ve vnější síti a naopak – typicky na rozhraní směrovače do vnější sítě, zvláštní případ tzv. **maškaráda** = překlad na 1 adresu (směrovače)
 - **vyhrazené rozsahy IP adres** pro uzavřené podnikové sítě (RFC1918): **10.0.0.0/8** (třída A), **172.16.0.0/12** (třída B), **192.168.0.0/16** (třída C) – použití dle libosti, **nesměřují se**
 - v praxi vyhrazený rozsah + NAT
- propojení dvou a více lokálních sítí:
 - směrovačem – jednoduché, nevýhoda komunikace přes směrovač
 - adresami sousední tvořící supersítě s kratší maskou, např. /23 pro /24
- **nečíslovaná síť**: „síť“ propojující dva směrovače (např. pomocí sériových linek), tvořící jeden „virtuální“ směrovač



Dynamické přidělování IP adres

- = oproti pevnému (statickému) podle potřeby při připojení uzlu do sítě
 - v lokální síti dnes aplikační protokol **DHCP** nahrazující dřívější protokoly RARP a BOOTP
 - v rozlehlé síti (typicky komutované telefonní) linkový protokol **PPP**

- = odesílání paketů na (následující) směrovač nebo cílový uzel (next hop), popř. do lokální sítě s cílovým uzlem
- **předávání (forwarding)** = předávání paketů v rámci směrovače mezi jeho síťovými rozhraními, na směrovači základ procesu směrování
 - umožňuje uzlu pracovat jako **směrovač** = pokud paket není adresován jemu, odešle (předá) ho dále (jiným rozhráním), stejně jako vlastní odchozí pakety
 - lze v OS povolit/zakázat za běhu, u MS Windows hodnota 1/0 v klíči IpEnableRouter v registru, u GNU/Linuxu v souboru `/proc/sys/net/ipv4/ip_forward`
 - děje se (zpravidla) bez vědomí vyšších vrstev, např. aplikační, konfiguruje se parametry (jádra) OS

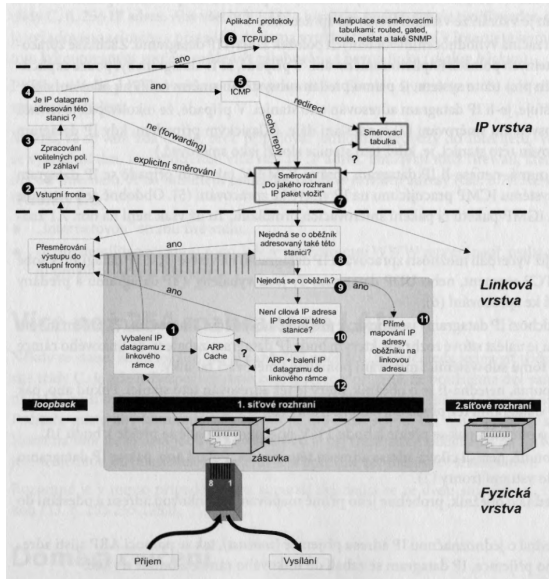
- děje se na **každém uzlu zapojeném do sítě**, nejen směrovačích! – odesílání paketů do (lokální) sítě nebo mimo ni (skrze směrovač)

Obrázek: Obrázek průvodce 186

Filtrace

- pakety mohou být před směrováním **filtrovány** – nastavením filtračních pravidel OS nebo pomocí aplikačního programu, na základě IP záhlaví (adres), TCP/UDP záhlaví (portů, příznaků) nebo aplikačního protokolu
- filtrace se často provádí (a doporučuje) i u koncových uzlů při příjmu paketů – posílení jejich ochrany a bezpečnosti
- filtrace bývá významnou funkcí tzv. **firewallů** – programů či stanic (směrovačů) chránících systém uzlu nebo (lokální) síť před útoky

Směrování (routing)



Směrovací tabulky

- pro odchozí paket se musí **rozhodnout**, kterým síťovým rozhraním jej odeslat dále (next hop)
- rozhoduje se pomocí **směrovací tabulky** se směry (cestami, **route**), např. (typická pro koncový uzel):

síť/uzel	maska	next hop (gateway)	rozhraní	metrika, vlajky aj.
158.194.92.0	255.255.255.0	0.0.0.0	Ethernet 1	...
158.194.80.0	255.255.255.0	158.194.80.1	Ethernet 2	...
(127.0.0.0	255.0.0.0	127.0.0.1	loopback	...)
10.0.0.0	255.255.0.0	0.0.0.0	Virtual Eth.	...
...
0.0.0.0	0.0.0.0	158.194.254.66	Ethernet 3	...

- **setříděna sestupně podle adresy sítě** (1. sloupec) – více specifická (s delší maskou) má přednost před obecnější v případě stejných směrů pro paket

Směrovací tabulky

■ rozhodování:

- 1 průchod tabulkou odshora dolů, log. **vynásobení cílové adresy paketu s maskou** v tabulce (2. sloupec)
- 2 pokud se výsledek **rovná adrese sítě**, popř. uzlu (maska samé 1) v tabulce (1. sloupec), paket se odešle skrze rozhraní (4. sloupec) na následující směrovač nebo cílový uzel (**next hop**, 3. sloupec), popř. do lokální sítě s cílovým uzlem (next hop = 0.0.0.0, tzv. **přímé směrování**), jinak další řádek
- 3 poslední řádek (adresa sítě i maska = 0.0.0.0) = **výchozí (implicitní) směr** pro paket nevyhovující žádnému předchozímu záznamu (žádné síti), typicky směr do Internetu
 - agregace záznamů tabulky u supersítí a autonomních systémů

Směrovací tabulky

- udržovány na každém uzlu v síti, směrovači i koncovém – odesílají pakety
- **naplnění tabulky** = „skutečný“ problém směrování:
 - staticky (**statické směrování**) ručně, automaticky při konfiguraci síťového rozhraní OS (nejčastější) nebo pomocí managementu sítě (např. aplikační protokol SNMP)
 - dynamicky (**dynamické směrování**) z ICMP zpráv (změny směrování) nebo **směrovacími aplikačními protokoly**
- výpis tabulky pomocí programu netstat, výpis a (statická) editace správcem OS pomocí programů route/ip (UNIX, GNU/Linux), "Směrování a vzdálený přístup" (MS Windows Server), ip route (CISCO) apod.

CVIČENÍ: výpis a editace směrovací tabulky (např. výmaz a vrácení směru default) programy netstat, route, ip apod.

- aplikační protokoly k vytvoření směrů, tj. k **dynamické aktualizaci směrovacích tabulek** směrovačů, NE k vlastnímu procesu směrování
- dělení: IGP (v rámci AS) a EGP (výměna směrovacích informací mezi AS, směrovací politiky), RVP a LSP (podle použitého směrovacího algoritmu)

RVP (Routing Vector Protocols)

- = algoritmus **DVA (Distance Vector Algorithm), Bellman-Fordův**: směrovač opakovaně **odešle** svou **směrovací tabulku** sousedním a z přijatých tabulek si do své dočasně (2-5 minut) doplní záznamy (vektory) pro neznámé sítě nebo s **menší vzdáleností** (metrikou, počet směrovačů na cestě) s navýšenou metrikou (typicky o 1), konec při max. metrice (např. 16) = nedostupná síť
- jednoduché, ale při výpadku připojení směrovače do sítě nebo v rozlehlějších sítích (při vyšší max. metrice) mohou tabulky **oscilovat** → nedoplňovat záznamy, které směrovač sám dříve odeslal
- např. RIP (pouze pro standardní masky), **RIP 2** (multicast 224.0.0.9), **RIPng** (pro IPv6), **IGRP**, **BGP**, program routed

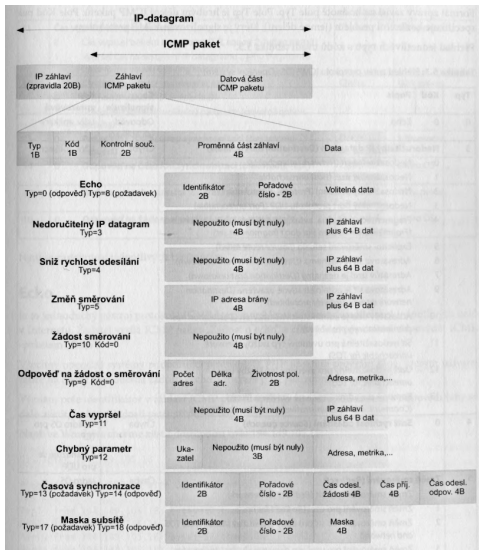
LSP (Link State Protocols)

- = algoritmus **LSA (Link State Algorithm)**: směrovač opakovaně ohodnotí (metrika) cesty k sousedním (např. podle odezvy) a jejich seznam spolu s ohodnoceními rozešle do celé rozlehlé sítě, ze získané **topologie celé sítě** si pak (dočasně) doplní/upraví záznamy v tabulce pro síť na základě nejkratších cest vypočtených **algoritmem nalezení nejkratších cest v grafu (SPF, Shortest Path First, Dijkstrův)**
- rozdělení rozlehlejších sítí na **oblasti (směrovací domény)**, z více směrovačů na jedné síti se vybere jeden
- oproti RVP méně dat, **stabilnější**, pružnější, ale složitější konfigurace
- např. **OSPF** (páteřní oblast, autentizace, IPv6 aj.), IS-IS, **EGP**, program gated

- Internet Control Message Protocol, RFC 777
- = služební protokol IP pro **diagnostiku a signalizaci mimořádných** (chybových) **stavů**
- OS většinou nepodporují všechny **zprávy**, směrovače mohou z bezpečnostních důvodů nějaké zahazovat
- **ICMP pakety** obsaženy v paketech IP, záhlaví (8B): typ (1B), kód (1B), kontrolní součet (2B) a proměnná část (4B), a data

Obrázek: Obrázek průvodce 135

CVIČENÍ: zachytávání a inspekce ICMP paketů generovaných programem ping nebo traceroute/tracert





Echo

- typ 8 (žádost, request) a 0 (odpověď, reply), kód 0
- použití pro **testování dosažitelnosti uzlu** pomocí programu ping – měří a vypisuje i čas mezi žádostí a odpovědí, tj. čas k uzlu a zpět (**Round Trip Time, RTT**), a použité TTL
- pole Identifikátor (v proměnné části záhlaví) pro spárování žádosti a odpovědi

CVIČENÍ: zjištění vzdálenosti (počtu směrovačů, hopů) uzlu pomocí programu ping se změnou TTL

Čas vypršel (Time exceeded)

- typ 11, kód 0 (TTL = 0 a IP paket bude zahozen) a 1 (IP paket nelze v určeném čase sestavit z fragmentů)
- zahozený/částečný IP paket (prvních 64 B) v datové části ICMP paketu
- použití (kód 0) pro **zjištění cesty** (směrovačů) **k uzlu** pomocí programu traceroute/tracert:
 - 1 na cílový uzel odeslána ICMP žádost Echo nebo UDP datagram (traceroute, port lze nastavit) s TTL = 1
 - 2 první směrovač na cestě signalizuje zahození paketu (sníží TTL na 0)
 - 3 získání adresy směrovače a změření času od odeslání k přijetí signalizace (čas ke směrovači a zpět, RTT), výpis obojího
 - 4 toto třikrát, pak s TTL = 2 (zahodí druhý směrovač) atd. až do přijetí ICMP odpovědi Echo nebo signalizace nedoručitelného IP paketu (typ 3) od cílového uzlu

CVIČENÍ: zjištění cesty (směrovačů) k uzlu pomocí programu traceroute/tracert, zjištění autonomních systémů na cestě pomocí programu whois

Nedoručitelný IP paket (Destination unreachable)

- typ 3, signalizace odesílateli, pokud paket nemůže být předán dál nebo doručen a je zahozen
- zahozený IP paket (prvních 64 B) v datové části ICMP paketu
- **důvody** (kódy): nedosažitelná síť (0), uzel (1), protokol (2), UDP port (3), fragmentace zakázána, ale nutná pro další přenos (4), neznámá adresátova síť (6), uzel (7) atd.

Další

- **sniž rychlost odesílání** (typ 4, kód 0) – odesílateli signalizuje směrovač, který není schopen IP paket předat dál (je zahlcený)
- **změň směrování** (typ 5, kódy 0-3), **žádost+odpověď o směrování** (typy 9, 10, kód 0) – doporučení změny ve směrovací tabulce odesílatele (pro tento směr) nebo zjištění směrovačů (žádost na všeobecnou adresu, směrovače odpoví)
- ... (mnoho)